Superadditivity of Quantum Channels

Sanjay Raman*

Department of Physics,

Massachusetts Institute of Technology

Abstract

A fundamental problem of quantum information theory is the calculation of the classical capacity of quantum channels. In particular, the possibility of entanglement over both inputs and measurements can potentially increase the channel capacity. The Holevo-Schumacher-Westmoreland (HSW) theorem [3] computes the classical capacity of a quantum channel without allowing for entangled inputs, but it was shown in [1] that entangled inputs can boost the capacity of a quantum channel beyond the HSW value (superadditivity). We discuss the intuition behind the superadditive quantum channel constructed in [1]. We analyze the various requirements for superadditivity, and we show that a class of potentially non-unital quantum channels in arbitrary dimensions must be additive. We develop bounds on the minimum output entropy of certain classes of quantum channels, and we formulate a conjecture constraining which channels can be superadditive.

* sanjayra@mit.edu

I. INTRODUCTION

Suppose Alice wants to send Bob information over some channel, subject to noise. How much information can Alice transmit to Bob? This is a fundamental question in information theory, and a quantitative answer to this question can be formulated in terms of the *channel capacity*. More precisely, the capacity of a channel is the number of bits that can be sent per channel use, asymptotically over many uses of the channel.

Consider now the introduction of quantum mechanics. A particularly interesting case is the transmission of classical information across a quantum channel. In particular, suppose Alice wishes to send classical information over a quantum channel $\mathcal{E}$ to Bob. Alice could prepare a state by some preparation method $\mathcal{P}$, send it over $\mathcal{E}$, and Bob could then deduce classical information by performing some positive operator-valued measurement $\mathcal{M}$ (POVM) on the state received.

When considering the transmission of classical information over a quantum channel $\mathcal{E}$, however, *entanglement* plays a role. More precisely, it becomes necessary to consider the possibility of entangling inputs over multiple channels. If a single channel is used multiple times, we could consider entangling the inputs over multiple uses of the channel [1]. Furthermore, we could also consider entangled *measurements* over multiple channels or multiple uses of the same channel (See Fig. **??**). There are in fact four different notions of classical capacity of a quantum channel, each making use of entanglement in a different way:

- $C_{1,1}$: Do not allow entanglement.
- $C_{\infty,1}$: Allow for inputs entangled over $\mathcal{E}^{\otimes n}$.
- $C_{1,\infty}$: Allow for a POVM entangled over $\mathcal{E}^{\otimes n}$.
- $C_{\infty,\infty}$: Allow for both entangled inputs and entangled measurements.

A fundamental question of quantum information theory is understanding how entanglement affects the classical capacity of a quantum channel. How can we compute the different capacities $C_{1,1}, C_{\infty,1}, C_{1,\infty}, C_{\infty,\infty}$, taking into account the effect of entanglement?

A. The Holevo-Schumacher-Westmoreland Theorem

As it turns out, there is an exact formula for $C_{1,\infty}$ [3, 5]:

Theorem I.1 (Holevo-Schumacher-Westmoreland [3]). *Suppose Bob receives states ρ_X corresponding to the letters X chosen by Alice with probability p_X, and let $\rho = \sum_X p_X \rho_X$. Then*

$$C_{1,\infty} = \max_{p(X)} \left[H(\rho) - \sum_X p_X H(\rho_X) \right]. \tag{1}$$

In other words, the Holevo-Schumacher-Westmoreland (HSW) theorem gives an explicit formula for the capacity of a quantum channel, provided that we may entangle measurements but *not* inputs. It is also clear that

$$C_{\infty,\infty} \geq C_{1,\infty} \tag{2}$$

so the HSW theorem provides a lower bound on the classical capacity of a quantum channel.

B. Entangled Inputs and Superadditivity

The HSW theorem gives us an important tool in determining the capacity of quantum channels. The question now becomes whether or not equality holds in (2). Can allowing for entanglement over inputs boost the capacity $C \equiv C_{\infty,\infty}$ beyond the value $C_{1,\infty}$? A simple restatement of this question is in terms of the *additivity* of quantum channels. More precisely, can we choose channels $\mathcal{E}_1$ and $\mathcal{E}_2$ such that

$$C(\mathcal{E}_1 \otimes \mathcal{E}_2) > C(\mathcal{E}_1) + C(\mathcal{E}_2)? \tag{3}$$

As shown in [4], this question may be equivalently reformulated as follows. Define the *minimum output entropy* $H_{\min}$ of a channel $\mathcal{E}$ as

$$H_{\min}(\mathcal{E}) = \min_{|\chi\rangle} H(\mathcal{E}(|\chi\rangle\langle\chi|)). \tag{4}$$

where the minimum is taken over all input pure states $|\chi\rangle$. Can we choose $\mathcal{E}_1, \mathcal{E}_2$ such that

$$H_{\min}(\mathcal{E}_1 \otimes \mathcal{E}_2) < H_{\min}(\mathcal{E}_1) + H_{\min}(\mathcal{E}_2)? \tag{5}$$

As it turns out, the answer is *yes.*

Theorem I.2 (Hastings [1]). *Choose a channel $\mathcal{E}$ such that*

$$\mathcal{E}(\rho) = \sum_{i=1}^{D} p_i U_i \rho U_i^\dagger \tag{6}$$

where ρ is an input density matrix over an N-dimensional Hilbert space $\mathcal{H}$, the $U_i \in U(N)$ are unitary matrices, and the p_i are probabilities summing to 1. As shown in the seminal work [1], we may then choose $\mathcal{E}$ such that

$$H_{\min}(\mathcal{E} \otimes \overline{\mathcal{E}}) < H_{\min}(\mathcal{E}) + H_{\min}(\overline{\mathcal{E}}) = 2H_{\min}(\mathcal{E}), \tag{7}$$

provided that $N \gg D \gg 1$. Here $\overline{\mathcal{E}}$ is the complex conjugate of the channel $\mathcal{E}$ (so that obviously $H_{\min}(\mathcal{E}) = H_{\min}(\overline{\mathcal{E}})$).

In this paper, our first goal will be to unpack and develop an intuitive picture the result of [1]. In particular will investigate the construction of the superadditive channel $\mathcal{E}$ in a bit of detail. From this, we will then consider the following interesting question: Where do the requirements $N \gg D \gg 1$ come from? What happens when the requirements $N \gg D \gg 1$ are relaxed? To study this question, we will consider a more general class of channels $\mathcal{F}$ with the following decomposition into Kraus operators:

$$\mathcal{F}(\rho) = \sum_{i=1}^{D} V_i \rho V_i^\dagger, \quad \sum_{i=1}^{D} V_i^\dagger V_i = 1. \tag{8}$$

Our goal will be to understand how such channels $\mathcal{F}$ behave when D is small: We will attempt to discuss the additivity of noisy quantum channels representable in terms of a small number of Kraus operators. We will show that the channels $\mathcal{F}$ *must* be additive for $D = 2$, regardless of the value of N. This result is interesting, since it shows that a large class of possibly non-unital quantum channels in an arbitrary number of dimensions must be additive. Finally, we will also prove an upper bound on the minimum output entropy of channels for $D > 2$, and we will develop a conjecture on the additivity of quantum channels for D small enough.

II. INTUITION FOR SUPERADDITIVE QUANTUM CHANNELS

In this section, we will discuss the intuition behind the result of Theorem I.2. Our goal here will be to intuitively understand this result and to outline a few key steps in the proof which will prove useful our later results.

To begin, we will define $\mathcal{E}$ as in (6), and we will compute the output entropy of the maximally entangled state $|\Psi_{\mathrm{ME}}\rangle$ given by

$$|\psi_{\mathrm{ME}}\rangle = \frac{1}{\sqrt{N}} \sum_{\alpha=1}^{N} |\alpha\rangle \otimes |\alpha\rangle. \tag{9}$$

through the channel $\mathcal{E} \otimes \overline{\mathcal{E}}$. (Here the $|\alpha\rangle$ are any basis of our Hilbert space $\mathcal{H}$.) The key realization is as follows. Consider the action of $U_i \otimes \overline{U_i}$ on $|\Psi_{\mathrm{ME}}\rangle$. We may take the $|\alpha\rangle$ to be the eigenbasis of U_i with eigenvalues λ_α. Since U_i is unitary, $|\lambda_\alpha| = 1$. We then have

$$U_i \otimes \overline{U_i}|\psi_{\mathrm{ME}}\rangle = \frac{1}{\sqrt{N}} \sum_{\alpha=1}^{N} U_i|\alpha\rangle \otimes \overline{U_i}|\alpha\rangle = \frac{1}{\sqrt{N}} \sum_{\alpha=1}^{N} |\lambda_\alpha|^2 \, |\alpha\rangle \otimes |\alpha\rangle = |\Psi_{\mathrm{ME}}\rangle. \tag{10}$$

Thus,

$$\mathcal{E} \otimes \overline{\mathcal{E}}(|\Psi_{\mathrm{ME}}\rangle \langle\Psi_{\mathrm{ME}}|) = \sum_{i=1}^{N} p_i^2 \, |\Psi_{\mathrm{ME}}\rangle \langle\Psi_{\mathrm{ME}}| + \sum_{1\leq i\neq j\leq N} p_i p_j U_i \otimes \overline{U_j} \, |\Psi_{\mathrm{ME}}\rangle \langle\Psi_{\mathrm{ME}}| \, U_i^\dagger \otimes \overline{U_j}^\dagger. \tag{11}$$

The key realization here is that the output $\mathcal{E} \otimes \overline{\mathcal{E}}(|\Psi_{\mathrm{ME}}\rangle \langle\Psi_{\mathrm{ME}}|)$ is *not* maximally mixed! In other words, there exists a substantial preference for the state $|\Psi_{\mathrm{ME}}\rangle \langle\Psi_{\mathrm{ME}}|$ in $\mathcal{E} \otimes \overline{\mathcal{E}}(|\Psi_{\mathrm{ME}}\rangle \langle\Psi_{\mathrm{ME}}|)$.

As it turns out, this non-maximal mixing of $\mathcal{E} \otimes \overline{\mathcal{E}}(|\Psi_{\mathrm{ME}}\rangle \langle\Psi_{\mathrm{ME}}|)$ is enough to bound the entropy $H(\mathcal{E} \otimes \overline{\mathcal{E}}(|\Psi_{\mathrm{ME}}\rangle \langle\Psi_{\mathrm{ME}}|))$ away from the maximal possible value of $2\log D$:

Lemma II.1 (Hastings [1])**.** Let $\mathcal{E}$ be a channel given as in (6), and let $\overline{\mathcal{E}}$ be the complex conjugate of $\mathcal{E}$. We then have

$$H_{\min}(\mathcal{E} \otimes \overline{\mathcal{E}}) \leq H(\mathcal{E} \otimes \overline{\mathcal{E}}(|\Psi_{\mathrm{ME}}\rangle \langle\Psi_{\mathrm{ME}}|)) \leq 2\log D - \frac{\log D}{D}. \tag{12}$$

The technical details of the proof of this lemma are relatively unimportant. The main takeaway here is the intuition: We are looking for entangled inputs that somehow cause the outputs to not be maximally mixed. The maximally entangled state $|\Psi_{\mathrm{ME}}\rangle$ provides exactly such an input to $\mathcal{E} \otimes \overline{\mathcal{E}}$.

This in itself is not enough, however. We require the individual entropies $H_{\min}(\mathcal{E}), H_{\min}(\overline{\mathcal{E}})$ to be large enough such that $H_{\min}(\mathcal{E}) + H_{\min}(\overline{\mathcal{E}}) > H_{\min}(\mathcal{E} \otimes \overline{\mathcal{E}})$. Intuitively speaking, we thus want the individual channel $\mathcal{E}$ to "mix" the input state as much as possible. Since there are only D Kraus operators, a maximally mixed output $\mathcal{E}(|\chi\rangle\langle\chi|)$ would have entropy close to $\log D$. The randomized construction in [1] shows that this is indeed possible. In particular, the following result was proven:

Lemma II.2 (Hastings [1])**.** Let $\mathcal{E}$ be a channel given as in (6). Then the U_i and p_i may be chosen such that

$$H_{\min}(\mathcal{E}) > \log D - \frac{c_1}{D} + p_1(D)\mathcal{O}\left(\sqrt{\log N/N}\right). \tag{13}$$

Here c_1 is a constant and p_1 is a polynomial function in D.

Continuing with our intuitive description, we thus see that a state $|\chi\rangle$ passed through the channel $\mathcal{E}$ can get "mixed enough" such that $H(\mathcal{E}(|\chi\rangle\langle\chi|))$ is close to $\log D$, regardless of the input state $|\chi\rangle$. Additionally, we see immediately where the requirement $N \gg D \gg 1$ comes from. The minimum output entropy $H_{\min}(\mathcal{E})$ can only be made close to $\log D$ provided that the two error terms in (13) are small.

In summary, Theorem I.2 can be viewed as follows. The minimum output entropy of an entangled input through $\mathcal{E} \otimes \overline{\mathcal{E}}$ can be bounded above, since entanglement can be used to enforce non-maximal mixing in the output, regardless of what we choose $\mathcal{E}$ to be. However, we can then choose $\mathcal{E}$ such that the channel $\mathcal{E}$ itself almost maximally mixes any input state $|\chi\rangle$. All of this can only occur provided that the Hilbert space is large enough and there are enough operators U_i ($N \gg D \gg 1$).

III. QUANTUM CHANNELS FOR SMALL VALUES OF D

In the previous section, we developed a picture for why exactly certain quantum channels are superadditive. However, a piece of the puzzle that remains is the requirement $N \gg D \gg 1$ in the construction. It is mathematically clear why this must hold true, given the bounds in Lemmas II.1 and II.2. However, the mathematical relations derived fail to give an intuitive picture for why this must be the case.

We can first consider the requirement $N \gg D$. This would seem to imply that superadditivity cannot arise in quantum channels of the form in (6) over small Hilbert spaces. Indeed, this accords well with our expectations. For example, it is already known [2] that any unital *qubit* channels (for $N = 2$), including the depolarizing channel, are necessarily additive. This result should give us confidence that the requirement $N \gg D$ is indeed plausible.

Nevertheless, the requirement $D \gg 1$ still remains enigmatic. Why should we expect to need a large number of Kraus operators in the decomposition of $\mathcal{E}$ for superadditivity to occur? We will develop some motivation for this in the following sections.

A. Quantum Channels with $D = 2$

A first step in considering the implications of the $D \gg 1$ requirement are to consider cases when D is a small integer. In particular, we let N be arbitrary, but we restrict our attention to the case where $D = 2$. Is it still possible to find superadditive quantum channels with $D = 2$? As it turns out, the answer is a general *no*:

Lemma III.1. Let $\mathcal{F}$ be a channel defined as below:

$$\mathcal{F}(\rho) = V_1 \rho V_1^\dagger + V_2 \rho V_2^\dagger, \quad V_1^\dagger V_1 + V_2^\dagger V_2 = 1., \tag{14}$$

where ρ is a density matrix over an N-dimensional Hilbert space (here N can be arbitrary). Then $\mathcal{F}$ is additive, in the sense that

$$H_{\min}(\mathcal{F} \otimes \Phi) = H_{\min}(\mathcal{F}) + H_{\min}(\Phi) \tag{15}$$

for any channel Φ. This in fact follows from the stronger result

$$H_{\min}(\mathcal{F}) = 0. \tag{16}$$

Proof. We consider the action of $\mathcal{F}$ on the pure state $|\chi\rangle$. We find

$$\mathcal{F}(|\chi\rangle\langle\chi|) = V_1|\chi\rangle\langle\chi|V_1^\dagger + V_2|\chi\rangle\langle\chi|V_2^\dagger. \tag{17}$$

There are now two cases.

Case I: Either V_1 or V_2 is invertible.

Suppose that either V_1 or V_2 is invertible. Suppose without loss of generality that V_1 is invertible, and consider the operator $V_1^{-1}V_2$. This operator has at least one eigenvector; call this eigenvector $|\chi\rangle$, and let the corresponding eigenvalue be λ. We then see that

$$V_2|\chi\rangle = V_1(V_1^{-1}V_2)|\chi\rangle = \lambda V_1|\chi\rangle. \tag{18}$$

Thus, we see that

$$\mathcal{F}(|\chi\rangle\langle\chi|) = (1+|\lambda|^2)V_1|\chi\rangle\langle\chi|V_1^\dagger. \tag{19}$$

Since V_1 is invertible, $V_1|\chi\rangle$ is a nonzero element of the Hilbert space $\mathcal{H}$. Hence, the output $\mathcal{F}(|\chi\rangle\langle\chi|)$ is necessarily a pure state. Since $\mathcal{F}$ is a valid quantum channel, its output is a density matrix, and $\mathcal{F}(|\chi\rangle\langle\chi|)$ is normalized, as required.

However, we see now that the eigenvalues of the density matrix of a pure state $\rho = |\psi\rangle\langle\psi|$ are either 0 or 1. In particular, this implies

$$H(\mathcal{F}(|\chi\rangle\langle\chi|)) = 0, \tag{20}$$

and we find

$$H_{\min}(\mathcal{F}) = 0, \tag{21}$$

and $\mathcal{E}$ is additive, as required.

Case II: Neither V_1 nor V_2 is invertible.

Suppose now that neither V_1 nor V_2 are invertible. In this case, V_1 must have at least one zero-eigenvalue; let $|\chi\rangle$ be a vector annihilated by V_1. Then $V_2|\chi\rangle \neq 0$, since

$$|\chi\rangle = (V_1^\dagger V_1 + V_2^\dagger V_2)|\chi\rangle = V_2^\dagger V_2|\chi\rangle. \tag{22}$$

We then see that

$$\mathcal{F}(|\chi\rangle\langle\chi|) = V_2 |\chi\rangle\langle\chi| V_2^\dagger. \tag{23}$$

Since $V_2|\chi\rangle \neq 0$, we see that $\mathcal{F}(|\chi\rangle\langle\chi|)$ is a pure state, and

$$H(\mathcal{F}(|\chi\rangle\langle\chi|)) = 0. \tag{24}$$

Thus, in this case, we also have

$$H_{\min}(\mathcal{F}) = 0, \tag{25}$$

and $\mathcal{E}$ is additive, as required. □

The above result makes no requirement that the channel $\mathcal{F}$ be unital. We thus see that a large class of non-unital channels must nevertheless be additive. Second, the result makes no requirement on the dimensionality N. In particular, N can be as large as we desire. Thus, we see that channels requiring $D = 2$ Kraus operators are in some sense very simple, regardless of the dimensionality of the space in which they live.

B. Quantum Channels with $D > 2$

We now proceed to the case where $D > 2$. For $D = 2$, there was a very simple and general result; however, the same will not hold true for larger values of D. An intuitive reason for this is that there need not exist a simultaneous eigenvector of $n > 1$ operators. Nevertheless, we may still develop an upper bound on the minimum output entropy of channels with $D > 2$:

Lemma III.2. Let $\mathcal{E}$ be a channel defined as below:

$$\mathcal{E}(\rho) = \sum_{i=1}^{D} p_i U_i \rho U_i^\dagger, \quad \sum_{i=1}^{D} p_i = 1 \tag{26}$$

where ρ is a density matrix over an N-dimensional Hilbert space (here N can be arbitrary) and the $U_i \in U(N)$ are unitary matrices. Then

$$H_{\min}(\mathcal{E}) \leq \log D - \frac{2 \log 2}{D}. \tag{27}$$

Proof. Choose U_i, U_j such that p_i, p_j are the largest possible. Without loss of generality, suppose that U_1, U_2 are thus chosen, and consider the operator $U_1^\dagger U_2$. This operator has N eigenvectors; call such an eigenvector $|\chi\rangle$, and let the corresponding eigenvalue be λ. Since $U_1^\dagger U_2$ is unitary, $|\lambda| = 1$. We then see that

$$U_2|\chi\rangle = U_1(U_1^\dagger U_2)|\chi\rangle = \lambda U_1|\chi\rangle. \tag{28}$$

Thus, we see that

$$\mathcal{E}(|\chi\rangle \langle\chi|) = (p_1 + p_2)U_1 |\chi\rangle \langle\chi| U_1^\dagger + \sum_{i=3}^{D} p_i U_i |\chi\rangle \langle\chi| U_i^\dagger. \tag{29}$$

We now bound the entropy of this density matrix. In particular, we notice that $\mathcal{E}(|\chi\rangle \langle\chi|)$ projects a state $|\psi\rangle$ to the subspace spanned by the following elements:

$$U_1|\chi\rangle, U_3|\chi\rangle, \cdots, U_D|\chi\rangle. \tag{30}$$

Thus, $\mathcal{E}(|\chi\rangle \langle\chi|)$ has at most $D - 1$ nonzero eigenvalues and eigenvectors; call these eigenvalues $q_1, \cdots, q_{D-1}$. Now, it can be shown that the entropy is maximized if the states $U_i|\chi\rangle$ are mutually orthogonal, implying

$$\mathcal{E}(|\chi\rangle \langle\chi|) \leq -(p_1 + p_2) \log(p_1 + p_2) - \sum_{i=3}^{D} p_i \log p_i. \tag{31}$$

Moreover, since p_1, p_2 are the *maximum* possible p_i, we have $p_1 + p_2 \geq \frac{2}{D}$. It can then be shown that the right hand side is then maximized when $p_1 + p_2 = 2/D$ and $p_i = 1/D$ for $i \geq 3$. We thus have

$$H(\mathcal{E}(|\chi\rangle \langle\chi|)) \leq -\frac{2}{D} \log \frac{2}{D} - \frac{D-2}{D} \log \frac{1}{D} = \log D - \frac{2 \log 2}{D}, \tag{32}$$

as claimed. $\square$

In fact, we may also deduce a slightly weaker result for a more general class of channels:

Lemma III.3. Let $\mathcal{F}$ be a channel defined as below:

$$\mathcal{F}(\rho) = \sum_{i=1}^{D} V_i \rho V_i^\dagger, \quad \sum_{i=1}^{D} V_i V_i^\dagger \tag{33}$$

where ρ is a density matrix over an N-dimensional Hilbert space (here N can be arbitrary). Then

$$H_{\min}(\mathcal{F}) \leq \log(D-1) = \log D - \frac{1}{D} + \mathcal{O}(D^{-2}). \tag{34}$$

Proof. See Appendix A. □

Intuitively speaking, the bounds in Lemmas III.2 and III.3 show us that it is *impossible* to guarantee a maximally mixed output of an arbitrary channel $\mathcal{E}$. However, we see also the effect of increasing D: As D is increased, the minimum output entropy of $\mathcal{E}$ can potentially get closer and closer to the maximally mixed value of $\log D$. This allows us to gain a clearer understanding of why exactly we might want to require that D be large: the output of $\mathcal{E}$ can be closer to maximally mixed. Theorem I.2, as proved in in [1], expresses the fact that there exists a construction of $\mathcal{E}$ that almost maximally mixes every input, provided that $N \gg D \gg 1$.

Now, consider a channel $\mathcal{E}$ given by

$$\mathcal{E}(\rho) = \sum_{i=1}^{D} p_i U_i \rho U_i^\dagger \tag{35}$$

where as usual the p_i are probabilities summing to 1 and the U_i are unitary matrices. As per Lemma II.1,

$$H_{\min}(\mathcal{E} \otimes \overline{\mathcal{E}}) \leq 2\log D - \frac{\log D}{D}. \tag{36}$$

Now, as per Lemma III.2,

$$H_{\min}(\mathcal{E}) + H_{\min}(\overline{\mathcal{E}}) = 2H_{\min}(\mathcal{E}) \leq 2\log D - \frac{4\log 2}{D}. \tag{37}$$

We see therefore that $H_{\min}(\mathcal{E}) + H_{\min}(\overline{\mathcal{E}})$ can only possibly exceed the upper bound on $H_{\min}(\mathcal{E} \otimes \overline{\mathcal{E}})$ provided that $D > 16$. We therefore make the following conjecture:

Conjecture 1. Let $\mathcal{E}$ be a quantum channel given by

$$\mathcal{E}(\rho) = \sum_{i=1}^{D} p_i U_i \rho U_i^\dagger \tag{38}$$

where as usual the p_i are probabilities summing to 1 and the U_i are unitary matrices. Then

$$H_{\min}(\mathcal{E} \otimes \overline{\mathcal{E}}) = H_{\min}(\mathcal{E}) + H_{\min}(\overline{\mathcal{E}}) \tag{39}$$

for $D \leq 16$.

Essentially, this conjecture postulates that the output entropy of the maximally entangled state $|\Psi_{\mathrm{ME}}\rangle$ through $\mathcal{E} \otimes \overline{\mathcal{E}}$ is in some sense optimal. That is, the minimum output entropy of $\mathcal{E} \otimes \overline{\mathcal{E}}$ is close enough (or equal to) the output entropy $H(\mathcal{E} \otimes \overline{\mathcal{E}}(|\Psi_{\mathrm{ME}}\rangle \langle \Psi_{\mathrm{ME}}|))$. This conjecture in some sense allows us to quantify the size of D at which we expect the additivity of these quantum channels $\mathcal{E}$ break down.

IV. DISCUSSION AND CONCLUSION

In this paper, we analyzed and developed an intuitive picture of the superadditivity of quantum channels due to entangled inputs. We saw that quantum channels representable using $D = 2$ Kraus operators are additive, and we developed an upper bound on the minimum output entropy of a channel $\mathcal{E}$ using $D > 2$ Kraus operators. Using our results, we developed a conjecture about the additivity of a certain class of quantum channels for small enough values of D.

Our work leaves open many possible questions. In particular, the explicit, non-random, construction of a superadditive quantum channel remains elusive; the randomized construction in [1] provides few hints as to how to find such an explicit construction. If Conjecture 1 were to hold true, we would expect to see superadditive quantum channels $\mathcal{E}$ only in 16-dimensional Hilbert spaces or larger (since $N \geq D$). This potentially explains the apparent difficulty encountered in actually constructing a explicit example of a superadditive quantum channel. To make matters worse, the requirement $N \gg D$ in [1] seems to imply that the dimension N of the requisite Hilbert space is much, much larger than 16, possibly exponentially more. This would appear to make the explicit construction of a superadditive quantum channel even more difficult.

Furthermore, it would be interesting to extend our analysis and provide sharper bounds on the minimum output entropy of channels for $D > 2$. In the future, we hope to conduct a more thorough study of the minimum output entropy bounds of the channels $\mathcal{E}$, perhaps extending the results of Lemmas II.1, III.2, and III.3.

---

[1] M. B. Hastings. Superadditivity of communication capacity using entangled inputs. *Nature Physics*, 5(4):255–257, March 2009.

[2] Christopher King. Additivity for unital qubit channels. *Journal of Mathematical Physics*, 43(10):4641, 2002.

[3] Benjamin Schumacher and Michael D. Westmoreland. Sending classical information via noisy quantum channels. *Phys. Rev. A*, 56:131–138, Jul 1997.

[4] Peter W. Shor. Equivalence of additivity questions in quantum information theory. 2003.

[5] John Watrous. *The Theory of Quantum Information*. Cambridge University Press, April 2018.

Appendix A: Proof of Lemma III.3

Lemma. *Let $\mathcal{F}$ be a channel defined as below:*

$$\mathcal{F}(\rho) = \sum_{i=1}^{D} V_i \rho V_i^\dagger, \quad \sum_{i=1}^{D} V_i V_i^\dagger \tag{A1}$$

where ρ is a density matrix over an N-dimensional Hilbert space (here N can be arbitrary). Then

$$H_{\min}(\mathcal{F}) \leq \log(D-1) = \log D - \frac{1}{D} + \mathcal{O}(D^{-2}). \tag{A2}$$

Proof. The proof proceeds similarly to the proof of Lemma 4.1, and we may also divide our analysis into two cases.

Case I: One of the V_i is invertible.

Suppose that one of the V_i is invertible. Suppose without loss of generality that V_1 is invertible, and consider the operator $V_1^{-1}V_2$. This operator has at least one eigenvector; call this eigenvector $|\chi\rangle$, and let the corresponding eigenvalue be λ. We then see that

$$V_2|\chi\rangle = V_1(V_1^{-1}V_2)|\chi\rangle = \lambda V_1|\chi\rangle. \tag{A3}$$

Thus, we see that

$$\mathcal{F}(|\chi\rangle\langle\chi|) = (1+|\lambda|^2)V_1|\chi\rangle\langle\chi|V_1^\dagger + \sum_{i=3}^{D} V_i|\chi\rangle\langle\chi|V_i^\dagger. \tag{A4}$$

We now bound the entropy of this density matrix. In particular, we notice that $\mathcal{F}(|\chi\rangle\langle\chi|)$ projects a state $|\psi\rangle$ to the subspace spanned by the following elements:

$$V_1|\chi\rangle, V_3|\chi\rangle, \cdots, V_D|\chi\rangle. \tag{A5}$$

Thus, $\mathcal{F}(|\chi\rangle\langle\chi|)$ has at most $D-1$ nonzero eigenvalues; call these eigenvalues $p_1, \cdots, p_{D-1}$. We then have

$$H(\mathcal{F}(|\chi\rangle\langle\chi|)) = -\sum_{i=1}^{D-1} p_i \log p_i. \tag{A6}$$

The right hand side is nonnegative and is maximized when all of the p_i are equal (maximal mixing). Since $\mathcal{F}(|\chi\rangle\langle\chi|)$ is a valid density matrix, the p_i sum to 1, so the maximal entropy is achieved when $p_i = \frac{1}{D-1}$.

$$H(\mathcal{F}(|\chi\rangle\langle\chi|)) = -\sum_{i=1}^{D-1} p_i \log p_i \leq \log(D-1), \tag{A7}$$

as claimed.

Case II: None of the V_i are invertible.

Suppose now that none of the V_i are invertible. Then, V_1 has at least one zero eigenvalue; let $|\chi\rangle$ be a vector annihilated by V_1. Then we have We then see that

$$\mathcal{F}(|\chi\rangle\langle\chi|) = \sum_{i=2}^{D} V_i |\chi\rangle\langle\chi| V_i^\dagger \tag{A8}$$

As in the previous case, $\mathcal{F}(|\chi\rangle\langle\chi|)$ has at most $D-1$ nonzero eigenvalues; call these eigenvalues $p_1, \cdots, p_{D-1}$. Proceeding exactly as in the previous case, we obtain the same bound:

$$H(\mathcal{F}(|\chi\rangle\langle\chi|)) = -\sum_{i=1}^{D-1} p_i \log p_i \leq \log(D-1), \tag{A9}$$

as claimed. □