Weak multiplicativity for random quantum channels

Ashley Montanaro[*]

August 28, 2012

Abstract

It is known that random quantum channels exhibit significant violations of multiplicativity of maximum output p-norms for any $p > 1$. In this work, we show that a weaker variant of multiplicativity nevertheless holds for these channels. For any constant $p > 1$, given a random quantum channel $\mathcal{N}$ (i.e. a channel whose Stinespring representation corresponds to a random subspace S), we show that with high probability the maximum output p-norm of $\mathcal{N}^{\otimes n}$ decays exponentially with n. The proof is based on relaxing the maximum output ∞-norm of $\mathcal{N}$ to the operator norm of the partial transpose of the projector onto S, then calculating upper bounds on this quantity using ideas from random matrix theory.

1 Introduction

For many years, some of the most vexatious open problems of quantum information theory have concerned maximum output p-norms of quantum channels. If $\mathcal{N}$ is a quantum channel (i.e. completely positive, trace-preserving map), the maximum output p-norm of $\mathcal{N}$ is defined as

$$\|\mathcal{N}\|_{1 \to p} := \max\{\|\mathcal{N}(\rho)\|_p, \ \rho \geq 0, \ \operatorname{tr} \rho = 1\},$$

where $\|X\|_p := (\operatorname{tr}|X|^p)^{1/p}$ is the Schatten p-norm. (The notation $\nu_p(\mathcal{N})$ is also used for $\|\mathcal{N}\|_{1 \to p}$. Technically, $\|\mathcal{N}\|_{1 \to p}$ is actually defined as $\sup\{\|\mathcal{N}(X)\|_p / \|X\|_1\}$, where the supremum is taken over all non-zero Hermitian operators X, but it can be shown that this definition is equivalent [2].) It was a long-standing conjecture in quantum information theory [3] that, for any two quantum channels $\mathcal{N}_1, \mathcal{N}_2$,

$$\|\mathcal{N}_1 \otimes \mathcal{N}_2\|_{1 \to p} \stackrel{?}{=} \|\mathcal{N}_1\|_{1 \to p}\|\mathcal{N}_2\|_{1 \to p},$$

at least for p fairly close to 1. This is equivalent to the question of additivity of *minimum* output Rényi p-entropies, which are defined in terms of maximum output p-norms as

$$H_p^{\min}(\mathcal{N}) := \frac{1}{1 - p} \log \|\mathcal{N}\|_{1 \to p}^p.$$

The minimum output (von Neumann) entropy $H^{\min}(\mathcal{N})$ is obtained by taking the limit $p \to 1$ [3]. This case of the additivity question was of particular interest due to its connections with many other additivity problems in quantum information theory [34].

---

[*]Centre for Quantum Information and Quantum Foundations, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, UK; `am994@cam.ac.uk`.

All of these multiplicativity/additivity conjectures are now known to be false. First, Werner and Holevo found a counterexample to multiplicativity for $p > 4.79$ [35]. Some years later, the conjecture was falsified in the range $p > 2$ by Winter [36], which was swiftly extended to $1 < p < 2$ by Hayden [26]. These works were combined as [27], which also includes the remaining case $p = 2$. One can generalise the conjectures to $p < 1$ (where $\|\cdot\|_p$ is of course no longer a norm), and in this setting Cubitt et al [21] falsified additivity of minimum output Rényi p-entropies for $p \approx 0$. Finally, Hastings showed that the minimum output entropy is not additive [25]. Following this, Aubrun, Szarek and Werner showed that the results of Hayden, Winter and Hastings can be obtained from Dvoretzky's theorem in the language of asymptotic geometric analysis [5, 6].

As well as the limit $p \to 1$, another important special case of the multiplicativity question is $p = \infty$, which turns out to be closely related to a number of other quantities studied in quantum information theory, as we now discuss. Any quantum channel performing a map from a d_A-dimensional quantum system A to a d_B-dimensional quantum system B can be written as $\mathcal{N}(\rho) = \mathrm{tr}_E V \rho V^\dagger$ for some isometry $V : \mathbb{C}^{d_A} \to \mathbb{C}^{d_B} \otimes \mathbb{C}^{d_E}$ (a form known as the Stinespring dilation). The operator $M = VV^\dagger$ projects onto a subspace $S \subseteq \mathbb{C}^{d_B} \otimes \mathbb{C}^{d_E}$. For our purposes, we can simply identify $\mathcal{N}$ with either V, S or M.

Let SEP $\subset \mathcal{B}(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B})$ be the set of $d_A \times d_B$-dimensional separable quantum states. For any operator $M \in \mathcal{B}(\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B})$ such that $0 \leq M \leq I$, the quantity

$$h_{\mathrm{SEP}}(M) := \max_{\rho \in \mathrm{SEP}} \mathrm{tr}\, M\rho$$

is known as the support function of the separable states, evaluated at M. This quantity has the following connection to maximum output p-norms:

Fact 1. *Let $\mathcal{N}$ be a quantum channel with corresponding isometry V, and set $M = VV^\dagger$. Then*

$$h_{\mathrm{SEP}}(M) = \|\mathcal{N}\|_{1 \to \infty}.$$

This fact can easily be proven using the Schmidt decomposition, and indeed can be generalised to arbitrary operators $0 \leq M \leq I$ [29] (see [24] for a proof). The quantity h_{SEP} is crucially important in the study of multiple-prover quantum Merlin-Arthur games [28, 24], which we now briefly discuss. The complexity class QMA(2) is informally defined as the class of decision problems which can be solved by a polynomial-time quantum verifier (Arthur) given access to two *unentangled* quantum states (or "proofs") produced by two all-powerful but potentially malicious provers (Merlin A and Merlin B). Consider an instance of a QMA(2) problem for which Arthur should output "no". If M denotes Arthur's measurement operator corresponding to a "yes" outcome, the maximal probability with which the two Merlins can convince Arthur to (incorrectly) output "yes" is precisely $h_{\mathrm{SEP}}(M)$. Therefore, projectors M such that $h_{\mathrm{SEP}}(M^{\otimes n}) = h_{\mathrm{SEP}}(M)^n$ correspond to measurement operators occurring in two-prover quantum Merlin-Arthur games which obey *perfect parallel repetition*, i.e. where Arthur can simply repeat the protocol n times in parallel to reduce a failure probability of s to a failure probability of s^n. The failure of multiplicativity for $\|\mathcal{N}\|_{1 \to \infty}$ implies that such a precise form of parallel repetition cannot hold in general; however, it could still be the case that a weaker form of parallel repetition holds, where $h_{\mathrm{SEP}}(M^{\otimes n})$ necessarily decreases exponentially with n.

h_{SEP} also turns out to have many other connections to important quantities in quantum information theory and tensor optimisation (see [24] for some examples).

1.1 Counterexamples to multiplicativity

The known counterexamples to multiplicativity fall into two classes: explicit and randomised. The known explicit counterexamples are a channel of Werner and Holevo [35], which acts on d-dimensional square matrices ρ by the map

$$\rho \mapsto \frac{1}{d-1} \left((\operatorname{tr} \rho) I - \rho^T \right),$$

and the channel whose corresponding subspace in the Stinespring form is the antisymmetric subspace of $\mathbb{C}^d \otimes \mathbb{C}^d$ [22]. This latter channel violates multiplicativity for all $p > 2$, and is of particular interest due both to its simplicity and the fact that it displays a very strong violation of multiplicativity when $p = \infty$. Indeed, if P_{anti} denotes the projector onto the antisymmetric subspace, it holds that

$$h_{\text{SEP}}(P_{\text{anti}}) = \frac{1}{2}, \text{ but } h_{\text{SEP}}(P_{\text{anti}}^{\otimes 2}) \geq \frac{1}{2} \left(1 - \frac{1}{d} \right).$$

The counterexamples of Hayden and Winter [36, 26, 27] are random constructions. In particular, the construction used to falsify p-norm multiplicativity for all $p > 1$ is to choose the first channel $\mathcal{N}$'s corresponding subspace $S \subset \mathbb{C}^d \otimes \mathbb{C}^d$ at random from the set of all subspaces of dimension $r = O(d^{1+1/p})$ (i.e. according to Haar measure on the unitary group), and to take $\bar{\mathcal{N}}$ as the second channel.

In the case $p = \infty$, the violation of multiplicativity displayed by this construction is also near-maximal. In other words,

$$\| \mathcal{N} \otimes \bar{\mathcal{N}} \|_{1 \to \infty} \approx \| \mathcal{N} \|_{1 \to \infty}.$$

(One can show that, for any channels $\mathcal{N}_1$ and $\mathcal{N}_2$, $\| \mathcal{N}_1 \otimes \mathcal{N}_2 \|_{1 \to p} \leq \| \mathcal{N}_1 \|_{1 \to p}$ [3].) While the example of the antisymmetric subspace implies that there exists a channel $\mathcal{N}$ such that one can achieve a much larger output p-norm by using an entangled state as input to $\mathcal{N}^{\otimes 2}$ than is possible using only product states, it leaves open the question of the general behaviour of $\| \mathcal{N}^{\otimes n} \|_{1 \to p}$ for larger n. To the author's knowledge, two extreme situations are still possibilities: on the one hand, it might hold that

$$\| \mathcal{N}^{\otimes n} \|_{1 \to p} \overset{?}{\leq} \| \mathcal{N} \|_{1 \to p}^{n/2}$$

for all $\mathcal{N}$; alternatively, there might be no universal constant α such that, for all channels $\mathcal{N}$,

$$\| \mathcal{N}^{\otimes n} \|_{1 \to p} \leq \| \mathcal{N} \|_{1 \to p}^{\alpha n}.$$

The former possibility would imply that the largest possible violation of multiplicativity is quite mild, and in the case $p = \infty$ that a form of parallel repetition holds for two-prover quantum Merlin-Arthur games; the latter would mean that severe violations are possible and parallel repetition fails.

Interestingly, for the antisymmetric subspace it turns out that the former possibility is closer to the truth. In a beautiful and technical recent work, Christandl, Schuch and Winter [10, 11] have shown that there is a constant $C > 0$ such that

$$h_{\text{SEP}}(P_{\text{anti}}^{\otimes n}) \leq 2^{-Cn},$$

implying that a weak variant of multiplicativity does indeed hold for this channel.

1.2 New results

The main result of this work is that, even though random quantum channels do not obey multiplicativity, their violations of multiplicativity are in some sense also very weak.

Definition 2. *A quantum channel $\mathcal{N}$ obeys weak p-norm multiplicativity with exponent α if, for all $n \geq 1$,*

$$\|\mathcal{N}^{\otimes n}\|_{1\to p} \leq \|\mathcal{N}\|_{1\to p}^{\alpha n}.$$

Observe that, for any $p > 1$,

$$\|\mathcal{N}^{\otimes n}\|_{1\to p} \leq \|\mathcal{N}^{\otimes n}\|_{1\to\infty}^{1-1/p},$$

which follows from the (matrix) Hölder inequality $\|X\|_p^p \leq \|X\|_1 \|X\|_\infty^{p-1}$ for any X. So, if $\mathcal{N}$ obeys weak ∞-norm multiplicativity with exponent α, $\mathcal{N}$ also obeys weak p-norm multiplicativity for any $p > 1$, with exponent $\alpha(1 - 1/p)$.

Our main result can be summarised informally as follows (a more technical version is given as Theorem 9 below).

Theorem 3. *Let $\mathcal{N}$ be a quantum channel whose corresponding subspace is a random dimension r subspace of $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$, set $m := \min\{r, d_A, d_B\}$, and assume $m \geq 2(\log_2 \max\{d_A, d_B\})^{3/2}$ and $r = o(d_A d_B)$. Then the probability that $\mathcal{N}$ does* not *obey weak ∞-norm multiplicativity with exponent $1/2 - o(1)$ (if $r \geq d_B/d_A$), or $1 - o(1)$ (if $r \leq d_B/d_A$) is exponentially small in m.*

Note that we have switched notation slightly from $\mathbb{C}^{d_B} \otimes \mathbb{C}^{d_E}$ to $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$, and will use this notation henceforth. In Theorem 3, and throughout the rest of the paper, our notion of a random dimension r subspace $S \subseteq \mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$ is as follows: the projector M onto S is formed by starting with the projector M_0 onto an arbitrary fixed dimension r subspace $S_0 \subseteq \mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$, and taking $M = U M_0 U^\dagger$, where U is a random unitary operator (i.e. U is picked according to the Haar measure on the unitary group $U(d_A d_B)$). We usually think of r, d_A and d_B as all growing, but not necessarily at the same rate.

Theorem 3 implies that, for random channels $\mathcal{N}$ satisfying some mild dimensionality constraints, with high probability

$$\|\mathcal{N}^{\otimes n}\|_{1\to p} \leq \|\mathcal{N}\|_{1\to p}^{(1/2-o(1))(1-1/p)n},$$

so random channels obey weak p-norm multiplicativity with exponent $(1/2 - o(1))(1 - 1/p)$. Note that the results of Hayden and Winter imply that when $r \approx d_A = d_B$, with high probability

$$\|\mathcal{N} \otimes \bar{\mathcal{N}}\|_{1\to\infty} \approx \|\mathcal{N}\|_{1\to\infty}.$$

It was also observed by Hastings [25] that picking the subspace corresponding to $\mathcal{N}$ by replacing the random unitary U with an operator picked according to Haar measure on the orthogonal group $O(d_A d_B)$ gives a channel such that, when $r \approx d_A = d_B$,

$$\|\mathcal{N}^{\otimes 2}\|_{1\to\infty} \approx \|\mathcal{N}\|_{1\to\infty}.$$

Thus Theorem 3 is perhaps essentially the strongest multiplicativity result one could expect for the maximum output ∞-norm of random quantum channels (although note that it is not in fact known whether multiplicativity is violated by two copies of a random channel $\mathcal{N}$).

We remark that our results also apply without change to the setting where, instead of taking n copies of a fixed random channel $\mathcal{N}$, one takes n independently picked random channels $\mathcal{N}_1, \ldots, \mathcal{N}_n$ (where n is fixed) and compares $\|\mathcal{N}_1 \otimes \cdots \otimes \mathcal{N}_n\|_{1\to p}$ with $(\|\mathcal{N}_1\|_{1\to p} \ldots \|\mathcal{N}_n\|_{1\to p})^\alpha$.

1.3 The case of the von Neumann entropy

In certain regimes, our results also imply a weak *additivity* result for the von Neumann entropy[1]. Recall that for a quantum channel $\mathcal{N}$, the minimum output Rényi ∞-entropy of $\mathcal{N}$ is defined by taking the limit as $H_\infty^{\min}(\mathcal{N}) = -\log \|\mathcal{N}\|_{1\to\infty}$. A corollary of Theorem 8 below is that there is a universal constant C such that the probability that the regularised minimum output ∞-entropy of a random quantum channel $\mathcal{N}$ does *not* satisfy

$$\frac{1}{n} H_\infty^{\min}(\mathcal{N}^{\otimes n}) \geq \begin{cases} \frac{1}{2}\left(\log d_A + \log d_B - \log r\right) - C & \text{if } r \geq d_B/d_A \\ \log d_A - C & \text{if } r \leq d_B/d_A \end{cases}$$

is exponentially small in $m := \min\{r, d_A, d_B\}$, assuming that $m \geq 2(\log_2 \max\{d_A, d_B\})^{3/2}$. As $\frac{1}{n} H^{\min}(\mathcal{N}^{\otimes n}) \geq \frac{1}{n} H_\infty^{\min}(\mathcal{N}^{\otimes n})$ for all channels $\mathcal{N}$ by monotonicity of Rényi entropies, any upper bound on $H^{\min}(\mathcal{N})$ implies a limitation on the extent of possible additivity violations for $H^{\min}(\mathcal{N})$. In particular, as $H^{\min}(\mathcal{N})$ is always upper bounded by $\log d_A$, we obtain a weak additivity result for the minimum output von Neumann entropy when r is not too large with respect to d_B. For example, if $r = d_A = d_B$, we obtain that

$$\frac{1}{n} H^{\min}(\mathcal{N}^{\otimes n}) \geq \frac{1}{2} H^{\min}(\mathcal{N}) - C$$

with high probability. If $r \leq d_B/d_A$, we get the stronger result that with high probability

$$\frac{1}{n} H^{\min}(\mathcal{N}^{\otimes n}) \geq H^{\min}(\mathcal{N}) - C.$$

1.4 Organisation and proof strategy

The remainder of this paper is devoted to the proof of Theorem 3. Conceptually, the proof is simple: we find a general upper bound on $h_{\mathrm{SEP}}(M)$ (for arbitrary M) which is multiplicative, and then show that this upper bound is not too far from the truth when M corresponds to a random quantum channel. The upper bound we use is the operator norm of the partial transpose of M, $\|M^\Gamma\|_\infty$, and the main technical contribution of this paper is to prove tail bounds on this quantity, which is carried out using the method of moments from random matrix theory. Note that we have not attempted to optimise the constants and lower order terms which occur in our bounds, which could probably be substantially improved.

The next section discusses the proof strategy and some related work, and states a more formal version of Theorem 3. The following section contains the proofs of the main technical results and lemmas. The paper finishes in Section 4 with some conclusions and suggestions for future work.

2 Maximum overlap with separable states and PPT

Our proof of Theorem 3 will be based on a general upper bound strategy for $h_{\mathrm{SEP}}(M)$. Maximising over the set of separable states is a daunting task, and a useful relaxation is to maximise over the larger set of PPT states (bipartite quantum states ρ such that $\rho^\Gamma \geq 0$, where $^\Gamma$ denotes the partial transpose operation, i.e. the transpose operation performed only on the second subsystem) and consider

$$h_{\mathrm{PPT}}(M) := \max_{\rho \in \mathrm{PPT}} \operatorname{tr} M\rho.$$

---

[1] I would like to thank Fernando Brandão for pointing this out.

An upper bound can be put on this quantity as follows.

Proposition 4. $h_{\text{PPT}}(M) \leq \|M^\Gamma\|_\infty$.

Proof. By definition, we have

$$h_{\text{PPT}}(M) = \max_{\substack{\rho,\,\rho \geq 0, \\ \rho^\Gamma \geq 0,\, \text{tr}\,\rho = 1}} \text{tr}\,M\rho = \max_{\substack{\sigma,\,\sigma \geq 0, \\ \sigma^\Gamma \geq 0,\, \text{tr}\,\sigma^\Gamma = 1}} \text{tr}\,M\sigma^\Gamma,$$

and for any density matrix σ, $\text{tr}\,M\sigma^\Gamma = \text{tr}\,M^\Gamma \sigma \leq \|M^\Gamma\|_\infty$. $\qquad\square$

This upper bound is often very weak (and for example may be much bigger than 1), but for some operators M it does give something non-trivial. Observe that if M projects onto a subspace $S \subseteq \mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$, $\|M^\Gamma\|_\infty$ is *not* the same as $\max_{|\psi\rangle \in S} \|\psi^\Gamma\|_\infty$. Indeed, for any state $|\psi\rangle$ we have

$$\|\psi^\Gamma\|_\infty \leq \|\psi^\Gamma\|_2 = \|\psi\|_2 = 1,$$

whereas $\|M^\Gamma\|_\infty$ may be considerably higher.

A key property of $\|M^\Gamma\|_\infty$ which we will use is that it is multiplicative.

Observation 5. *For any operators M, N, $\|(M \otimes N)^\Gamma\|_\infty = \|M^\Gamma \otimes N^\Gamma\|_\infty = \|M^\Gamma\|_\infty \|N^\Gamma\|_\infty$.*

Thus, if we can show that $\|M^\Gamma\|_\infty \leq \delta$ for some δ, we immediately have that $h_{\text{SEP}}(M^{\otimes n}) \leq \delta^n$. If δ is small enough, this can be used to prove that M obeys weak ∞-norm multiplicativity. We formalise this as the following observation.

Observation 6. *For any projector M, if $\|M^\Gamma\|_\infty \leq h_{\text{SEP}}(M)^\alpha$ for some constant $0 < \alpha \leq 1$, M's corresponding quantum channel obeys weak ∞-norm multiplicativity with exponent α.*

Of course, for many operators M, $\|M^\Gamma\|_\infty$ is much larger than 1. However, we will see that for random quantum channels, with high probability this strategy does lead to a non-trivial bound. Let M be the projector onto a subspace picked at random from the set of dimension r subspaces of $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$, i.e. M is formed by taking the projector M_0 onto a fixed dimension r subspace $S_0 \subseteq \mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$, and conjugating M_0 by a Haar-random unitary. Then we have the following easy *lower* bound on $h_{\text{SEP}}(M)$.

Proposition 7. *Let M be the projector onto an r-dimensional subspace of $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$. Then*

$$h_{\text{SEP}}(M) \geq \max\left\{ \frac{r}{d_A d_B}, \frac{1}{d_A} \right\}.$$

Proof. For the first part, pick a random product state $|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$ by choosing each of $|\psi_A\rangle$ and $|\psi_B\rangle$ uniformly at random (according to Haar measure). On average this achieves

$$\mathbb{E}_\psi[\text{tr}\,M\psi] = \text{tr}\,M(\mathbb{E}_{\psi_A}\psi_A) \otimes (\mathbb{E}_{\psi_B}\psi_B) = \frac{r}{d_A d_B}.$$

The second part follows from the correspondence with quantum channels. Any state output from the channel which corresponds to M must have largest eigenvalue at least $1/d_A$, so $h_{\text{SEP}}(M) \geq 1/d_A$. $\qquad\square$

On the other hand, we will prove the following upper bound.

6

Theorem 8. *Fix $d_A \leq d_B$ and r such that $m := \min\{d_A, r\} \geq 2(\log_2 d_B)^{3/2}$, and let M be the projector onto a random dimension r subspace of $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$. Then there exists a universal constant C such that, if $r \geq d_B/d_A$,*

$$\mathbb{E}\|M^\Gamma\|_\infty \leq \frac{Cr^{1/2}}{d_A^{1/2} d_B^{1/2}},$$

and if $r \leq d_B/d_A$,

$$\mathbb{E}\|M^\Gamma\|_\infty \leq \frac{C}{d_A}.$$

Further, for any $\delta > 0$, there exists a universal constant C' such that if $r \geq d_B/d_A$ then

$$\Pr\left[\|M^\Gamma\|_\infty \geq \delta \frac{2^8 r^{1/2}}{d_A^{1/2} d_B^{1/2}}\right] \leq C' m^{16/3} \delta^{-(m/2)^{2/3}},$$

and if $r \leq d_B/d_A$ then

$$\Pr\left[\|M^\Gamma\|_\infty \geq \delta \frac{2^8}{d_A}\right] \leq C' m^{16/3} \delta^{-(m/2)^{2/3}}.$$

Combining Proposition 7 and Theorem 8, we obtain Theorem 3, which we now state more formally.

Theorem 9. *Fix $d_A \leq d_B$ and r such that $m := \min\{d_A, r\} \geq 2(\log d_B)^{3/2}$, and let M be the projector onto a random dimension r subspace of $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$. Then there exists a universal constant C such that if $r \geq d_B/d_A$, then*

$$\Pr\left[\|M^\Gamma\|_\infty \geq h_{\mathrm{SEP}}(M)^{1/2-\epsilon}\right] \leq C m^{16/3} 2^{-(m/2)^{2/3}},$$

where $\epsilon = \frac{9}{\log_2(d_A d_B/r)}$, and if $r \leq d_B/d_A$, then

$$\Pr\left[\|M^\Gamma\|_\infty \geq h_{\mathrm{SEP}}(M)^{1-\epsilon'}\right] \leq C m^{16/3} 2^{-(m/2)^{2/3}},$$

where $\epsilon' = \frac{9}{\log_2 d_A}$.

Proof. If $r \geq d_B/d_A$, using Proposition 7 and taking $\delta = 2$ in Theorem 8, we obtain

$$
\begin{aligned}
\Pr[\|M^\Gamma\|_\infty \geq h_{\mathrm{SEP}}(M)^{1/2-\epsilon}] &\leq \Pr\left[\|M^\Gamma\|_\infty \geq \left(\frac{r}{d_A d_B}\right)^{1/2-\epsilon}\right] \\
&= \Pr\left[\|M^\Gamma\|_\infty \geq \frac{2^9 r^{1/2}}{d_A^{1/2} d_B^{1/2}}\right] \\
&\leq C' m^{16/3} 2^{-(m/2)^{2/3}},
\end{aligned}
$$

and a similar argument holds for $r \leq d_B/d_A$. $\square$

2.1 Related work

Important intuition that some version of Theorem 8 should hold comes from recent work by Aubrun [4], who studied the asymptotic spectrum of partially transposed Wishart matrices (also see [19] for some interesting related work). Such matrices, which are a natural model for random mixed quantum states, can be formed as follows. Let G be a $d \times r$ matrix whose entries are picked independently from the complex normal distribution $N(0,1)$, and set $W = \frac{1}{d}GG^\dagger$; we say that W is a (d,r)-Wishart matrix[2]. Then one of Aubrun's results can be stated as follows. For a fixed constant $0 < \alpha < 1$, let W_d be a $(d^2, \lfloor \alpha d^2 \rfloor)$-Wishart matrix, understood as acting on the bipartite space $\mathbb{C}^d \otimes \mathbb{C}^d$, and set $Y_d = W_d^\Gamma$. Then

$$\lim_{d \to \infty} \Pr[|\lambda_{\max}(Y_d) - \sqrt{\alpha}(2 + \sqrt{\alpha})| > \epsilon] = 0.$$

As $r = \alpha d^2$, this implies that with high probability, $\lambda_{\max}(Y_d) = O(\sqrt{r}/d)$ for large d. As the columns of G are approximately orthogonal for large d, one might expect that the operator norm of the partial transpose of the projector onto a random r-dimensional subspace of $\mathbb{C}^d \otimes \mathbb{C}^d$ should behave similarly to that of Y_d; the present work demonstrates that this is indeed true.

As well as the final bound we obtain being of a similar order to that of Aubrun, the reader familiar with [4] will notice that we need to prove some analogous combinatorial lemmas. However, it does not seem clear that the results given here could be obtained as a formal consequence of [4], or indeed vice versa; it would be interesting to determine whether this is the case. We also note two minor technical differences between this work and [4]: Aubrun's result is only stated for $d_A = d_B = d$ (some recent work by Banica and Nechita removes this restriction [8]), and only for α constant (so r grows as a constant fraction of d). By contrast, here d_A, d_B and r can be essentially arbitrary, although the bound we obtain becomes trivial if r is too large as a fraction of $d_A d_B$.

Some very recent work by Collins, Fukuda and Nechita [13] also uses related techniques to those which we use here to prove Theorem 8 (e.g. calculations with Weingarten functions [20], see Section 3.2 below). The goal of [13] was to find the state which, when input to a tensor product of two random quantum channels (either the same channel, or a channel and its conjugate), achieves minimal output entropy. A sequence of recent papers by Collins and Nechita [15, 16, 17, 18] carries out a number of interesting quantum information-theoretic calculations using Weingarten functions. This method seems to be a powerful tool which may be expected to find many other applications in quantum information.

The rest of this paper is devoted to the proof of Theorem 8, which we now begin.

3 Proof of Theorem 8: moments of partially transposed projectors

As before, let M_0 be the projector onto an arbitrary fixed r-dimensional subspace of $\mathbb{C}^{d_A} \otimes \mathbb{C}^{d_B}$, where $d_A \leq d_B$, and let $M = UM_0U^\dagger$ be the result of applying a Haar-random unitary operator to M_0. Also let

$$D_d(\pi) := \sum_{i_1,\dots,i_k=1}^{d} |i_{\pi(1)}\rangle |i_{\pi(2)}\rangle \dots |i_{\pi(k)}\rangle\langle i_1| \dots \langle i_k|$$

be the representation of the permutation $\pi \in S_k$ which acts on k systems of dimension d by permuting the systems; in the case $d = d_A d_B$, we simply omit d and write $D(\pi) := D_{d_A d_B}(\pi)$. Let

---

[2]Note that we use a different normalisation to [4].

8

$\kappa \in S_k$ be the permutation that maps $i \mapsto i+1$ for $i = 1, \ldots, k-1$, and maps $k \mapsto 1$. For any permutation π, let $c(\pi)$ be the number of cycles in π.

In order to put good upper bounds on $\|M^\Gamma\|_\infty$, it suffices to understand $\mathbb{E}\,\mathrm{tr}[(M^\Gamma)^k]$ for arbitrary even k. Observe that

$$
\begin{aligned}
\mathbb{E}\,\mathrm{tr}[(M^\Gamma)^k] &= \mathbb{E}_U\,\mathrm{tr}[((UM_0U^\dagger)^\Gamma)^k] \\
&= \mathrm{tr}[\mathbb{E}_U D(\kappa)((U^{\otimes k}M_0^{\otimes k}(U^\dagger)^{\otimes k})^\Gamma)] \\
&= \mathrm{tr}[D(\kappa)^\Gamma M^{(k)}],
\end{aligned}
$$

where for brevity we write

$$
M^{(k)} := \mathbb{E}_U[U^{\otimes k}M_0^{\otimes k}(U^\dagger)^{\otimes k}].
$$

The second equality above is the observation that $\mathrm{tr}[X^k] = \mathrm{tr}\,D(\kappa)X^{\otimes k}$ (indeed, this holds for any permutation $\pi \in S_k$ such that $c(\pi) = 1$).

Theorem 8 will follow easily from the following result.

Theorem 10. *There is a universal constant C such that, for any k satisfying $2k^{3/2} \leq \min\{d_A, r\}$,*

$$
\mathrm{tr}[D(\kappa)^\Gamma M^{(k)}] \leq
\begin{cases}
Ck^8 2^{6k} r^{k/2} d_A^{-k/2+1} d_B^{-k/2+1} & \text{if } r \geq d_B/d_A \\
Ck^8 2^{6k} d_A^{-k+1} d_B & \text{otherwise.}
\end{cases}
$$

Proof of Theorem 8 (assuming Theorem 10). For any even $k \geq 2$,

$$
\mathbb{E}\|M^\Gamma\|_\infty \leq (\mathbb{E}[\mathrm{tr}(M^\Gamma)^k])^{1/k},
$$

and taking k to be the largest even number smaller than $(m/2)^{2/3}$, where $m := \min\{d_A, r\}$ and we assume that $m \geq 2(\log_2 d_B)^{3/2}$, we obtain that there is a universal constant C' such that

$$
\mathbb{E}\|M^\Gamma\|_\infty \leq \frac{C' r^{1/2}}{d_A^{1/2} d_B^{1/2}} \text{ if } r \geq d_B/d_A,
$$

and

$$
\mathbb{E}\|M^\Gamma\|_\infty \leq \frac{C'}{d_A} \text{ otherwise,}
$$

which is the first part of Theorem 8. (Note that we took $m \geq 2(\log_2 d_B)^{3/2}$ in order to kill off terms of the form $d_B^{1/k}$.) The concentration bound follows from Markov's inequality, which implies that, for any even $k \geq 2$ and any $x > 0$,

$$
\Pr[\|M^\Gamma\|_\infty \geq x] = \Pr\left[\|M^\Gamma\|_\infty^k \geq x^k\right] \leq \frac{\mathbb{E}[\mathrm{tr}(M^\Gamma)^k]}{x^k};
$$

and once again taking k to be the largest even number smaller than $(m/2)^{2/3}$, we obtain

$$
\Pr\left[\|M^\Gamma\|_\infty \geq \delta \frac{2^8 r^{1/2}}{d_A^{1/2} d_B^{1/2}}\right] \leq \frac{Ck^8 d_A d_B}{(4\delta)^k} \leq C' m^{16/3} \delta^{-(m/2)^{2/3}}
$$

if $r \geq d_B/d_A$, and similarly

$$
\Pr\left[\|M^\Gamma\|_\infty \geq \delta \frac{2^8}{d_A}\right] \leq \frac{Ck^8 d_A d_B}{(4\delta)^k} \leq C' m^{16/3} \delta^{-(m/2)^{2/3}}
$$

if $r \leq d_B/d_A$. $\qquad\square$

We now proceed to prove Theorem 10. By Schur-Weyl duality, as $M^{(k)}$ commutes with local unitaries and permutations of the k systems, it can be expanded in terms of permutations $D(\pi)$ as

$$M^{(k)} = \sum_{\pi \in S_k} \alpha_\pi D(\pi)$$

for some coefficients α_π. When k is small with respect to $d_A d_B$, the matrices $\{D(\pi)\}$ are almost orthonormal with respect to the normalised Hilbert-Schmidt inner product, i.e.

$$\frac{1}{(d_A d_B)^k} \operatorname{tr}[D(\pi)^\dagger D(\sigma)] \approx 0 \text{ if } \pi \neq \sigma.$$

It has been pointed out by Harrow that one can prove a number of interesting results in quantum information based on this philosophy [23]. In our case, because of this near-orthonormality we ought to have

$$\alpha_\pi \approx \frac{\operatorname{tr}[M^{(k)} D(\pi^{-1})]}{\operatorname{tr}[D(\pi^{-1}) D(\pi)]} = \frac{r^{c(\pi)}}{(d_A d_B)^k};$$

the following key technical lemma, which we prove in Section 3.2, makes this approximate equality quantitative.

Lemma 11. *Assume $k \leq (r/2)^{2/3}$. Then there is a universal constant C such that*

$$|\alpha_\pi| \leq Ck2^{4k} \frac{r^{c(\pi)}}{(d_A d_B)^k}.$$

Now it holds that

$$
\begin{aligned}
\operatorname{tr}[D(\kappa)^\Gamma M^{(k)}] &= \sum_{\pi \in S_k} \alpha_\pi \operatorname{tr}[D(\kappa)^\Gamma D(\pi)] = \sum_{\pi \in S_k} d_A^{c(\kappa\pi)} d_B^{c(\kappa^{-1}\pi)} \alpha_\pi \\
&\leq Ck2^{4k} \sum_{\pi \in S_k} d_A^{c(\kappa\pi)-k} d_B^{c(\kappa^{-1}\pi)-k} r^{c(\pi)},
\end{aligned}
\tag{1}
$$

where the second equality is the fact that

$$
\begin{aligned}
\operatorname{tr}[D(\kappa)^\Gamma D(\pi)] &= \operatorname{tr}[(D_{d_A}(\kappa) \otimes D_{d_B}(\kappa)^T)(D_{d_A}(\pi) \otimes D_{d_B}(\pi))] \\
&= \operatorname{tr}[D_{d_A}(\kappa) D_{d_A}(\pi)] \operatorname{tr}[D_{d_B}(\kappa^{-1}) D_{d_B}(\pi)] \\
&= d_A^{c(\kappa\pi)} d_B^{c(\kappa^{-1}\pi)},
\end{aligned}
$$

and the inequality is Lemma 11. In order to upper bound the quantity (1), we will use the following lemma[3], which we also prove afterwards, in Section 3.1.

Lemma 12. *For any $\pi, \sigma \in S_k$,*

$$c(\pi^{-1}\sigma) + c(\sigma) \leq k + c(\pi).$$

Further, for any $\pi \in S_k$ and any integer $\delta \geq 0$, there are at most $4^{k-1} k^{3\delta/2+1}$ permutations $\sigma \in S_k$ such that

$$c(\pi^{-1}\sigma) + c(\sigma) = k + c(\pi) - \delta.$$

---

[3]Related combinatorial results appear in the literature, for example in [4]. However, the precise statement we need here does not seem to have been written down.

The exact constants in this lemma are not so important; what matters is that, when $i + j \approx k + c(\pi)$, there are not too many permutations σ such that $c(\pi^{-1}\sigma) = i$, $c(\sigma) = j$ (i.e. there are $O(4^k \operatorname{poly}(k))$ of them, rather than $\Omega(k!)$). Now define

$$N(a, b, c) := |\{\pi \in S_k : c(\kappa\pi) = a, c(\kappa^{-1}\pi) = b, c(\pi) = c\}|$$

and rewrite (1) as

$$\operatorname{tr}[D(\kappa)^\Gamma M^{(k)}] \leq \frac{Ck2^{4k}}{d_A^k d_B^k} \sum_{a,b,c \in \{1,\ldots,k\}} N(a, b, c) d_A^a d_B^b r^c.$$

By Lemma 12, we have $N(a, b, c) = 0$ whenever

$$a + b > k + 2, \ a + c > k + 1, \ \text{or} \ b + c > k + 1.$$

Call the triple (a, b, c) *valid* when none of these events occur. Further, we have the bound from Lemma 12 that for all valid triples,

$$N(a, b, c) \leq 4^{k-1} k^{(3/2)(k+2-\max\{a+b,a+c,b+c\})+1}.$$

We therefore obtain the upper bound

$$\operatorname{tr}[D(\kappa)^\Gamma M^{(k)}] \leq \frac{Ck^4 2^{4k}}{d_A^k d_B^k} \max_{(a,b,c) \text{ valid}} \left\{ 4^k k^{(3/2)(k-\max\{a+b,a+c,b+c\})+4} d_A^a d_B^b r^c \right\}.$$

Now observe that, as long as $k^{3/2} \leq \min\{d_A, r\}$, this maximum will be achieved when at least one of the validity inequalities is saturated, because otherwise we could increase the maximum by at least a factor of $\min\{d_A, r\}/k^{3/2}$, by increasing at least one of (a, b, c). Thus the upper bound simplifies to

$$\operatorname{tr}[D(\kappa)^\Gamma M^{(k)}] \leq \frac{Ck^8 2^{6k}}{d_A^k d_B^k} \max_{(a,b,c) \text{ valid}} \left\{ d_A^a d_B^b r^c \right\}. \tag{2}$$

We have seen that the triple (a, b, c) satisfies the linear inequalities $2 \leq a+b \leq k+2$, $2 \leq a+c \leq k+1$, $2 \leq b+c \leq k+1$. Therefore, the optimal value of the following simple linear program gives an upper bound on $\log \max_{(a,b,c) \text{ valid}} d_A^a d_B^b r^c$:

$$
\begin{aligned}
\text{maximise} \quad & (\log d_A)a + (\log d_B)b + (\log r)c \\
\text{subject to} \quad & 2 \leq a + b \leq k + 2 \\
& 2 \leq a + c \leq k + 1 \\
& 2 \leq b + c \leq k + 1, \\
& a, b, c \geq 0.
\end{aligned}
$$

Performing standard manipulations gives the dual:

$$
\begin{aligned}
\text{minimise} \quad & (k+1)u^+ + (k+1)v^+ + (k+2)w^+ - 2(u^- + v^- + w^-) \\
\text{subject to} \quad & u + w \geq \log d_A \\
& v + w \geq \log d_B \\
& u + v \geq \log r,
\end{aligned}
$$

where the notation w^+ (resp. w^-) is used for the positive (resp. negative) part of x, i.e. $w^+ = \max\{w, 0\}$, $w^- = -\min\{w, 0\}$. Observe that the following dual solution saturates all three inequalities:

$$u = \frac{1}{2}\left(\log d_A - \log d_B + \log r\right), \ v = \frac{1}{2}\left(-\log d_A + \log d_B + \log r\right), \ w = \frac{1}{2}\left(\log d_A + \log d_B - \log r\right).$$

As $r \leq d_A d_B$ and $d_A \leq d_B$, in this solution v and w are always non-negative. If it additionally holds that $r \geq d_B/d_A$, so $\log d_B \geq \log d_A + \log r$, u is also non-negative, so we achieve an objective value of

$$(k+1)u + (k+1)v + (k+2)w = \frac{1}{2}\left((k+2)\log d_A + (k+2)\log d_B + k\log r\right).$$

On the other hand, if $r \leq d_B/d_A$, consider the solution

$$u = 0, \ v = \log d_B - \log d_A, \ w = \log d_A.$$

It is easy to verify that $u, v, w \geq 0$ and this solution achieves an objective value of

$$\log d_A + (k+1)\log d_B.$$

These two solutions correspond to upper bounds in (2) of

$$\mathrm{tr}[D(\kappa)^\Gamma M^{(k)}] \leq \begin{cases} Ck^8 2^{6k} r^{k/2} d_A^{-k/2+1} d_B^{-k/2+1} & \text{if } r \geq d_B/d_A \\ Ck^8 2^{6k} d_A^{-k+1} d_B & \text{otherwise.} \end{cases}$$

This completes the proof of Theorem 10.

3.1 Proof of Lemma 12: combinatorics of permutations

Our next task is to prove the above combinatorial lemma, which we restate for convenience.

Lemma 12 (restated). *For any $\pi, \sigma \in S_k$,*

$$c(\pi^{-1}\sigma) + c(\sigma) \leq k + c(\pi).$$

Further, for any $\pi \in S_k$ and any integer $\delta \geq 0$, there are at most $4^{k-1}k^{3\delta/2+1}$ permutations $\sigma \in S_k$ such that

$$c(\pi^{-1}\sigma) + c(\sigma) = k + c(\pi) - \delta.$$

The proof will rely on some previously known combinatorial results regarding permutations; we first review some basic ideas in this area. Consider the Cayley graph whose vertices are elements of S_k and where two vertices π, σ are connected if and only if there exists a transposition τ such that $\tau\pi = \sigma$. For any permutations π, σ, let $d(\pi, \sigma)$ be the shortest-path metric with respect to this graph (i.e. the minimum number of transpositions required to change π into σ). This is indeed a metric as $d(\pi, \sigma) \leq d(\pi, \rho) + d(\rho, \sigma)$ for any $\rho \in S_k$. Also observe that $d(\pi, \sigma) = d(\sigma, \pi)$ and $d(\pi, \sigma) = d(\pi^{-1}\sigma, e)$, where e denotes the identity permutation.

Let τ be a transposition exchanging elements i and j. If π is given in cycle notation as $\pi = (c_1)(c_2)\dots(c_\ell)$, where each c_p is a sequence of integers, then the permutation $\tau\pi$ takes one of two forms depending on whether the transposed elements are in the same cycle or not. Let $c = (c_1, \dots, c_m)$ and $d = (d_1, \dots, d_{m'})$ be two cycles in π. If elements c_i and c_j are transposed, c splits

12

into two cycles $(c_1, \ldots, c_{i-1}, c_j, \ldots, c_m)$ and $(c_i, \ldots, c_{j-1})$. If elements c_i and d_j are transposed, cycles c and d are joined to produce the cycle $(c_1, \ldots, c_{i-1}, d_j, d_{j+1}, \ldots, d_{m'}, d_1, \ldots, d_{j-1}, c_i, \ldots, c_m)$. Thus performing a transposition τ always either increases or decreases the number of cycles by exactly 1, and in each case the result preserves the ordering of elements within the original cycles.

This implies that, for any $\pi \in S_k$, $d(\pi, e) = k - c(\pi)$, because $\ell - 1$ transpositions are required to split a cycle of length ℓ into ℓ cycles of length 1, and this can be achieved. This immediately implies the first part of Lemma 12 (which is well-known). Simply use the triangle inequality:

$$c(\pi^{-1}\sigma) + c(\sigma) = k - d(\pi, \sigma) + k - d(e, \sigma) \leq 2k - d(e, \pi) = k + c(\pi). \tag{3}$$

Also note that, if $(\tau_1, \ldots, \tau_m)$ is a sequence of transpositions such that

$$\tau_m \ldots \tau_1 \pi = \sigma,$$

it holds that $m - d(\pi, \sigma)$ is a multiple of 2.

We now turn to the second part of Lemma 12. We first observe that the special case $\delta = 0$ has a simple (and well-known, e.g. see [32]) proof.

Lemma 13. *For any permutation $\pi \in S_k$,*

$$|\{\sigma : c(\pi^{-1}\sigma) + c(\sigma) = k + c(\pi)\}| \leq C_k,$$

where C_k is the k'th Catalan number

$$C_k := \frac{1}{k+1}\binom{2k}{k}.$$

Proof. The number of permutations $\sigma \in S_k$ such that $c(\pi^{-1}\sigma) + c(\sigma) = k + c(\pi)$ is equal to

$$|\{\sigma : d(\pi, \sigma) + d(\sigma, e) = d(\pi, e)\}|, \tag{4}$$

or in other words the same as the number of permutations which lie on a shortest path ("geodesic") between π and e. Any such permutation must be obtained from π by a sequence of transpositions, each of which splits a cycle of π in two. Permutations on the shortest path between a cycle of length ℓ and the identity are known to be isomorphic to non-crossing partitions of $\{1, \ldots, \ell\}$ [9], which are counted by the Catalan numbers C_ℓ. Therefore, an overall upper bound on (4) is

$$\prod_{i=1}^{c(\pi)} C_{c_i} \leq C_k,$$

where c_i is the length of the i'th cycle of π and we use the simple upper bound $C_a C_b \leq C_{a+b}$, valid for all integers $a, b \geq 1$. $\qquad\square$

For the more complicated case of $\delta > 0$, we will rely on a result of Adrianov [1]. This work enumerates the so-called bicoloured unicellular maps with m white and n black vertices and k edges, one of which is marked; this is known to be equivalent to the problem of counting pairs of permutations $\pi, \sigma \in S_k$ such that $c(\pi) = m$, $c(\sigma) = n$ and $\pi\sigma$ is a fixed cycle of length k [1]. For $\pi \in S_k$, write

$$N_g(\pi) := |\{\sigma : c(\pi\sigma^{-1}) + c(\sigma) = k + c(\pi) - 2g\}| = |\{\sigma : d(\pi, \sigma) + d(\sigma, e) = d(\pi, e) + 2g\}|.$$

By Lemma 13 and the above arguments, $N_g(\pi) = 0$ for all $g < 0$ and for non-integer g. Also set $B_g(k) := N_g(\kappa)$, where $\kappa \in S_k$ is an arbitrary cycle of length k. Then the following result holds.

13

Theorem 14 (Adrianov [1], Corollary 3). $B_g(k)$ *satisfies the recurrence*

$$(k+1)B_g(k) = 2(2k-1)B_g(k-1) + (k-2)(k-1)^2 B_{g-1}(k-2)$$

with initial conditions

$$B_0(1) = 1, \ B_0(2) = 2, \ B_g(k) = 0 \ \textit{for } g < 0.$$

Using Theorem 14, we now complete the proof of Lemma 12. We have

$$B_g(k) \le 4B_g(k-1) + (k-1)^2 B_{g-1}(k-2),$$

and we now apply induction on g to upper bound this quantity. For $g = 0$, we have $B_0(k) \le 4^{k-1}$ (agreeing with the known exact values $B_0(k) = C_k$). For any $g > 0$ we have $B_g(1) = 0$, and more generally

$$
\begin{aligned}
B_g(k) &\le \ 4\left(4B_g(k-2) + (k-2)^2 B_{g-1}(k-3)\right) + (k-1)^2 B_{g-1}(k-2) \\
&\le \ \cdots \le \sum_{i=1}^{k-1} 4^{k-i-1} i^2 B_{g-1}(i-1).
\end{aligned}
$$

We now use the inductive hypothesis that $B_{g-1}(k) \le 4^{k-1} k^{3(g-1)}$ for all k, which implies that

$$
\begin{aligned}
B_g(k) &\le \ \sum_{i=1}^{k-1} 4^{k-i-1} i^2 4^{i-2} (i-1)^{3(g-1)} \le 4^{k-3} \sum_{i=1}^{k-1} i^{3g-1} \\
&\le \ 4^{k-3} \int_1^k x^{3g-1} dx \le 4^{k-3} \frac{k^{3g}}{3g} \le 4^{k-1} k^{3g}.
\end{aligned}
$$

We use this to show that, for any $k \ge 2$, any $\pi \in S_k$ and any permutation $\kappa \in S_k$ which consists of one cycle of length k,

$$\sum_{h=0}^{g} N_h(\pi) \le \sum_{h=0}^{g} N_h(\kappa) \le 4^{k-1} k^{3g+1}. \tag{5}$$

The argument for the first inequality is as follows. Let κ' be a cycle of length k which minimises $d(\kappa', \pi)$. Then, for any permutation σ such that $d(\pi, \sigma) + d(\sigma, e) = d(\pi, e) + 2g$, it holds that

$$d(\kappa', \sigma) + d(\sigma, e) \le d(\kappa', \pi) + d(\pi, \sigma) + d(\sigma, e) = d(\kappa', \pi) + d(\pi, e) + 2g = d(\kappa', e) + 2g,$$

where the second equality holds because π is on the shortest path between κ' and e. Thus σ contributes to $N_h(\kappa')$ for some $h \le g$. Just looking at one term in the sum on the left-hand side of (5), we have that $N_g(\pi) \le 4^{k-1} k^{3g+1}$ for all $\pi \in S_k$. To complete the proof of Lemma 12, simply observe that

$$|\{\sigma : c(\pi^{-1}\sigma) + c(\sigma) = k + c(\pi) - \delta\}| = N_{\delta/2}(\pi)$$

for even δ, and $|\{\sigma : c(\pi^{-1}\sigma) + c(\sigma) = k + c(\pi) - \delta\}| = 0$ for odd δ.

3.2 Proof of Lemma 11: permutations and Weingarten functions

Let A be the symmetric matrix defined by $A_{\pi\sigma} = d^{c(\pi^{-1}\sigma)-k}$, for $\pi, \sigma \in S_k$. Given some matrix M such that $M = \sum_{\pi \in S_k} \alpha_\pi D_d(\pi)$, A determines the coefficients α_π as follows:

$$\operatorname{tr} MD_d(\sigma) = \sum_{\pi \in S_k} \alpha_\pi d^{c(\pi\sigma)} = d^k \sum_{\pi \in S_k} A_{\sigma^{-1}\pi} \alpha_\pi.$$

Letting v and w be the vectors defined by

$$v_\sigma = \frac{1}{d^k} \operatorname{tr} MD_d(\sigma^{-1}), \ w_\pi = \alpha_\pi,$$

this is equivalent to the claim that $Aw = v$. Thus, if A^{-1} exists, we can determine the α_π coefficients by computing $A^{-1}v$. Note that A is approximately equal to the identity when d is large with respect to k, as its off-diagonal entries rapidly decay [23].

In order to evaluate the entries of A^{-1}, we define the *Weingarten function* [20]

$$\operatorname{Wg}(\pi) := \frac{1}{(k!)^2} \sum_{\lambda \vdash k} \frac{(f^\lambda)^2}{s_\lambda(1^{\times d})} \chi^\lambda(\pi).$$

This expression uses standard notation from the representation theory of the symmetric group (see [7] for an accessible introduction). The sum is over partitions λ of $\{1, \ldots, k\}$; f^λ is the number of standard Young tableaux with shape λ; $\chi^\lambda(\pi)$ is the character of the symmetric group corresponding to partition λ, evaluated at π; and finally $s_\lambda(1^{\times d})$ is the Schur function corresponding to the partition λ, which has the explicit expression

$$s_\lambda(1^{\times d}) = \frac{f^\lambda}{k!} \prod_{(i,j) \in \lambda} (d + j - i).$$

Then we have the following lemma, which is well-known but we prove for completeness in Appendix A.

Lemma 15.
$$A_{\pi\sigma}^{-1} = d^k \operatorname{Wg}(\pi^{-1}\sigma).$$

For small k, one can use this expression to calculate the coefficients α_π exactly, but as k increases the Weingarten function becomes cumbersome to work with. In order to address this, we now give a general upper bound on this function.

Lemma 16. *For any $k \leq d^{2/3}$,*

$$|\operatorname{Wg}(\pi)| \leq \frac{3 \, C_{k-1}}{2} \, d^{c(\pi) - 2k},$$

where C_{k-1} is the $(k-1)$'th Catalan number.

The reader familiar with the work [20] may wonder why we do not use the seemingly tighter bound given there that

$$\operatorname{Wg}(\pi) = d^{c(\pi)-2k}(\operatorname{Moeb}(\pi) + O(d^{-2})),$$

where $\operatorname{Moeb}(\pi)$ is a function which can be shown to be upper bounded by C_{k-1}; the reason is that the $O(d^{-2})$ term in this bound hides an unspecified dependence[4] on k. Also note that the very recent work [14] gives an elegant alternative upper bound on the Weingarten function, but this bound does not seem to suffice for us to obtain tight results.

To prove Lemma 16, we will use a result of Matsumoto and Novak [33, 30, 31] which states that the Weingarten function can be expanded in terms of primitive factorisations in the symmetric group. This will allow us to use combinatorial bounds on such factorisations to obtain corresponding bounds on the Weingarten function. A primitive factorisation of a permutation $\pi \in S_k$ is a sequence of transpositions $(s_1, t_1), \ldots, (s_\ell, t_\ell)$ such that the product $(s_1, t_1) \ldots (s_\ell, t_\ell) = \pi$ and $t_1 \leq t_2 \leq \cdots \leq t_\ell$. Define $w_\ell(\pi)$ to be the number of primitive factorisations of π into ℓ transpositions.

---

[4] I would like to thank Carlos González Guillén for pointing this out.

Theorem 17 ([33, 30, 31]). *For any $k \leq d$, and any $\pi \in S_k$,*

$$\mathrm{Wg}(\pi) = \frac{1}{d^k} \sum_{\ell=0}^{\infty} w_\ell(\pi) \left(\frac{-1}{d}\right)^\ell.$$

Note that this expansion implies that $w_\ell(\pi)$ depends only on the cycle type of π. Also observe that, by the same argument as used in Section 3.1, $w_\ell(\pi) = 0$ unless $\ell = k - c(\pi) + 2g$, for integer $g \geq 0$. In the case where π is a cycle of length k (called a k-cycle in what follows), $\mathrm{Wg}(\pi)$ can be evaluated explicitly.

Theorem 18 ([12, 31]). *Fix $k \leq d$ and let $\kappa \in S_k$ be an arbitrary k-cycle. Then*

$$\mathrm{Wg}(\kappa) = \frac{(-1)^{k+1} C_{k-1}}{d(d^2 - 1^2)\ldots(d^2 - (k-1)^2)},$$

where C_{k-1} is the $(k-1)$'th Catalan number.

Proof of Lemma 16. We first show that, for arbitrary $\pi \in S_k$, and arbitrary integer $g \geq 0$,

$$w_{k-c(\pi)+2g}(\pi) \leq w_{k-1+2g}(\kappa), \tag{6}$$

where κ is an arbitrary k-cycle. For any permutation π, there exists a k-cycle κ_π such that κ_π can be obtained from π using $d(\kappa_\pi, \pi)$ primitive transpositions (simply apply the sequence of transpositions $(j_1, k)(j_2, k)\ldots(j_{c(\pi)-1}, k)$ to π, where $\{j_1, \ldots, j_{c(\pi)-1}\}$ contains one element from each of the cycles of π, except the cycle containing k). Thus each distinct primitive factorisation of π of length a gives a distinct primitive factorisation of κ_π of length exactly $a + c(\pi) - 1$, which proves inequality (6). Hence we have

$$\begin{aligned}
|\mathrm{Wg}(\pi)| &= \frac{1}{d^k} \sum_{\ell=1}^{\infty} w_\ell(\pi) d^{-\ell} = d^{c(\pi)-2k} \sum_{g=0}^{\infty} w_{k-c(\pi)+2g} d^{-2g} \leq d^{c(\pi)-2k} \sum_{g=0}^{\infty} w_{k-1+2g}(\kappa) d^{-2g} \\
&= d^{c(\pi)-1} |\mathrm{Wg}(\kappa)|,
\end{aligned}$$

where the inequality is (6). Now we can upper bound

$$d^{2k-1}|\mathrm{Wg}(\kappa)| = \frac{C_{k-1}}{(1 - 1^2/d^2)\ldots(1 - (k-1)^2/d^2)} \leq \frac{C_{k-1}}{1 - \frac{1}{d^2}\sum_{i=1}^{k-1} i^2} \leq \frac{3\,C_{k-1}}{2},$$

where we use $k^3 \leq d^2$. This completes the proof of Lemma 16. $\qquad\square$

Lemma 16 gives that

$$|A_{\pi\sigma}^{-1}| \leq \frac{3\,C_{k-1}}{2}\, d^{c(\pi^{-1}\sigma)-k} \tag{7}$$

for all $\pi, \sigma \in S_k$. We now use this to obtain an upper bound on the coefficients occurring in the decomposition of the specific matrix $M^{(k)}$ in terms of permutations, and hence prove Lemma 11. In fact, we prove the following lemma, from which Lemma 11 immediately follows by observing that $C_{k-1} \leq 4^{k-1}/k$.

Lemma 19. *Let M_0 be the projector onto a fixed r-dimensional subspace of $\mathbb{C}^d$, let k be an arbitrary integer satisfying $1 \leq k \leq (r/2)^{2/3}$, and set*

$$M^{(k)} = \mathbb{E}_U[U^{\otimes k} M_0^{\otimes k} (U^\dagger)^{\otimes k}].$$

Write $M^{(k)} = \sum_{\pi \in S_k} \alpha_\pi D_d(\pi)$. Then

$$|\alpha_\pi| \leq 3\,C_{k-1} k^2 4^{k-1} \frac{r^{c(\pi)}}{d^k}.$$

Proof. As $\operatorname{tr}[M^{(k)} D_d(\pi^{-1})] = r^{c(\pi)}$ for all $\pi \in S_k$, by (7) we have

$$\alpha_\pi = \frac{1}{d^k} \sum_{\sigma \in S_k} A_{\pi\sigma}^{-1} r^{c(\sigma)} \leq \frac{3\,C_{k-1}}{2} \sum_{\sigma \in S_k} d^{c(\pi^{-1}\sigma) - 2k} r^{c(\sigma)}.$$

Rewriting the sum and using Lemma 12, we obtain

$$
\begin{aligned}
|\alpha_\pi| &\leq \frac{3\,C_{k-1}}{2} d^{-2k} \sum_{i=1}^{k} d^i \sum_{\sigma, c(\pi^{-1}\sigma) = i} r^{c(\sigma)} \\
&= \frac{3\,C_{k-1}}{2} d^{-2k} \sum_{i=1}^{k} d^i \sum_{j} r^j |\{\sigma : c(\pi^{-1}\sigma) = i, c(\sigma) = j\}| \\
&\leq \frac{3\,C_{k-1}}{2} d^{-2k} \sum_{i=1}^{k} d^i \sum_{j=1}^{k+c(\pi)-i} r^j 4^{k-1} k^{(3/2)(k+c(\pi)-(i+j))+1} \\
&= 3\,C_{k-1} d^{-2k} 2^{2k-3} k^{(3/2)(k+c(\pi))+1} \sum_{i=1}^{k} \left(\frac{d}{k^{3/2}}\right)^i \sum_{j=1}^{k+c(\pi)-i} \left(\frac{r}{k^{3/2}}\right)^j \\
&\leq 3\,C_{k-1} d^{-2k} 4^{k-1} k^{(3/2)(k+c(\pi))+1} \sum_{i=1}^{k} \left(\frac{d}{k^{3/2}}\right)^i \left(\frac{r}{k^{3/2}}\right)^{k+c(\pi)-i} \\
&= 3\,C_{k-1} d^{-2k} k 4^{k-1} r^{k+c(\pi)} \sum_{i=1}^{k} \left(\frac{d}{r}\right)^i \\
&\leq 3\,C_{k-1} d^{-2k} k^2 4^{k-1} r^{k+c(\pi)} \left(\frac{d}{r}\right)^k \\
&= 3\,C_{k-1} d^{-k} k^2 4^{k-1} r^{c(\pi)}.
\end{aligned}
$$

$\square$

Most of these steps are self-explanatory; in the third inequality we use the fact that $r/2 \geq k^{3/2}$, and in the fourth the fact that $r \leq d$.

4 Conclusion

We have shown that random channels obey weak multiplicativity with high probability. When combined with the results of Christandl, Schuch and Winter [10, 11] on the antisymmetric subspace, this implies that two of the constructions of channels which display the strongest known

two-copy multiplicativity violations are in fact weakly multiplicative when the number of copies increases. This naturally leads one to conjecture that in fact *all* channels satisfy some form of weak multiplicativity (see [25] for a similar conjecture). However, note that the proof strategy used here fails badly for the antisymmetric subspace, as

$$P_{\text{anti}}^{\Gamma} = \frac{1}{2}\left(I - F^{\Gamma}\right) = \frac{1}{2}\left(I - d\,\Phi\right),$$

where F is the flip (swap) operator which interchanges two systems and Φ is a maximally entangled state. Hence we have $\|P_{\text{anti}}^{\Gamma}\|_{\infty} = (d-1)/2$, which does not provide a useful upper bound on $h_{\text{SEP}}(P_{\text{anti}})$. It would be very interesting to find a single proof strategy which works in both cases.

It is also interesting to observe that our main result that random channels obey weak p-norm multiplicativity with exponent $(1/2 - o(1))(1 - 1/p)$ becomes weaker as $p \to 1$, whereas known violations of multiplicativity are strongest at $p = \infty$. It is therefore tempting to conjecture that this exponent could be improved to $1/2 - o(1)$ for all $p > 1$, or even to $(1/2 - o(1))f(p)$ for some function $f(p)$ which *decreases* with p.

Acknowledgements

A Proof of Lemma 15: Exact expression for A^{-1}

We finally prove Lemma 15.

Lemma 15 (restated).

$$A_{\pi\sigma}^{-1} = \frac{d^k}{(k!)^2}\sum_{\lambda \vdash k}\chi^{\lambda}(\pi^{-1}\sigma)\frac{(f^{\lambda})^2}{s_{\lambda}(1^{\times d})} = d^k\,\mathrm{Wg}(\pi^{-1}\sigma).$$

Proof. We simply verify that

$$
\begin{aligned}
(A^{-1}A)_{\pi\sigma} &= \sum_{\tau \in S_k} A_{\pi\tau}^{-1}A_{\tau\sigma} \\
&= \frac{1}{(k!)^2}\sum_{\tau \in S_k}\left(\sum_{\lambda \vdash k}\chi^{\lambda}(\pi^{-1}\tau)\frac{(f^{\lambda})^2}{s_{\lambda}(1^{\times d})}\right)d^{c(\tau^{-1}\sigma)} \\
&= \frac{1}{(k!)^2}\sum_{\lambda,\mu \vdash k}\frac{(f^{\lambda})^2}{s_{\lambda}(1^{\times d})}s_{\mu}(1^{\times d})\sum_{\tau \in S_k}\chi^{\lambda}(\pi^{-1}\tau)\chi^{\mu}(\tau^{-1}\sigma) \\
&= \frac{1}{k!}\sum_{\lambda,\mu \vdash k}\frac{f^{\lambda}}{s_{\lambda}(1^{\times d})}s_{\mu}(1^{\times d})\chi^{\lambda}(\pi^{-1}\sigma)\delta_{\lambda\mu} \\
&= \frac{1}{k!}\sum_{\lambda \vdash k}f^{\lambda}\chi^{\lambda}(\pi^{-1}\sigma) = \delta_{\pi\sigma}.
\end{aligned}
$$

18

The third equality above is the well-known relation $\sum_{\mu \vdash k} s_\mu(1^{\times d}) \chi^\mu(\pi) = d^{c(\pi)}$ [7], and the fourth and sixth are character orthogonality relations (observing that $f^\lambda = \chi^\lambda(e)$). $\qquad\square$

References

[1] N. Adrianov. An analog of the Harer-Zagier formula for unicellular bicolored maps. *Functional Analysis and Its Applications*, 31(3), 1997.

[2] G. Amosov and A. Holevo. On the multiplicativity conjecture for quantum channels. *Theor. Probab. Appl.*, 47:143146, 2002. `math-ph/0103015`.

[3] G. Amosov, A. Holevo, and R. Werner. On some additivity problems in quantum information theory, 2000. `math-ph/0003002`.

[4] G. Aubrun. Partial transposition of random states and non-centered semicircular distributions, 2010. `arXiv:1011.0275`.

[5] G. Aubrun, S. Szarek, and E. Werner. Non-additivity of Renyi entropy and Dvoretzky's theorem. *J. Math. Phys.*, 51:022102, 2010. `arXiv:0910.1189`.

[6] G. Aubrun, S. Szarek, and E. Werner. Hastings' additivity counterexample via Dvoretzky's theorem. *Comm. Math. Phys.*, 305:85–97, 2011. `arXiv:1003.4925`.

[7] K. Audenaert. A digest on representation theory of the symmetric group, 2006. Available at `http://personal.rhul.ac.uk/usah/080/QITNotes_files/Irreps_v06.pdf`.

[8] T. Banica and I. Nechita. Asymptotic eigenvalue distributions of block-transposed Wishart matrices, 2011. `arXiv:1105.2556`.

[9] P. Biane. Some properties of crossings and partitions. *Discrete Mathematics*, 175:41–53, 1997.

[10] M. Christandl, N. Schuch, and A. Winter. Entanglement of the antisymmetric state, 2009. `arXiv:0910.4151`.

[11] M. Christandl, N. Schuch, and A. Winter. Highly entangled states with almost no secrecy. *Phys. Rev. Lett.*, 104:240405, 2010. `arXiv:1101.4522`.

[12] B. Collins. Moments and cumulants of polynomial random variables on unitary groups, the Itzykson-Zuber integral and free probability. *Int. Math. Res. Not.*, 17:953–982, 2003. `math-ph/0205010`.

[13] B. Collins, M. Fukuda, and I. Nechita. Towards a state minimizing the output entropy of a tensor product of random quantum channels, 2011. `arXiv:1111.6269`.

[14] B. Collins, C. González-Guillén, and David Pérez-García. Matrix product states, random matrix theory and the principle of maximum entropy, 2012. `arXiv:1201.6324`.

[15] B. Collins and I. Nechita. Eigenvalue and entropy statistics for products of conjugate random quantum channels. *Entropy*, 12:1612–1631, 2010. `arXiv:1006.3247`.

[16] B. Collins and I. Nechita. Random quantum channels I: Graphical calculus and the Bell state phenomenon. *Comm. Math. Phys.*, 297(2):345–370, 2010. `arXiv:0905.2313`.

[17] B. Collins and I. Nechita. Gaussianization and eigenvalue statistics for random quantum channels (III). *Ann. Appl. Probab.*, 21(3):1136–1179, 2011. `arXiv:0910.1768`.

[18] B. Collins and I. Nechita. Random quantum channels II: Entanglement of random subspaces, Renyi entropy estimates and additivity problems. *Advances in Mathematics*, 226(2):1181–1201, 2011. `arXiv:0906.1877`.

[19] B. Collins, I. Nechita, and D. Ye. The absolute positive partial transpose property for random induced states, 2011. `arXiv:1108.1935`.

[20] B. Collins and P. Śniady. Integration with respect to the Haar measure on unitary, orthogonal and symplectic group. *Comm. Math. Phys.*, 264:773–795, 2006. `math-ph/0402073`.

[21] T. Cubitt, A. W. Harrow, D. Leung, A. Montanaro, and A. Winter. Counterexamples to additivity of minimum output p-Renyi entropy for p close to 0. *Comm. Math. Phys.*, 284:281–290, 2008. `arXiv:0712.3628`.

[22] A. Grudka, M. Horodecki, and L. Pankowski. Constructive counterexamples to additivity of minimum output Rényi entropy of quantum channels for all $p > 2$. *J. Phys. A: Math. Gen.*, 43:425304, 2010. `arXiv:0911.2515`.

[23] A. Harrow. Permutations are sort of orthogonal, 2012.

[24] A. Harrow and A. Montanaro. An efficient test for product states, with applications to quantum Merlin-Arthur games. In *Proc. 51st Annual Symp. Foundations of Computer Science*, pages 633–642, 2010. `arXiv:1001.0017`.

[25] M. B. Hastings. Superadditivity of communication capacity using entangled inputs. *Nature Physics*, 5:255, 2009. `arXiv:0809.3972`.

[26] P. Hayden. The maximal p-norm multiplicativity conjecture is false, 2007. `arXiv:0707.3291`.

[27] P. Hayden and A. Winter. Counterexamples to the maximal p-norm multiplicativity conjecture for all p > 1. *Comm. Math. Phys.*, 284(1):263–280, 2008.

[28] H. Kobayashi, K. Matsumoto, and T. Yamakami. Quantum Merlin-Arthur proof systems: are multiple Merlins more helpful to Arthur? In *Proc. ISAAC '03*, pages 189–198, 2003. `quant-ph/0306051`.

[29] Keiji Matsumoto. Some new results and applications of additivity problem of quantum channel. Poster at QIP'05 conference, 2005.

[30] S. Matsumoto and J. Novak. Unitary matrix integrals, primitive factorizations, and Jucys-Murphy elements. In *Discrete Math. Theor. Comput. Sci., FPSAC 2010*, pages 403–412, 2010. `arXiv:1005.0151`.

[31] S. Matsumoto and J. Novak. Jucys-Murphy elements and unitary matrix integrals. *International Mathematics Research Notices*, 2012. `arXiv:0905.1992`.

[32] A. Nica and R. Speicher. *Lectures on the combinatorics of free probability*. Volume 335 of London Mathematical Society Lecture Note Series. Cambridge University Press, 2006.

[33] J. Novak. Jucys-Murphy elements and the unitary Weingarten function. *Banach Center Publ.*, 89:231–235, 2010.

[34] P. W. Shor. Equivalence of additivity questions in quantum information theory. *Comm. Math. Phys.*, 246(3):453–472, 2004. `quant-ph/0305035`.

[35] R. Werner and A. Holevo. Counterexample to an additivity conjecture for output purity of quantum channels, 2002. `quant-ph/0203003`.

[36] A. Winter. The maximum output p-norm of quantum channels is not multiplicative for any p>2, 2007. `arXiv:0707.0402`.