

# Additivity counterexamples in quantum information theory via random matrix methods.

C. King, Northeastern University  
Joint work with M. Fukuda, UC Davis

Harvard Random Matrix Seminar, November 2009

The longstanding conjecture that *quantum channel capacity is additive* was finally disproved in 2008 by Matt Hastings using random matrix methods.

The goal of this talk is to introduce the problem, explain some aspects of Hastings' paper related to random matrix techniques, and show how his ideas lead to a general result on 'worst-case' entanglement of random subspaces. Some open problems will also be discussed.

The general setting concerns finite-dimensional quantum channels. A quantum channel can be defined in terms of an isometric embedding of a subspace in a product space, and we start with this viewpoint. We review some ideas about entanglement of bipartite states, and then use this approach to define the notion of a channel.

# Entanglement of a bipartite state

Consider a product of Hilbert spaces  $A \otimes B$ . The *entanglement* of a state  $|\psi\rangle \in A \otimes B$  is defined to be

$$E(|\psi\rangle) = S(\text{Tr}_A |\psi\rangle\langle\psi|) = S(\text{Tr}_B |\psi\rangle\langle\psi|)$$

where  $S(\cdot)$  is the von Neumann entropy:

$$S(\rho) = -\text{Tr} \rho \log \rho$$

and  $\text{Tr}_B |\psi\rangle\langle\psi|$  is the reduced density matrix of the orthogonal projector onto  $|\psi\rangle$ . It satisfies the bounds

$$0 \leq E(|\psi\rangle) \leq \log \min\{\dim(A), \dim(B)\}$$

$E(|\psi\rangle) = 0$  if and only if  $|\psi\rangle$  is a product state.

$E(|\psi\rangle) = \log \min\{\dim(A), \dim(B)\}$  if and only if  $|\psi\rangle$  is a maximally entangled state.

# Entanglement of a subspace

Now consider a subspace  $C$  of  $A \otimes B$ . The entanglement of  $C \subset A \otimes B$  is defined to be

$$E(C) = \inf_{|\psi\rangle} E(|\psi\rangle) \quad (1)$$

where the infimum runs over normalized states  $|\psi\rangle$  in  $C$ . Note that  $E(C) \geq 0$  with equality if and only if  $C$  contains a product state.

Note that again

$$0 \leq E(C) \leq \log \min\{\dim(A), \dim(B)\}$$

It is easy to find subspaces with small entanglement, for example  $E(C) = 0$  if  $C$  is a product of subspaces in  $A$  and  $B$ .

We will be interested in finding subspaces whose entanglement is *large*. In such a subspace all states are close to a maximally entangled state. It is not obvious that there are high-dimensional subspaces with large entanglement.

The approach to this question uses a randomizing argument. Subspaces with large entanglement are proved to exist without explicitly constructing them.

First we make the link to quantum channels.

Let  $d = \text{Dim } A$ ,  $n = \text{Dim } B$  and  $s = \text{Dim } C$ . Then the subspace  $C \subset A \otimes B$  is defined by an isometric embedding  $W : \mathbb{C}^s \rightarrow \mathbb{C}^d \otimes \mathbb{C}^n$  satisfying  $W^*W = I$ , where  $C$  is the image of  $W$ .

Use  $W$  to define the map

$$\begin{aligned}\Phi_W &: \mathbb{C}^{s \times s} \rightarrow \mathbb{C}^{n \times n} \\ \Phi_W(\rho) &= \text{Tr}_{\mathbb{C}^{d \times d}} W \rho W^*\end{aligned}$$

The map  $\Phi_W$  is a *quantum channel*: it is trace-preserving and completely positive, meaning that  $\Phi_W \otimes I_k$  is positivity-preserving for any  $k$ . Every quantum channel arises in this way from an isometric embedding (the representation is not unique).

There is another channel defined by  $W$ , namely the 'complementary' channel

$$\begin{aligned}\Phi_W^C &: \mathbb{C}^{s \times s} \rightarrow \mathbb{C}^{d \times d} \\ \Phi_W(\rho) &= \text{Tr}_{\mathbb{C}^{n \times n}} W \rho W^*\end{aligned}$$

There is a one-to-one correspondence between isometric embeddings  $W$  and pairs of complementary channels  $(\Phi_W, \Phi_W^C)$ .

Every channel  $\Phi_W$  has a Kraus representation with  $d$  operators:

$$\Phi_W(\rho) = \sum_{k=1}^d A_k \rho A_k^*$$

where  $A_k : \mathbb{C}^s \rightarrow \mathbb{C}^n$  and  $\sum_k A_k^* A_k = I$ .

# Minimum output entropy

The *minimum output entropy* of a channel  $\Phi_W$  is defined by

$$\begin{aligned} S_{\min}(\Phi_W) &= \inf_{|\phi\rangle} \left\{ S(\Phi_W(|\phi\rangle\langle\phi|)) \right\} \\ &= \inf_{|\phi\rangle} \left\{ S(\text{Tr}_{\mathbb{C}^{d \times d}} W |\phi\rangle\langle\phi| W^*) \right\} \\ &= \inf_{|\psi\rangle \in \mathcal{C}} \left\{ S(\text{Tr}_{\mathbb{C}^{d \times d}} |\psi\rangle\langle\psi|) \right\} \\ &= E(\mathcal{C}) \end{aligned}$$

where  $\mathcal{C} = \text{Image}(W)$ . It follows that

$$E(\mathcal{C}) = S_{\min}(\Phi_W) = S_{\min}(\overline{\Phi}_W) \quad (2)$$

Thus a subspace with large entanglement defines a channel with large minimum output entropy, that is a *noisy* channel.

**Conjecture 1** For any quantum channels  $\Phi$  and  $\Omega$ ,

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

The right side can be achieved by restricting the infimum on the left side to product states, hence it is always true that

$$S_{\min}(\Phi \otimes \Omega) \leq S_{\min}(\Phi) + S_{\min}(\Omega).$$

So the question is whether entangled input states can yield a smaller output entropy.

## Conjecture 2

Define the Holevo quantity for a channel:

$$\chi(\Phi) = \sup_{\rho_i, |\psi_i\rangle} \left\{ S \left( \sum_i \rho_i \Phi(|\psi_i\rangle\langle\psi_i|) \right) - \sum_i \rho_i S(\Phi(|\psi_i\rangle\langle\psi_i|)) \right\}$$

We will see later how this is related to the channel capacity of  $\Phi$ .

Conjecture 2 For any quantum channels  $\Phi$  and  $\Omega$ ,

$$\chi(\Phi \otimes \Omega) = \chi(\Phi) + \chi(\Omega)$$

Conjectures 1,2 are equivalent, and were recently shown to be false. However the proof is non-constructive, and to date there are no known explicit examples of channels which violate either conjecture.

One of the main ingredients in the proof is the demonstration that there are channels  $\Phi$  which are sufficiently noisy, in the sense that they have large minimum output entropy  $S_{\min}(\Phi)$ . As noted before, this is equivalent to the assertion that there are  $s$ -dimensional subspaces of  $\mathbb{C}^n \otimes \mathbb{C}^d$  with large entanglement. The precise statement follows.

For fixed dimensions  $s, n, d$ , define the maximum entanglement of a  $s$ -dimensional subspace of  $\mathbb{C}^n \otimes \mathbb{C}^d$ :

$$E_{\max}(s, d, n) = \sup\{E(W) \mid W : \mathbb{C}^s \rightarrow \mathbb{C}^d \otimes \mathbb{C}^n\}$$

We will assume that  $d \leq n$  so that  $E_{\max}(s, d, n) \leq \log d$ .

## Theorem

Let  $d \geq 2$ , and  $0 < r_1 \leq r_2$ . For  $h$  sufficiently large, there is  $n_0 < \infty$  such that for  $n \geq n_0$ , and all  $s$  satisfying  $r_1 \leq s/n \leq r_2$ ,

$$E_{\max}(s, d, n) > \log d - h\left(\frac{s}{nd}\right) \quad (3)$$

As mentioned, the proof of this Theorem uses a randomizing argument. In fact the result will imply that the probability that a randomly selected subspace satisfies this bound will converge to 1 as  $n, s \rightarrow \infty$ . Before explaining the ideas of the proof, we show how it implies the existence of channels which violate additivity.

# Violations of additivity

Hastings (2008) considered the case  $s = n$ , so input and output spaces have the same dimension, and  $d$  is the dimension of the environment. He proved the existence of unital channels  $\Phi$  for which

$$S_{\min}(\Phi \otimes \bar{\Phi}) < S_{\min}(\Phi) + S_{\min}(\bar{\Phi})$$

Here  $\bar{\Phi}$  is the complex conjugate channel, defined by

$$\bar{\Phi}(\rho) = \overline{\Phi(\bar{\rho})}$$

The proof uses two main ideas: first, the channel  $\Phi \otimes \bar{\Phi}$  is shown to always have a relatively 'pure' output state and hence a relatively small  $S_{\min}$ .

This observation is due to A. Winter, and exploits a property of the maximally entangled state in  $\mathbb{C}^n \otimes \mathbb{C}^n$

$$|\psi_m\rangle = \frac{1}{\sqrt{n}} \sum_{i=1}^n |i\rangle \otimes |i\rangle$$

namely that for any matrix  $A$

$$(A \otimes I)|\psi_m\rangle = (I \otimes A^T)|\psi_m\rangle$$

Hastings' channel has the form

$$\Phi(\rho) = \sum_{i=1}^d q_i U_i \rho U_i^*, \quad \bar{\Phi}(\rho) = \sum_{i=1}^d q_i \bar{U}_i \rho \bar{U}_i^*$$

where  $\sum q_i = 1$  and where  $U_i$  are  $n \times n$  unitary matrices. Hence when  $\Phi \otimes \bar{\Phi}$  acts on the maximally entangled state  $|\psi_m\rangle\langle\psi_m|$ , the 'diagonal' terms leave the state invariant:

$$\begin{aligned} (U_i \otimes \bar{U}_i) |\psi_m\rangle\langle\psi_m| (U_i^* \otimes \bar{U}_i^*) &= (I \otimes \bar{U}_i U_i^T) |\psi_m\rangle\langle\psi_m| (I \otimes \bar{U}_i^* \bar{U}_i) \\ &= |\psi_m\rangle\langle\psi_m| \end{aligned}$$

This means that  $\Phi \otimes \bar{\Phi}(|\psi_m\rangle\langle\psi_m|)$  has a large component in the pure state  $|\psi_m\rangle\langle\psi_m|$ , and therefore has a small entropy.

In fact the argument is more general and implies a universal bound for the output entropy of any product channel  $\Phi \otimes \bar{\Phi}$ , namely (again assuming  $s = n$ )

$$S_{\min}(\Phi \otimes \bar{\Phi}) \leq 2 \log d - \frac{1}{d} (\log d - 1)$$

Now we use the subspace entanglement bound of Theorem 1, taking  $s = n$ . This guarantees the existence of channels  $\Phi$  for which

$$S_{\min}(\Phi) = S_{\min}(\bar{\Phi}) > \log d - \frac{h}{d}$$

for sufficiently large  $n$ . Thus

$$\begin{aligned} \Delta S(\Phi) &= 2S_{\min}(\Phi) - S_{\min}(\Phi \otimes \bar{\Phi}) \\ &\geq 2 \log d - \frac{2h}{d} - 2 \log d + \frac{1}{d} (\log d - 1) \\ &= \frac{1}{d} \left[ \log d - 2h - 1 \right] \end{aligned}$$

By taking  $d > \exp(2h + 1)$  we get  $\Delta S(\Phi) > 0$  and thus a channel which exhibits non-additivity of minimal output entropy.

# Minimal dimensions

Note that the size of the violation provided by the bound goes to zero as  $d \rightarrow \infty$ . It would be interesting to find the smallest dimensions which allow violations, as well as the size of the largest possible violation. Following this line of reasoning we define

$$d_{\min} = \inf \left\{ d : \exists n, \exists s, \exists \Phi \text{ s.t. } \Delta S(\Phi) > 0 \right\}$$

$$n_{\min} = \inf \left\{ n : \exists d, \exists s, \exists \Phi \text{ s.t. } \Delta S(\Phi) > 0 \right\}$$

$$\Delta S_{\max} = \sup_{s,n,d} \sup_{\Phi} \Delta S(\Phi)$$

The next result gives some bounds on these quantities.

## Proposition

$$d_{\min} < 3.9 \times 10^4 \sim 2^{15}$$

$$n_{\min} < 7.8 \times 10^{32} \sim 2^{110}$$

$$\Delta S_{\max} > 9.5 \times 10^{-6}$$

These bounds are surely not optimal, however they may indicate the delicacy of the non-additivity effect.

# Random subspaces

Theorem 1 is proved by estimating the probability that the bound holds for a randomly selected subspace, or equivalently a random embedding. Let  $W_0 : \mathbb{C}^s \rightarrow \mathbb{C}^d \otimes \mathbb{C}^n$  be a fixed embedding, then every other embedding can be written as  $W = U W_0$  for some unitary  $nd \times nd$  matrix. Thus we can obtain random embeddings (and hence also random channels  $\Phi, \Phi^C$ ) by using random unitary matrices  $U$  (with normalized Haar measure).

We write  $\mathcal{R}(s, n, d)$  for the set of all embeddings  $W : \mathbb{C}^s \rightarrow \mathbb{C}^d \otimes \mathbb{C}^n$ , with  $W^* W = I$ . This defines a probability measure  $\mathcal{P}_{s,n,d}$  on the set of embeddings  $\mathcal{R}(s, n, d)$ .

# Random states

Let  $|\psi\rangle$  be a random unit vector in  $\mathbb{C}^n \otimes \mathbb{C}^d$ . The reduced density matrix

$$\rho = \text{Tr}_{\mathbb{C}^{n \times n}} |\psi\rangle\langle\psi|$$

is a random state in  $\mathbb{C}^{d \times d}$ . The distribution of its eigenvalues  $\mu_{d,n}$  has a Wishart form: for any event  $A$

$$\mu_{d,n}(A) = Z(n, d)^{-1} \int_A \prod_{1 \leq i < j \leq d} (w_i - w_j)^2 \prod_{i=1}^d w_i^{n-d} \delta\left(\sum_{i=1}^d w_i - 1\right) [dw]$$

where  $Z(n, d)$  is a normalization factor.

Estimation of  $Z$  leads to the following bound.

### Lemma

*For all  $d$ , for  $n$  sufficiently large, and for any event  $A$ ,*

$$\mu_{d,n}(A) \leq \frac{1}{(d-1)!} \exp \left[ d^2 \log n + (n-d) d \log d \right. \\ \left. + (n-d) \sup_{w \in A} \sum_{i=1}^d \log w_i \right]$$

# The entanglement bounds

Recall that we want to show that there are embeddings  $W$  for which the entanglement is large, i.e.

$$E(W) > \log d - h\left(\frac{s}{nd}\right)$$

It will be convenient to use both the notation of channels and of embeddings from now on. Recall the relation:

$$W : \mathbb{C}^s \rightarrow \mathbb{C}^d \otimes \mathbb{C}^n, \quad W^* W = I$$

$$\Phi(\rho) = \text{Tr}_{\mathbb{C}^{d \times d}} W \rho W^* \in \mathbb{C}^{n \times n}$$

$$\Phi^C(\rho) = \text{Tr}_{\mathbb{C}^{n \times n}} W \rho W^* \in \mathbb{C}^{d \times d}$$

Hastings' proof relies on three important properties of random channels. These properties are 'generic' in the sense that they hold with high probability for a randomly selected channel. Thus we can find channels for which all three properties hold simultaneously and this does the job.

For convenience: if  $S(\rho) > \log d - h(s/nd)$  say that entropy of  $\rho$  is 'large', otherwise entropy of  $\rho$  is 'small'.

Our goal: prove existence of  $W$  such that the entropy of  $\Phi^C(|\phi\rangle\langle\phi|)$  is large for *all* input states  $|\phi\rangle$ .

**Property 1** For any fixed input state  $|\phi\rangle$  and a random channel  $\Phi$ , the output state  $\Phi^C(|\phi\rangle\langle\phi|)$  is very likely to be concentrated near the maximally mixed state:

$$\left\| \Phi^C(|\phi\rangle\langle\phi|) - \frac{1}{d}I \right\|_{\infty} \leq 2 \sqrt{\frac{\log n}{n}}$$

The entropy of such an output state is very close to the maximum  $\log d$ , and thus is large.

In fact, with high probability (approaching 1 as  $n \rightarrow \infty$ ) this bound will hold for at least one half of all input states (using the uniform measure on the input states).

The main idea is to consider pairs  $(W, |\phi\rangle)$  where  $W$  is a random embedding and  $|\phi\rangle$  is a random input state. The state  $W|\phi\rangle$  is a random bipartite state in  $\mathbb{C}^n \otimes \mathbb{C}^s$ , and thus the eigenvalues of  $\Phi^C(|\phi\rangle\langle\phi|)$  have the Wishart distribution above. This allows to prove that for any eigenvalue  $q_i$

$$P\left(|q_i - 1/d| > 2\sqrt{\frac{\log n}{n}}\right) \rightarrow 0$$

as  $n \rightarrow \infty$ . This estimate is used to show that there is a large set of channels for which at least one half of all input states will satisfy the bound. Thus there is a set of *typical* channels for which the bound holds, and as  $s, n \rightarrow \infty$  this set approaches measure .

So for a randomly selected typical channel  $\Phi$ , we just have to worry about the remaining half of input states which might get mapped outside this ball.

**Property 2** To handle the remaining half of the input states, Hastings made the following key observation: for a randomly selected embedding, the set of all output states

$$\{\Phi^C(|\phi\rangle\langle\phi|) : |\phi\rangle \in \mathbb{C}^S\}$$

is *sticky*. If an output state pulls away from the maximally mixed state  $\frac{1}{d}I$ , then it must drag along a tail of output states with it. This stickiness prevents the occurrence of isolated states far away from the maximally mixed state, which would have correspondingly small entropy and thus spoil the bound.

To be precise, with every output state  $\rho$  we associate a *tube* which points halfway toward the maximally mixed state: define the line segment

$$L(\rho) = \left\{ r\rho + (1-r)\frac{1}{d}I : 1/2 \leq r \leq 1 \right\}$$

then the tube is

$$\text{Tube}(\rho) = \left\{ \sigma \in \mathbb{C}^{d \times d} : \text{dist}(\sigma, L(\rho)) \leq 2\sqrt{\frac{\log n}{n}} + 13d\sqrt{\frac{\log d}{s}} \right\}$$

where

$$\text{dist}(\sigma, L(\rho)) = \inf_{\tau \in L(\rho)} \|\sigma - \tau\|_{\infty}$$

The next result shows that for any output state  $\rho$  in the image of  $\Phi^C$ , there is a uniform lower bound for the probability that a randomly chosen state belongs to the tube at  $\rho$ .

### Lemma

*Let  $\Phi$  be a typical channel, then for  $n$  sufficiently large, and for all  $\rho \in \text{Im}(\Phi^C)$ ,*

$$P\left(|\phi\rangle : \Phi^C(|\phi\rangle\langle\phi|) \in \text{Tube}(\rho)\right) \geq 2^{-s-1} \quad (5)$$

This explains the previous statement about stickiness: wherever the state  $\rho$  happens to be, there is a tail of input states connecting it to the maximally mixed state.

Main idea: by assumption there is an input state  $|\phi\rangle$  for which  $\Phi^C(|\phi\rangle\langle\phi|)$  has small entropy. Take another random input state  $|\theta\rangle$ . Then with high probability  $|\theta\rangle$  will be almost orthogonal to  $|\phi\rangle$ . If  $x = \langle\phi|\theta\rangle$  then

$$P(|x| > t) = (1 - t^2)^{s-1}$$

Thus if we write

$$|\theta\rangle = x|\phi\rangle + \sqrt{1 - |x|^2} |z\rangle$$

where  $|z\rangle$  is orthogonal to  $|\phi\rangle$ , then  $|\theta\rangle$  will be almost equal to  $|z\rangle$ . Precisely,

$$P(|\theta\rangle : \|\theta\rangle - |z\rangle\|_2 > t) \leq \left(1 - \frac{t^2}{2}\right)^{s-1}$$

It follows that

$$|\theta\rangle\langle\theta| = |x|^2 |\phi\rangle\langle\phi| + (1 - |x|^2) |z\rangle\langle z| + \sqrt{1 - |x|^2} (x |\phi\rangle\langle z| + \bar{x} |z\rangle\langle\phi|)$$

Write  $r = |x|^2$ , then

$$\begin{aligned} \Phi^C(|\theta\rangle\langle\theta|) &= \left( r\Phi^C(|\phi\rangle\langle\phi|) + (1 - r)\frac{1}{d}I \right) \\ &= (1 - r) \left( \Phi^C|z\rangle\langle z| - \frac{1}{d}I \right) \\ &\quad + \sqrt{r(1 - r)}\Phi^C \left( e^{i\xi} |\phi\rangle\langle z| + e^{-i\xi} |z\rangle\langle\phi| \right) \end{aligned}$$

**Property 3** The last property uses again the observation that if we randomly select an embedding  $W$  and randomly select an input state  $|\phi\rangle$ , then  $W|\phi\rangle = U W_0|\phi\rangle = |\psi\rangle$  is a random state in  $\mathbb{C}^d \otimes \mathbb{C}^n$ . So the eigenvalues of the reduced density matrix have the Wishart distribution. Using the explicit form of the distribution we can obtain very precise estimates for the probability that such a density matrix has a small entropy (take  $s = n$  for convenience):

$$P\left(S(\text{Tr}_2|\psi\rangle\langle\psi|) < \log d - h/d\right) \leq \alpha(d) \exp[d^2 \log n - n m_d(h)]$$

Here  $m_d$  is a large deviation rate function, and satisfies

$$m_d(h) \rightarrow \infty \quad \text{as} \quad h \rightarrow \infty$$

By combining these properties we can finish the proof. Take  $s = n$  for convenience. Consider the event consisting of those embeddings which do not satisfy the desired bound:

$$B = \left\{ W : E(W) \leq \log d - h/d \right\} = \left\{ \text{'bad' embeddings} \right\}$$

We want to show that  $P(B) < 1$  for  $h$  sufficiently large (in fact we show that  $P(B) \rightarrow 0$  as  $n \rightarrow \infty$ ). Given  $W \in B$  let  $\rho_W$  be an output state with low entropy (such a state must exist):

$$S(\rho_W) \leq \log d - h/d \Rightarrow \textit{small}$$

Consider the joint event

$$\text{Prob}\left((W, |\phi\rangle) : W \in B, \Phi^C(|\phi\rangle\langle\phi|) \in \text{Tube}(\rho_W)\right)$$

Then from Property 2 we get

$$\begin{aligned} \text{Prob}\left((W, |\phi\rangle) : W \in B, \Phi^C(|\phi\rangle\langle\phi|) \in \text{Tube}(\rho_W)\right) & \\ &= \text{Prob}_{W \in B}\left(\text{Prob}_{|\phi\rangle}(\Phi^C(|\phi\rangle\langle\phi|) \in \text{Tube}(\rho_W))\right) \\ &\geq \text{Prob}_{W \in B}\left(2^{-n-1}\right) \\ &= 2^{-n-1} \text{Prob}(B) \end{aligned}$$

Furthermore we get an upper bound for this event using Property 3:

$$\begin{aligned}
 \text{Prob}\left((W, |\phi\rangle) : W \in B, \Phi^C(|\phi\rangle\langle\phi|) \in \text{Tube}(\rho_W)\right) \\
 \leq \text{Prob}\left((W, |\phi\rangle) : S(\Phi^C(|\phi\rangle\langle\phi|)) \text{ is small}\right) \\
 = \text{Prob}_{|\psi\rangle}\left(S(\text{Tr}_{\mathbb{C}^{n \times n}}|\psi\rangle\langle\psi|) \text{ is small}\right) \\
 \leq \alpha(d) \exp[d^2 \log n - n m_d(h)]
 \end{aligned}$$

Combining these gives

$$\text{Prob}(B) \leq \alpha(d) \exp[d^2 \log n - n m_d(h) + n \log 2]$$

Thus for  $h$  sufficiently large we have  $P(B) \rightarrow 0$  as  $n \rightarrow \infty$ .

This is the proof of violation of additivity of minimal output entropy. As noted this implies violation of additivity of the Holevo quantity  $\chi(\Phi)$ . This latter result is important because of its relation to the channel capacity  $C(\Phi)$ . The capacity is defined operationally as the maximum rate at which classical information can be sent through the channel  $\Phi$ , using input states to encode the information and measurements to decode at the output. The capacity is known to equal

$$C(\Phi) = \lim_{n \rightarrow \infty} \frac{1}{n} \chi(\Phi^{\otimes n})$$

Non-additivity of  $\chi$  means that  $C(\Phi) > \chi(\Phi)$ , and thus there is no 'single-letter' formula for the capacity.

Open questions:

- (1) find useful bounds for the channel capacity  $C(\Phi)$
- (2) find a new compact 'single-letter' formula for  $C(\Phi)$
- (3) settle the additivity question for  $C(\Phi)$