Additivity and Distinguishability of Random Unitary Channels

Bill Rosgen

*Institute for Quantum Computing and School of Computer Science, University of Waterloo, Canada**

(Dated: October 15, 2008)

A random unitary channel is one that is given by a convex combination of unitary channels. It is shown that the conjectures on the additivity of the minimum output entropy and the multiplicativity of the maximum output p-norm can be equivalently restated in terms of random unitary channels. This is done by constructing a random unitary approximation to a general quantum channel. This approximation can be constructed efficiently, and so it is also applied to the computational problem of distinguishing quantum circuits. It is shown that the problem of distinguishing random unitary circuits is as hard as the problem of distinguishing general mixed state circuits, which is complete for the class of problems that have quantum interactive proof systems.

I. INTRODUCTION

A quantum channel is *random unitary* if it can be decomposed into the probabilistic application of one of a finite set of unitary operations. More formally, Φ is random unitary if there exist unitary operators $U_1, \ldots, U_n$ and a probability distribution $p_1, \ldots, p_n$ such that

$$\Phi(X) = \sum_{i=1}^{n} p_i U_i X U_i^*. \tag{1}$$

It has been shown by Gregoratti and Werner[1] that the random unitary channels describe exactly the noise processes that can be corrected using classical information obtained by measuring the environment. For channels on qubits the random unitary channels are exactly the unital channels, but for larger dimensions this is not the case[2,3,4]. Audenaert and Scheel have recently provided necessary and sufficient conditions for a channel to be random unitary[5]. Buscemi has also provided an upper bound on the number of unitaries needed for a random unitary decomposition[6], as in Equation (1).

A natural question arises from this class of channels: is the additivity conjecture simplified when restricted to the random unitary channels? In the present paper this question is answered in the negative. This is done using a method to approximate an arbitrary quantum channel by a random unitary one. A recent survey on the additivity conjecture and a few related conjectures that we will also consider can be found in Ref. 7. One such conjecture is the question of the additivity of the minimum output entropy. The approximation scheme constructed here is also used to show that this conjecture can be restricted to the random unitary channels with no loss of generality, extending the results of Fukuda[8] on unital channels. In addition to these results, this approximation scheme implies the computational hardness of distinguishing mixed-state quantum circuits that implement random unitary channels.

All Hilbert spaces considered here are finite dimensional and denoted by calligraphic letters $\mathcal{H}, \mathcal{K}, \ldots$. The set of all (bounded) linear operators on a space $\mathcal{H}$ is denoted by $\mathbf{L}(\mathcal{H})$. The set of mixed states, or density operators, which are the positive semidefinite operators with unit trace on the space $\mathcal{H}$, is denoted $\mathbf{D}(\mathcal{H})$. The set $\mathbf{D}(\mathcal{H})$ is compact and convex. The extreme points of $\mathbf{D}(\mathcal{H})$ are called pure states, which are given by the rank-one projectors $|\psi\rangle\langle\psi|$ onto states of $\mathcal{H}$. The notation $\mathbf{T}(\mathcal{H}, \mathcal{K})$ is used for the set of admissible maps from $\mathbf{L}(\mathcal{H})$ to $\mathbf{L}(\mathcal{K})$. An admissible map, hereafter be called a channel, is one that is completely positive and trace preserving. The notation $\tilde{I}_{\mathcal{H}}$ will be used to denote the maximally mixed state on the space $\mathcal{H}$, i.e. $\tilde{I}_{\mathcal{H}} = I_{\mathcal{H}} / \dim \mathcal{H}$.

The entropy of a quantum state ρ is given by $S(\rho) = -\operatorname{tr}\rho\log\rho$. This quantity can be seen as a measure of purity, as ranges between $S(|\psi\rangle\langle\psi|) = 0$ for any pure state and $S(\tilde{I}_{\mathcal{H}}) = \log\dim\mathcal{H}$ for the maximally mixed state. A description of many of the fundamental properties of the von Neumann entropy can be found in Ref. 9. Of particular significance here is the concavity, which is given by $S(\sum_i p_i\rho_i) \geq \sum_i p_i S(\rho_i)$.

The classical capacity of a single use of a channel Φ is given by the χ-capacity[10,11]

$$C_\chi(\Phi) = \max\left[S(\Phi(\sum_i p_i\rho_i)) - \sum_i p_i S(\Phi(\rho_i))\right],$$

where the maximum is taken over all convex mixtures $\sum_i p_i\rho_i$ of quantum states. This quantity is also referred to as the "one-shot" or "one-step" capacity of Φ. A central question in quantum information theory is whether is quantity is *additive*, i.e. does entangling inputs across multiple uses of the channel increase the capacity? This question was first raised in Ref. 12, and the standing conjecture is that

$$C_\chi(\Phi\otimes\Psi) \stackrel{?}{=} C_\chi(\Phi) + C_\chi(\Psi), \tag{2}$$

which is the statement that entangled inputs do not increase the classical information carrying capacity of quantum channels.

Closely related to the additivity of the χ-capacity is the question of the additivity of the minimum output entropy, defined by $S_{\min}(\Phi) = \min_\rho S(\Phi(\rho))$, where the minimization is over all density operators. The additivity of this quantity, given by

$$S_{\min}(\Phi\otimes\Psi) \stackrel{?}{=} S_{\min}(\Phi) + S_{\min}(\Psi), \tag{3}$$

was first studied by King and Ruskai[13], who attribute this conjecture to Shor. This conjecture is connected to the additivity of the χ-capacity by a result of Shor[14] that shows that both of these conjectures are equivalent to a third conjecture: the strong superadditivity of the entanglement of formation.

One potential path to resolving these conjectures lies in yet another conjecture. This is the conjectured multiplicativity of the maximum output p-norm, first stated by Amosov, Holevo, and Werner[15]. This conjecture involves the maximum output p-norm of a quantum channel $\Phi \in \mathbf{T}(\mathcal{H},\mathcal{K})$. This quantity, for $p\in[1,\infty)$ is given by

$$\nu_p(\Phi) = \max_{\rho\in\mathbf{D}(\mathcal{H})} \|\Phi(\rho)\|_p = \max_{\rho\in\mathbf{D}(\mathcal{H})} \left(\operatorname{tr}|\Phi(\rho)|^p\right)^{\frac{1}{p}},$$

which is simply the p-norm of the singular values of $\Phi(\rho)$, maximized over all inputs ρ. This is extended to the case of $p=\infty$ in the usual way by replacing the sum in the p-norm with a maximization over the singular values of $\Phi(\rho)$. The conjecture of Amosov, Holevo, and Werner[15] corresponding to this quantity is that it is multiplicative with respect to the tensor product of two channels, i.e. that

$$\nu_p(\Phi\otimes\Psi) \stackrel{?}{=} \nu_p(\Phi)\,\nu_p(\Psi). \tag{4}$$

This conjecture implies the additivity of the minimum output entropy, given in Equation (3), as the derivative of $\nu_p(\Phi)$ for p approaching one gives an expression for $S_{\min}(\Phi)$. The multiplicativity conjecture is known to fail for any fixed $p > 1$[16]. This does not eliminate interest in this quantity, however, as the conjecture can be weakened to ask if, for given Φ,Ψ, does there exist a sequence $\{p_n\}$ converging to one with $p_n > 1$ for which Equation (4) holds? This weakened conjecture still implies the additivity conjectures given by Equations (2) and (3).

There are two main approaches to resolving these conjectures. The first of these is to consider restricted classes of channels for which the conjectures can be shown true, in the hope that such a strategy will yield some insight into the general problem. The entanglement breaking channels are one such class, with the additivity of the minimum output entropy shown by Shor[17] and the multiplicativity of the p-norm shown by King[18]. This class of channels contains many important channels, such as the completely depolarizing channel. Another class of channels for which the additivity and multiplicativity conjectures are known to hold is the class of unital channels on qubits[19], which is particularly interesting in the context of this paper, as the unital qubit channels are exactly the random unitary channels on qubits.

The second approach taken toward resolving these conjectures is to show that they remain equivalent when restricted to certain classes of channels, in the hope that the restrictions will aid search for either a proof or a counterexample. One such equivalence, due to Fukuda[8], shows that the conjectures on the additivity of the minimum output entropy and the multiplicativity of the p-norm lose no generality when they are restricted to unital channels, which are the channels $\Phi \in \mathbf{T}(\mathcal{H}, \mathcal{H})$ satisfying $\Phi(I_{\mathcal{H}}) = I_{\mathcal{H}}$. The same approach is taken by Fukuda and Wolf[20], who show that, with no loss in generality, these conjectures can be restricted to two copies of the same channel, i.e. letting $\Psi = \Phi$ in Equations (3) and (4).

In the present paper we take the second approach, extending the results of Fukuda[8] on unital channels to the random unitary case. This is done by constructing a random unitary approximation to an arbitrary channel in Section II, and showing that this approximation is not too far from the original channel in Section III. In Section IV it is shown how this approximation can be used to restrict the multiplicativity conjecture to the random unitary case. This is of less interest than the same result for additivity conjecture on the minimum output entropy, which appears in Section V, but the argument for the p-norm case is presented first, as it is both similar to and simpler than the argument for the minimum output entropy. Finally, the paper concludes with a discussion of a circuit implementation of the approximation scheme, in Section VI, which is used in Section VII to show that the computational problem of distinguishing two mixed state quantum circuits is made no easier by adding the restriction that the circuits implement random unitary transformations.

II. RANDOM UNITARY APPROXIMATION

Stinespring's Dilation Theorem[21] states that any quantum channel Φ can be written as

$$\Phi(X) = \mathrm{tr}_{\mathcal{B}}\, U(|0\rangle\langle 0| \otimes X)U^*, \tag{5}$$

for U a unitary operation. There are two operations in this representation that are not random unitary, as defined by Equation (1). These operations are the partial trace over the system $\mathcal{B}$, and the introduction of the ancillary system in the state $|0\rangle$. To find an approximation to Φ that is random unitary, we will need to deal with both of these operations.

Fixing notation, let Φ be a completely positive and trace preserving map from $\mathbf{L}(\mathcal{H})$ to $\mathbf{L}(\mathcal{K})$. Representing Φ as in Equation 5, let $\mathcal{A}$ be the space containing the ancillary space starting in the $|0\rangle$ state, and let $\mathcal{B}$ be the space that is traced out. This implies that U is a unitary map from $\mathcal{A} \otimes \mathcal{H}$ to $\mathcal{K} \otimes \mathcal{B}$.

To avoid tracing out the system in the space $\mathcal{B}$ the partial trace may be replaced by the operation $N_{\mathcal{B}}$ that takes the state in $\mathcal{B}$ to the completely mixed state. This operation can be implemented as a random unitary operation as the uniform mixture of the discrete Weyl operators)[22,23,24]. The discrete Weyl operators are also known as the generalized Pauli operators, as they are one generalization of the Pauli X and Z operators to higher dimensional systems. It is not difficult to

see that for ρ a density matrix on $\mathcal{K} \otimes \mathcal{B}$,

$$N_{\mathcal{B}}(\rho) = (\mathrm{tr}_{\mathcal{B}}\, \rho) \otimes \tilde{I}_{\mathcal{B}}. \tag{6}$$

This implies that if the system to be traced out instead has $N_{\mathcal{B}}$ applied to it, the resulting state is the same, up to a tensor factor of a maximally mixed state in the space $\mathcal{B}$. This factor will change both the minimum output entropy and the maximum output p-norm by a fixed value that will not affect the additivity or multiplicativity of these quantities.

Replacing the introduction of the ancillary space $\mathcal{A}$ with a random unitary operation is more complicated. The strategy employed is to expand the input of the transformation to include the space $\mathcal{A}$. The input state of this system will not, in general, be the desired state $|0\rangle$, and so an additional operation is needed to force this to be the case for any input that minimizes the output entropy. As we are only interested in the minimum output entropy and the maximum output p-norm, those inputs on which the resulting channel produces an output with high entropy can be ignored, as they will be far from those inputs that achieve the minimum (resp. maximum).

To this end, the ideal operation to perform this forcing does not alter any input state of the form $|0\rangle\langle 0| \otimes \sigma$, but takes any orthogonal state to the completely mixed state $\tilde{I}_{\mathcal{A}\otimes\mathcal{H}}$. This operation is, unfortunately, not random unitary, as it is not unital. A closely related strategy that is random unitary is to project the input state either onto the subspace $S_0 = |0\rangle \otimes \mathcal{H}$ or the orthogonal subspace $S_0^{\perp} = |0\rangle^{\perp} \otimes \mathcal{H}$. This projection is then be followed by a mixing operation on the subspace $S_0^{\perp}$. This mixing process is be introduced first. It is given by the channel M that does not affect the subspace S_0 but completely mixes $S_0^{\perp}$. More concretely, on a state $\rho = q\rho_{S_0} + (1-q)\rho_{S_0^{\perp}}$ where $\rho_{S_0} = |0\rangle\langle 0| \otimes \sigma$ is a density operator on S_0 and $\rho_{S_0^{\perp}}$ a density operator on $S_0^{\perp}$, the output of M is given by

$$\begin{aligned} M(\rho) = qM(\rho_{S_0}) + (1-q)M(\rho_{S_0^{\perp}}) &= q\rho_{S_0} + (1-q)\tilde{I}_{S_0^{\perp}} \\ &= q|0\rangle\langle 0| \otimes \sigma + (1-q)\frac{I_{\mathcal{A}} - |0\rangle\langle 0|}{\dim \mathcal{A} - 1} \otimes \tilde{I}_{\mathcal{H}}. \end{aligned} \tag{7}$$

The channel M can be implemented as a random unitary channel in the same way as the completely depolarizing channel: a uniform mixture of the discrete Weyl operators, except here these operators are taken over the subspace $S_0^{\perp}$.

This channel is not exactly the desired one. If the output of M on $\rho_{S_0^{\perp}}$ in Equation (7) were the completely mixed state on $\mathcal{A} \otimes \mathcal{H}$ and not the subspace $S_0^{\perp}$ then this process would create an essentially error-free random unitary approximation of the original channel (for the purpose of minimizing the output entropy). Fortunately, the error involved at this step will be shown, in Lemma 4, to be $O(1/\dim \mathcal{A})$, and so by taking the space $\mathcal{A}$ large enough we will be able to approximate the ideal case.

There is one further convenient property that this mixing channel does not satisfy: it does not remove coherences between the subspaces S_0 and $S_0^{\perp}$. If the channel M had this property, then an equation similar to Equation (7) would hold for all input states ρ, not just those states that have no entanglement between the subspaces S_0 and $S_0^{\perp}$. This property will be essential to the analysis that follows, and so the additional operation that decoheres these two subspaces needs to be applied before the mixing operation M. This operation, D, can be implemented by leaving the state unchanged with probability one half and applying a unitary U with probability one half, where the action of U on basis states is given by

$$U|i\rangle = \begin{cases} |i\rangle & \text{if } |i\rangle \in S_0, \\ -|i\rangle & \text{if } |i\rangle \in S_0^{\perp}. \end{cases}$$

In other words, U applies a phase of -1 to states in $S_0^\perp$ and does not change states in S_0. When U is applied with probability one half the result is complete dephasing between the two subspaces. This can be seen by observing that this is the restriction of the completely dephasing channel, as considered in Ref. 25, to a system with only two orthogonal states, which are here given by the subspaces S_0 and $S_0^\perp$. Alternately, when this is applied to a density matrix expressed in the computational basis, the result is, by a simple calculation, the zeroing of the off-diagonal elements of the first row and column. When this operation, D, is applied to a density operator $\rho \in \mathbf{D}(\mathcal{A} \otimes \mathcal{H})$, the result is

$$D(\rho) = q\rho_{S_0} + (1-q)\rho_{S_0^\perp} = q|0\rangle\langle 0| \otimes \sigma + (1-q)\rho_{S_0^\perp}, \tag{8}$$

where $\rho_{S_0} = |0\rangle\langle 0| \otimes \sigma$ is a density operator on the subspace $S_0 = |0\rangle \otimes H$, $\rho_{S_0^\perp}$ is a density operator on $S_0^\perp$, and $0 \leq q \leq 1$.

Combining Equations (7) and (8), the output of D followed by M on a density operator ρ on $\mathcal{A} \otimes \mathcal{H}$ is given by a state of the form

$$(M \circ D)(\rho) = qM(|0\rangle\langle 0| \otimes \sigma) + (1-q)M(\rho_{S_0^\perp}) = q|0\rangle\langle 0| \otimes \sigma + (1-q)\frac{I_{\mathcal{A}} - |0\rangle\langle 0|}{\dim \mathcal{A} - 1} \otimes \tilde{I}_{\mathcal{H}}.$$

This operation $M \circ D$ will be used as a way to force any input that results in a low output entropy to be close to the subspace S_0 of inputs having the 'ancilla' space $\mathcal{A}$ in the desired $|0\rangle$ state. On these inputs the constructed random unitary channel will behave in a similar way to the original channel that is being approximated. On inputs that are far from this subspace, the resulting state has high entropy, and so it will not be close to a state minimizing the output entropy. As $M \circ D$ mixes the input, conditional on the state being in the subspace $S_0^\perp$, this operation will be referred to as the conditional mixing procedure.

Putting all of these pieces together, given a channel $\Phi(\rho) = \mathrm{tr}_{\mathcal{B}} U(\rho \otimes |0\rangle\langle 0|)U^*$, the random unitary approximation Φ' is constructed as

$$\Phi'(\rho) = N_{\mathcal{B}}\left(U\left[(M \circ D)(\rho)\right]U^*\right), \tag{9}$$

which, more plainly, is simply the application of the conditional mixing procedure, the unitary operation from a Stinespring dilation of Φ, and finally the completely mixing channel to the space that would have been traced out by Φ. As the composition of random unitary transformations remains random unitary, the channel Φ' will be a random unitary channel.

It will be useful to observe that the channel Φ' specified in Equation (9) can be used to simulate the channel Φ. This occurs when the input $|0\rangle\langle 0| \otimes \sigma$, i.e. an input in the space S_0, is provided to Φ'. This is argued in the following proposition.

Proposition 1. *Let Φ be a quantum channel from $\mathbf{L}(\mathcal{H})$ to $\mathbf{L}(\mathcal{K})$. If Φ' is the random unitary channel mapping $\mathbf{L}(\mathcal{A} \otimes \mathcal{H})$ to $\mathbf{L}(\mathcal{K} \otimes \mathcal{B})$ that is constructed from Φ in Equation* (9)*, then*

$$\Phi'(|0\rangle\langle 0| \otimes \sigma) = \Phi(\sigma) \otimes \tilde{I}_{\mathcal{B}}.$$

Proof. Notice that both D and M do not affect this input: the decoherence operation D does not affect the state as it is in the subspace S_0 and M does not affect the state by Equation (7). Thus, the output of the channel Φ' is

$$\begin{aligned}
\Phi'(|0\rangle\langle 0| \otimes \sigma) &= N_{\mathcal{B}}\left(U\left[(M \circ D)(|0\rangle\langle 0| \otimes \sigma)\right]U^*\right) \\
&= N_{\mathcal{B}}\left(U(|0\rangle\langle 0| \otimes \sigma)U^*\right) \\
&= \mathrm{tr}_{\mathcal{B}}\left(U(|0\rangle\langle 0| \otimes \sigma)U^*\right) \otimes \tilde{I}_{\mathcal{B}} \\
&= \Phi(\sigma) \otimes \tilde{I}_{\mathcal{B}},
\end{aligned}$$

where the penultimate equality is an application of Equation (6). □

Combining this proposition with Equation (8) that demonstrates the effect of the $M \circ D$ on states not of this form, and the observation that applying $M \circ D$ twice has no further effect than applying it once, the output of Φ' on an arbitrary input state ρ is given by

$$\Phi'(\rho) = p\Phi'(|0\rangle\langle 0| \otimes \sigma) + (1-p)\Phi'(\rho_{S_0^\perp}) = p\Phi(\sigma) \otimes \tilde{I}_{\mathcal{B}} + (1-p)\Phi'(\rho_{S_0^\perp}), \tag{10}$$

where as in Equation (8) $\rho_{S_0^\perp}$ is a density operator on the subspace $S_0^\perp$ of inputs orthogonal to those with the state $|0\rangle$ on the space $\mathcal{A}$. The major technical portion of the results that follow lies in bounding the distance from the maximally mixed state of the second term in this equation, from which most of the results follow.

III. PROPERTIES OF THE CONSTRUCTED CHANNEL

In this section some basic results on the random unitary approximation of a channel are shown. Throughout this section, and the following two sections Φ will represent the original transformation and Φ' will represent the random unitary transformation constructed from it as in Equation (9).

As a first step to showing that Φ' approximates Φ it is shown that random unitary transformations cannot increase the distance of a state from the completely mixed state. This lemma shows that the output of a random unitary transformation cannot be more pure than the input. The extra space $\mathcal{B}$ appearing in this lemma will correspond to a reference system needed for the results in Section VII.

Lemma 2. *Let $|||\cdot|||$ be a unitarily invariant norm on $\mathbf{L}(\mathcal{A} \otimes \mathcal{B})$. If $\Psi \in \mathbf{T}(\mathcal{A}, \mathcal{A})$ is random unitary, then for any $\rho \in \mathbf{D}(\mathcal{A} \otimes \mathcal{B})$*

$$\left|\left|\left|(\Psi \otimes I_{\mathcal{B}})(\rho) - \tilde{I}_{\mathcal{A}} \otimes \mathrm{tr}_{\mathcal{A}}\, \rho\right|\right|\right| \leq \left|\left|\left|\rho - \tilde{I}_{\mathcal{A}} \otimes \mathrm{tr}_{\mathcal{A}}\, \rho\right|\right|\right|$$

Proof. As Ψ is random unitary, let $\Psi(X) = \sum_i p_i U_i X U_i^*$ with the U_i unitary, $0 \leq p_i \leq 1$, and $\sum_i p_i = 1$. Using this decomposition

$$\begin{aligned}\left|\left|\left|(\Psi \otimes I_{\mathcal{B}})(\rho) - \tilde{I}_{\mathcal{A}} \otimes \mathrm{tr}_{\mathcal{A}}\, \rho\right|\right|\right| &= \left|\left|\left|\sum_i p_i (U_i \otimes I)\rho(U_i^* \otimes I) - \tilde{I}_{\mathcal{A}} \otimes \mathrm{tr}_{\mathcal{A}}\, \rho\right|\right|\right| \\ &\leq \sum_i p_i \left|\left|\left|(U_i \otimes I)\rho(U_i^* \otimes I) - \tilde{I}_{\mathcal{A}} \otimes \mathrm{tr}_{\mathcal{A}}\, \rho\right|\right|\right|.\end{aligned}$$

Using the fact that $U_i \tilde{I}_{\mathcal{A}} U_i^* = \tilde{I}_{\mathcal{A}}$, and the unitary invariance of the norm, this becomes

$$\begin{aligned}\sum_i p_i \left|\left|\left|(U_i \otimes I)(\rho - \tilde{I}_{\mathcal{A}} \otimes \mathrm{tr}_{\mathcal{A}}\, \rho)(U_i^* \otimes I)\right|\right|\right| &= \sum_i p_i \left|\left|\left|\rho - \tilde{I}_{\mathcal{A}} \otimes \mathrm{tr}_{\mathcal{A}}\, \rho\right|\right|\right| \\ &= \left|\left|\left|\rho - \tilde{I}_{\mathcal{A}} \otimes \mathrm{tr}_{\mathcal{A}}\, \rho\right|\right|\right|.\end{aligned}$$

Combining these equations yields the statement of the lemma. □

This lemma can be used to show not only that the conditional mixing procedure sends states in the subspace $S_0^\perp$ of states where the ancillary space is not in the $|0\rangle$ state to states that are almost completely mixed, but that the channel Φ' also has this behaviour. Before doing this, however, the Lemma is extended to the case of the von Neumann entropy, where the proof is essentially identical, with the exception that the triangle inequality is replaced by concavity.

Corollary 3. *If $\Psi \in \mathbf{T}(\mathcal{A}, \mathcal{A})$ is random unitary, and $\rho \in \mathbf{D}(\mathcal{A})$, then $S(\rho) \leq S(\Psi(\rho))$.*

Proof. Let $\Psi(\rho) = \sum_i p_i U_i \rho U_i^*$ as in Lemma 2. Using this notation, and the concavity of the von Neumann entropy

$$S(\Psi(\rho)) = S(\sum_i p_i U_i \rho U_i^*) \geq \sum_i p_i S(U_i \rho U_i^*) = \sum_i p_i S(\rho) = S(\rho),$$

where the unitary invariance of the von Neumann entropy has been used in the penultimate equality. □

The next lemma shows that when the input is in the subspace $S_0^\perp$ the output of Φ' is very close to completely mixed. The distance measure used in the lemma is the trace norm, but this can be applied to the case of the maximum output p-norm due to the fact that $\|\rho\|_{\mathrm{tr}} = \|\rho\|_1 \geq \|\rho\|_p$ for all $p \in [1, \infty]$. This lemma forms a significant portion of the proof of the main results on the additivity and multiplicativity conjectures.

Lemma 4. *On input states $\rho \in S_0^\perp$ the output of Φ' satisfies*

$$\left\| \Phi'(\rho) - \tilde{I}_{\mathcal{A}\otimes\mathcal{H}} \right\|_{\mathrm{tr}} \leq \frac{2}{\dim \mathcal{A}}.$$

Proof. On input $\rho \in S_0^\perp$ the operation D that introduces decoherence between S_0 and $S_0^\perp$has no effect. This implies that the output of $M \circ D$ on ρ is given by setting $q = 0$ in Equation (7), which is

$$\frac{1}{\dim \mathcal{A} - 1} \left(I_{\mathcal{A}} - |0\rangle\langle 0| \right) \otimes \tilde{I}_{\mathcal{H}}, \tag{11}$$

Setting $d = \dim \mathcal{A}$, we can then compute the distance from the completely mixed state as

$$\left\| \frac{I_{\mathcal{A}} - |0\rangle\langle 0|}{d-1} \otimes \tilde{I}_{\mathcal{H}} - \frac{I_{\mathcal{A}}}{d} \otimes \tilde{I}_{\mathcal{H}} \right\|_{\mathrm{tr}} = \left\| \frac{I_{\mathcal{A}} - d|0\rangle\langle 0|}{d(d-1)} \right\|_{\mathrm{tr}} \leq \frac{d-1}{d(d-1)} + \frac{d-1}{d(d-1)} = \frac{2}{d}. \tag{12}$$

Finally, by noting that the remainder of the transformation Φ' is random unitary, an application of Lemma 2 yields the desired bound. □

Once again we can extend this result to the case of the von Neumann entropy. In this case we do not simply repeat the same method of proof, but instead extend the result to the entropy using the relationship between the trace distance and the entropy given by Fannes' inequality. This extension requires that $\dim \mathcal{A} \geq \dim \mathcal{H}$, but this can be assured by considering only those dilations of Φ with this property. In the following corollary we set $m = \log \dim \mathcal{A}$ for convenience, but no assumption is made that this value is an integer.

Corollary 5. *Let $m = \log \dim \mathcal{A}$. If $\dim \mathcal{A} \geq \dim \mathcal{H}$, $m \geq 3$, and $\rho \in S_0^\perp$, then*

$$S(\Phi'(\rho)) \geq S(\tilde{I}_{\mathcal{A}\otimes\mathcal{H}}) - \frac{m}{2^{m-3}}.$$

Proof. Let $\hat{\rho}$ be the state in Equation (11) of the proof of Lemma 4. This is the state after the conditional mixing procedure of Φ' has been applied to the input. For convenience, set $\delta = \left\| \hat{\rho} - \tilde{I}_{\mathcal{A}\otimes\mathcal{H}} \right\|_{\mathrm{tr}}$. By Equation (12), this trace distance between satisfies $\delta \leq 2/\dim \mathcal{A} = 2^{-(m-1)}$. Applying Fannes' inequality[26] (see also Ref. 9) yields

$$\begin{aligned} \left| S(\hat{\rho}) - S(\tilde{I}_{\mathcal{A}\otimes\mathcal{H}}) \right| &\leq (\log \dim \mathcal{H} \otimes \mathcal{A})\delta - \delta \log \delta \\ &\leq \frac{\log \dim \mathcal{H} + m}{2^{m-1}} + \frac{m}{2^{m-1}} \\ &\leq \frac{m}{2^{m-3}}, \end{aligned}$$

where the first inequality is by Fannes' inequality, and the second inequality follows from fact that $-x \log x$ is monotone for $x \in [0, 1/e]$, and $\delta \leq 2^{-(m-1)} < 1/e$ whenever $m \geq 3$.

By Corollary 3 applying the remainder of Φ' to the state $\hat{\rho}$ cannot decrease the entropy, as this portion of Φ' is random unitary, and so the previous equation implies that

$$S(\Phi'(\rho)) \geq S(\hat{\rho}) \geq S(\tilde{I}_{\mathcal{A}\otimes\mathcal{H}}) - \frac{m}{2^{m-3}},$$

as in the statement of the lemma. □

IV. MULTIPLICATIVITY OF RANDOM UNITARY TRANSFORMATIONS

In this section the construction of Section II is used to show some results about the multiplicativity of the maximum output p-norm and random unitary channels. The main result is that, for $p < \infty$, the p-norm of the tensor product of two channels is multiplicative if and only if the p-norm is multiplicative on the random unitary approximations to these channels.

As a first step towards this theorem, it is shown that the random unitary channel Φ' constructed from Φ in Equation (9) is a good approximation with respect to the p-norm.

Theorem 6. *If $\Phi \in \mathbf{T}(\mathcal{H}, \mathcal{K})$, then the random unitary $\Phi' \in \mathbf{T}(\mathcal{A} \otimes \mathcal{H}, \mathcal{K} \otimes \mathcal{B})$ satisfies*

$$\nu_p(\Phi) \leq \frac{\nu_p(\Phi')}{\left\|\tilde{I}_{\mathcal{B}}\right\|_p} \leq \nu_p(\Phi) + \frac{2 \dim \mathcal{B}}{\dim \mathcal{A}}.$$

Proof. For convenience, let $d = \dim \mathcal{A}$. The first inequality is simple: $\Phi'(|0\rangle\langle 0| \otimes \rho) = \Phi(\rho) \otimes \tilde{I}_{\mathcal{B}}$ by Proposition 1, and so it is clear that $\nu_p(\Phi)\left\|\tilde{I}_{\mathcal{B}}\right\|_p \leq \nu_p(\Phi')$, by the multiplicativity of $\|\cdot\|_p$ with respect to the tensor product of states.

To prove the second inequality let $\rho \in \mathbf{D}(\mathcal{A} \otimes \mathcal{H})$ be a state such that

$$\nu_p(\Phi') = \left\|\Phi'(\rho)\right\|_p.$$

Such a state exists by the compactness of $\mathbf{D}(\mathcal{A} \otimes \mathcal{H})$. The output of Φ' on ρ is given by Equation (10), applying the triangle inequality to this yields

$$\left\|\Phi'(\rho)\right\|_p = \left\|q\Phi(\sigma) \otimes \tilde{I}_{\mathcal{B}} + (1-q)\Phi'(\rho_{S_0^\perp})\right\|_p \leq q\left\|\Phi(\sigma) \otimes \tilde{I}_{\mathcal{B}}\right\|_p + (1-q)\left\|\Phi'(\rho_{S_0^\perp})\right\|_p.$$

Applying Lemma 4 to this gives

$$\left\|\Phi'(\rho)\right\|_p \leq q\left\|\Phi(\sigma) \otimes \tilde{I}_{\mathcal{B}}\right\|_p + (1-q)\left(\left\|\tilde{I}_{\mathcal{K}\otimes\mathcal{B}}\right\|_p + \frac{2}{d}\right).$$

Then, as the norm $\|\cdot\|_p$ is multiplicative with respect to the tensor product of states, and $\left\|\tilde{I}_{\mathcal{K}}\right\|_p \leq \|\xi\|_p$ for any state $\xi \in \mathbf{D}(\mathcal{K})$,

$$\left\|\Phi'(\rho)\right\|_p \leq q\left\|\Phi(\sigma)\right\|_p\left\|\tilde{I}_{\mathcal{B}}\right\|_p + (1-q)\left(\left\|\tilde{I}_{\mathcal{K}}\right\|_p\left\|\tilde{I}_{\mathcal{B}}\right\|_p + \frac{2}{d}\right) \leq \left\|\Phi(\sigma)\right\|_p\left\|\tilde{I}_{\mathcal{B}}\right\|_p + \frac{2}{d}.$$

Finally, by the choice of the input ρ

$$\nu_p(\Phi') = \left\|\Phi'(\rho)\right\|_p \leq \nu_p(\Phi)\left\|\tilde{I}_{\mathcal{B}}\right\|_p + \frac{2}{d},$$

which completes the proof of the theorem, as $\left\|\tilde{I}_{\mathcal{B}}\right\|_p = \dim \mathcal{B}^{1/p-1}$. □

With this approximation result, the main theorem on the maximum output p-norm can be shown. This extends part of the work done by Fukuda[8] on unital channels to the random unitary case.

Theorem 7. *If $\Phi, \Psi \in \mathbf{T}(\mathcal{H}, \mathcal{K})$ and $p \in [1, \infty)$, then*

$$\nu_p(\Phi \otimes \Psi) = \nu_p(\Phi)\, \nu_p(\Psi)$$

if

$$\nu_p(\Phi'_d \otimes \Psi) = \nu_p(\Phi'_d)\, \nu_p(\Psi),$$

for all sufficiently large d, where Φ'_d is the random unitary extension of the channel Φ obtained by applying the construction of Section II to a Stinespring dilation of Φ using a d-dimensional ancillary space.

Proof. As adding ancillary space to Φ' increases both $\dim \mathcal{A}$ and $\dim \mathcal{B}$, by taking $d = \dim \mathcal{A}$ large enough is can be assumed that $\dim \mathcal{B} \leq 2d$. Let $\epsilon > 0$, and choose d so that $2 \dim \mathcal{B}^{1-1/p}/d \leq 2/d^{1/p} < \epsilon$. Then, as $\Phi'_d(|0\rangle\langle 0| \otimes \rho) = \Phi(\rho) \otimes \tilde{I}_{\mathcal{B}}$ by Proposition 1,

$$\nu_p(\Phi \otimes \Psi) \leq \frac{\nu_p(\Phi'_d \otimes \Psi)}{\left\| \tilde{I}_{\mathcal{B}} \right\|_p}.$$

By assumption, this second quantity is multiplicative, so that

$$\nu_p(\Phi \otimes \Psi) \leq \frac{\nu_p(\Phi'_d \otimes \Psi)}{\left\| \tilde{I}_{\mathcal{B}} \right\|_p} = \frac{\nu_p(\Phi'_d)\, \nu_p(\Psi)}{\left\| \tilde{I}_{\mathcal{B}} \right\|_p} \leq \left[\nu_p(\Phi) + \frac{2}{d^{1/p}} \right] \nu_p(\Psi) < \nu_p(\Phi)\, \nu_p(\Psi) + \epsilon,$$

where the penultimate inequality is an application of Theorem 6. As epsilon was chosen arbitrarily, the multiplicativity of $\nu_p(\Phi'_d)$ for all large enough d implies the multiplicativity of $\nu_p(\Phi)$. □

V. MINIMUM OUTPUT ENTROPY AND RANDOM UNITARY CHANNELS

These results on the multiplicativity of the p-norm can be extended to the additivity of the minimum output entropy. This is done using a similar method of proof as the results of the previous section. The following theorem demonstrates that the random unitary channel Φ' constructed in Equation (9) forms a good approximation of the original channel Φ, from which the result on the additivity will follow directly.

Theorem 8. *If $\Phi \in \mathbf{T}(\mathcal{H}, \mathcal{K})$, then the random unitary $\Phi' \in \mathbf{T}(\mathcal{A} \otimes \mathcal{H}, \mathcal{K} \otimes \mathcal{B})$ satisfies*

$$S_{\min}(\Phi) \geq S_{\min}(\Phi') - \log \dim \mathcal{B} \geq S_{\min}(\Phi) - \frac{8 \log \dim \mathcal{A}}{\dim \mathcal{A}}.$$

Proof. Exactly as in Theorem 6, Proposition 1 implies the first inequality, as $\Phi'(|0\rangle\langle 0| \otimes \rho) = \Phi(\rho) \otimes \tilde{I}_{\mathcal{B}}$.

Let ρ be a state minimizing $S(\Phi'(\rho))$ and for convenience let $\delta = 8 \log \dim \mathcal{A} / \dim \mathcal{A}$. Equation (10) gives the output of Φ' on ρ. Applying the concavity of the entropy to this, we obtain

$$S_{\min}(\Phi') = S(\Phi'(\rho)) \geq q S(\Phi(\sigma) \otimes \tilde{I}_{\mathcal{B}}) + (1 - q) S(\Phi'(\rho_{S_0^{\perp}})).$$

Applying Corollary 5 this becomes

$$S_{\min}(\Phi') \geq qS(\Phi(\sigma) \otimes \tilde{I}_{\mathcal{B}}) + (1-q)(S(\tilde{I}_{\mathcal{A}\otimes\mathcal{H}}) - \delta).$$

Notice that since Φ' is random unitary, it is the case that $\mathcal{A} \otimes \mathcal{H}$ is isomorphic to $\mathcal{K} \otimes \mathcal{B}$. This implies that $S(\tilde{I}_{\mathcal{A}\otimes\mathcal{H}}) = S(\tilde{I}_{\mathcal{K}\otimes\mathcal{B}})$. Two additional properties of the entropy will be useful: $S(\sigma \otimes \xi) = S(\sigma) + S(\xi)$ for any σ, ξ and $S(\xi) \leq \log \dim \mathcal{K} = S(\tilde{I}_{\mathcal{K}})$ for all $\xi \in \mathbf{D}(\mathcal{K})$. Using these three observations, in order, we find that

$$\begin{aligned} S_{\min}(\Phi') &\geq qS(\Phi(\sigma) \otimes \tilde{I}_{\mathcal{B}}) + (1-q)(S(\tilde{I}_{\mathcal{K}\otimes\mathcal{B}}) - \delta) \\ &= q(S(\Phi(\sigma)) + S(\tilde{I}_{\mathcal{B}})) + (1-q)(S(\tilde{I}_{\mathcal{K}}) + S(\tilde{I}_{\mathcal{B}}) - \delta) \\ &\geq q(S(\Phi(\sigma)) + S(\tilde{I}_{\mathcal{B}})) + (1-q)(S(\Phi(\sigma)) + S(\tilde{I}_{\mathcal{B}}) - \delta) \\ &\geq S(\Phi(\sigma)) + S(\tilde{I}_{\mathcal{B}}) - \delta. \end{aligned}$$

Finally, since $S(\tilde{I}_{\mathcal{B}}) = \log \dim \mathcal{B}$ and $S_{\min}(\Phi) \leq S(\Phi(\xi))$ for any ξ, we have

$$S_{\min}(\Phi') \geq S_{\min}(\Phi) + \log \dim \mathcal{B} - \delta,$$

which completes the proof of the theorem. □

The proof that the additivity conjecture can be equivalently restricted to random unitary channels follows from the previous theorem in a way that is identical to the proof of Theorem 7, with the exception that the p-norm has been replaced by the minimum output entropy.

Theorem 9. *If $\Phi, \Psi \in \mathbf{T}(\mathcal{H}, \mathcal{K})$, then*

$$S_{\min}(\Phi \otimes \Psi) = S_{\min}(\Phi) + S_{\min}(\Psi)$$

if

$$S_{\min}(\Phi'_d \otimes \Psi) = S_{\min}(\Phi'_d) + S_{\min}(\Psi),$$

for all sufficiently large d, where Φ'_d is the random unitary extension of the channel obtained by applying the construction of Section II to Stinespring dilation for Φ using an ancillary space of dimension d.

Proof. Let $\epsilon > 0$, and choose d so that $8(\log d)/d < \epsilon$. Then, as $\Phi'_d(|0\rangle\langle 0| \otimes \rho) = \Phi(\rho) \otimes \tilde{I}_{\mathcal{B}}$,

$$S_{\min}(\Phi \otimes \Psi) \geq S_{\min}(\Phi'_d \otimes \Psi) - \log \dim \mathcal{B}.$$

By assumption, this second quantity is additive, so that

$$\begin{aligned} S_{\min}(\Phi \otimes \Psi) &\geq S_{\min}(\Phi'_d \otimes \Psi) - \log \dim \mathcal{B} \\ &= S_{\min}(\Phi'_d) + S_{\min}(\Psi) - \log \dim \mathcal{B} \\ &\geq S_{\min}(\Phi) - \frac{8 \log d}{d} + S_{\min}(\Psi) \\ &> S_{\min}(\Phi) + S_{\min}(\Psi) - \epsilon \end{aligned}$$

where the penultimate inequality is an application of Theorem 8. As ϵ was chosen arbitrarily, the additivity of Φ'_d for all large enough d implies the additivity of Φ. □

A direct corollary of this theorem generalizes a result of Fukuda[8] on the additivity of the minimum output entropy of unital channels. This implies that in the search for either a proof of this conjecture or a counterexample to it, only random unitary channels need to be considered.

Corollary 10. *The additivity of the minimum output entropy, given by*

$$S_{\min}(\Phi \otimes \Psi) = S_{\min}(\Phi) + S_{\min}(\Psi)$$

is true for all channels Φ *and* Ψ *if and only if*

$$S_{\min}(\Phi \otimes \Phi) = S_{\min}(\Phi) + S_{\min}(\Phi)$$

is true for all random unitary channels Φ*.*

Proof. By a result of Fukuda and Wolf[20] the additivity conjecture can be equivalently restricted to the case that $\Psi = \Phi$, i.e. that the two channels are the same. Applying Theorem 9 twice results in the statement of corollary. □

VI. CIRCUIT CONSTRUCTIONS

In this section an efficient circuit construction is provided for the random unitary approximation described in Section II. This construction is used to extend the hardness of computationally distinguishing quantum circuits to the case of random unitary circuits.

Before constructing these circuits, it will be important to specify the circuit models that are being used. The circuit model used to define the quantum circuit distinguishability problem is the *mixed state quantum circuit* model of Aharonov, Kitaev, and Nisan[27]. Circuits in this model can include unitary gates as well as measurements and other non-unitary operations, but as shown in Ref. 27, we may assume that all such circuits first introduce any necessary ancillary qubits, then perform a unitary operation, and finally trace out those qubits that are not part of the output. This approach is equivalent to building a circuit for the Stinespring dilation of a channel. As all unitary transformations can be (approximately) implemented using one and two qubit gates there is no loss in generality in assuming that the unitary transformations implemented in such a circuit are composed of gates from some finite basis of one and two qubit gates. Circuits in this model can represent any physically realizable quantum operation.

The second model of quantum circuits we consider is the model of *random unitary quantum circuits*. These circuits consist of one and two qubits gates as well as random unitary gates, which implement a unitary gate with probability one half. More formally, the application of such a gate takes the state ρ to the state $(1/2)U\rho U^* + (1/2)\rho$, where U is a one or two qubit unitary gate. This is an extremely simple model that does not appear to be universal for the class of transformations that implement random unitary operations. It is not clear what is the correct definition of the random unitary circuit model, and since the aim of the present paper is to prove a hardness result, an extremely weak definition has been chosen so that the result will apply to as large a class of circuit models as possible.

One drawback of this weak model is that it is not clear that the exact construction used in Section II can be implemented. Specifically, the operation D that decoheres the subspaces S_0 and $S_0^\perp$ seems to require a unitary operation that cannot be decomposed into a series of one and two qubit gates, applied with probability one half. A similar situation occurs for the discrete Weyl operators on the subspace $S_0^\perp$. These operations can be implemented in a random unitary way in a more permissive circuit model, but in order to keep the hardness result on distinguishing random unitary circuits as general as possible, a modified construction is presented here. This modified

construction is built from pieces similar to those used in Section II, but the specific building blocks are not exactly the same. The construction in this section can also be applied to the additivity problems, but it is somewhat more complicated that the construction already presented.

In order to approximate a given circuit with a random unitary circuit we once again make use of three components. We once again use N, D, and M to refer to these components as they play the same roles as the components used in Section II, though they are not exactly the same. The first two of these components, N the completely noisy channel and D the complete dephasing channel, are easy to implement as random unitary operations in the chosen circuit model. More difficult to implement is the channel M, which performs a function similar to the channel described by Equation (7).

The complete dephasing channel D is the channel that sets to zero all of the off-diagonal elements of a density matrix. This is stronger than the operation considered in Section II, but it is easier to implement as a random unitary circuit. The action of this operator applied to the space $\mathcal{A}$, for an input ρ on $\mathcal{A} \otimes \mathcal{H}$ is given by

$$D_{\mathcal{A}}(\rho) = \sum_{i=0}^{\dim \mathcal{A}-1} p_i |i\rangle\langle i| \otimes \rho_i, \tag{13}$$

where the p_i form a probability distribution. This operation is equivalent to measuring the space $\mathcal{A}$ in the computational basis and forgetting the result. The operation $D_{\mathcal{A}}$ can be implemented as a random unitary circuit by applying the Pauli Z operation to each qubit of $\mathcal{A}$ independently with probability $1/2$, as described in Ref. 28. This will have the effect of negating the off-diagonal elements of a density matrix with probability $1/2$, so that the resulting state is diagonal in the computational basis.

The completely noisy channel N is also simple to implement as a random unitary circuit. This channel can be realized by performing a uniform mixture of the Pauli operators on each qubit. This mixture can be implemented by, independently on each qubit, applying the Pauli Z operation with probability $1/2$, followed by applying the Pauli X operation with probability $1/2$, as shown in Ref. 23. Intuitively, the Z operations will zero the off-diagonal elements of a density matrix (viewed in the computational basis), and the X operations will scramble the diagonal, resulting in the completely mixed state, $I/2$, on each qubit.

In Section II the channel M was implemented as a completely depolarizing channel on the subspace $S_0^{\perp}$. While the same channel suffices for the circuit case, it is not clear how this can be implemented using only two-qubit random unitary gates. To avoid this difficulty a more complicated construction is used. This construction is intuitively the same: it does not affect states in the subspace S_0 of inputs with the $|0\rangle$ state in the space $\mathcal{A}$, and it applies depolarizing noise to states in the space $S_0^{\perp}$. The difference is exactly how this noise is applied. The circuit that is constructed will implement the operation M given by

$$M(|i\rangle\langle i| \otimes \rho) = \begin{cases} \frac{1}{\dim \mathcal{A}}(I_{\mathcal{A}} - |0\rangle\langle 0| + |\psi_i\rangle\langle\psi_i|) \otimes \tilde{I}_{\mathcal{H}} & \text{if } i \neq 0, \\ |0\rangle\langle 0| \otimes \rho & \text{if } i = 0, \end{cases} \tag{14}$$

where $|\psi_i\rangle$ is a nonzero state that depends on i, the exact specification of which will not be significant.

As might be expected, the transformation M can be implemented using only controlled-mixing operations. Before describing this implementation, notice that the controlled application of the completely depolarizing channel N to a single qubit can be described by a random unitary circuit. This is because the above implementation of N as a mixture of Pauli Z and X operations consists only of single qubit gates. Adding a control qubit to each of these gates results in two qubit gates,

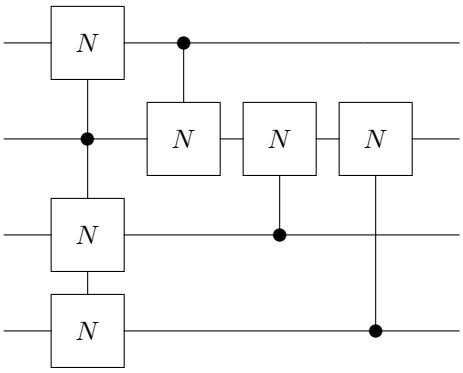

FIG. 1: One stage of the mixing procedure on the ancillary qubits. The mixing operations applied to the qubits in the space $\mathcal{H}$ are not shown.

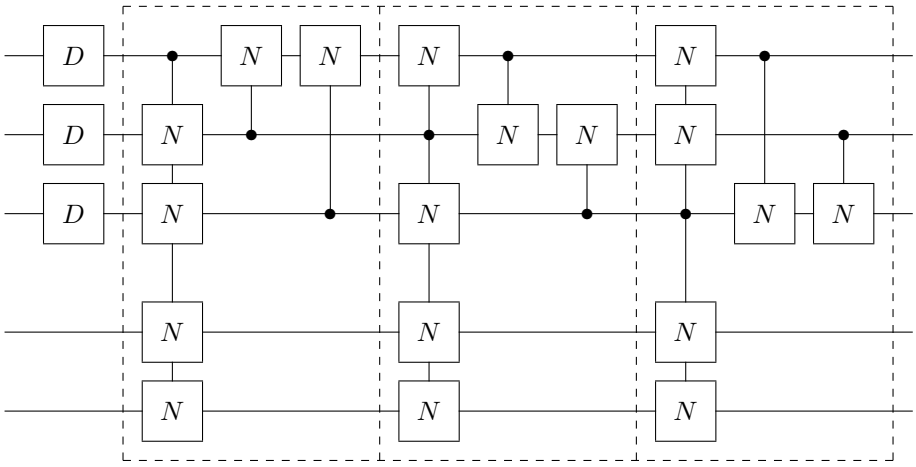

FIG. 2: Circuit performing the conditional mixing procedure $M \circ D$. The top three qubits are simulating the ancillary qubits of the original circuit in the space $\mathcal{A}$, and the bottom two are simulating the input to the original circuit in the space $\mathcal{H}$. The dashed lines separate the stages of the mixing procedure.

which fit into the model of random unitary circuits used here. It is not clear that general controlled random unitary operations can be implemented as random unitary circuits in this model, but the only controlled operation that will be needed for this construction is the completely depolarizing channel.

Let m be the number of qubits in the space $\mathcal{A}$ that are given as part of the input to M, i.e. the number of ancillary qubits used to represent the ancillary space used by the original channel. The implementation of M consists of m stages, with the jth stage testing that the jth qubit of the space $\mathcal{A}$ is in the $|0\rangle$ state, and mixing the qubits if this is not the case. An example of one stage of the circuit is given in Figure 1. The jth stage consists first of an application of the controlled N operation from the jth qubit to each other qubit of $\mathcal{A} \otimes \mathcal{H}$. After these operations, stage j is completed by $m - 1$ further controlled N operations: each with the jth qubit as the target qubit and one of the other qubits of $\mathcal{A}$ as the control qubit. An example of this construction with $m = 3$ is presented in Figure 2.

Given these circuit implementations of the three channels D, N, M, the random unitary circuit C that approximates a given circuit Q is constructed in exactly the same was as in Equation (9). More concretely, let Q be a circuit implementing the operation

$$Q(\rho) = \mathrm{tr}_{\mathcal{B}}\, U(|0\rangle\langle 0| \otimes \rho)U^*,$$

where the ancillary qubits are in the space $\mathcal{A}$. The circuit C that approximates it is then given by

$$C(\rho) = N_{\mathcal{B}} \left(U \left[(M \circ D_{\mathcal{A}})(\rho) \right] U^* \right). \tag{15}$$

This circuit C is given by a random unitary circuit, since it is the composition of smaller random unitary circuits. As the operations $D_{\mathcal{A}}$ and M do not affect inputs of the form $|0\rangle\langle 0| \otimes \rho$, the proof of Proposition 1 holds also for the circuit case, so that

$$C(|0\rangle\langle 0| \otimes \sigma) = Q(\sigma) \otimes \tilde{I}_{\mathcal{B}}. \tag{16}$$

Combining this with equation (13) and the fact that applying $D_{\mathcal{A}}$ twice has no further effect, the output of C on an arbitrary input state ρ is of the form

$$C(\rho) = \sum_{i=0}^{\dim \mathcal{A}-1} p_i C(|i\rangle\langle i| \otimes \rho_i) = p_0 Q(\rho_0) \otimes \tilde{I}_{\mathcal{B}} + \sum_{i=1}^{\dim \mathcal{A}-1} p_i C(|i\rangle\langle i| \otimes \rho_i). \tag{17}$$

In the remainder of the paper it is shown that this construction does not significantly alter the distinguishability properties of quantum circuits.

As a first step towards this, it is shown that the above circuit construction correctly implements the channel M described by Equation 14. Much of the proof of this lemma is similar to the proof of Lemma 4, but the operation M considered in this section is slightly different and we also need to extend the lemma to the case that there is an additional reference system. This system, given by the space $\mathcal{F}$, is needed in the case of distinguishability, as a party attempting to distinguish two channels is permitted to use a portion of a larger entangled state as input to the channels. This can be seen from the definition of the distinguishability problem, which is given in the next section.

Lemma 11. *On input states of the form $|k\rangle\langle k| \otimes \rho \in \mathbf{D}(\mathcal{A} \otimes \mathcal{H} \otimes \mathcal{F})$ for $|k\rangle\langle k| \in \mathbf{D}(\mathcal{A})$ with $0 < k \leq 2^m - 1$, the output of C satisfies*

$$\left\| (C \otimes I_{\mathcal{F}})(|k\rangle\langle k| \otimes \rho) - \tilde{I}_{\mathcal{A}\otimes\mathcal{H}} \otimes \operatorname{tr}_{\mathcal{H}} \rho \right\|_{\mathrm{tr}} \leq \frac{1}{2^{m-1}},$$

where m is the number of ancillary qubits used by the circuit Q.

Proof. On input of the form $|k\rangle\langle k| \otimes \rho$ the decoherence operations that are applied to the qubits in $\mathcal{A}$ can be ignored, as they have no effect on qubits in a state of the computational basis. As $k \neq 0$ at least one qubit is in the state $|1\rangle$, and so the controlled mixing operations in the implementation of the channel M will have an effect. Let the first nonzero qubit among the qubits of $\mathcal{A}$ be the jth one. The first controlled N operation with nonzero control qubit that effects the jth qubit will be at the jth stage of the mixing process, where the jth qubit is the control qubit. As this qubit is not modified before this stage (as any previous qubits are in the state $|0\rangle$ by choice of j), the first $m-1$ gates in the jth stage will mix the remaining qubits, so that the state after these gates is, using Equation (6),

$$|1\rangle\langle 1| \otimes \tilde{I}_{\mathcal{A}'} \otimes \tilde{I}_{\mathcal{H}} \otimes \operatorname{tr}_{\mathcal{H}} \rho,$$

where for notational convenience the jth qubit has been written first, and $\mathcal{A}'$ is the space of all but the jth qubit of $\mathcal{A}$. The remainder of the jth stage of the mixing process consists of $m-1$ controlled N gates with the jth qubit as the target, each controlled by one of the $m-1$ qubits in $\mathcal{A}'$. Considering the state $I/2^{m-1}$ on $\mathcal{A}'$ in the computational basis, the only term for which qubit

j is not mixed by these operations is the all zero term. With this observation, the state after the jth stage is

$$\frac{1}{2^{m-1}}\left[|1\rangle\langle 1| \otimes (|0\rangle\langle 0|)^{\otimes m-1} + \frac{|0\rangle\langle 0| + |1\rangle\langle 1|}{2} \otimes (I_{\mathcal{A}'} - (|0\rangle\langle 0|)^{\otimes m-1})\right] \otimes \tilde{I}_{\mathcal{H}} \otimes \mathrm{tr}_{\mathcal{H}}\, \rho$$
$$= \frac{I_{\mathcal{A}} + |1\rangle\langle 1| \otimes (|0\rangle\langle 0|)^{\otimes m-1} - (|0\rangle\langle 0|)^{\otimes m}}{2^m} \otimes \tilde{I}_{\mathcal{H}} \otimes \mathrm{tr}_{\mathcal{H}}\, \rho.$$

This proves that the circuit implementing the channel M does so correctly, as this quantity is exactly the state given in Equation (17) with the addition of $\mathrm{tr}_{\mathcal{H}}\, \rho$ in the reference system.

As in the proof of Lemma 4, let this state be σ. Computing the distance from this state to the desired one, we have

$$\left\| \sigma - \tilde{I}_{\mathcal{A}} \otimes \tilde{I}_{\mathcal{H}} \otimes \mathrm{tr}_{\mathcal{H}}\, \rho \right\|_{\mathrm{tr}} = \frac{1}{2^m} \left\| |1\rangle\langle 1| \otimes (|0\rangle\langle 0|)^{\otimes m-1} - (|0\rangle\langle 0|)^{\otimes m} \right\|_{\mathrm{tr}} = \frac{1}{2^{m-1}}.$$

Finally, by noting that the remainder of the circuit C is random unitary, an application of Lemma 2 yields the desired bound. □

VII. QIP-COMPLETENESS OF DISTINGUISHING RANDOM UNITARY CIRCUITS

The construction outlined in the previous section can be used to show that the problem of distinguishing random unitary quantum circuits is QIP-complete, where QIP is the class of problems having quantum interactive proof systems. The basic idea is to reduce an instance of the quantum circuit distinguishability problem to one with random unitary circuits that has the same distinguishability properties. This will be done by taking the instance (Q_1, Q_2) and constructing the instance (C_1, C_2) by applying the construction of Section VI to each of these circuits. The quantum circuit distinguishability problem is given by

Quantum Circuit Distinguishability. *For constants $0 \leq b < a \leq 2$, the input consists of quantum circuits Q_1 and Q_2 that implement transformations from $\mathcal{H}$ to $\mathcal{K}$. The promise problem is to distinguish the two cases:*

Yes $\|Q_1 - Q_2\|_\diamond \geq a$,

No $\|Q_1 - Q_2\|_\diamond \leq b$.

This problem was introduced and shown to be complete for the complexity class QIP in Ref. 29. The norm used in the definition of the problem is the diamond norm, which can be defined on a channel $\Phi \in \mathbf{T}(\mathcal{H}, \mathcal{K})$ by

$$\|\Phi\|_\diamond = \sup_{\|X\|_{\mathrm{tr}}=1} \|(\Phi \otimes I_{\mathcal{F}})(X)\|_{\mathrm{tr}},$$

where the space $\mathcal{F}$ has dimension at least as large as $\mathcal{H}$. A more thorough definition as well as some properties of this norm can be found in Ref. 30. It may be helpful to note that the diamond norm of Φ is just the completely bonded norm of the adjoint channel Φ^*, where the adjoint is taken with respect to the Hilbert-Schmidt inner product. It is shown in Ref. 29 that the maximum of this norm on the difference of two completely positive transformations is achieved by a density matrix, and so we can restrict the supremum in the definition to $\mathbf{D}(\mathcal{H} \otimes \mathcal{F})$.

Here we consider this distinguishability problem with the added restriction that the input circuits are random unitary circuits in the model defined in Section VI. The following theorem states

that the constructed circuits C_1 and C_2 have almost the same distinguishability characteristics as the original circuits Q_1 and Q_2. As the circuit distinguishability problem is defined as a promise problem, this theorem shows immediately that the problem of distinguishing random unitary circuits is QIP-complete, as the construction of the circuits C_1 and C_2 can be performed efficiently.

Theorem 12. *For any $\epsilon > 0$,*

$$\|Q_1 - Q_2\|_\diamond \leq \|C_1 - C_2\|_\diamond \leq \|Q_1 - Q_2\|_\diamond + \epsilon,$$

where the circuits C_1 and C_2 use $O(\log 1/\epsilon)$ ancillary qubits.

Proof. The first inequality is not hard to show. Once again, if the state $(|0\rangle\langle 0|)^{\otimes m} \otimes \rho$ is given as input to the circuit C_i, then by Equation 16, the output is a simulation of Q_i, so that the distinguishability of Q_1 and Q_2 cannot be greater than the distinguishability of C_1 and C_2. More formally, note that

$$\|Q_1 - Q_2\|_\diamond = \sup_{\rho \in \mathbf{D}(\mathcal{H} \otimes \mathcal{F})} \|(Q_1 \otimes I_\mathcal{F})(\rho) - (Q_2 \otimes I_\mathcal{F})(\rho)\|_{\mathrm{tr}},$$

and fix $\delta > 0$ and ρ as a state achieving a value within δ of this supremum. By Equation 16, if the state $(|0\rangle\langle 0|)^{\otimes m} \otimes \rho$ is given as input to the circuit C_i, then the output is given by $(Q_i \otimes I_\mathcal{F})(\rho)$. Using this we have

$$\begin{aligned}
\|C_1 - C_2\|_\diamond &\geq \left\|(C_1 \otimes I_\mathcal{F})((|0\rangle\langle 0|)^{\otimes m} \otimes \rho) - (C_2 \otimes I_\mathcal{F})((|0\rangle\langle 0|)^{\otimes m} \otimes \rho)\right\|_{\mathrm{tr}} \\
&= \|(Q_1 \otimes I_\mathcal{F})(\rho) - (Q_2 \otimes I_\mathcal{F})(\rho)\|_{\mathrm{tr}} \\
&\geq \|Q_1 - Q_2\|_\diamond - \delta.
\end{aligned}$$

Since this is true for any $\delta > 0$, it must be the case that $\|Q_1 - Q_2\|_\diamond \leq \|C_1 - C_2\|_\diamond$.

The second inequality requires somewhat more work. Let m be the number of ancillary qubits and let n be the number of input qubits used by the circuits Q_i, so that $m = \lceil \log \dim \mathcal{A} \rceil$ and $n = \lceil \log \dim \mathcal{H} \rceil$. Without loss of generality let $2^{-(m-3)} < \epsilon$, by adding at most $3 + \log(1/\epsilon)$ extra (unused) ancillary qubits to Q_1 and Q_2. Let $\rho \in \mathbf{D}(\mathcal{A} \otimes \mathcal{H} \otimes \mathcal{F})$ be a state such that

$$\|C_1 - C_2\|_\diamond - \epsilon/2 \leq \|(C_1 \otimes I_\mathcal{F})(\rho) - (C_2 \otimes I_\mathcal{F})(\rho)\|_{\mathrm{tr}},$$

and note that the reference system $\mathcal{F}$ need not have the same dimension as the space of the same name considered in the proof of the previous inequality. The first gates applied in the circuit C_i are the decoherence gates applied to $\mathcal{A}$. These gates produce a state of the form $\sum_{i=0}^{2^m-1} p_i |i\rangle\langle i| \otimes \sigma_i$, and since applying these gates twice has no further effect, the output of the circuits C_1 and C_2 is the same on ρ as it is on this state. Applying the triangle inequality, the quantity of interest is

$$\|C_1 - C_2\|_\diamond - \epsilon/2 \leq \sum_{i=0}^{2^m-1} p_i \|(C_1 \otimes I_\mathcal{F})(|i\rangle\langle i| \otimes \sigma_i) - (C_2 \otimes I_\mathcal{F})(|i\rangle\langle i| \otimes \sigma_i)\|_{\mathrm{tr}} \tag{18}$$

Then, by applying Lemma 11 to each term with $i \neq 0$ the states in the norm can be replaced with completely mixed states on $\mathcal{A} \otimes \mathcal{H}$ plus a small correction factor. Doing this for each of these terms we have

$$\begin{aligned}
&p_i \|(C_1 \otimes I_\mathcal{F})(|i\rangle\langle i| \otimes \sigma_i) - (C_2 \otimes I_\mathcal{F})(|i\rangle\langle i| \otimes \sigma_i)\|_{\mathrm{tr}} \\
&\quad \leq p_i \left[\frac{2}{2^{m-1}} + \left\|\tilde{I}_{\mathcal{A}\otimes\mathcal{H}} \otimes \mathrm{tr}_\mathcal{H}\, \sigma_i - \tilde{I}_{\mathcal{A}\otimes\mathcal{H}} \otimes \mathrm{tr}_\mathcal{H}\, \sigma_i\right\|_{\mathrm{tr}}\right] \\
&\quad = p_i / 2^{m-2} < p_i \epsilon / 2.
\end{aligned}$$

Applying this to Equation (18) we have

$$\|C_1 - C_2\|_\diamond - \epsilon/2 \leq p_0 \|(C_1 \otimes I_{\mathcal{F}})(|0\rangle\langle 0| \otimes \sigma_0) - (C_2 \otimes I_{\mathcal{F}})(|0\rangle\langle 0| \otimes \sigma_0)\|_{\mathrm{tr}} + \sum_{i=1}^{2^m-1} p_i \epsilon/2.$$

By Equation 16 the output of the circuit C_i on this input can be replaced the output of the circuit Q_i and a maximally mixed state. When this is done to the previous equation, the desired bound is given by

$$\|C_1 - C_2\|_\diamond \leq p_0 \|(Q_1 \otimes I_{\mathcal{F}})(\sigma_0) - (Q_2 \otimes I_{\mathcal{F}})(\sigma_0)\|_{\mathrm{tr}} + (1-p_0)\epsilon/2 + \epsilon/2 \leq \|Q_1 - Q_2\|_\diamond + \epsilon.$$

This completes the proof of the theorem, as $0 \leq p_0 \leq 1$. □

VIII. CONCLUSION

A method for approximating a quantum channel with one that is random unitary has been provided. This approximation yields the equivalence of several important problems when restricted to random unitary channels. These results raise the open problem of how far these equivalences extend. What other problems can be restricted to the random unitary case without loss of generality, and what problems are simplified when restricted to this class of channels?

Acknowledgements

I would like to thank John Watrous for several helpful discussions, as well as Michael Wolf and the anonymous referees for their comments, including a suggestion by one referee that simplified the argument in Section II. Canada's NSERC and MITACS have supported this research.

---

* Electronic address: wrosgen@iqc.ca

[1] M. Gregoratti and R. F. Werner, Journal of Modern Optics **50**, 915 (2003).
[2] S. L. Tregub, Soviet Mathematics (Iz. VUZ) **30**, 105 (1986).
[3] B. Kümmerer and H. Maassen, Communications in Mathematical Physics **109**, 1 (1987).
[4] L. J. Landau and R. F. Streater, Linear Algebra and its Applications **193**, 107 (1993).
[5] K. M. R. Audenaert and S. Scheel, New Journal of Physics **10**, 023011 (2008).
[6] F. Buscemi, Physics Letters A **360**, 256 (2006).
[7] A. S. Holevo, in *Proceedings of the International Congress of Mathematicians* (2006), vol. 3, pp. 1000–1017.
[8] M. Fukuda, Quantum Information Processing **6**, 179 (2007).
[9] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, 2000).
[10] A. S. Holevo, IEEE Transactions on Information Theory **44**, 269 (1998).
[11] B. Schumacher and M. D. Westmoreland, Physical Review A **56**, 131 (1997).
[12] C. H. Bennett, C. A. Fuchs, and J. A. Smolin, in *Quantum Communication, Computing, and Measurement: Proceedings of the Third International Conference on Quantum Communication and Measurement, 1996*, edited by O. Hirota, A. S. Holevo, and C. M. Caves (Plenum Press, New York, 1997), pp. 79–88.
[13] C. King and M. B. Ruskai, IEEE Transactions on Information Theory **47**, 192 (2001).
[14] P. W. Shor, Communications in Mathematical Physics **246**, 453 (2004).
[15] G. G. Amosov, A. S. Holevo, and R. F. Werner, Problems of Information Transmission **36**, 305 (2000).
[16] P. Hayden and A. Winter, *Counterexamples to the maximal p-norm multiplicativity conjecture for all* $p > 1$, arXiv:0807.4753v1 [quant-ph] (2008).

[17] P. W. Shor, Journal of Mathematical Physics **43**, 4334 (2002).
[18] C. King, Quantum Information and Computation **3**, 186 (2003).
[19] C. King, Journal of Mathematical Physics **43**, 4641 (2002).
[20] M. Fukuda and M. M. Wolf, Journal of Mathematical Physics **48**, 072101 (2007).
[21] W. F. Stinespring, Proceedings of the American Mathematical Society **6**, 211 (1955).
[22] A. Ambainis, M. Mosca, A. Tapp, and R. de Wolf, in *Proceedings of the 41st IEEE Symposium on Foundations of Computer Science* (2000), pp. 547–553.
[23] P. O. Boykin and V. Roychowdhury, Physical Review A **67**, 042317 (2003).
[24] P. Hayden, D. W. Leung, P. Shor, and A. Winter, Communications in Mathematical Physics **250**, 371 (2004).
[25] N. Datta, M. Fukuda, and A. S. Holevo, Quantum Information Processing **5**, 179 (2006).
[26] M. Fannes, Communications in Mathematical Physics **31**, 291 (1973).
[27] D. Aharonov, A. Kitaev, and N. Nisan, in *Proceedings of the 30th ACM Symposium on the Theory of Computing* (1998), pp. 20–30.
[28] I. L. Chuang and Y. Yamamoto, Physical Review A **55**, 114 (1997).
[29] B. Rosgen and J. Watrous, in *Proceedings of the 20th Conference on Computational Complexity* (2005), pp. 344–354.
[30] A. Y. Kitaev, A. H. Shen, and M. N. Vyalyi, *Classical and Quantum Computation*, vol. 47 of *Graduate Studies in Mathematics* (American Mathematical Society, 2002).