Contemporary Mathematics

Remarks on the Additivity Conjectures for Quantum Channels

Christopher King

Abstract. In this article we present the statements of the additivity conjectures for quantum channels, together with some brief history and context. The conjectures were recently shown to be false in general, and we include a review of the state of knowledge concerning counterexamples. The article concludes with a short list of current open questions and topics of research arising out of the additivity problem.

1. Introduction

Quantum information theory (QIT) has emerged in the last two decades as a vibrant and exciting field of research. As well as providing a novel perspective on quantum theory, the field has generated new conjectures and results in mathematics. The particular focus of this article is the family of related questions known as the additivity conjectures. These conjectures emerged from the attempt to find a closed form expression for the information-carrying capacity of a noisy quantum channel. There has been much progress toward this goal, but there are still many interesting questions which are the subject of current research. The reader is referred to the papers [**8, 32, 33**] for a fuller account of the early history of this topic.

The most significant recent result is Hastings' proof of the existence of counterexamples to the additivity conjectures [**27**]. It might have been expected that this result would kill the problem, but in fact it has stimulated further research toward finding explicit counterexamples, and has generated new questions about the extent and typicality of the additivity violations that can occur. The purpose of this article is to state and explain the additivity conjectures, and to indicate some directions of current research. There is no attempt to provide a completely comprehensive survey of all research in this field, however it is hoped that enough references are provided to allow easy access to the literature. Progress is rapid, so

1991 *Mathematics Subject Classification.* Primary 81P45, 94A40.

Key words and phrases. Quantum information theory, channel capacity, additivity conjectures.

the state of current research presented here is a snapshot, and may be soon out of date.

2. Quantum channels

The basic object of study in quantum information theory is the quantum channel. Quantum channels arise naturally in several different ways; as a physical description of decoherence, as the quantum analog of Shannon's notion of a discrete memoryless channel, and as a natural class of maps on matrix algebras. We shall quickly review these three descriptions below, starting with the mathematical definition.

2.1. Quantum channel as a completely positive map on a matrix algebra. Stinespring defined completely positive maps in the context of C^*-algebras [**48**]. Here we restrict attention to maps on finite-dimensional matrix algebras (see [**11**] for more details). Let M_n denote the algebra of complex $n \times n$ matrices.

DEFINITION 2.1. *A linear map $\Phi : M_n \to M_m$ is called* completely positive *(or CP) if $\Phi \otimes I_k : M_n \otimes M_k \to M_m \otimes M_k$ is positivity preserving for every $k \geq 1$, where I_k is the identity map on M_k.*

Various characterizations of CP maps are known, we will describe several. The first is the existence of a Kraus representation. Namely, a map $\Phi : M_n \to M_m$ is CP if and only if there are $m \times n$ matrices $A_1, \dots, A_K$ such that for all $M \in M_n$,

$$\Phi(M) = \sum_{i=1}^{K} A_i \, M \, A_i^* \tag{2.1}$$

The matrices $\{A_i\}$ are known as Kraus operators for the map Φ. The Kraus representation is not unique, however if $A_1, \dots, A_K$ and $B_1, \dots, B_L$ are both Kraus representations for the same map Φ with $K \leq L$ then there is a $K \times L$ matrix $W = (w_{ij})$ such that

$$A_i = \sum_j w_{ij} B_j, \quad WW^* = I_K \tag{2.2}$$

Furthermore there is a Kraus representation with $K \leq nm$.

Stinespring's original definition of a CP map was expressed in the Heisenberg representation, namely as an operator on the observables of a quantum system. In the following we will use the Schrödinger representation, and regard a CP map as acting on the states of a quantum system. We first review some definitions of quantum systems and states.

Let $\mathcal{H}$ be the Hilbert space of a finite-dimensional quantum system, thus $\mathcal{H} = \mathbb{C}^n$ for some n. We denote by $\mathcal{S}(\mathcal{H}) \subset \mathcal{M}_n$ the set of states on $\mathcal{H}$, that is the convex set of positive semidefinite operators with trace one. A state is pure if it has rank one, otherwise it is mixed. The pure states are the extreme points of $\mathcal{S}(\mathcal{H})$.

If ρ is a quantum state and Φ is a CP map acting on $\mathcal{S}(\mathcal{H})$ then $\Phi(\rho)$ is required to also be a state, that is a positive semidefinite matrix with trace 1. The CP condition ensures that $\Phi(\rho) \geq 0$, but in order to preserve probabilities the map Φ is also required to be trace-preserving (TP). Thus a quantum channel is finally defined to be a CPTP map between matrix algebras.

Definition 2.2. *A linear map $\Phi : M_n \to M_m$ is called a* quantum channel *if Φ is completely positive and trace preserving (CPTP).*

For a CPTP map Φ, the matrices in the Kraus representation (2.1) must satisfy the condition

$$\sum_{i=1}^{K} A_i^* A_i = I_n \tag{2.3}$$

The Stinespring Dilation Theorem [**48**] implies a second way to describe a quantum channel, namely as an isometric embedding followed by a partial trace. Let $W : \mathbb{C}^n \to \mathbb{C}^d \otimes \mathbb{C}^m$ be an isometric embedding, that is a linear map satisfying $W^* W = I$, and define $\Phi : \mathcal{M}_n \to \mathcal{M}_m$ by

$$\Phi(\rho) = \mathrm{Tr}_{\mathbb{C}^d} W \rho W^* \tag{2.4}$$

Every quantum channel $\Phi : \mathcal{M}_n \to \mathcal{M}_m$ can be described in this way with some choice of W and d. In particular, if the channel Φ is presented by a Kraus representation as in (2.1), then taking $d = K$ and

$$W = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_K \end{pmatrix} \tag{2.5}$$

gives the representation (2.4).

2.2. Quantum channels as models for decoherence. This section describes a physical interpretation of quantum channels (this material is explained in detail in many places, two excellent sources being the text by Nielsen and Chuang [**42**], and the online notes of J. Preskill's course at CalTech). Consider a bipartite quantum model with state space $\mathcal{H} \otimes \mathcal{E}$, where $\mathcal{H}$ is the state space of our system (which we can control and measure, at least to some extent), and where $\mathcal{E}$ is the state space of the environment, which is outside our control. In the Schrödinger picture, the dynamics of the coupled system is governed by an interaction Hamiltonian H leading to the state evolution

$$\theta \mapsto e^{-iHt/\hbar}\, \theta\, e^{iHt/\hbar} = U \theta U^* \tag{2.6}$$

where $U = e^{-iHt/\hbar}$ (we henceforth suppress the time dependence as we are not concerned with dynamics, but rather with the description of the system at a fixed time). In quantum theory an observable is represented by a Hermitian operator acting on the state space. A *local observable* A of the system $\mathcal{H}$ acts on the coupled system as $A \otimes I$, where I is the identity operator on $\mathcal{E}$. The expected value of the result of a measurement of $A \otimes I$ in the state $U \theta U^*$ is

$$\mathrm{Tr}\ [(A \otimes I)\, U \theta U^*] = \mathrm{Tr}_{\mathcal{H}} A \rho \tag{2.7}$$

where ρ is the reduced density matrix of the system given by

$$\rho = \mathrm{Tr}_{\mathcal{E}}\ [U \theta U^*] \tag{2.8}$$

Here $\mathrm{Tr}_{\mathcal{E}}$ is the partial trace over the environment state space, and $\mathrm{Tr}_{\mathcal{H}}$ is the partial trace over the system. Thus ρ contains all information about the state of the coupled system which can be accessed by local measurements.

Suppose now that the system is prepared in a state σ (this assumes that the experimenter can isolate the system from its environment for long enough to prepare the state). Then the initial state of the coupled system will be a product $\sigma \otimes \omega$ where ω is some state of the environment. Thus ρ in (2.8) can be viewed as the result of a linear map applied to the initial state σ:

$$\rho = \mathrm{Tr}_{\mathcal{E}} \left[U\,(\sigma \otimes \omega)\,U^*\right] = \Phi(\sigma) \tag{2.9}$$

which serves as the definition of the map Φ. In general Φ maps pure states into mixed states, which corresponds to our view of decoherence as introducing noise into a system through entanglement with the environment. From the definition (2.9) it follows that Φ is trace-preserving and completely positive, and thus Φ is a quantum channel. In the case where ω is a pure state it can be seen that (2.9) is equivalent to the formulation (2.4), with the same input and output spaces (if ω is a mixed state then Φ can still be written in the form (2.4) but with a larger environment).

2.3. Quantum channel as an information device. Shannon's model of a discrete memoryless channel [**44**] is based on the notion that an information source can be viewed as a stochastic process, producing strings of random letters drawn from a source alphabet. Such a string is then transmitted through a channel, and the output of the channel is another stochastic process which is correlated with the input. The simplest assumption to make is that the channel acts independently on each letter, randomly changing it according to a fixed transition matrix $\{p_{ij}\}$. That is, letting X denote the input and Y the output letters,

$$P(Y = j|X = i) = p_{ij}, \quad \sum_j p_{ij} = 1 \tag{2.10}$$

If X is a random variable with distribution $\pi_i = P(X = i)$, then the distribution of Y is given by

$$q_j = P(Y = j) = \sum_i \pi_i\, p_{ij} \tag{2.11}$$

From this point of view a discrete memoryless channel is a linear map T on probability distributions $\pi = (\pi_1, \pi_2, \dots)$:

$$T\,:\,\pi \mapsto q, \quad q_j = T(\pi)_j = \sum_i \pi_i\, p_{ij} \tag{2.12}$$

The channel acts independently on successive letters in the input string:

$$P(Y_1 = j_1, \dots, Y_m = j_m|X_1 = i_1, \dots, X_m = i_m) = p_{i_1,j_1} \cdots p_{i_m,j_m} \tag{2.13}$$

This can be expressed as the action of the m-fold product map $T^{\times m}$ on input product distributions: letting $\pi(X)$ denote the distribution of the input letter X,

$$T^{\times m}(\pi(X_1) \times \cdots \times \pi(X_m)) = T(\pi(X_1)) \times \cdots \times T(\pi(X_m)) \tag{2.14}$$

By linearity the map $T^{\times m}$ extends uniquely to a map on the set of all probability distributions on m-letter input strings, and it is straightforward to check that this map is also a channel.

This viewpoint on a classical channel leads directly to the definition of a quantum channel as a linear map on quantum states. Namely, let $\mathcal{H}_{in}$ and $\mathcal{H}_{out}$ be the input and output state spaces of the channel. Then the set of states $\mathcal{S}(\mathcal{H}_{in})$

is the quantum analog of the set of input probability distributions for the classical channel T, and similarly $\mathcal{S}(\mathcal{H}_{out})$ is the analog of the output distributions. Thus a quantum channel is a linear map

$$\Phi \,:\, \mathcal{S}(\mathcal{H}_{in}) \to \mathcal{S}(\mathcal{H}_{out}) \tag{2.15}$$

Physical considerations imply that Φ should be trace-preserving and positivity preserving. Furthermore, as in the case of a classical channel, the quantum channel acts independently on successive states in a string of inputs; letting $\rho_1, \ldots, \rho_m$ denote the input states, the output string is

$$\Phi^{\otimes m}(\rho_1 \otimes \cdots \otimes \rho_m) = \Phi(\rho_1) \otimes \cdots \otimes \Phi(\rho_m) \tag{2.16}$$

As in the classical case, by linearity the map $\Phi^{\otimes m}$ extends to a map on the states of the full tensor product of input spaces $\mathcal{H}_{in}^{\otimes m}$. However, unlike in the classical case, it does *not* follow automatically that $\Phi^{\otimes m}$ is itself a quantum channel. This requires the additional assumption that Φ is completely positive.

3. The capacity of a quantum channel

3.1. Classical channel. Shannon defined the capacity of a discrete memoryless channel as the maximum rate for transmission of information through the channel [**44**]. This maximum rate is approached by encoding the information in input strings which are sufficiently different that the resulting output strings can be reliably distinguished. By using longer and longer strings to encode the information, the maximum rate can be asymptotically approached. Furthermore, Shannon provided an explicit formula for this maximum rate. Suppose that the input letters X have distribution $\{\pi_i\}$ and the channel matrix is $\{p_{ij}\}$, then the mutual information of the input and output is defined as

$$I(X,Y) = \sum_{i,j} \pi_i \, p_{ij} \, \log \frac{p_{ij}}{q_j} \tag{3.1}$$

where again $\{q_j\}$ is the distribution of the channel output Y. Shannon proved that the maximum rate for information transmission using the channel Φ with input source X is $I(X,Y)$. Following this, the classical channel capacity C_{class} is defined to be the maximum of $I(X,Y)$ over all possible distributions of X, that is

$$C_{class}(T) = \sup_{\pi_i} I(X,Y) \tag{3.2}$$

3.2. Shannon capacity of quantum channel. The capacity of a quantum channel is defined by viewing it as a particular realization of a classical channel. That is, one considers the use of the channel for transmission of a signal composed of a string of letters drawn from a finite alphabet. Transmission is achieved by first encoding the input signal in a quantum state, then allowing the channel to act on the state, and finally measuring the output state in order to recover the information.

In the simplest protocol each letter i of the input alphabet is encoded as a quantum state ρ_i in the input space $\mathcal{H}_{in}$. This input state is mapped by the channel to an output state $\Phi(\rho_i)$. At the output a measurement is performed. Recall that in quantum theory a measurement is defined by a POVM, that is a collection of positive semidefinite matrices $E_1, \ldots, E_k$ satisfying $\sum_{j=1}^{k} E_j = I$. When applied to the output state $\Phi(\rho_i)$, this measurement returns the index j with probability

$\operatorname{Tr} E_j\,\Phi(\rho_i)$. The classical channel is thus constructed by choosing a set of states $\{\rho_i\}$ to encode the input letters, and choosing a POVM $\{E_j\}$ to measure the output. The transition matrix of the channel is

$$p_{ij} = \operatorname{Tr} E_j\,\Phi(\rho_i) \tag{3.3}$$

Now the formula (3.2) provides the capacity of this channel. The *Shannon capacity* of Φ is then defined to be the maximum of this capacity taken over all choices of input encoding states and output measurements, that is

$$C_{\mathrm{Shan}}(\Phi) = \sup_{\rho_i, E_j} C_{class}(T) \tag{3.4}$$

where T is the classical channel with transition matrix (3.3). Note that the number of input states and the number of POVM elements is not fixed on the right side of (3.4), and the supremum includes a search over all sizes of these sets (though there are dimension-dependent bounds for the number of states and POVM elements needed).

3.3. Entangled inputs and outputs. However this is not the end of the story. In the formula (3.4) there is an implicit assumption about the way that input strings are encoded, namely as products drawn from a fixed set of states. For example, suppose that the alphabet is $\{0,1\}$, and we want to efficiently transmit the four strings $\{00, 01, 10, 11\}$. Using the above product state protocol we would select two input states ρ_0, ρ_1 and encode these strings as the product states

$$00 \mapsto \rho_0 \otimes \rho_0, \quad 01 \mapsto \rho_0 \otimes \rho_1, \quad 10 \mapsto \rho_1 \otimes \rho_0, \quad 11 \mapsto \rho_1 \otimes \rho_1 \tag{3.5}$$

Then at the output we select a POVM $\{E_0, E_1\}$ which tries to distinguish the states $\Phi(\rho_0)$ and $\Phi(\rho_1)$. The average error probability for these four strings will be

$$p_e = \frac{1}{4} \sum_{i,j=0}^{1} [1 - \operatorname{Tr} E_i \Phi(\rho_i)\ \operatorname{Tr} E_j \Phi(\rho_j)] \tag{3.6}$$

However there may be another way to encode and decode the strings that produces a smaller error probability. For example, the four input strings could be encoded using the four Bell states: these are defined as [**42**]

$$\begin{aligned}
|\beta_{00}\rangle &= \frac{1}{\sqrt{2}}\,(|00\rangle + |11\rangle) \\
|\beta_{01}\rangle &= \frac{1}{\sqrt{2}}\,(|01\rangle + |10\rangle) \\
|\beta_{10}\rangle &= \frac{1}{\sqrt{2}}\,(|00\rangle - |11\rangle) \\
|\beta_{11}\rangle &= \frac{1}{\sqrt{2}}\,(|01\rangle - |10\rangle)
\end{aligned}$$

The encoding would then be

$$00 \mapsto |\beta_{00}\rangle\langle\beta_{00}|, \quad 01 \mapsto |\beta_{01}\rangle\langle\beta_{01}|, \quad 10 \mapsto |\beta_{10}\rangle\langle\beta_{10}|, \quad 11 \mapsto |\beta_{11}\rangle\langle\beta_{11}| \tag{3.7}$$

Furthermore at the output we may also select a POVM that projects onto states which are entangled across the outputs. So we choose a POVM $\{E_{00}, E_{01}, E_{10}, E_{11}\}$

where $\sum_{i,j} E_{ij} = I \otimes I$. Then the average error probability becomes

$$p_e' = \frac{1}{4} \sum_{i,j=0}^{1} [1 - \operatorname{Tr} E_{ij} (\Phi \otimes \Phi)(|\beta_{ij}\rangle\langle\beta_{ij}|)] \tag{3.8}$$

If $p_e' < p_e$ for any choice of single-letter protocols $\{\rho_i, E_i\}$ then we may increase the channel capacity beyond $C_{\text{Shan}}(\Phi)$ by encoding input strings with this entangled protocol, thus implying that

$$C_{\text{Shan}}(\Phi \otimes \Phi) > 2\, C_{\text{Shan}}(\Phi) \tag{3.9}$$

The existence of channels satisfying the superadditivity property (3.9) was demonstrated in Holevo's 1973 paper [**31**].

In general, it is possible to encode an input string using a state which is entangled across multiple channel inputs, and it is possible to use a POVM which uses operators which are entangled across the channel outputs. When such entangled encodings and measurements are considered over n uses of the channel the resulting capacity is

$$\frac{1}{n}\, C_{\text{Shan}}(\Phi^{\otimes n}) \tag{3.10}$$

(The factor $1/n$ is needed because we consider information transfer per channel use). By allowing n to increase arbitrarily we reach the 'ultimate' capacity which is given by

$$C_{ult}(\Phi) = \lim_{n \to \infty} \frac{1}{n}\, C_{\text{Shan}}(\Phi^{\otimes n}) \tag{3.11}$$

Note that for a classical channel T superadditivity does not occur, and thus $C_{ult}(T) = C_{class}(T)$.

3.4. The Holevo capacity. The Holevo capacity of the channel Φ is defined as [**31, 32, 33**]

$$\chi(\Phi) = \sup_{p_i, \rho_i} \left[S\left(\Phi(\sum_i p_i\, \rho_i) \right) - \sum_i p_i S(\Phi(\rho_i)) \right] \tag{3.12}$$

where the sup on the right side runs over all input ensembles for the channel, and where $S(\cdot)$ is the von Neumann entropy (the function inside the sup is convex and hence the ensemble may be assumed to consist of pure states). Holevo proved the following bound for the Shannon capacity:

$$C_{\text{Shan}}(\Phi) \leq \chi(\Phi) \tag{3.13}$$

(this bound had also appeared in earlier work [**25**], [**40**]). The Holevo capacity was given an operational meaning through the later work of Hausladen et al [**28**], Holevo [**33**], and Schumacher and Westmoreland [**43**], who proved that $\chi(\Phi)$ is equal to a restricted version of the capacity $C_{ult}(\Phi)$. The restricted version is obtained by allowing entangled measurements at the output for multiple channel uses, but allowing only product input states. It follows that

$$\chi(\Phi) \leq C_{ult}(\Phi) \tag{3.14}$$

Furthermore the entangled input states may be re-introduced by considering multiple channel uses, thus leading to

$$C_{ult}(\Phi) = \lim_{n\to\infty} \frac{1}{n}\, \chi(\Phi^{\otimes n}) \tag{3.15}$$

4. The additivity conjectures

4.1. The additivity conjecture for Holevo capacity. The original additivity conjecture [**7**] was to the effect that the regularization is unnecessary in (3.15), meaning that it can be replaced by the simpler 'one-shot' formula

$$C_{ult}(\Phi) = \chi(\Phi) \tag{4.1}$$

Equivalently, the function χ is additive over n-fold tensor products:

$$\chi(\Phi^{\otimes n}) = n\, \chi(\Phi) \tag{4.2}$$

A slightly generalized version of this soon became the standard additivity conjecture: for any two quantum channels Φ and Ω,

$$\chi(\Phi \otimes \Omega) = \chi(\Phi) + \chi(\Omega) \tag{4.3}$$

There is an operational meaning for this conjecture. It says that the channel capacity is achieved using coding on product states only, in other words using entangled input states for the channel does not increase the capacity. It was known that entangled measurements at the output are necessary to achieve the Holevo capacity, and hence also the full channel capacity, but this conjecture implies that the input states can always be chosen from an ensemble consisting only of product states.

4.2. Equivalence to other additivity conjectures. In a quest for new approaches to the additivity problem, the minimal output entropy and minimal output Renyi entropy were studied. These are:

$$S_{\min}(\Phi) = \inf_{\rho} S(\Phi(\rho)) \tag{4.4}$$

and for $p > 1$

$$S_{p,\min}(\Phi) = \inf_{\rho} \frac{1}{1-p}\, \log \mathrm{Tr}\,(\Phi(\rho)^p) \tag{4.5}$$

Note that $\lim_{p\to 1} S_{p,\min}(\Phi) = S_{\min}(\Phi)$. The additivity conjecture is: for all channels Φ and Ω, and all $p \geq 1$

$$S_{p,\min}(\Phi \otimes \Omega) = S_{p,\min}(\Phi) + S_{p,\min}(\Omega) \tag{4.6}$$

The question of additivity of minimal output entropy was posed in the paper [**39**], where it was conjectured that this would provide an indirect way to attack the additivity problem for Holevo capacity. This approach was confirmed in 2002 by Shor [**47**], who proved the equivalence of several additivity conjectures, including additivity of Holevo capacity and additivity of minimal output von Neumann entropy (this result involved also the entanglement of formation but we will not consider that quantity here).

Following an influential article by Amosov, Holevo and Werner [**4**], it was believed that a promising method for proving additivity of minimal output entropy was to prove first (4.6) for $p > 1$, and then hope to recover additivity in the limit $p \to 1$. For some special classes of channels this turned out to be a fruitful approach, and led to proofs of additivity, as the following list of papers shows: [**1**], [**2**], [**3**], [**10**], [**16**], [**17**], [**18**], [**19**], [**20**], [**23**], [**24**], [**34**], [**35**], [**36**], [**37**], [**38**], [**41**], [**45**], [**46**].

5. The counterexamples

As mentioned above the additivity conjectures are now known to be false. Historically the minimal Renyi entropy was first shown to be non-additive [**49**] for large values of p, with successive counterexamples lowering the value, until finally it was shown for all $p > 1$ [**30**]. The final breakthrough came when Hastings [**27**] proved the existence of counterexamples at $p = 1$, thereby disproving the original conjectures.

The first family of counterexamples was discovered by Werner and Holevo [**49**]. These channels are highly structured and symmetric, and this suggested that the search for counterexamples should be directed toward similarly special classes. However the breakthrough came with A. Winter's 2007 paper [**50**] where counterexamples for all $p > 2$ were proven using random channels. The channels were random unitary channels of the form

$$\Phi(\rho) = \frac{1}{N} \sum_{i=1}^{N} U_i \, \rho \, U_i^* \tag{5.1}$$

where U_i are randomly selected $d \times d$ unitary matrices. Winter's key observation was that for any choice of random unitaries the product channel $\Phi \otimes \overline{\Phi}$ when applied to the maximally entangled state has a large eigenvalue, and that this in turn gives a useful upper bound for $S_{\min}(\Phi \otimes \overline{\Phi})$. When combined with a lower bound for $S_{\min}(\Phi) = S_{\min}(\overline{\Phi})$ this provides the contradiction to additivity. The hard part of the proof is finding a good lower bound for $S_{\min}(\Phi)$. Winter's method was non-constructive, and used a randomized argument to imply the existence of such channels. This randomizing argument was extended by Hayden and Winter in the paper [**30**], where it was used to prove the existence of counterexamples for all $p > 1$. There was a brief hope that additivity might hold for $p \leq 1$, but counterexamples to this conjecture were also found [**15**].

Shortly afterwards Hastings [**27**] extended the reach of the counterexamples by introducing some new ideas and techniques. He adopted the same general approach of looking at product channels of the form $\Phi \otimes \overline{\Phi}$ where Φ is a random unitary channel. His main contribution was to find improved lower bounds for $S_{\min}(\Phi)$. Again the argument is based on a randomized technique and is non-constructive. By exploiting the explicit form of the eigenvalue distribution for the reduced density matrix of a random pure bipartite state, combined with a novel idea for estimating the probability of low entropy output states, Hastings was able to prove the existence of channels for which minimal output entropy is non-additive.

6. Current directions of research and open problems

6.1. Explicit counterexamples. As mentioned before, Werner and Holevo found explicit channels which violate additivity of minimal Renyi entropy for all $p > 4.79$. Recently, Grudka, Horodecki and Pankowski [**26**] have found explicit channels which violate additivity of Renyi entropy for all $p > 2$. It is very tempting to believe now that explicit examples at $p = 1$ may be found soon.

6.2. Random subspaces and channels. Hayden, Leung and Winter [**29**] investigated the entanglement properties of random subspaces, using concentration of measure arguments and other tools from random matrix analysis. More recently, several authors have developed new approaches to finding bounds for entanglement, some based on the new methods introduced by Hastings [**9**], [**12**], [**13**], [**14**], [**22**].

6.3. Bounds for capacity. Since additivity fails, the convenient 'one-shot' formula (4.1) for channel capacity does not hold, and instead the more awkward regularized formula (3.15) must be used. This raises the question of finding useful bounds for the capacity, which is related to the question of finding bounds for the size of the violation of additivity. The Hayden-Winter examples provide large violations of additivity for all $p > 1$, however the dimensions of the spaces diverge as p approaches 1. In contrast, the method of proof used by Hastings produces a small violation of additivity at $p = 1$ [**21**], and it remains an open question whether larger violations are possible.

6.4. Additivity of capacity over different channels. The question is whether $C_{ult}(\Phi \otimes \Omega) = C_{ult}(\Phi) + C_{ult}(\Omega)$ for two different channels Φ and Ω. Based on experience with the additivity conjectures, it seems reasonable to expect that this is false.

6.5. Geometrical approach. Recently, Aubrun, Szarek and Werner [**5**] have used Dvoretzky's Theorem from convex geometry to give a new proof of existence of counterexamples for all $p > 1$. Dvoretzky's Theorem concerns the existence of almost spherical cross-sections of high-dimensional convex bodies, and the additivity problem for Renyi entropy can be restated in precisely this form. The dimensions of the counterexamples diverge as $p \downarrow 1$, however in another paper [**6**] the same authors have used related methods to prove existence of counterexamples at $p = 1$.

6.6. Non-unital qubit channels. Qubit channels are the simplest quantum channels [**39**]. It is known that additivity holds for unital qubit channels for all $p \geq 1$ [**34**], and for non-unital channels at $p = 2$ and $p \geq 4$ [**37**], [**20**]. However the additivity question for non-unital channels for $p < 2$ is still open, and the additivity of channel capacity is still open. There is no evidence that qubit channels can violate additivity, however it seems worthwhile to settle this question.

References

[1] R. Alicki and M. Fannes, "Note on multiple additivity of minimal entropy output of extreme $SU(d)$-covariant channels", *Open Systems and Information Dynamics* **11**, 339 – 342 (2004).

[2] G. G. Amosov, "On Weyl channels being covariant with respect to the maximum commutative group of unitaries", *J. Math. Phys.* **48**, 2104 – 2117 (2007).

[3] G. G. Amosov, "The strong superadditivity conjecture holds for the quantum depolarizing channel in any dimension", *Phys. Rev. A* **75**, 2104 – 2117 (2007).

[4] G. G. Amosov, A. S. Holevo and R. F. Werner, "On some additivity problems in quantum information theory", *Problems in Information Transmission* **36**, 305 – 313 (2000).

[5] G. Aubrun, S. Szarek and E. Werner, "Non-additivity of Renyi entropy and Dvoretzky's Theorem", *J. Math. Phys.* **51**, 022102 (2010).

[6] G. Aubrun, S. Szarek and E. Werner, "Hastings' additivity counterexample via Dvoretzky's theorem", arXiv:1003.4925.

[7] C. H. Bennett, C. A. Fuchs and J. A. Smolin, "Entanglement-enhanced classical communication on a noisy quantum channel", *Quantum Communication, Computing and Measurement*, eds. O. Hirota, A. S. Holevo and C. M. Caves (Plenum Press, NY 1997), 79 – 88.

[8] C. H. Bennett and P. W. Shor, "Quantum Information Theory", *IEEE Trans. Info. Theory* **44**, 2724 – 2742 (1998).
[9] F. G. S. L. Brandao and M. Horodecki, "On Hastings' counterexamples to the minimum output entropy additivity conjecture", arXiv:0907.3210.
[10] D. Bruss, L. Faoro, C. Macchiavello and M. Palma, "Quantum entanglement and classical communication through a depolarizing channel", *J. Mod. Opt.* **47** 325 (2000).
[11] M. Choi, "Completely Positive Linear Maps on Complex matrices", *Linear Algebra and Its Applications* **12**, 95 – 100 (1975).
[12] B. Collins and I. Nechita, "Random quantum channels I: graphical calculus and the Bell state phenomenon", arXiv:0905.2313
[13] B. Collins and I. Nechita, "Random quantum channels II: Entanglement of random subspaces, Renyi entropy estimates and additivity problems", arXiv:0906.1877
[14] B. Collins and I. Nechita, "Gaussianization and eigenvalue statistics for Random quantum channels (III)", arXiv:0910.1768
[15] 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", *Commun. Math. Phys.* **284**, 281 – 290 (2008).
[16] N. Datta and M. B. Ruskai, "Maximal output purity and capacity for asymmetric unital qudit channels", *J. Phys. A: Math. Gen.* **38**, 9785 – 9802 (2005).
[17] N. Datta, M. Fukuda and A. S. Holevo, "Complementarity and additivity for covariant channels", *Quant. Info. Proc.* **5**, 179 – 207 (2006).
[18] B. Dierckx and M. Fannes, "Additivity of the renyi entropy of order 2 for positive-partial-transpose-inducing channels", *Phys. Rev. A*, **77**, art.nr. 060302 (2008).
[19] A. Fujiwara and T. Hashizume, "Additivity of the capacity of depolarizing channels", *Phys. Lett. A* **299**, 469 – 475 (2002).
[20] M. Fukuda, "An application of decomposable maps in proving multiplicativity of low dimensional maps", *J. Math. Phys.* **51**, 022201 (2010).
[21] M. Fukuda, C. King and D. Moser, "Comments on Hastings' additivity counterexamples", *Commun. Math. Phys.* **296**, no. 1, 111 (2010).
[22] M. Fukuda and C. King, "Entanglement of random subspaces via the Hastings bound", *J. Math. Phys.* **51**, 042201 (2010).
[23] V. Giovannetti and S. Lloyd, "Additivity properties of a Gaussian channel", *Phys. Rev. A* **69** (6): Art. No. 062307, (2004).
[24] V. Giovannetti, S. Lloyd and M. B. Ruskai, "Conditions for multiplicativity of maximal l_p-norms for fixed integer p", *J. Math. Phys.* **46**, 042105 (2005).
[25] J. P. Gordon, "Noise at optical frequencies: Information theory", *Proceedings of the International School of Physics "Enrico Fermi"*, ed. P. A. Mills (Academic Press, New York, 1964), 156 – 181.
[26] A. Grudka, M. Horodecki and L. Pankowski, "Constructive counterexamples to additivity of minimum output Renyi entropy of quantum channels for all $p > 2$", arXiv:0911.2515 [quant-ph].
[27] M. B. Hastings, "Superadditivity of communication capacity using entangled inputs", *Nature Phys.* **5**, 255 – 257 (2008), arXiv:0809.3972.
[28] P. Hausladen, R. Jozsa, B, Schumacher, M. Westmoreland, W. Wootters, "Classical information capacity of a quantum channel", *Phys Rev A* **54**, no. 3, 1869 – 1876 (1996).
[29] P. Hayden, D. Leung and A. Winter, "Aspects of generic entanglement", *Commun. Math. Phys.* **265**, 95 – 117 (2007).
[30] P. Hayden and A. Winter, "Counterexamples to the maximal p-norm multiplicativity conjecture for all $p > 1$", *Commun. Math. Phys.* **284**, 263 – 280 (2008).
[31] A. S. Holevo, "Bounds for the quantity of information transmitted by a quantum communication channel", *Probl. Inf. Transm. (USSR)* **9**, 177 – 183 (1973).
[32] A. S. Holevo, "The capacity of the quantum channel with general signal states", *IEEE Trans. Info. Theory* **44**, 269 – 273 (1998).
[33] A. S. Holevo, "Coding theorems for quantum channels", Research Reviews of Tamagawa University, N4, (1998) (available at arXiv:quant-ph/9809023).
[34] C. King, "Additivity for unital qubit channels", *J. Math. Phys.* **43**, no.3, 1247 – 1260 (2002).
[35] C. King, "The capacity of the quantum depolarizing channel", *IEEE Trans. Info. Theory* **49**, no.1, 221 – 229 (2003).

[36] C. King, “Maximal p-norms of entanglement breaking channels”, *Quantum Info. and Comp.* **3**, no.2, 186 – 190 (2003).
[37] C. King and N. Koldan, “New multiplicativity results for qubit maps”, *J. Math. Phys.* **47**, 042106 (2006).
[38] C. King, M. Nathanson and M. B. Ruskai, “Multiplicativity properties of entrywise positive maps”, *Lin. Alg. and its Appl.* **404**, 367 – 379 (2005).
[39] C. King and M. B. Ruskai, “Minimal entropy of states emerging from noisy quantum channels”, *IEEE Trans. Info. Theory* **47**, 192 – 209 (2001).
[40] L. B. Levitin, “On the quantum measure of the amount of information”, *Proceedings of the Fourth all-union conference on Information Theory (in Russian)*, Tashkent, 111 – 115 (1969).
[41] S. Michalakis, “Multiplicativity of the maximal output 2-norm for depolarized Werner-Holevo channels”, *J. Math. Phys.* **48**, no. 12, 122102, (2007).
[42] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press (2000).
[43] B. Schumacher and M. Westmoreland, “Sending classical information via a noisy quantum channel”, *Phys. Rev. A* **56**, 131 – 138 (1997).
[44] C. E. Shannon, “A mathematical theory of communication”, *Bell System Tech. J.* **27**, 379 – 423, 623 – 656 (1948).
[45] M. E. Shirokov, “The Holevo capacity of infinite dimensional channels and the additivity problem”, *Commun. Math. Phys.* **262**, 137 – 159 (2006).
[46] P. W. Shor, “Additivity of the classical capacity of entanglement-breaking quantum channels”, *J. Math. Phys.* **43**, 4334 – 4340 (2002).
[47] P. W. Shor, “Equivalence of additivity questions in quantum information theory”, *Comm. Math. Phys.* **246**, 453 – 472 (2004).
[48] W. F. Stinespring, “Positive Functions on C^*-algebras”, *Proceedings of the American Mathematical Society* **6**, 211 – 216 (1955).
[49] R. F. Werner and A. S. Holevo, “Counterexample to an additivity conjecture for output purity of quantum channels”, *J. Math. Phys.* **43**, no.9, 4353 – 4357 (2002).
[50] A. Winter, “The maximum output p-norm of quantum channels is not multiplicative for any $p > 2$”, ArXiv:0707.0402.

Department of Mathematics, Northeastern University, Boston, Massachusetts 02115
E-mail address: `c.king@neu.edu`