The global and local additivity problems in quantum information theory

Shmuel Friedland
Univ. Illinois at Chicago

IQIS Seminar, University of Calgary
9 November, 2011

Overview

- Capacity of a classical noisy channel
- Capacity of a quantum channel
- Three additivity conjectures
- Hastings's counterexample to additivity conjectures
- Local additivity of the minimum entropy output

Capacity of a classical noisy channel

$\mathcal{X} := \{x_1, \ldots, x_m\}, \mathcal{Y} = \{y_1, \ldots, y_n\}$ two finite alphabets.

$p(x, y)$-probability sending $x \in \mathcal{X}$ and receiving $y \in \mathcal{Y}$.

$P(\mathcal{X}, \mathcal{Y}) := [p(x, y)]_{x\in\mathcal{X}, y\in\mathcal{Y}} \in \mathbb{R}_+^{m\times n}$ stochastic matrix

Noiseless channel: $\mathcal{X} = \mathcal{Y}, P(\mathcal{X}, \mathcal{Y}) = I_m$

Completely noisy channel: $p(x, y) = \frac{1}{n}$ for all $x \in \mathcal{X}, y \in \mathcal{Y}$

Capacity of a channel: $\mathrm{Cap}(P(\mathcal{X}, \mathcal{Y})) :=$

$$\frac{\text{the length of the message to transmit}}{\text{the length of the actual message transmitted to overcome the noise}}$$

Capacity of noiseless channel is 1,
Capacity of completely noisy channel 0.

Classical entropies

X, Y two random variables $X \in \mathcal{X}, Y \in \mathcal{Y}$

$\pi(x, y) := \mathrm{Prob}(X = x, Y = y)$

$\pi_X(x) := \mathrm{Prob}(X = x) = \sum_{y \in \mathcal{Y}} \pi(x, y), \pi_Y(y) := \sum_{x \in \mathcal{X}} \pi(x, y)$

Entropy: $H(X) := -\sum_{x \in \mathcal{X}} \pi_X(x) \log \pi_X(x)$

$H(X, Y) := -\sum_{x \in \mathcal{X}, y \in \mathcal{Y}} \pi(x, y) \log \pi(x, y)$

Conditional entropy: $0 \leq H(X|Y) := H(X, Y) - H(Y)$ $(\leq H(X, Y))$

Mutual information:
$H(X : Y) = H(Y : X) = H(X) + H(Y) - H(X, Y) = H(X) - H(X|Y) \geq 0$

Shannon's formula for capacity of noisy channel

$\kappa(\pi_X) = \kappa(X) := H(\pi_X P(\mathcal{X},\mathcal{Y})) - \sum_{x\in\mathcal{X}} \pi(x) H(\delta_x \mathcal{P}(\mathcal{X},\mathcal{Y}))$
$\mathrm{Cap}(P(\mathcal{X},\mathcal{Y})) = \max_{\pi_X\in\Pi(\mathcal{X})} \kappa(X)$

$\Pi(\mathcal{X})$ the simplex of probability vectors $\pi = \pi_X$, $H(\pi) := H(X)$

$\delta_x \in \Pi(\mathcal{X})$ probability vector concentrated on $x \in \mathcal{X}$

$H(\pi)$ strictly concave on $\Pi(\mathcal{X})$, $\kappa(\pi)$ is concave.

Critical points of $\kappa(\pi)$ are the intersection of a hyperplane and $\Pi(\mathcal{X})$

induced by the kernel of $P(\mathcal{X},\mathcal{Y})^\top$

the critical value of $\kappa(\pi)$ is unique, and is the maximum

the crit. pt. unique when $\ker P((\mathcal{X},\mathcal{Y}))^\top = \{\mathbf{0}\}$ - Generic if $m \le n$

Shannon's capacity is additive under the tensor product:
$\mathrm{Cap}(P(\mathcal{X},\mathcal{Y}) \otimes P(\mathcal{X}',\mathcal{Y}')) = \mathrm{Cap}(P(\mathcal{X},\mathcal{Y})) + \mathrm{Cap}(P(\mathcal{X},'\mathcal{Y}'))$

As the tensor product of critical points is critical

Von Neumann and related quantum entropies

$\mathrm{H}_n \supset \mathrm{H}_{n,+,1}$ the space of hermitian matrices and density matrices

acting on $\mathbb{C}^n$ with the inner product $\langle \mathbf{x}, \mathbf{y} \rangle := \mathbf{y}^*\mathbf{x}$

$\mathrm{S}^{2n-1} := \{\mathbf{u} \in \mathbb{C}^n,\ \mathbf{u}^*\mathbf{u} = 1\}$

$\mathbf{u}\mathbf{u}^*, \mathbf{u} \in \mathrm{S}^{2n-1}$ (pure states) is the set of extreme points of $\mathrm{H}_{n,+,1}$

$\boldsymbol{\lambda}(\rho) := (\lambda_1(\rho) \geq \ldots \geq \lambda_n(\rho))$ eigenvalues of $\rho \in \mathrm{H}_{n,+,1}$

Von Neumann Entropy: $S(\rho) := H(\boldsymbol{\lambda}(\rho)) = -\operatorname{tr}(\rho \log \rho)$

strictly concave on $\mathrm{H}_{n,+,1}$

$S(\rho \otimes \sigma) = S(\rho) + S(\sigma)$ analog of independent random variable

$S(A, B) = -\operatorname{tr}(\rho^{AB} \log \rho^{AB}),\ \rho^{AB} \in \mathrm{H}_{mn,+,1}$

$S(A|B) = S(A, B) - S(B)$ -conditional entropy, but may be negative!

$S(A : B) := S(A) + S(B) - S(A, B)\ (\geq 0)$

Holevo's bound and channel capacity of QC

Alice prepares a state $\rho_X \in \mathrm{H}_{mn,+,1}$

$X = 1, \ldots, k, \mathrm{Prob}(X = i) = \pi_i, \rho_X = \sum_{i=1}^k \pi_i \rho_i$

Bob performs POVM measurements $E_1, \ldots, E_l$ with outcome Y.

$S(X : Y) \le \chi(\sum_{i=1}^k \pi_i \rho_i) := S(\sum_{i=1}^k \pi_i \rho_i) - \sum_{i=1}^k \mathbf{p}_i S(\rho_i)$

Quantum channel $\Phi : \mathrm{H}_{m,+,1} \to \mathrm{H}_{n,+,1}$

$\Phi(\rho) = \sum_{j=1}^k A_j^* \rho A_j,\ A_j \in \mathbb{C}^{m \times n},\ \sum_{j=1}^k A_j A_j^* = I_m$

Stinespring dilatation theorem:

$\Phi(\rho) = \mathrm{tr}_B\, U(\rho) = \mathrm{tr}_B\, U \rho U^*, U^* = [A_1\ A_2 \ldots A_k] \in \mathbb{C}^{m \times kn}$

$\mathrm{Cap}_\chi(\Phi) := \max_{\pi \in \Pi_{m^2}, \mathbf{u}_i \in S^{m-1}} \chi(\sum_{i=1}^{m^2} \pi_i \Phi(\mathbf{u}_i \mathbf{u}_i^*))$ [2]

Entanglement of formation and channel capacity

Entanglement of formation

$E_F(\rho) := \inf_{M\in\mathbb{N}} \min\{\sum_{i=1}^{M} \pi_i S(\Phi(\mathbf{u}_i\mathbf{u}_i^*)),\ \pi \in \Pi_M\}$,
$\sum_{i=1}^{M} \pi_i \mathbf{u}_i\mathbf{u}_i^* = \rho \in \mathrm{H}_{m,+,1}$ Equivalent to

$E_F(\rho) := \min\{\sum_{i=1}^{m^2} \pi_i S(\Phi(\mathbf{u}_i\mathbf{u}_i^*)),\ \pi \in \Pi_{m^2}, \sum_{i=1}^{m^2} \pi_i \mathbf{u}_i\mathbf{u}_i^* = \rho \in \mathrm{H}_{m,+,1}\}$

$E_F(\rho)$ convex (in view of the minimum characterization)

$\chi_N(\rho) := S(\Phi(\rho)) - E_F(\rho)$ - concave, but not smooth!

$\mathrm{Cap}_\chi(\Phi) := \max_{\rho\in\mathrm{H}_{m,+,1}} \chi_N(\rho)$

can be found in polynomial time?

Additivity conjectures for channel capacity and entanglement of formation

$\mathrm{Cap}_\chi(\Phi_1 \otimes \Phi_2) \geq Cap_\chi(\Phi_1) + Cap_\chi(\Phi_1)$

As $S(\sigma_1 \otimes \sigma_2) = S(\sigma_1) + S(\sigma_2)$ and

$(\sum_i \pi_{i,1}\mathbf{u}_i\mathbf{u}_i^*) \otimes (\sum_j \pi_{j,2}\mathbf{v}_j\mathbf{v}_j^*) = \sum_{i,j} \pi_{i,1}\pi_{j,2}(\mathbf{u}_i \otimes \mathbf{v}_j)(\mathbf{u}_i \otimes \mathbf{v}_j)^*$

$E_F(\sigma_1 \otimes \sigma_2) \geq E_F(\sigma_1) + E_F(\sigma_2)$

Additivity conjectures

$\mathrm{Cap}_\chi(\Phi_1 \otimes \Phi_2) = Cap_\chi(\Phi_1) + Cap_\chi(\Phi_1)$

$E_F(\sigma_1 \otimes \sigma_2) = E_F(\sigma_1) + E_F(\sigma_2)$

The minimum entropy output of a quantum channel

$S_{\min}(\Phi) := \min_{\rho\in \mathrm{H}_{m,+,1}} S(\Phi(\rho))$

Minimum of concave function is achieved on extreme points

$S_{\min}(\Phi) = \min_{\mathbf{u}\in \mathrm{S}^{2m-1}} S(\Phi(\mathbf{u}\mathbf{u}^*))$

Computation of $S_{\min}(\Phi)$ seems to be NP hard theoretically?

$S_{\min}(\Phi_1 \otimes \Phi_2) \le S_{\min}(\Phi_1) + S_{\min}(\Phi_2)$

Additivity conjecture: $S_{\min}(\Phi_1 \otimes \Phi_2) = S_{\min}(\Phi_1) + S_{\min}(\Phi_2)$

All three additivity conjectures are equivalent

Main steps in establishing the equivalence of the three additivity conjectures

1. Stinespring dilatation theorem relates a constrained version of Holevo capacity formula to the entanglement of formation MSW [1]

2. Using the duality of a corresponding linear programming problem AD [1]

3. P. Shor put all the above ingredients together [3]

Hastings' counterexample to additivity conjectures 09

Φ random unitary channel

$\Phi_1(\rho) = \Phi(\rho) := \sum_{i=1}^d w_i U_i \rho U_i^*, \ \Phi_2(\rho) = \bar{\Phi}(\rho) = \sum_{i=1}^d w_i \bar{U}_i \rho U_i^\top$

$U_1, \ldots, U_d$ randomly chosen unitary matrices in $\mathbf{U}(n) \subset \mathbb{C}^{n\times n}$

$(w_1, \ldots, w_k) \in \Pi_k$ randomly chosen probability vector–uniformly

For $n \gg d \gg 1$ there exists unitary random channel with

$\Delta S(\Phi) := S_{\min}(\Phi) + S_{\min}(\bar{\Phi}) - S_{\min}(\Phi \otimes \bar{\Phi}) > 0$ Hastings [5]

Estimates for the counterexample Fukuda-King-Moser [4]

$d_{\min} < 3.910^4, \quad n_{\min} < 7.810^{32}, \quad \Delta S_{\max} > 9.510^{-6}$

Outline of the counterexample steps

$\mathcal{R}_d(n)$ The set of unitary random channels with parameters n, d

Lemma 1: $S_{min}(\Phi \otimes \bar{\Phi}) \le 2\log d - \frac{\log d}{d}$ for all $\Phi \in \mathcal{R}_d(n)$

Theorem 2: There is $h_{\min} < \infty$, such that for all $h > h_{\min}$, all d satisfying $d \log d \ge h$ and all n sufficiently large, there is $\Phi \in \mathcal{R}_d(n)$ satisfying

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

Choosing d large so that $2h_{\min} < \log d$ we deduce the existence of $\Phi \in \mathcal{R}_d(n)$

$S_{min}(\Phi) > d - \frac{\log d}{2d}$

Since $S_{\min}(\bar{\Phi}) = S_{\min}(\Phi)$ we deduce from Lemma 1 the inequality

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

Minimum entropy output of a matrix subspace

$\mathrm{S}(p,n) := \{A \in \mathbb{C}^{p\times n}, \|A\|_F^2 = \mathrm{tr}(A^*A) = 1\}$

$\phi : \mathrm{S}(p,n) \to \mathrm{H}_{n,+,1}, \phi(A) = A^*A$

$\mathbf{V}$ m-dimensional subspace of $\mathbb{C}^{p\times n}$, denote $\mathrm{S}(\mathbf{V}) := \mathbf{V} \cap \mathrm{S}(p,n)$

$S_{\min}(\mathbf{V}) := \min_{A\in \mathrm{S}(\mathbf{V})\cap\mathbf{V}} S(A^*A)$

Claim $S_{\min}(\mathbf{V}) = S_{\min}(\Phi)$ for corresponding QC $\Phi : \mathrm{H}_{m,+,1} \to \mathrm{H}_{n,+,1}$

Choose an orthonormal base $C_1, \ldots, C_m \in \mathbf{V}$:
$\mathrm{tr}(C_i^* C_j) = \delta_{ij}, i, j = 1, \ldots, m$

$\Phi(B) = \sum_{i,j=1}^m b_{ij} C_i^* C_j, \quad B = [b_{ij}] \in \mathrm{H}_m$

$\Phi : \mathrm{H}_m \to \mathrm{H}_n$

$\Phi(\mathbf{u}\mathbf{u}^*) = C^*C \succeq 0, \quad C = \sum_{j=1}^m \bar{\mathbf{u}}_j C_j, \quad \mathrm{tr}(\Phi(\mathbf{u}\mathbf{u}^*)) = \sum_{j=1}^m |u_j|^2$

So Φ is a QC and $S_{\min}(\mathbf{V}) = S_{\min}(\Phi)$

All QC can be obtained this way

Critical points of the subspace entropy function

$f : \mathbf{V} \to \mathbb{R}, f(A) = S(A^*A) = -\operatorname{tr}(A^*A \log(A^*A))$.

B is a critical point of $f|_{\mathrm{S}(\mathbf{V})}$ if $Df|_{\mathrm{S}(\mathbf{V})}(B) = 0$.

One has to be careful with derivative of $S(A^*A)$

as $(-x \log x)' = -1 - \log x$ not defined at $x = 0$

$$T(A) := \begin{bmatrix} 0_{p\times n} & A \\ A^* & 0_{n\times p} \end{bmatrix} \in \mathrm{H}_{p+n}$$

$\sigma_1(A) = \lambda_1(T(A)) \geq \ldots \geq \sigma_r(A) = \lambda_r(T(A)) >$

$0 = \lambda_{r+1}(T(A)) = \ldots = \lambda_{p+n-r}(T(A)) >$

$\lambda_{p+n-r+1} = -\sigma_1(A) \geq \ldots \geq \lambda_{p+n}(A) = -\sigma_1(A)$

$S(A^*A) = -\frac{1}{2}\sum_{i=1}^{p+n} \lambda_i(T(A))^2 \log \lambda_i(T(A))^2$

As $-x^2 \log x^2$ is continuously differentiable $f|_{\mathbf{V}\cap\mathrm{S}(p,n)}$

is continuously differentiable critical point is well defined

Tensor product of critical points is critical

Lemma AIM workshop July 21-25, 2008

If B_i is a critical point of $f|_{\mathrm{S}(\mathbf{V}_i)}$ for $i = 1, 2$ then

$B_1 \otimes B_2$ is a critical point of $f|_{\mathrm{S}(\mathbf{V}_1 \otimes V_2)}$

First and second variations

$B, C \in \mathrm{S}(\mathbf{V}), \mathrm{tr}(BC^*) = 0,\ \mathrm{S}(\mathbf{V}) \ni B(t) = \frac{1}{\sqrt{1+t^2}}(B + tC) =$

$B + tC - \frac{1}{2}t^2 B + O(t^3)$

$X, Y \in \mathrm{H}_N, \quad X = T(B), Y = T(C)$

$F : \mathrm{H}_N \to \mathrm{H}_N$, special case $F_0(X) = -X^2 \log X^2, f(B) = \frac{1}{2} \mathrm{tr}\, F(T(B))$.

$F(X + tY) = F(P) + tL_X(Q) + t^2 Q_X(Y) + O(t^3)$ - Caution for F_0

$L_X : \mathrm{H}_N \to \mathrm{H}_N$ a linear map

$Q_X : \mathrm{H}_N \to \mathrm{H}_N$ is a quadratic map in the entries of Y

Criticality of F at X: $\mathrm{tr}(L_X(Y)) = 0$ if $\mathrm{tr}(XY) = 0$

Local strict minimum at critical X if the second derivative of

$\mathrm{tr}(F_0(\frac{1}{\sqrt{1+t^2}}(X + tY)))$ at $t = 0$ positive or $+\infty$ for all above Y

Simple formulas for first and second variation

Assume that $X = \mathrm{diag}(\xi_1, \ldots, \xi_N)$, $Y = [\eta_{ij}] \in \mathrm{H}_N$ then

$L_X(Y) = [\frac{F(\xi_i)-F(\xi_j)}{\xi_i-\xi_j} y_{ij}],\ \mathrm{tr}(L_X(Y)) = \sum_{i=1}^N F'(x_i)y_{ii}$. (*)

$[Q_X(Y)]_{ij} = \sum_{k=1}^n \triangle^2 F(\xi_i, \xi_k, \xi_j) y_{ik} y_{kj}$

$\mathrm{tr}(Q_X(Y)) = \sum_{i,j=1}^n \frac{F'(\xi_i)-F'(\xi_j)}{2(\xi_i-\xi_j)} y_{ij} y_{ji}$

Linearization:

$\mathrm{tr}(F_0(\frac{1}{\sqrt{1+t^2}}(X + tY))) = \mathrm{tr}(F_0(X + tY)) + t^2(1 - \mathrm{tr}(F_0(X))) + O(t^4)$

Why critical points of f are stable under the tensor product?

Use (*), $\log(ab) = \log a + \log b$ and the structure of

the tangent hyperplane of $\mathrm{S}(\mathbf{V}_1 \otimes \mathbf{V}_2)$ at $B_1 \otimes B_2$

The minimum entropy output of a quantum channel is locally additive

Theorem Gour-Friedland [3]

Assume that $B_i \in \mathbf{V}_i \subset \mathbb{C}^{p_i \times n_i}$ is a local strict minimum of $S|\mathrm{S}(\mathbf{V}_i)$ for $i = 1, 2$.

Then $B_1 \otimes B_2$ is a local strict minimum of $S|\mathrm{S}(\mathbf{V}_1 \otimes \mathbf{V}_2)$

The proof is tour de force

The assumption that each subspace $\mathbf{V}_i$ is complex valued is crucial

Open problems

1. When one can expect generically that all local minimal points of $S|\mathbb{S}(\mathbf{V})$ are strict local minima?

Necessary condition $\dim \mathbf{V} \le (p-1)(n-1)+1$

If $\dim V > (p-1)(n-1)+1$ then $\mathbf{V} \subset \mathbb{C}^{p\times n}$ contains a variety of matrices of rank one of dimension $\dim \mathbf{V} - (p-1)(n-1) - 1$, hence $S_{\min}(\mathbf{V}) = 0$ is attained at a connected set of rank one matrices in $\mathbf{V}$.

2. Suppose that $S_{\min}(\mathbf{V}_i) = S(B_i)$ and B_i is unique and strict minimum for $i = 1, 2$

Is $S_{\min}(\mathbf{V}_1 \otimes \mathbf{V}_2) = S(B_1 \otimes B_2)$?

References 1

K.M.R. Audenaert and S.L. Braunstein, On strong superadditivity of the entanglement of formation, *Commun. Math. Phys.* 246 (2004) , 443–452.

S. Friedland, G. Gour and A. Roy, Local extrema of entropy functions under tensor products, arXiv:1105.5380, to appear in *Quantum Info. Comp.*

G. Gour and S. Friedland, The minimum entropy output of a quantum channel is locally additive, arXiv:1105.6122.

M. Fukuda, C. King and D.K. Moser, Comments on Hastings' additivity counterexamples, *Commun. Math. Phys.* 296 (2010), 111–143.

M.B. Hastings, A counterexample to additivity of minimum output entropy, *Nature Physics* 5 (2009),255–257.

References 2

- K. Matsumoto, T. Shimono and A. Winter, Remarks on additivity of the Holevo capacity and of entanglement of formation, *Commun. Math. Phys.* 246 (2004), 427–442.
- B. Schumacher and M.D. Westmoreland, Optimal signal ensembles, *Phys. Rev. A* 63 (2001), 022308.
- P. Shor, Equivalence of additivity questions in quantum information theory, *Commun. Math. Phys.* 246 (2004), 453–472.