Properties of Quantum Entropy and Related Convex Trace Functions

Mary Beth Ruskai
marybeth.ruskai@tufts.edu

Tufts University

June 2009: Short Course at ICTP, Trieste, Italy

Lecture Plan

1. Introduction and Background on Quantum Information
2. Properties of Entropy and Relative Entropy
3. Simple Proof of Joint Convexity of Relative Entropy
4. a) Noise in Quantum Information

   b) More Inequalities and conjectures

I. Introduction and Background on Quantum Information

1. Preliminaries
2. Dirac bra and ket notation
3. Overview of quantum information
4. Some quantum mechanics basics
5. Quantum entropy
6. Tensor products and entanglement
7. Aside on SVD and “Schmidt” decompostion
8. Quantum relative entropy

Hilbert space set-up

Full quantum theory associates a particle, e.g., electron with

$$\mathcal{H} = L_2(\mathbf{R}_3) \otimes \mathbf{C}_d \qquad d = 2s+1, \; s \text{ denotes "spin"}$$

Quantum info suppress "spatial" part in $L_2(\mathbf{R}_3)$ and focus on spin,

typically spin $\frac{1}{2}$ or $\mathbf{C}_2$. Many particles or qubits use $\mathbf{C}_2^{\otimes n}$

can also consider $d > 2$ and $\mathbf{C}_d \otimes \mathbf{C}_{d'} \otimes \ldots$ etc.

Will work with $\mathcal{H} = \mathbf{C}_d$ or tensor products of these

and $\mathcal{B}(\mathcal{H}) = M_d =$ space of $d \times d$ matrices

$M_d =$ also a Hilbert space with inner product $\langle A, B \rangle = \operatorname{Tr} A^* B$

will also consider linear maps $\Phi : M_d \mapsto M_{d'}$

Most results extend to ∞ dim. $\mathcal{H}$ in suitable way

and can even replace M_d by a von Neumann algebra

Basic notation

A, B, C self-adjoint matrices

$A > B$ means $A - B$ positive definite $\langle v, (A - B)v \rangle > 0 \quad \forall\, v \neq 0$

$A \geq B$ or $A \gneq B$ means $A - B$ positive semi-definite, but $A \neq B$.

$A \geqq B$ means $A - B$ positive semi-definite (with $A = B$ allowed)

difference rarely significant $\begin{cases} f \quad \text{obviously diverges when } B = 0 \\ \lim f(B + \epsilon I) \text{ well defined as } \epsilon \to 0 \end{cases}$

operator (or matrix) inequality has one of above forms

trace inequality has form $\operatorname{Tr} AC > \operatorname{Tr} BC$ or $\operatorname{Tr} AC \geq \operatorname{Tr} BC$

with usual meanings for $>, \geq$ when $\operatorname{Tr} X \in \mathbf{R}$ (usually $[0, \infty)$

study trace inequalities related to maps $M_d \mapsto R$ that are convex

Functions of operators

For $A = UDU^*$, define $f(A) = U\, f(D)\, U^*$

$$A = U \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & 0 & \lambda_m \end{pmatrix} U^* \qquad f(A) = U \begin{pmatrix} f(\lambda_1) & 0 & \ldots & 0 \\ 0 & f(\lambda_2) & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \ldots & 0 & f(\lambda_m) \end{pmatrix} U^*$$

equiv. to any reasonable def using power series, integral rep., etc.

also applies to operators, e.g., L_Q acting on M_d space of matrices

For example, for $L_A(X) = AX$ find $L_{\log A}(X) = \log L_A(X)$

Examples of trace functions

$R, Q > 0$ but often extend to $R, Q \geq 0$

applications often use $\text{Tr}\, R = \text{Tr}\, Q = 1$ but not essential

$S(R) = -\text{Tr}\, R \log R$ quantum entropy (concave)

$H(R, Q) = \text{Tr}\, R\big(\log R - \log Q\big)$ relative entropy (jointly convex)

$$K \text{ fixed } \text{Tr}\, K^* R^p K Q^{1-p} \text{ jointly } \begin{cases} \text{concave} & 0 < p < 1 \quad \text{(Lieb)} \\ \text{linear} & p = 0, 1 \\ \text{convex} & 1 < p < 2 \quad \text{(Ando)} \end{cases}$$

(i) math – lead to many interesting matrix inequalities

(ii) original interest in quantum statistical mechanics

now important in quantum information theory

Dirac notation: $\mathbf{u}, \mathbf{v}$ in finite-dim vector space $\mathbf{C}_d$

Inner product $\langle v, u \rangle = \mathbf{v}^* \mathbf{u} = \begin{pmatrix} & \overline{v} & \end{pmatrix} \begin{pmatrix} \\ u \\ \\ \end{pmatrix}$

Reverse order $|u\rangle\langle v| = \mathbf{u}\ \mathbf{v}^* = \begin{pmatrix} \\ u \\ \\ \end{pmatrix} \begin{pmatrix} & \overline{v} & \end{pmatrix}$

get $n \times n$ matrix of map $w \mapsto \langle v, w \rangle u$

$P_u = \frac{1}{\|u\|^2} uu^* = \frac{|u\rangle\langle u|}{\|u\|^2}$ projection onto 1-dim subspace span$\{u\}$.

"ket" $|u\rangle \leftrightarrow \mathbf{u} \leftrightarrow$ col vec   dual "bra" $\langle u| \leftrightarrow \mathbf{u}^* \leftrightarrow$ row vec

- Can be justified by Riesz rep. theorem.
- Put complex conj. on "left" i.e., $\langle v, u \rangle$ linear in u; anti-lin in v.
- Use any convenient label, e.g. $|\lambda_k\rangle$ or $|k\rangle$ for eigenfctn v_k of λ_k

conventions

in view of above $\langle u, v \rangle = \mathbf{u}^* \mathbf{v}$ anti-linear in u and linear in v

$$a\langle u, v \rangle = \langle u, av \rangle = \langle \overline{a}u, v \rangle$$

math	physics	MBR	this lecture	
A^*	$A^\dagger$	$A^\dagger$	A^*	adjoint
$\overline{a}$	a^*	$\overline{a}$	$\overline{a}$	complex conjugate
		a^*		never – too confusing
		$\widehat{\Phi}$	$\widehat{\Phi}$	adjoint wrt H-S inner prod

$\Phi : M_d \mapsto M_{d'}$ linear map

$$\mathrm{Tr}\, A^* \Phi(B) = \langle A, \Phi(B) \rangle = \langle \widehat{\Phi}(A), B \rangle = \mathrm{Tr}\, [\widehat{\Phi}(A)]^* B$$

Quantum vs classical information

Classical – "bit" takes values in $\mathbf{Z}_2 = \{0, 1\}$, e..g, "on" or "off"

encode info in strings of 0 & 1, elements of $\mathbf{Z}_2^{\otimes n} = \mathbf{Z}_2 \otimes \mathbf{Z}_2 \ldots \otimes \mathbf{Z}_2$

Quantum – "qubit" takes values in $\mathbf{C}_2$ (up to norm. and phase)

$$0 \sim \begin{pmatrix} 1 \\ 0 \end{pmatrix} \equiv |0\rangle_z \quad \uparrow \quad \text{spin "up"} \qquad \text{or} \quad \uparrow \quad \text{vertical polar}$$

$$1 \sim \begin{pmatrix} 0 \\ 1 \end{pmatrix} \equiv |1\rangle_z \quad \downarrow \quad \text{spin "down"} \quad \text{or} \quad \rightarrow \quad \text{horiz polar}$$

Isomorphism between $\mathbf{Z}_2^{\otimes n}$ and O.N. prod basis for $\mathbf{C}_2^{\otimes n}$

"Computational basis", e.g., $|0110\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

Embed class in quant — but can do much more

$$0 \sim \begin{pmatrix} 1 \\ 0 \end{pmatrix} \equiv |0\rangle_z \quad \uparrow \quad \text{spin ``up''} \qquad \text{or} \quad \uparrow \quad \text{vertical polar}$$

$$1 \sim \begin{pmatrix} 0 \\ 1 \end{pmatrix} \equiv |1\rangle_z \quad \downarrow \quad \text{spin ``down''} \quad \text{or} \quad \rightarrow \quad \text{horiz polar}$$

Now consider $\begin{pmatrix} 1 \\ \pm 1 \end{pmatrix}$ "spin" right $\rightarrow$ or left $\leftarrow$ (in x-direction)

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \pm 1 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}} (|0\rangle_z + |1\rangle_z)$$

Then measure spin in z-direction. Get either $\uparrow$ or $\downarrow$ (i.e., 0 or 1)

each with probability $\frac{1}{2}$. But not classical prob, i.e.,

Not a classical mixture but a superposition of vectors.

In some sense $2^{-n/2} \begin{pmatrix} 1 \\ 1 \end{pmatrix}^{\otimes n} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \otimes \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \otimes \ldots \otimes \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

contains all 2^n strings of 00110101...., each with prob 2^{-n}

Can one encode more than $\{0,1\}$ in qubit ??

4 states $\begin{pmatrix}1\\0\end{pmatrix} \begin{matrix}\uparrow\\ \uparrow\end{matrix} \ \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix} \begin{matrix}\rightarrow\\ \nearrow\end{matrix} \ \begin{pmatrix}0\\1\end{pmatrix} \begin{matrix}\downarrow\\ \rightarrow\end{matrix} \ \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix} \begin{matrix}\leftarrow\\ \searrow\end{matrix}$??

Most general qubit $|v\rangle = a|0\rangle + b|1\rangle = \begin{pmatrix}a\\b\end{pmatrix} = \begin{pmatrix}\sin\theta\\ e^{i\varphi}\cos\theta\end{pmatrix}$

use to encode $[0,\pi]$ or $[-1,1]$ or more ??

NO — will see can only reliably distinguish orthog states

Thm: (Holevo bound) Accessible Information or max info/qubit one can extract ≤ 1. Will give formal math thm and proof

non-orthog encoding can't reliably distinguish BUT advantages!

- Noisy communication (use quant particles to sed class info)

  non-orthog input may yield more distinguishable outputs

- Quant cryptography — sacrifice info to detect eavesdropper

Quantum basics and von Neumann measurement

Fund Postulate of Q.M.: Observable represented by self-adj op A

spectral decomp $A = \sum_k a_k E_k = \sum_k a_k |\alpha_k\rangle\langle\alpha_k|$

Measurement of A with system in some state ψ.

(i) get some e-value (only possibility)

(ii) leave system in e-state α_k

(iii) probability is $|\langle\alpha_k, \psi\rangle|^2 = \text{Tr}\, E_k |\psi\rangle\langle\psi|$

Write $|\psi\rangle = \sum_k c_k |\alpha_k\rangle$ as a superposition of e-states, $c_k = \langle\alpha_k, \psi\rangle$

Coefficients c_k in superpos. give probs $|c_k|^2$ not classical

Average result of meas in state $|\psi\rangle$ is $\langle\psi, A\psi\rangle = \text{Tr}\, A|\psi\rangle\langle\psi|$

set $\{E_k\}$ orthog projections $E_j E_k = E_k \delta_{jk}$ with $\sum_k E_k = I$ called

von Neumann measurement or projection valued measure (PVM)

corresponds to "yes-no" experiment (e.g., polarization filter)

Density matrices: mixed vs. pure states

pure state often rep by vector $|\psi\rangle \in \mathcal{H}$ up to phase, with $\|\psi\| = 1$.

better to use rank-1 projection $|\psi\rangle\langle\psi|$ (with $\|\psi\| = 1$).

mixed state $\rho = \sum_k p_k |\phi_k\rangle\langle\phi_k|$ is convex comb of pure states

$p_k > 0, \sum_k p_k = 1$ and $\|\phi_k\| = 1$ but ϕ_k not nec orthog

ρ called density matrix (D.M.) or density operator

$\rho \in M_d$ is D.M. if and only if $\rho \geq 0$ and $\operatorname{Tr} \rho = 1$.

Interp: a) ensemble in quantum statistical mechanics

b) know only part (subsystem) $\rho = \rho_A = \operatorname{Tr}_B \rho_{AB}$

c) $A \mapsto \operatorname{Tr} A \rho$ positive linear functional on M_d.

two kinds of probability in Q.M. (i) p_k traditonal prob interp,

but (ii) $|\phi_k\rangle$ can be superposiition with different interp.

von Neumann's quantum entropy

von Neumann (1927) defined mixed quantum state and its entropy

$$S(\rho) \equiv -\text{Tr}\,\rho \log \rho = -\sum_k \lambda_k \log \lambda_k$$

where ρ spectral decomp $\rho = \sum_k \lambda_k |\chi_k\rangle\langle\chi_k|$ so λ_k e-vals

Density matrix $\rho > 0$ and $\text{Tr}\,\rho = 1 \quad \Rightarrow \quad S(\rho) \geq 0$

also find $\rho = |\psi\rangle\langle\psi|$ pure $\Leftrightarrow$ $\rho^2 = \rho$ $\Leftrightarrow$ $S(\rho) = 0$

But $S(P)$ well-defined for and concave any pos semi-def ops

will give three (3) proofs

$S(\rho) \geq 0$ is result of normalization and/or phys interp

Shannon (1948): classical info with entropy equiv. to diag matrix

Next Time: more, including subadditivity properties

Back to Measurement: Role of non-commutativity

Now consider two non-commuting observables

$$A = \sum_j a_j |\alpha_j\rangle\langle\alpha_j| = \sum_j a_j E_j, \qquad B = \sum_k b_k |\beta_k\rangle\langle\beta_k| = \sum_k b_k F_k$$

start in $|\psi\rangle$ measure A, then B ends in e-state $|\beta_k\rangle$ of B

start in $|\psi\rangle$ measure B, then A ends in e-state $|\alpha_j\rangle$ of A

in mixed $\rho = \sum_k p_k |\phi_k\rangle\langle\phi_k|$ average result of measuring A

is $\sum_k p_k \langle\phi_k, A\phi_k\rangle = \operatorname{Tr} \rho A$

Define map $\Omega_{\mathcal{M}}$ describes result of PVM or vN measurement

$$\Omega_{\mathcal{M}} : \rho \mapsto \sum_j E_j \rho E_j = \sum_j |\alpha_j\rangle\langle\alpha_j, \rho\,\alpha_j\rangle\langle\alpha_j| = \sum_j |\alpha_j\rangle\langle\alpha_j| \operatorname{Tr} \rho E_j$$

Measure B, then A ends with $F_k \mapsto \Omega_{\mathcal{M}}(F_k) = \sum_j E_j F_k E_j$

Quantum measurement: POVM

$$\sum_{jk} E_j F_k E_j = \sum_j E_j I E_j = I$$

$\{E_j F_k E_j\}$ example of POVM positive operator valued measurement

Def: (Davies and Lewis) POVM $\mathcal{M} = \{G_b\} \quad G_b > 0, \; \sum_b G_b = I$

Result of POVM depends on order in which G_b performed

QC map for von Neumann measurement

$$\Omega_{\mathcal{M}} : \gamma \mapsto \sum_j (\mathrm{Tr}\, \gamma E_j) |\alpha_j\rangle\langle\alpha_j| \qquad E_j = |\alpha_j\rangle\langle\alpha_j| \;\; \text{O.N.}$$

QC map for POVM using instrument with "pointer" $|f_b\rangle$

$$\Omega_{\mathcal{M}} : \gamma \mapsto \sum_j (\mathrm{Tr}\, \gamma G_b) |\phi_b\rangle\langle\phi_b| \otimes |f_b\rangle\langle f_b| \text{ where } \; |f_b\rangle \;\; \text{O.N.}$$

Tensor products and entanglement

Quant Info typical $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \ldots \otimes \mathcal{H}_n$ or $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B \otimes \ldots$

- Quantum Computer $\mathcal{H} = \mathbf{C}_2^{\otimes n}$ for n qubits
- Quantum Communication $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_E$

A = sender "Alice" B = receiver "Bob"

E = "Eve" (Eavesdropper – sexist) , but also

E = Environment (can be Evil or Friendly)

Partial trace $\text{Tr}_B : \mathcal{B}(\mathcal{H}_A \otimes \mathcal{H}_B) \mapsto \mathcal{B}(\mathcal{H}_A)$ or $M_{d_A} \otimes M_{d_B} \mapsto M_{d_A}$

$\text{Tr}_B\, A \otimes B = A(\text{Tr}\, B)$ extend by linearity

Formal partial inner product $|\phi_k\rangle$ O.N. basis for $\mathcal{H}_B$

$\text{Tr}_B X_{AB} = \sum_k \langle \phi_k, X_{AB}\, \phi_k \rangle_B$ means $\forall\, \chi, \psi \in \mathcal{H}_A$

$X_A = \text{Tr}_B X_{AB}$ iff $\langle \chi, X_A\, \psi \rangle = \sum_k \langle \chi \otimes \phi_k, X_{AB}\, \psi \otimes \phi_k \rangle_B$

Aside: entanglement measure

pure state $|\psi\rangle \in \mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ called "entangled" if it is

not a tensor prod, i.e., can not be written $|\psi\rangle \neq \phi_A \otimes \phi_B$

Example max entang Bell states on $\mathbf{C}_2 \otimes \mathbf{C}_2$

$$\tfrac{1}{\sqrt{2}}\big(|00\rangle \pm |11\rangle\big) \qquad \tfrac{1}{\sqrt{2}}\big(|01\rangle \pm |10\rangle\big)$$

have non-classical correlations long regarded as mysterious

modern quant info attitude: accept as way world is and ask

What nifty new things can we do with entanglement?

Measure entanglement of pure state

$E(\psi) = S\big(\mathrm{Tr}_B\, |\psi\rangle\langle\psi|\big) = S\big(\mathrm{Tr}_A\, |\psi\rangle\langle\psi|\big)$ essent unique

mixed ρ_{AB} separable if convex comb of pure product states

many (industry) of inequiv. entang measures for mixed states

Aside: SVD and "Schmidt" decomposition

Singular Value Decomposition: Recall $B^*B = \sum_k \mu_k^2 |b_k\rangle\langle b_k| \equiv |B|^2$

Then $B = U|B| = \sum_k \mu_k |a_k\rangle\langle b_k| \qquad |a_k\rangle = U|b_k>$

U partial isometry – restriction to $(\ker B)^\perp$ unique unitary

Isomorphism $\mathcal{B}(\mathcal{H}) \simeq \mathcal{H} \otimes \mathcal{H} \qquad |v\rangle\langle w| \leftrightarrow |v \otimes w\rangle$

apply SVD + iso to $|\psi\rangle \in \mathcal{H} \otimes \mathcal{H}$ $|\psi\rangle = \sum_k \mu_k |\alpha_k \otimes \beta_k\rangle$

pure $\rho_{AB} = |\psi\rangle\langle\psi| \Rightarrow$ reduced density matrices $\rho_A \equiv \text{Tr}_B\, \rho_{AB}$ etc.

$\rho_A = \sum_k |\mu_k|^2 |\alpha_k\rangle\langle\alpha_k| \quad \rho_B = \sum_k |\mu_k|^2 |\beta_k\rangle\langle\beta_k|$

Cor 1: $\rho_{AB} = |\psi\rangle\langle\psi|$ pure $\Rightarrow \rho_A, \rho_B$ have same non-zero e-vals

Cor 2: $\rho_{AB} = |\psi\rangle\langle\psi|$ pure $S(\rho_A) = S(\rho_B)$

Can reverse to get "purification" start with $\rho = \sum_k \lambda_k |\phi_k\rangle\langle\phi_k|$

Define $|\psi\rangle = \sum_k \sqrt{\lambda_k}\, |\phi_k \otimes \phi_k\rangle \in \mathcal{H} \otimes \mathcal{H}$ $\quad \mathrm{Tr}_B |\psi\rangle\langle\psi| = \rho$

Quant Info view: mystical result of Schmidt about tensor products

SVD for matrices started 1870's (Horn and Johnson, Chap. 3)

Schmidt(1907) equiv. result interp $K(x, y)$ as kernal of operator

$$g(y) \mapsto f(x) = \int K(x, y) g(y) dy$$

Rediscovered by quantum chemists called Carleson-Keller (1961)

John Coleman (1963) pointed out due to Schmidt

OK interp for $\psi(x_1 \ldots x_m, y_1 \ldots y_n) \in L_2(\mathbf{R}^{m+n})$ wave function

wrong-headed to look for extension to higher order tensor products

More info: See Appendix A of King & Ruskai

IEEE Trans. Info. Theory, 47, 192209 (2001) quant-ph/9911079

Aside: G Homogenous of degree one

Consequences of $G(\lambda A) = \lambda G(A)$

- $G(A)$ convex $\Leftrightarrow$ $G(A+B) \leq G(A) + G(B)$

convex $\Rightarrow$ "subadditive" [in sense $G(A+B) \leq G(A) + G(B)$]

$$\tfrac{1}{2}G(A+B) = G(\tfrac{1}{2}A + \tfrac{1}{2}B) \leq \tfrac{1}{2}G(A) + \tfrac{1}{2}G(B)$$

"subadditive" $\Rightarrow$ convex

$$\begin{aligned} G\big[xA + (1-x)B\big] &\leq G(xA) + G[(1-x)B] \\ &= xG(A) + (1-x)G(B) \end{aligned}$$

- $G(x)$ convex $\Rightarrow$ $\lim_{x\to 0} \frac{1}{x}\big[G(A+xB) - G(A)\big] \leq G(B)$

$$\lim_{x\to 0} \frac{G(A+xB) - G(A)}{x} \leq \lim_{x\to 0} \frac{G(A) + xG(B) - G(A)}{x}$$

Relative Entropy

Def: $H(\rho, \gamma) = \mathrm{Tr}\, \rho\big(\log \rho - \log \gamma\big)$

ρ, γ pair of density matrices, but can be any pos semi-def ops.

Use Greek $\rho, \gamma, \ldots$ for density matrices,

Use Roman R, Q for arb pos semi-def matrices

Joint Convexity $$H\Big(\sum_j R_j, \sum_j Q_j \Big) \; \leq \; \sum_j H(R_j, Q_j)$$

Cor: Monotone under Partial Trace (MPT)

$$H(R_1, Q_2) \leq H(R_{12}, Q_{12})$$

$$R_{12}, Q_{12} \in \mathcal{B}(\mathcal{H}_1 \otimes \mathcal{H}_2) \qquad R_1 = \mathrm{Tr}_2\, R_{12} \text{ etc.}$$

II. Properties of Entropy and Relative Entropy

1. Fundamental Properties of Entropy
   1.1 Concavity and subadditivity
   1.2 Minor properties
   1.3 Strong subadditiivty (SSA)
   1.4 First proof of concavity of $S(\rho)$
   1.5 Triangle inequality
2. Klein's inequality and second proof of concavity of $S(\rho)$
3. Fundamental Properties of Relative Entropy
4. Aside on Information Theory Expressions
5. Aside on Monotone and Convex Operator Functions
6. Third proof of concavity of $S(\rho)$

Fundamental Properties of Quantum Entropy

Def: $S(R) = -\text{Tr}\, R \log R$ for $R \geq 0$ pos semi-def $(0 \log 0 \equiv 0)$

- Concave: $x\, S(R_1) + (1-x)S(R_2) \leqq S\big(xR_1 + (1-x)R_2\big)$

but $\leq x\, S(R_1) + (1-x)S(R_2) + x \log x + (1-x) \log(1-x)$

- Subadditive: $S(R_{AB}) \leqq S(R_A) + S(R_B)$

  with $= \Leftrightarrow \; R_{AB} = R_A \otimes R_B$

- Strongly Subadditive $S(R_B) + S(R_{ABC}) \leqq S(R_{AB}) + S(R_{BC})$

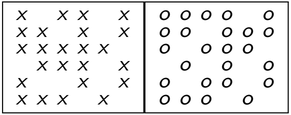
All indep of norm. if consistent, e.g., $\text{Tr}\, R_B = \text{Tr}\, R_{AB} = \text{Tr}\, R_{ABC}$

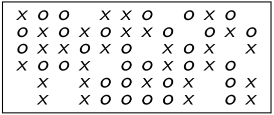
Also have (a) $S(\mu R) = \mu S(R) - \mu \log \mu \text{Tr}\, R$

(b) $R \in M_d$ and $\text{Tr}\, R = 1 \;\Rightarrow\; 0 \leq S(R) \leq \log d$

Concave: $x\,S(\rho_1) + (1-x)S(\rho_2) \leq S\big(x\rho_1 + (1-x)\rho_2\big)$

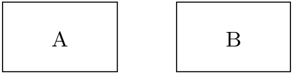
refers to mixture





Subadditive: $S(\rho_{AB}) \leq S(\rho_A) + S(\rho_B)$

refers to regions or subsystems



SSA: $S(\rho_B) + S(\rho_{ABC}) \leq S(\rho_{AB}) + S(\rho_{BC})$

overlapping regions

Subadditive $\Rightarrow$ Concave

Proof # 1 of $S(\rho)$ concave once subadd shown (easy)

$$\rho_{AB} = \begin{pmatrix} x\rho_1 & 0 \\ 0 & (1-x)\rho_2 \end{pmatrix}$$

$$\rho_A = x\rho_1 + (1-x)\rho_2, \qquad \rho_B = \left(\begin{smallmatrix} x & 0 \\ 0 & 1-x \end{smallmatrix}\right)$$

$$\begin{aligned} S(\rho_{AB}) &= S(x\rho_1) + S\big((1-x)\rho_2\big) \\ &= xS(\rho_1) - x\log x + (1-x)S(\rho_2) - (1-x)\log(1-x) \end{aligned}$$

$$\begin{aligned} S(\rho_{AB}) &\leq S(\rho_A) + S(\rho_B) \\ &= S\big(x\rho_1 + (1-x)\rho_2\big) - x\log x - (1-x)\log(1-x) \end{aligned}$$

$$\Rightarrow \qquad xS(\rho_1) + (1-x)S(\rho_2) \leq S\big(x\rho_1 + (1-x)\rho_2\big)$$

Triangle inequality for $S(\rho)$

$$S(\rho_{AB}) \leqq S(\rho_A) + S(\rho_B)$$

Given ρ_{AB} can find $\mathcal{H}_C$ and $|\psi_{ABC}\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$

s.t. $\text{Tr}_C\, \rho_{ABC} = \text{Tr}_C\, |\psi_{ABC}\rangle\langle\psi_{ABC}| = \rho_{AB}$ purification

$\rho_{AB} = \sum_k \lambda_k |\phi_k\rangle\langle\phi_k|$ $\quad |\psi_{ABC}\rangle = \sum_k \sqrt{\lambda_k}\, |\phi_k \otimes e_k\rangle$

$(\mathcal{H}_A \otimes \mathcal{H}_B) \otimes \mathcal{H}_C$ $S(\rho_{AB}) = S(\rho_C)$ $\quad \mathcal{H}_A(\otimes \mathcal{H}_B \otimes \mathcal{H}_C)$ $S(\rho_A) = S(\rho_{BC})$

subs above $\quad S(\rho_C) \leq S(\rho_{BC}) + S(\rho_B)$

$$\Rightarrow \quad S(\rho_C) - S(\rho_B) \leq S(\rho_{BC})$$

Reverse B $\leftrightarrow$ C and combine with subadd to get

$$\big|S(\rho_A) - S(\rho_B)\big| \leq S(\rho_{AB}) \leqq S(\rho_A) + S(\rho_B)$$

Klein's inequality:

$$\mathrm{Tr}\, A \log A - \mathrm{Tr}\, A \log B \geq \mathrm{Tr}\,(A - B) \quad \text{with} \;=\; \text{iff } A = B$$

g convex means diff quotients increase

$$\Rightarrow \quad \tfrac{g(b)-g(a)}{b-a} \leq g'(b) \quad \text{for} \quad a < b$$

$$\Rightarrow \quad g(b) - g(a) \leq (b-a)g'(b) \text{ for all } \; a, b$$

$\mathrm{Tr}\left[g'(B)(B-A) - g(B) + g(A)\right]$ $\qquad |\beta_k\rangle$ norm e-vec of B

$$= \sum_k \left[g'(b_k)\big(b_k - \langle \beta_k, A\,\beta_k\rangle\big) - g(b_k) + \langle \beta_k, g(A)\beta_k\rangle\right]$$

Jensen $g\big(\langle \beta_k, A\,\beta_k\rangle\big) \leq \langle \beta_k, g(A)\beta_k\rangle$

$$\geq \sum_k \left[g'(b_k)\big(b_k - \langle \beta_k, A\beta_k\rangle\big) - g(b_k) + g\big(\langle \beta_k, A\beta_k\rangle\big)\right] \geq 0$$

For $g(x) = x \log x, \quad g'(x) = 1 + \log x$

$$\mathrm{Tr}\left[(B-A)(I + \log B) - B\log B + A \log A\right] \geq 0$$

Entropy concave : Proof # 2

$A = R_1$ and $B = R = xR_1 + (1-x)R_2$ in Klein

$$\mathrm{Tr}\, A \log A - \mathrm{Tr}\, A \log B \geq \mathrm{Tr}\,(A - B)$$

Get $\quad \mathrm{Tr}\, R_1 \log R_1 - R_1 \log R \geq \mathrm{Tr}\,(R_1 - R)$

Repeat for R_2 and Mult by x and $1 - x$

$$\begin{aligned} x\,\mathrm{Tr}\, R_1 \log R_1 - \ x\, R_1 \log R &\geq \mathrm{Tr}\left[\ x\, R_1 - R\right] \\ (1-x)\mathrm{Tr}\, R_2 \log R_2 - (1-x) R_2 \log R &\geq \mathrm{Tr}\left[(1-x) R_2 - R\right] \end{aligned}$$

add to get

$$\begin{aligned} -xS(R_1) - (1-x)S(R_2) - \left[xR_1 + (1-x)R_2\right] \log R &\geq \mathrm{Tr}\,(R - R) \\ -xS(R_1) - (1-x)S(R_2) + S(R) &\geq 0 \end{aligned}$$

Relative entropy

Def: $H(R, Q) \;=\; \text{Tr}\, R \log R - \text{Tr}\, R \log Q$

Klein's ineq: $H(R, Q) \geq \text{Tr}\,(R - Q) \qquad \geq 0$ if $\text{Tr}\, R = \text{Tr}\, Q$

assume $R, Q > 0$ strictly pos — well-def if $\ker(Q) \subset \ker(R)$

Can obtain entropy from rel ent. $\quad H(R, \frac{1}{d} I) = -S(R) + \log d$

Or simply $H(R, I) = -S(R)$

homogenous of degree one $\qquad H(\lambda R, \lambda Q) = \lambda H(R, Q)$

Recall $\Rightarrow$ convexity equiv to simple form $F(A + B) \leq F(A) + F(B)$

Fundamental Properties of Relative Entropy

Joint Convexity $H(R_1 + R_2, Q_1 + Q_2) \leq H(R_1, Q_1) + H(R_2, Q_2)$

Cor: (MPT) $H(R_A, Q_A) \leq H(R_{AB}, Q_{AB})$

Monotone under Partial Trace

Ibinson-Winter: that's all folks!

Special Case of MPT $R_{AB} \to \rho_{ABC}, \; Q_{AB} \to I_A \otimes \rho_{BC}$

gives strong subadditivity $H(\rho_{AB}, \rho_B) \leq H(\rho_{ABC}, \rho_{BC})$

SSA $-S(\rho_{AB}) + S(\rho_B) \leq -S(\rho_{ABC}) + S(\rho_{BC})$

Will prove JC and then show $\Rightarrow$ MPT

Aside: Information Theory Expressions

Mutual Information: (always positive)

$$S(A:B) = S(\rho_A) + S(\rho_B) - S(\rho_{AB}) = H(\rho_{AB}, \rho_A \otimes \rho_B) \geq 0$$

Conditional Information: (always positive for classical systems)

$$\begin{aligned} S(A|B) &= S(\rho_{AB}) - S(\rho_B) = -H(\rho_{AB}, \rho_B) \\ &= -H(\rho_{AB}, \tfrac{1}{d} I_A \otimes \rho_B) + \log d \end{aligned}$$

Max entangled state $\quad S(\rho_{AB}) - S(\rho_B) = -\log d < 0$

$$\rho_{AB} = |\psi_{AB}\rangle\langle\psi_{AB}| \quad \psi_{AB} = \textstyle\sum_k \frac{1}{d} |\phi_k \otimes \phi_k\rangle$$

Conditional Info $S(\rho_{AB}) - S(\rho_B)$ concave

surprising since diff of concave functions – equiv. to SSA

Weak monotonicity of conditional Information

Classical $\quad S(\rho_{AB}) - S(\rho_B) \geq 0$

Quantum given any ρ_{ABC}

Purify $\rho_{ABC} = \sum_k \lambda_k |\phi_k\rangle\langle\phi_k|$ to ρ_{ABCD} s.t. $\text{Tr}_B\, \rho_{ABCD} = \rho_{ABC}$

$$\rho_{ABCD} = |\psi\rangle\langle\psi| \qquad |\psi_{ABCD}\rangle = \sum_k \sqrt{\lambda_k}\, |\phi_l \otimes f_k\rangle$$

where $|f_k\rangle$ O.N. in space iso to $(\ker \rho_{ABC})^\perp \subseteq \mathcal{H}_{ABC}$.

SSA:
$$S(\rho_{ABC}) + S(\rho_B) \;\leq\; S(\rho_{AB}) + S(\rho_{BC})$$
$$S(\rho_D) + S(\rho_B) \;\leq\; S(\rho_{CD}) + S(\rho_{BC})$$
$$S(\rho_{CD}) - S(\rho_D) + S(\rho_{BC}) - S(\rho_B) \;\geq\; 0$$

two cond entropies with common subsystem can't both be negative

Negative Conditional Information

QIT uses EPR states $\frac{1}{\sqrt{2}}\big(|00\rangle + |11\rangle\big)$ to transmit

a) classical information by process called "dense coding"

b) unknown quantum states by process called "teleportation"

HOW (M. Horodecki, J. Oppenheim, A. Winter) interpretation:

Nature **436**, 673–676 (2005); CMP **269** , (2007). quant-ph/0512247.

Cond info measures # of bits Alice needs to transmit message to

Bob when he has partial info – same interp class. and quant info

When negative, gives # of EPR pairs A and B have left for future

Aside: Operator-Monotone and -Convex Functions

$$A > B > 0 \;\Rightarrow\; \sqrt{A} > \sqrt{B} > 0 \;\text{ but }\; \not\Rightarrow \; A^2 > B^2$$

Def: $f : (a,b) \mapsto \mathbf{R}$ is operator-monotone if (a.k.a. Pick or Herglotz)

$$A > B > 0 \;\Rightarrow\; f(A) > f(B).$$

Thm: Let $f : (0,\infty) \mapsto \mathbf{R}$. TFAE the following are equivalent

a) f is operator monotone

b) f can be anal cont into UHP and maps UHP into UHP

$$\text{c) } f(x) = ax + \int_0^\infty \tfrac{xu-1}{x+u}\,\nu(u)du \qquad \text{with} \quad \nu(u) \geq 0$$

$$= a'x - \int_0^\infty \tfrac{1}{x+u}(1+u^2)\nu(u)du \quad \text{if} \quad \int u\,\nu(u)du < \infty$$

where UHP $= z : \mathrm{Im} z > 0$ (b) $\Leftrightarrow$ (c) Nevanlinna's Thm.

Operator-Monotone and -Convex Functions continued

$A > B > 0 \;\Rightarrow\; X = B^{-1/2}AB^{-1/2} > I \;\Rightarrow\;$ e-vals of $X > 1$

e-vals of $X^{-1} < 1 \;\Rightarrow\; B^{1/2}A^{-1}B^{1/2} < I \;\Rightarrow\; 0 < A^{-1} < B^{-1}$

So $f(x) = -(x+u)^{-1}$ is op-mon which gives (c) $\Rightarrow$ (a).

Def: $g : (0,\infty) \mapsto \mathbf{R}$ operator-convex if

$$g\big(xA_1 + (1-x)A_2\big) \le xg(A_1) + (1-x)g(A_2)$$

Roughly g is op-convex iff suitable diff quot is op-mon.

$g : [0,\infty) \mapsto \mathbf{R}$ and $g(0) = 0$ is op-convex iff

$$\frac{g(x) - g(0)}{x - 0} = \frac{g(x)}{x} \text{ op-mon}$$

Theory due to Löwner   Ando 197? notes

Entropy concave : Proof # 3

$g : [0,\infty) \mapsto \mathbf{R}$ and $g(0) = 0$ is op-convex iff $\frac{g(x)}{x}$ op-mon

Apply to $g(x) = x \log x$

$$\frac{g(x)}{x} = \log x : \text{UHP} \mapsto \{z : \mathrm{Im} z \in (0, 2\pi)\} \subset \text{UHP}$$

Find: $g(x) = x \log x$ op-convex $g(x) = -x \log x$ op concave

$\Rightarrow$ and stronger than $S(R) = -R \log R$ concave C. Davis

$f(x) = x^p$	$-1 < p < 0$	op-convex	and	op-mon-dec
	$0 < p < 1$	op-concave	and	op-mon
	$1 < p \leq 2$	op-convex	but	not op-mon

$f(x) = x^p$ neither op-convex nor op-mon for $p > 2$.

III: Simple Proof of Joint Convexity of Relative Entropy

1. Background and WYD
2. Unified proof of joint conv. of $\text{Tr}\, K^* R^p K Q^{1-p}$ and $H(R, Q)$
    2.1 Pedestian modular operator via left and right mult.
    2.2 Integral representations
    2.3 Convexity of $J_p(K, A, B)$
3. Comments on Schwarz inequalities
4. $q \neq 1 - p$
5. Monotonicity under partial traces
6. More history ?
7. Lieb's golden corollary
8. Equality conditions

Background to proofs

Original proof of SSA based on Lieb's result below

WYD skew entropy $\frac{1}{2}\text{Tr}\,[K,\gamma^p][K,\gamma^{1-p}]$

for $K = K^*$ and γ a density matrix

Wigner-Yanase introduced for $p = \frac{1}{2}$ and proved concave in γ.

Dyson suggested $p \in (0,1)$ – led to conjecture

Conj: $\gamma \mapsto \text{Tr}\,K\gamma^p K\gamma^{1-p} - \text{Tr}\,K\gamma K$ concave

Lieb dropped linear term and proved generalization

$(A,B) \mapsto \text{Tr}\,K^* A^p K B^{1-p}$ concave for $p \in (0,1)$

Claim: But advantage to retaining linear term !!

Retaining the linear term

$$J_p(K, A, B) = \frac{1}{p(1-p)}\left[\mathrm{Tr}\, K^* A K - \mathrm{Tr}\, K^* A^p K B^{1-p}\right]$$

Note: well-def for $p > 0$ and factor $(1-p)$ changes sign at $p = 1$

Thm: $(A, B) \mapsto J_p(K, A, B)$ is convex for $p \in (0, 2)$

$\Rightarrow$ $\mathrm{Tr}\, K^* A^p K B^{1-p}$ concave for $p \in (0, 1)$

$\Rightarrow$ $\mathrm{Tr}\, A(\log A - \log B)$ convex $p = 1$ – extend by cont $K = I$

$\Rightarrow$ $\mathrm{Tr}\, K^* A^p K B^{1-p}$ convex for $p \in (1, 2]$ and $p \in [-1, 0)$

will give proof, which is elementary, short, and sweet

$\mathrm{Tr}\, A = \mathrm{Tr}\, B \quad \Rightarrow \quad J_p(I, A, B) \geq 0$ with equality $\Leftrightarrow$ $A = B$

pseudo-metric in same sense as relative entropy $H(A, B)$ gen Klein

Pedestrian modular operator

$d \times d$ matrices form Hilbert space with $\langle A, B \rangle = \text{Tr}\, A^* B$

Def. Left and Right mult as linear operators on this vector space

$$L_A(X) = AX \qquad \text{and} \qquad R_B(X) = XB$$

a) L_A and R_B commute $L_A[R_B(X)] = AXB = R_B[L_A(X)]$

b) $A = A^* \;\Rightarrow\; L_A, R_A$ self-adjoint wrt H-S inner prod

For $A, B > 0$ positive definite

c) L_A, R_A pos def $\langle X, R_A(X) \rangle = \text{Tr}\, X^* X A = \text{Tr}\, X A X^* \geq 0$

d) $(L_A)^{-1} = L_{A^{-1}}, \quad (R_B)^{-1} = R_{B^{-1}}$

e) $f(L_A) = L_{f(A)} \quad f(R_B) = R_{f(B)}$, e.g., $L_A^p = L_{A^p}, R_A^p = R_{A^p}$

simple form of deep idea: Araki $\Delta_{AB} = L_A R_B^{-1}$ relative modular op

Back to $J_p(K, A, B)$

$$g_p(x) = \begin{cases} \frac{1}{p(1-p)}(x - x^p) & p \neq 1 \\ x \log x & p = 1 \end{cases}.$$

well-defined for $x > 0$ and $p \neq 0$, but $p \in [\frac{1}{2}, 2]$ would suffice

$$J_p(K, A, B) \equiv \operatorname{Tr} \sqrt{B} K^* g_p\big(L_A R_B^{-1}\big)(K\sqrt{B})$$

$$= \begin{cases} \frac{1}{p(1-p)}\big(\operatorname{Tr} K^*AK - \operatorname{Tr} K^*A^pKB^{1-p}\big) & p \in (0,1) \cup (1,2) \\ \operatorname{Tr} KK^*A \log A - \operatorname{Tr} K^*AK \log B) & p = 1 \\ -\frac{1}{2}\big(\operatorname{Tr} K^*AK - \operatorname{Tr} AKB^{-1}K^*A\big) & p = 2 \end{cases}$$

$$J_1(I, A, B) = \operatorname{Tr} A(\log A - \log B) = H(A, B)$$

Aside: extend to $[-1, 0)$

not quite symmetric around $p = \frac{1}{2}$ $\quad p \leftrightarrow 1 - p$

$$\widetilde{g}_p(x) = w\, g_{1-p}(w^{-1}) = \begin{cases} \frac{1}{p(1-p)}(1 - w^p) & p \neq 0 \\ -\log w & p = 0 \end{cases} \qquad p \in [-1, 1)$$

$$J_p(K, B, A) = \widetilde{J}_{1-p}(K^*, A, B)$$

$\widetilde{J}_p(K, A, B)$ jointly convex for $p \in [-1, 1)$

$$\widetilde{J}_0(I, A, B) = \operatorname{Tr} B(\log B - \log A) = H(B, A)$$

Integral representations

$$\frac{g_p(x)}{x} = \begin{cases} \frac{1}{p(1-p)}(1 - x^{p-1}) & p \neq 1 \\ \log x & p = 1 \end{cases}$$

well-def for $x \in (0, \infty)$ and operator monotone for $p \in (0, 2]$, or,

anal cont to upper half of complex plane and UHP $\mapsto$ UHP

$\Rightarrow g_p(x)$ has integral rep of form

$$\begin{aligned} g_p(x) &= ax^2 + \int_0^\infty \frac{x^2 t - x}{x + t}\,\nu(t)\,dt \\ &= ax^2 + \int_0^\infty \left[\frac{x^2}{x+t} - \frac{1}{t} + \frac{1}{x+t}\right] t\nu(t)\,dt \end{aligned}$$

with $\nu(t) \geq 0$

Specific integrals – elementary

$$\int_0^\infty \frac{x^{p-1}}{x+1} = \frac{\pi}{\sin p\pi} \qquad 0 < p < 1 \qquad c_p = \frac{\sin p\pi}{\pi}$$

allows us to give the following explicit representations

$$g(x) = \begin{cases} \frac{1}{p(1-p)}\left[x + c_p \int_0^\infty \left(\frac{t}{x+t} - 1\right) t^{p-1} dt\right] & p \in (0,1) \\ \int_0^\infty \left(\frac{x^2}{x+t} - 1 + \frac{t}{x+t}\right) \frac{1}{1+t} dt & p = 1 \\ \frac{1}{p(1-p)}\left[x - c_{p-1} \int_0^\infty \frac{x^2}{x+t} t^{p-2} dt\right] & p \in (1,2) \\ \frac{1}{2}(-x + x^2) & p = 2 \end{cases}$$

Important: For $p \in (0,2)$ integrand supported on $(0,\infty)$.

Integal representation using L_A and R_B

Recall $J_p(K,A,B) \;=\; \mathrm{Tr}\sqrt{B}K^* g_p\big(L_A R_B^{-1}\big)(K\sqrt{B})$

$$g_p(x) = ax + \int_0^\infty \left[\tfrac{x^2}{x+t} - \tfrac{1}{t} + \tfrac{1}{x+t}\right] t\nu(t)\,dt$$

$$\begin{aligned}\mathrm{Tr}\sqrt{B}K^*\frac{1}{L_A R_B^{-1}+tI}(K\sqrt{B}) &= \mathrm{Tr}\sqrt{B}K^*\frac{R_B}{L_A+tR_B}(K\sqrt{B}) \\ &= \mathrm{Tr}\,BK^*\frac{1}{L_A+tR_B}(KB)\end{aligned}$$

$$\begin{aligned}J_p(K,A,B) &= \mathrm{Tr}\,K^*AK - \mathrm{Tr}\,KBK^*\int_0^\infty \nu(t)\,dt \\ &+ \int_0^\infty \left[\mathrm{Tr}\,K^*A\frac{1}{L_A+tR_B}(AK) + \mathrm{Tr}\,BK^*\frac{1}{L_A+tR_B}(KB)\right] t\,\nu(t)\,dt\end{aligned}$$

Suffices to show $(A,B,X) \mapsto \mathrm{Tr}\,X^*\dfrac{1}{L_B+tR_A}(X)$ jointly convex

Proof:

Note: $\mathrm{Tr}\,(\lambda X)^* \dfrac{1}{L_{\lambda B} + tR_{\lambda A}}(\lambda X) = \lambda \,\mathrm{Tr}\, X^* \dfrac{1}{L_B + tR_A}(X)$

Homo of degree 1 $\Rightarrow$ suffices to prove "subadditivity" (omit x_k)

Let: $M = (\;)^{-1/2}(X) - (\;)^{1/2}(\Lambda)$

$$\begin{aligned} \mathrm{Tr}\, M^* M &= \langle M, M \rangle \\ &= \langle [(\;)^{-1/2}(X) - (\;)^{1/2}(\Lambda)]\,,\,[(\;)^{-1/2}(X) - (\;)^{1/2}(\Lambda)] \rangle \\ &= \langle X, (\;)^{-1}(X) \rangle - \langle X, \Lambda \rangle - \langle \Lambda, X \rangle + \langle \Lambda, (\;)(\Lambda) \rangle \end{aligned}$$

Choose $M = (L_A + tR_B)^{-1/2}(X) - (L_A + tR_B)^{1/2}(\Lambda)$

$$\mathrm{Tr}\, M^* M =$$
$$\mathrm{Tr}\, X^*(L_A + tR_B)^{-1}(X) - \mathrm{Tr}\, X^*\Lambda - \mathrm{Tr}\,\Lambda^* X + \mathrm{Tr}\,\Lambda^*(L_A + tR_B)(\Lambda)$$

Let $M_j = (L_{A_j} + tR_{B_j})^{-1/2}(X_j) - (L_{A_j} + tR_{B_j})^{1/2}(\Lambda)$. Then

$$
\begin{aligned}
0 \;&\leq\; \sum_j \operatorname{Tr} M_j^* M_j \;=\; \sum_j \operatorname{Tr} X_j^* (L_{A_j} + tR_{B_j})^{-1}(X_j) \\
&\qquad -\operatorname{Tr}\big(\textstyle\sum_j X_j^*\big)\Lambda - \operatorname{Tr}\Lambda^*\big(\textstyle\sum_j X_j\big) + \operatorname{Tr}\Lambda^* \textstyle\sum_j \big(L_{A_j} + tR_{B_j}\big)\Lambda \\
&= \sum_j \operatorname{Tr} X_j^* \frac{1}{(L_{A_j}+tR_{B_j})}(X_j) - \operatorname{Tr} X^*\Lambda - \operatorname{Tr}\Lambda^* X + \operatorname{Tr}\Lambda^*(L_A + tR_B)\Lambda
\end{aligned}
$$

Choose $\Lambda = \dfrac{1}{L_A + tR_B}(X)$ $\qquad X = \sum_j X_j, \ \sum_j L_{A_j} = L_{\sum_j A_j} = L_A$

$$
\operatorname{Tr}\Lambda^* \textstyle\sum_j \big(L_{A_j} + tR_{B_j}\big)\Lambda \;=\; \operatorname{Tr} X^* \frac{1}{L_A + tR_B} X \;=\; \operatorname{Tr} X \Lambda \;=\; \operatorname{Tr}\Lambda^* X
$$

$$
0 \;\leq\; \sum_j \operatorname{Tr} X_j^* \frac{1}{L_{A_j}+tR_{B_j}}(X_j) \;-\; \operatorname{Tr} X^* \frac{1}{L_A + \, tR_B}(X)
$$

compare elementary C-S ineq:

$$\Big|\sum_k \overline{v}_k w_k\Big|^2 \;\leq\; \sum_k |v_k|^2 \;\sum_k |w_k|^2$$

For $a_k > 0$ let $v_k = a_k^{1/2}, \; w_k = a_k^{-1/2} x_k$

$$\Big|\sum_k x_k\Big|^2 \;\leq\; \sum_k a_k \;\sum_k \overline{x}_k \frac{1}{a_k} x_k$$

Rewrite $\quad \Big(\sum_k \overline{x}_k\Big) \dfrac{1}{\sum_k a_k} \Big(\sum_k x_k\Big) \;\leq\; \sum_k \overline{x}_k \dfrac{1}{a_k} x_k$

Lieb and Ruskai (1973) proved operator version

$$\Big(\sum_k X_k^*\Big) \frac{1}{\sum_k A_k} \Big(\sum_k X_k\Big) \;\leq\; \sum_k X_k^* \frac{1}{A_k} X_k$$

Not suff. for SSA — need Araki rel mod op hidden in L_A and R_B.
Compare proof: $\big|\sum_k v_k + t w_k\big|^2 \geq 0 \;\forall\; t$ choose t to minimize

Remarks on $q \neq 1 - p$

$p, q > 0, \; p + q < 1 \qquad \mathrm{Tr}\, K^* A^p K B^{1-p}$ concave

Write $\mathrm{Tr}\, K^* A^p K B^q = \mathrm{Tr}\, K^* A^p K (B^s)^{1-p} \qquad 0 < s = \frac{1}{1-p} < 1$

B^s is op monotone and op concave for $s \in (0,1)$

$$\big(\lambda B_1 + (1-\lambda) B_2\big)^s > \lambda B_1^s + (1-\lambda) B_2^s$$

$$\begin{aligned} \mathrm{Tr}\, K^* A^p K B^q &= \mathrm{Tr}\, K^* A^p K \Big[\big(\lambda B_1 + (1-\lambda) B_2\big)^s\Big]^{1-p} \\ &> \mathrm{Tr}\, K^* A^p K \big(\lambda B_1^s + (1-\lambda) B_2^s\big)^{1-p} \\ &\geq \lambda \mathrm{Tr}\, K^* A_1^p K (B_1^s)^{1-p} + (1-\lambda) \mathrm{Tr}\, K^* A_2^p K (B_2^s)^{1-p} \\ &= \lambda \mathrm{Tr}\, K^* A_1^p K B_1^q + (1-\lambda) \mathrm{Tr}\, K^* A_2^p K B_2^q \end{aligned}$$

Note: $f(x)$ strictly concave and op concave $\Rightarrow$ strict op ineq

get equal only for trivial cases, $B_1 = B_2$ or $\lambda = 0, 1$.

Monotonicity under partial traces

Define generalized Pauli (Weyl-heisenberg) operators,

$$Z|e_n\rangle = e^{2\pi i n/d}|e_n\rangle \qquad X|e_n\rangle = |e_{n+1}\rangle$$

$$\sum_j Z^j A Z^{-j} = d\, A_{\rm diag} \qquad \sum_j X^j A_{\rm diag} X^{-j} = (\operatorname{Tr} A) I$$

$$\tfrac{1}{d}\sum_j \sum_k X^j Z^k\, A\, (X^j Z^k)^* = (\operatorname{Tr} A)\, I = \tfrac{1}{d}\sum_n W_n A W_n^*$$

$W_n = X^j Z^k$ in some ordering $n = 1, 2, \ldots d^2$, e.g., $n = j + d(k-1)$

$$\tfrac{1}{d_2}\sum_n (I_1 \otimes W_n)\, A_{12}\, (I_1 \otimes W_n)^* = A_1 \otimes I_2$$

Discrete version of Uhlmann's observation that partial trace can be obtained by integrating over $SU(n)$ using Haar measure.

$$J_p(K_2, A_2, B_2) = J_p(I_1 \otimes K_2, \tfrac{1}{d_1} I_1 \otimes A_2, \tfrac{1}{d_1} I_1 \otimes B_2)$$

$$= \tfrac{1}{d_1^2} J_p\Big(K_{12}, \textstyle\sum_n (W_n \otimes I_2) A_{12} (W_n \otimes I_2)^*, \sum_n (W_n \otimes I_2) B_{12} (W_n \otimes I_2)^*\Big)$$

$$\leq \tfrac{1}{d_1^2} \sum_n J_p(I_1 \otimes K_2, (W_n \otimes I_2) A_{12} (W_n \otimes I_2)^*, (W_n \otimes I_2) B_{12} (W_n \otimes I_2)^*)$$

$$= \tfrac{1}{d_1^2} \sum_n J_p(I_1 \otimes K_2, A_{12}, B_{12}) = J_p(I_1 \otimes K_2, A_{12}, B_{12})$$

used $J_p(I_1 \otimes K_2, A_{12}, B_{12})$ wrote $K_{12} = I_1 \otimes K_2$

$$= J_p\big(I_1 \otimes K_2, (W_n \otimes I_2) A_{12} (W_n \otimes I_2)^*, (W_n \otimes I_2) B_{12} (W_n \otimes I_2)^*\big)$$

$J_1(I, A_2, B_2) \leq J_1(I, A_{12}, B_{12})$ gives $H(A_2, B_2) \leq H(A_{12}, B_{12})$

Cor: SSA $H(A_{23}, A_2) \leq H(A_{123}, A_{13})$

“no transparent proof of SSA is known”

p. 645 of *Quantum Computation and Quantum Information*

Michael A. Nielsen and Isaac L. Chuang (Cambridge Press, 2000)

based on B. Simon’s version adapted from Uhlmann (1977) of

“elementary” proof of $(A, B) \mapsto \text{Tr}\, K^* A^p K B^{1-p}$ concave

similar argument in Wehrl *Rev. Mod. Phys* (1978). BUT

- MBR, “Lieb’s simple proof of concavity . . .” quant-ph/0404126

Int. J. Quant Info. **3**, 579–590 (2005) Schwarz + max mod

- Ando’s argument described in Carlen’s Tucson notes

- Petz – uses Δ_{AB} in book; elem version in quant-ph/0408130

- Proof here based on Schwarz ineq. using L_A, R_B really elem.

based on Lesniewski and Ruskai, JMP; and MBR quant-ph/0604206

Aside: Equality conditions in $J_p(K, A, X)$ convex

$$\int_0^\infty \text{Tr}\, K^* A \frac{1}{L_A + tR_B}(AK)\nu(t)dt \leq \int_0^\infty \sum_j \text{Tr}\,(A_j K)^* \frac{1}{L_{A_j} + tR_{B_j}}(A_j K)\nu(t)$$

$$= \sum_j \int_0^\infty \text{Tr}\,(A_j K)^* \frac{1}{L_{A_j} + tR_{B_j}}(A_j K)\nu(t)d$$

Equal $\Leftrightarrow$ equal for each term in integ , i.e., $M_j = 0 \qquad \forall\, j, \ \forall\, t$

$$(L_{A_j} + tR_{B_j})^{-1}(X_j) = (L_A + tR_B)^{-1}(X) \qquad \forall\, j, \ \forall\, t$$

equality conditions independent of $p \in (0, 2)$

$X = AK \qquad (I + t\Delta_{A_j B_j}^{-1})^{-1}(K) = (I + t\Delta_{AB}^{-1})^{-1}(K) \qquad \forall\, j, \ \forall\, t$

$X = BK \qquad (\Delta_{A_j B_j} + tI)^{-1}(K) = (\Delta_{AB} + tI)^{-1}(K) \qquad \forall\, j, \ \forall\, t$

Recall $\Delta_{AB} = L_A R_B^{-1} > 0$ prod of commuting pos def ops

$$(\Delta_{A_jB_j} + tI)^{-1}(K) = (\Delta_{AB} + tI)^{-1}(K) \qquad \forall\, j,\ \forall\, t$$

$\Delta_{AB} > 0 \quad \Rightarrow \quad (\Delta_{AB} + tI)^{-1}$ anal cont to $\mathbf{C} \backslash (-\infty, 0]$

can apply Cauchy integral Thm. to get

$\Rightarrow\ G(\Delta_{A_jB_j})(K) = G(\Delta_{AB})(K) \quad \forall\, j \quad G$ anal on $\mathbf{C} \backslash (-\infty, 0]$

allows several useful formulations

$\Rightarrow\ (\Delta_{AB} + tI)$ and $(I + \Delta_{AB}^{-1} t)$ forms equiv.

Equivalent equality conditions

Thm: For fixed K, and $A = \sum_j A_j, B = \sum_j B_j$ TFAE

a) $J_p(K, A, B) = \sum_j J_p(K, A_j, B_j)$ for all $p \in (0, 2)$.

b) $J_p(K, A, B) = \sum_j J_p(K, A_j, B_j)$ for some $p \in (0, 2)$.

c) $(\Delta_{A_j B_j} + tI)^{-1}(K) = (\Delta_{AB} + tI)^{-1}(K) \quad \forall\, j$ and $\forall\, t > 0$.

d) $A_j^{it} K B_j^{-it} = A^{it} K B^{-it} \quad \forall\, j$ and $\forall\, t > 0$.

e) $(\log A - \log A_j)K = K(\log B - \log B_j) \quad \forall\, j$.

In addition when $K = I$, equiv to

f) There are $D_j > 0$ such that $[A_j, D_j] = [B_j, D_j] = 0$, and

$$A_j = AD^{-1}D_j, \quad B_j = BD^{-1}D_j \text{ with } D = \sum_j D_j$$

neccessity of (f) uses sufficient subalgebra – developed by Petz

formulation here from Jenčová and Petz, CMP, **263**, 259–276 (2006).

Equality conditions for SSA

use form $\log A_{123} - \log A_{12} - \log A_{23} + \log A_2 = 0$

Easy to see $A_{123} = A_1 \otimes A_{23}$ or $A_{12} \otimes A_3$ will suffice

If $\mathcal{H}_2 = \mathcal{H}_{2_L} \otimes \mathcal{H}_{2_R}$ then $A_{123} = A_{12_L} \otimes A_{2_R 3}$ will suffice

Thm: Equality holds in SSA if and only if

$$\mathcal{H}_2 = \bigoplus_n \mathcal{H}_n^L \otimes \mathcal{H}_n^R \quad \text{and} \quad A_{123} = \bigoplus_n A_n^L \otimes A_n^R$$

with $A_n^L \in \mathcal{B}(\mathcal{H}_1 \otimes \mathcal{H}_n^L), \quad A_n^R \in \mathcal{B}(\mathcal{H}_n^R \otimes \mathcal{H}_3)$

Cor: Equality in $\mathrm{Tr}\, A_{23}^p A_2^{1-p} \leq \mathrm{Tr}\, A_{123}^p A_{12}^{1-p}$ iff same cond

$p \in (0,1)$ $\leq$ $\qquad p \in (1,2)$ $\geq$

Other approaches to JC of $\mathrm{Tr}\, K^* A^p K B^{1-p}$

Lieb; based on maximum modulus principle –

basic result rel easy in finite dim.

Epstein; based on op-monotone or Herglotz functions

deep and not easy – paper must be read backwards !

Uhlmann - Simon elementary, but long and not insightful

gave bad rep – see Nielsen-Chuang quote

Ando; used iso between $\mathcal{B}(\mathcal{H}_B, \mathcal{H}_A) \simeq \mathcal{H}_A \otimes \mathcal{H}_B$ to observe

$$\mathrm{Tr}\, K^* A^p K B^{1-p} = \langle K, A^p \otimes B^{1-p} K \rangle$$

where K interp as vec in $\mathcal{H}_A \otimes \mathcal{H}_B$

to give linear algebra proof — first with no complex anal.

Differentiatie $\log A$ for matrix

$$\begin{aligned}
\log(R + xQ) - \log R &= \int_0^\infty \left(\frac{1}{R + uI} - \frac{1}{R + xQ + uI} \right) du \\
&= \int_0^\infty \frac{1}{R+uI} \big[(R + xQ + uI) - (R + uI) \big] \frac{1}{R + xQ + uI} du \\
&= \int_0^\infty \frac{1}{R + uI} xQ \frac{1}{R + xQ + uI} du
\end{aligned}$$

Can take $\lim_{x \to 0} \frac{1}{x} (\log(R + xQ) - \log R$ and find

$$\log(R + xQ) = \log R + x \int_0^\infty \frac{1}{R + uI} Q \frac{1}{R + uI} du + O(x^2)$$

Lieb's golden inequality

$R \mapsto G(R) \equiv \text{Tr}\, e^{A+\log R}$ concave and $G(\lambda R) = \lambda G(R)$

- Lieb extracted from JC of $\text{Tr}\, K^* A^p K B^{1-p}$ by chain of ineq.
- Epstein: easy to check map UHP $\mapsto$ UHP – at end of paper

Recall $\Rightarrow \lim_{x \to 0} \frac{1}{x} \big[G(R + xQ) - G(R) \big] \geq G(Q)$

Use results above, let $A = \log T - \log R$, after a bit of work

$$\text{Tr}\, e^{\log Q - \log R + \log T} \leq \text{Tr} \int_0^\infty T \frac{1}{R + uI} Q \frac{1}{R + uI} du$$

Golden -Thompson-Symanzik $\qquad \text{Tr}\, e^{A+B} \leq \text{Tr}\, e^A e^B$

Above subs for $\quad \text{Tr}\, e^{A+B+C} \not\leq |\text{Tr}\, e^A e^B e^C|$

Used for orig proof of SSA described in Exercise 14

Epstein and Lieb-Thirring

Lieb $R \mapsto \operatorname{Tr} e^{A+\log R}$ $\quad A = A^*$ concave

Epstein – proof via integral reps of Herglotz (op-mon) fctn

Compare using Lie-Trotter $\operatorname{Tr} e^{A+\log R} = \lim_{n\to\infty} (e^{A/2n} R^{1/n} (e^{A/2n})^n$

Lieb conj and Epstein showed

$A \mapsto \operatorname{Tr} (B\, A^{1/p} B)^p$ concave for $p \geq 1, A, B \geq 0$

Pisier (op space theory) Lemma 1.14 in *Non-comm L_p ....* (2002).

Cor: (Lieb-Thriing): $\operatorname{Tr} \left(\sqrt{B} A \sqrt{B}\right)^p \leqq \operatorname{Tr} A^p B^p$ for $p > 1$

Later proofs by H. Araki, B. Simon (esp. simple) Theorem I.4.9 in

The Statistical Mechanics of Lattice Gases (Princeton, 1993)

IV. Quantum Noise and More Inequalities

1. Quantum noise, channels and CPT maps
2. Stinespring, Kraus and Choi representations
3. Holevo bound
4. Carlen Lieb inequalites for mixed (p, q) "norms"
5. Two generalization of SSA
6. Horn's Lemma
7. Audenaert-Ruskai conjecture on block matrix generalization

Quantum Noise

Noise model: Hilbert space of system $\mathcal{H}_C$ and environment $\mathcal{H}_E$

Hamiltonian $H_{CE} = H_C \otimes I_E + I_C \otimes H_E + V_{CE}$ on $\mathcal{H}_C \otimes \mathcal{H}_E$

pure state $|\psi_C^0\rangle$ means prod $|\psi_C^0 \otimes \psi_E^0\rangle \equiv |\Psi_{CE}^0\rangle$ which evolves

in time to $|\Psi_{CE}(t)\rangle = U_{CE}(t)|\psi_C^0 \otimes \psi_E^0\rangle = \sum_k c_k\, |\psi_C^k \otimes \psi_E^k\rangle$

where $U_{CE}(t)$ unitary evolution of Ham, e.g., $e^{-iH_{CE}t}$ if time-indep

fix time t and obtain state on $\mathcal{H}_C$ by partial trace – mixed in gen

$$\mathrm{Tr}_E\,(|\Psi_{CE}\rangle\langle\Psi_{CE}|) = \sum_k |c_k|^2 |\psi_C^k\rangle\langle\psi_C^k|$$

Extend by linearity, to any (mixed) state – called quantum channel

$$\rho_C \mapsto \Phi(\rho_C) \equiv \mathrm{Tr}_E[U_{CE}(t)\,\rho_C \otimes \gamma_E\, U_{CE}^\dagger(t)]$$

Φ is CPT under assump system not entangled with env at start

fixed t means “snapshot” not full dynamics

Quantum channels and CPT maps

linear $\Phi : M_{d_A} \to M_{d_B}$ or, more gen $\Phi : \mathcal{A}_1 \to \mathcal{A}_2$ where $\mathcal{A}$ op alg

Def: Φ positivity-preservering (or "positive") $X \geq 0 \;\Rightarrow\; \Phi(X) \geq 0$

If Φ also trace-preserving (TP) maps then D.M. $\mapsto$ D.M.

Recall: density matrix (D.M.) ρ means $\rho \geq 0$ and $\operatorname{Tr} \rho = 1$

Not enough because mixed state D.M. have hidden part in $\mathcal{H}_E$.

Φ is k-positive if $(\Phi \otimes \mathcal{I}_k)(X)$ is pos-pres on $M_{d_A} \otimes M_{d_E}$

Φ is completely positive (CP) is k-pos $\forall\, k$ in finite dims equiv to

$X_{AE} \geq 0 \;\Rightarrow (\Phi \otimes \mathcal{I})(X_{AE}) \geq 0 \quad X_{AE} \in \mathcal{H}_A \otimes \mathcal{H}_A$, i.e. $d_E = d_A$

CPT map Φ called quantum channel – reps noise at "snapshot"

of time when system init unentangled with environment

$\widehat{\Phi}$ adjoint $\operatorname{Tr} A^* \Phi(B) = \langle A, \Phi(B) \rangle = \langle \widehat{\Phi}(A), B \rangle = \operatorname{Tr} [\widehat{\Phi}(A)]^* B$

TFAE for $\Phi : \mathcal{A}_1 \to \mathcal{A}_2$ or $\Phi : M_{d_A} \to M_{d_B}$

- Completely Positive (Op Alg) and Trace-Preserving (CPT)

- Stinespring: can find reps π of $\mathcal{A}_1$ such that

$$\widehat{\Phi}[X] = V^*\pi(X)V, \quad V^*V = I_A$$

- (ancilla) When $d_A = d_B$ can find $\mathcal{H}_E$, unitary U_{AE} s.t.

$$\Phi(R) = \text{Tr}_E \left[U_{AE}\, R \otimes |\phi\rangle\langle\phi|\, U_{AE}^*\right]$$

- Kraus (also Choi): can find F_k s.t.

$$\Phi(P) = \sum_k F_k P F_k^* \quad \text{with} \quad \sum_k F_k^* F_k = I$$

non-unique, but can get canonical via e-vecs of Choi matrix

- Choi: $(\Phi \otimes \mathcal{I})(|\beta\rangle\langle\beta|)$ matrix with blocks $\Phi(E_{jk})$ pos semi-def

where $E_{jk} = |e_j\rangle\langle e_k|$ has els $\delta_{ij}\delta_{\ell k}$. TP cond uses partial trace

Stinespring in finite dims

For $\Phi : M_{d_A} \mapsto M_{d_B}$ recall adjoint $\widehat{\Phi} : M_{d_B} \mapsto M_{d_A}$

can choose $\pi(R) = R \otimes I_E$ with $d_E \leq d_A d_B$.

$$V : \mathcal{H}_A \mapsto \mathcal{H}_B \otimes \mathcal{H}_E \qquad \widehat{\Phi}(R) = V^*(X \otimes I_E)V =$$

$$\begin{pmatrix} F_1^* & F_2^* \ldots F_{d_E}^* \end{pmatrix} \begin{pmatrix} X & 0 & \ldots & \ldots & 0 \\ 0 & X & \ldots & \ldots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & & \ldots & \ldots & X \end{pmatrix} \begin{pmatrix} F_1 \\ F_2 \\ \vdots \\ F_{d_E} \end{pmatrix} = \sum_k F_k^* \, X \, F_k$$

If Φ is also TP, $\widehat{\Phi}(I_B) = \sum_k F_k^* F_K = I_A$

CP rep of $\Phi : M_{d_A} \mapsto M_{d_B}$ recall adjoint $\widehat{\Phi} : M_{d_B} \mapsto M_{d_A}$

$$VRV^* = \begin{pmatrix} F_1 \\ \vdots \\ F_{d_E} \end{pmatrix} R \begin{pmatrix} F_1^* & \dots & F_{d_E}^* \end{pmatrix}$$

$$= \begin{pmatrix} F_1 R F_1^* & F_1 R F_2^* & \dots & F_1 R F_{d_E}^* \\ F_2 R F_1^* & F_2 R F_2^* & \dots & F_2 R F_{d_E}^* \\ \vdots & & \ddots & \vdots \\ F_{d_E} R F_1^* & F_{d_E} R F_2^* & F_{d_E} R F_{d_E}^* & \end{pmatrix}$$

$$\begin{aligned} \langle \widehat{\Phi}(X), R \rangle &= \mathrm{Tr}_A [\widehat{\Phi}(X)]^* R = \mathrm{Tr}_A \textstyle\sum_k (F_k^* X F_k) R \\ &= \mathrm{Tr}_B\, X \left(\textstyle\sum_k F_k R F_k^* \right) \\ \langle \widehat{\Phi}(X), R \rangle_A &= \langle X, \Phi(R) \rangle_B \end{aligned}$$

rue $\forall\ X \in M_{d_B}, R \in M_{d_A} \ \Rightarrow\ \Phi(R) = \sum_k F_k R F_k^*$

TP $\Rightarrow\ \sum_k F_k^* F_k$

Choi matrix

Given linear map $\Phi : M_{d_A} \mapsto M_{d_B}$ define Choi matrix using blocks

$$\begin{pmatrix} \Phi(|e_1\rangle\langle e_1| & \Phi(|e_1\rangle\langle e_2|) & \ldots & & \Phi(|e_1\rangle\langle e_{d_A}|) \\ \Phi(|e_2\rangle\langle e_1|) & \Phi(|e_2\rangle\langle e_2| & \Phi(|e_2\rangle\langle e_3|) & \ldots & \Phi(|e_2\rangle\langle e_{d_A}|) \\ \vdots & & & & \vdots \\ \vdots & & \ddots & & \vdots \\ \vdots & & & & \vdots \\ \Phi(|e_{d_A}\rangle\langle e_1|) & \ldots & & \ldots & \Phi(|e_{d_A}\rangle\langle e_{d_A}| \end{pmatrix}$$

When $d_A = d_B$ Choi matrix is $(\Phi \otimes \mathcal{I})(|\beta\rangle\langle\beta|)$ with

$|\beta\rangle = \frac{1}{\sqrt{d}} \sum_k |e_k \otimes e_k\rangle$ max entangled

NOT the same as matrix rep of Φ in product basis $|e_j \otimes e_j\rangle$

need to "reshuffle", e.g. els of block $\Phi(|e_1\rangle\langle e_1|)$ give first col

Choi matrix has blocks $\Phi(|e_j\rangle\langle e_k|)$

Thm: (Choi) Φ is CP if and only if Choi matrix is pos semi-def.

"stack" e-vecs $|g_k\rangle$ ($d_A d_B \times 1$) to get $d_B \times d_A$ matrix G_k

Then let $F_k \equiv \sqrt{\lambda_k} G_k$ find $\Phi(R) = \sum_k F_k A F_k^*$.

recover Kraus ops for Φ (use only non-zero e-vals)

$\sum_k F_k^* F_k = \widehat{\Phi}(I)$ Thus for CPT Φ $\sum_k F_k^* F_k = I_A$

For U unitary and $\widetilde{F}_j = \sum_k u_{jk} F_k$ find $\sum_j \widetilde{F}_j R \widetilde{F}_j^* = \Phi(R)$ also

Choi-rank is rank of Choi matrix = minimum number of Kraus ops.

Extreme points of CPT maps

Fix Y_A. The set of CP maps s.t. $\widehat{\Phi}(I_B) = Y_A$ is convex

Thm: (Choi) Φ is extreme iff $\{F_k^* F_j\}$ is lin indep in M_{d_A}.

Cor: If Φ is extreme then Choi rank $\leq d_A$

Thm: The closure $\overline{\mathcal{E}(d_A, d_B)}$ of the set of extreme points of CPT maps $\Phi : M_{d_A} \mapsto M_{d_B}$ is the set of such maps with Choi rank $\leq d_A$.

Often want $\overline{\mathcal{E}(d_A, d_B)}$ rather than just true ext pts.

Montonicity of $H(R, Q)$ under CPT maps

Lindblad proved using ancilla rep. and SSA; later indep Uhlmann

$\Phi(R) = \text{Tr}_E\, VRV^*$ $\quad V : \mathcal{H}_A \mapsto \mathcal{H}_B \otimes \mathcal{H}_E$ $\quad V^*V = I_A$

$$\begin{aligned} H\big[\Phi(R), \Phi(Q)\big] &= H\big[\text{Tr}_E\, VRV^*,\, \text{Tr}_E\, VQV^*\big] \\ &\leq H\big[\, VRV^*,\, VQV^*\big] = H(R, Q) \end{aligned}$$

Note: $S(\rho) - S[\Phi(\rho)] = S(\rho_{BE}) - S(\rho_B)$ concave

form of mutual information with $\rho_{BE} = V\rho V^*$

Aside: CPT map $\Phi^C : M_{d_A} \mapsto M_{d_E}$ $\quad \Phi^C(R) = \text{Tr}_B VRV^*$
called complementary channel – gives envrionment's view of noise

Measurement revisited

Remark – two ways in which quantum systems change

- unitary evolution – for which rep "snapshot" by CPT Φ
- measurement POVM $\mathcal{M} = \{G_m\}$ $G_m > 0, \; \sum_m G_m = I$

QC map for POVM using instrument with "pointer" $|m\rangle$ O.N.

$$\begin{aligned} \Omega_{\mathcal{M}} : \rho \;\; &\mapsto \;\; \sum_j (\operatorname{Tr} \rho G_m) |\phi_m\rangle\langle\phi_m| \otimes |m\rangle\langle m| \\ &= \;\; \sum_j (\operatorname{Tr} \rho G_m) |f_m\rangle\langle f_m| \end{aligned}$$

simplify to $|f_m >$ some O.N. set

POVM can be rep as v.N. meas. on larger space, similar to CP map

actually Naimark's Thm. QC = quantum-classical

Holevo bound on accessible information

To access info must measure $\Omega_{\mathcal{M}}(\rho) = \sum_m (\text{Tr}\, \rho G_m)\, |f_m\rangle\langle f_m|$.

Ensemble: $\{\pi_j, \rho_j\}$ with $\pi_j > 0$, $\sum_j \pi_j = 1$ and ρ_j dens. matrix

Holevo bound on accessible info or max info can get from ensemble

mutual info between ensemble and measurement outcome

$$S[\Omega_{\mathcal{M}}(\rho_{\rm av})] - \sum_j \pi_j S[\Omega_{\mathcal{M}}(\rho_j)] \;\leq\; S(\rho_{\rm av}) - \sum_j \pi_j S(\rho_j)$$

RHS $\quad S(\rho_{\rm av}) - \mathrm{Av}\big[S(\rho_j)\big] \equiv \chi(\mathcal{E}) \quad$ Holeov χ-quantity

$\rho_{\rm av} = \sum_j \pi_j \rho_j \qquad$ with $=$ if and only if all ρ_j commute

Cor: Access info $\leq\; S(\rho_{\rm av}) \leq \log d \quad = n$ if $d = 2^n$, i.e. n qubits

Holevo: direct proof in 1973 (same as SSA) without using SSA.

Will give 3 simple proofs of Holevo bound based on SSA

I. formal mutual info $\gamma_{AB} = \begin{pmatrix} \pi_1\rho_1 & 0 & \dots & \dots & 0 \\ 0 & \pi_2\rho_2 & \dots & \dots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & \dots & 0 & \pi_m\rho_m \end{pmatrix}$

Then $\gamma_A = \sum_j \pi_j\rho_j = \rho_{\rm av}$, and $\gamma_B = \sum_j |j\rangle\langle j|\, \pi_j$

$$\begin{aligned} H[\gamma_{AB}, \gamma_A \otimes \gamma_B] &= -S(\gamma_{AB}) + S(\gamma_A) + S(\gamma_B) \\ &= -\sum_j S(\pi_j\rho_j) + S(\rho_{\rm av}) + S[\pi_j] \\ &= -\sum_j \pi_j S(\rho_j) + \sum_j \pi_j \log \pi_j + S(\rho_{\rm av}) + S[\pi_j] \\ &= S(\rho_{\rm av}) - \sum_j \pi_j S(\rho_j) \end{aligned}$$

$$H[(\Omega_{\mathcal{M}} \otimes I)(\gamma_{AB}), (\Omega_{\mathcal{M}} \otimes I)(\gamma_A \otimes \gamma_B)] \leq H[\gamma_{AB}, \gamma_A \otimes \gamma_B]$$

LHS is mutual info between ensemble and POVM outcome

II. Yuen-Ozawa (1993): Rewrite

$$\begin{aligned} S(\rho_{\rm av}) - \sum_j \pi_j S(\rho_j) &= \sum_j \pi_j \operatorname{Tr} \rho_j \log \rho_j - \sum_j \operatorname{Tr} \pi_j \rho_j \log \rho_{\rm av} \\ &= \sum_j \pi_j \operatorname{Tr} \rho_j \big(\log \rho_j - \log \rho_{\rm av} \big) \\ &= \sum_j \pi_j H(\rho_j, \rho_{\rm av}) \end{aligned}$$

Then by monotonicity of relative entropy

$$\begin{aligned} S[\Omega_{\mathcal{M}}(\rho_{\rm av})] - \sum_j \pi_j S[\Omega_{\mathcal{M}}(\rho_j)] &= \sum_j \pi_j H\big[\Omega_{\mathcal{M}}(\rho_j), \Omega_{\mathcal{M}}(\rho_{\rm av})\big] \\ &\leq \sum_j \pi_j H(\rho_j, \rho_{\rm av}) \\ &= S(\rho_{\rm av}) - \sum_j \pi_j S(\rho_j) \end{aligned}$$

III. Lieb-Seiringer

observed: $\rho \mapsto S(\rho) - S[\Omega_{\mathcal{M}}(\rho)]$ concave means

$$S(\rho_{\rm av}) - S[\Omega_{\mathcal{M}}(\rho_{\rm av})] \quad \geq \quad \sum_j \pi_j \big(S(\rho_j) - S[\Omega_{\mathcal{M}}(\rho_j)] \big)$$

equiv to

$$S(\rho_{\rm av}) - \sum_j \pi_j S(\rho_j) \quad \geq \quad S[\Omega_{\mathcal{M}}(\rho_{\rm av})] - S[\Omega_{\mathcal{M}}(\rho_j)]$$

All three proofs exend to partial measurement

$$\Omega_{\mathcal{M}_B} : \gamma_{AB} \mapsto \sum_j |j\rangle\langle j| \, \mathrm{Tr}_B \, \gamma_{AB} \, I_A \otimes F_j \qquad \sum_j F_j = I_B$$

Carlen Lieb Inequalities

$$\widehat{\Upsilon}_{p,1}(K,A) = \tfrac{1}{(p-1)}\left[\mathrm{Tr}\, K^* A^p K)^{1/p} - \tfrac{1}{p}\mathrm{Tr}\, K^* A K\right]$$
$$= \inf\left\{J_p(K,A,X) + \tfrac{1}{p}\mathrm{Tr}\, X : X > 0\right\}$$

$$\widehat{\Phi}_{p,1}\Big(\sum_k A_k\Big) = \widehat{\Phi}_{p,1}(\mathcal{A}) = \tfrac{1}{(p-1)}\left[\mathrm{Tr}\,\Big(\sum_k A_k^p\Big)^{1/p} - \tfrac{1}{p}\mathrm{Tr}\,\sum_k A_k\right]$$

$$\widehat{\Psi}_{p,1}(\mathcal{A}_{12}) = \tfrac{1}{(p-1)}\left[\mathrm{Tr}_1\big(\mathrm{Tr}_2\,\mathcal{A}_{12}^p\big)^{1/p} - \tfrac{1}{p}\mathrm{Tr}_{12}\,\mathcal{A}_{12}\right]$$

All convex for $0 < p \le 2$. $\widehat{\Phi}(\mathcal{A})$ is block diag case of $\widehat{\Psi}(\mathcal{A}_{12})$

conditional entropy $\widehat{\Psi}_{(1,1)}(\mathcal{A}_{12}) = S(\mathcal{A}_1) - S(\mathcal{A}_{12})$

Aside on Motivation: mixed (p, q) norms

Try obvious $\Psi_{(p.q)}(A) = \left[\mathrm{Tr}_1 \left(\mathrm{Tr}_2\, A_{12}^p\right)^{q/p}\right]^{1/q}$, but not a norm

For $X = X^*$, Carlen-Lieb define

$$|||X|||_{(p,q)} \equiv \inf\{\Psi_{(p,q)}(A) + \Psi_{(p,q)}(A+X) : A, A+X \geq 0\}$$

extend to non s.a. via $\left|\left|\left| Y \begin{pmatrix} 0 & Y \\ Y^* & 0 \end{pmatrix} \right|\right|\right|_{(p,q)}$

norm is non-montone, i.e. $A \geq B \geq 0 \;\not\Rightarrow\; |||A|||_{(p,q)} \geq |||B|||_{(p,q)}$

compare Pisier: used powerful machinery of operator spaces and interpolation to define mixed (p, q) norms. But have nice duality, monotone props., triple Minkowski, etc.

theory has CB (completely bounded) norms, complete isometry, . . .

Koldan and King (arXiv:0904.1710) gave some bounds and showed norms are inequivalent in one case.

$$|\mathbb{1}\rangle = (1, 1, \ldots, 1) \qquad |e_1\rangle = (1, 0, \ldots, 0)$$

$$\mathcal{K} = \tfrac{1}{d} I \otimes |\mathbb{1}\rangle\langle e_1| = \begin{pmatrix} I & 0 & \ldots & 0 \\ I & 0 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ I & 0 & \ldots & 0 \end{pmatrix}$$

$$\mathcal{A} = \sum_k A_k \otimes |e_k\rangle\langle e_k| = \begin{pmatrix} A_1 & 0 & 0 & \ldots & 0 \\ 0 & A_2 & 0 & \ldots & 0 \\ 0 & 0 & A_1 & \ldots & 0 \\ \vdots & & & \ddots & \vdots \end{pmatrix}$$

$$\mathcal{K}^* \mathcal{A}^p \mathcal{K} = \big(\textstyle\sum_k A_k^p \big) \otimes |e_1\rangle\langle e_1| \quad \Rightarrow \quad \widehat{\Phi}_{p,1}(\mathcal{A}) = \widehat{\Upsilon}_{p,1}(K, A)$$

MPT and equality conds in Carlen-Lieb

monotonicity: $\quad 0 < p < 2 \qquad \widehat{\Psi}_{(p,1)}(\mathcal{A}_{23}) \;\leq\; \widehat{\Psi}_{(p,1)}(\mathcal{A}_{123})$

conditional entropy: $\;\widehat{\Psi}_{(1,1)}(\mathcal{A}_{23}) \;=\; S(\mathcal{A}_2) - S(\mathcal{A}_{23})$

$p = 1 \qquad S(\mathcal{A}_2) - S(\mathcal{A}_{23}) \leq S(\mathcal{A}_{12}) - S(\mathcal{A}_{123})$ SSA

Carlen-Lieb Minkowski:

$$\begin{aligned} \mathrm{Tr}_3\left[\mathrm{Tr}_2(\mathrm{Tr}_1\,\mathcal{A}_{123})^p\right]^{1/p} &= \Psi_{(p,1)}(\mathcal{A}_{32}) \\ &\leq \Psi_{(p,1)}(\mathcal{A}_{132}) \;=\; \mathrm{Tr}_3\,\mathrm{Tr}_1\left(\mathrm{Tr}_2\mathcal{A}_{123}^p\right)^{1/p} \end{aligned}$$

for $1 < p \leq 2$ and reverse ineq $\geq$ for $0 < p < 1$.

$p = 1$ equality conditions same as for SSA – extend to all $p \in (0,2)$

Generalizations of SSA

Uniform treatment led to two distinct generalizations of SSA

pseudo p-metric based on MPT of $J_p(1, A_{123}, A_{23})$

$$\operatorname{Tr} A_{23}^p A_2^{1-p} \quad \begin{matrix} \leq \\ \geq \end{matrix} \quad \operatorname{Tr} A_{123}^p A_{12}^{1-p} \qquad \begin{matrix} p \in (0,1) \\ p \in (1,2) \end{matrix}$$

pseudo p-norm based on MPT of $\widehat{\Psi}(\mathcal{A}_{123})$

$$\operatorname{Tr}_2(\operatorname{Tr}_3 \mathcal{A}_{23}^p)^{1/p} \quad \begin{matrix} \geq \\ \leq \end{matrix} \quad \operatorname{Tr}_{12}(\operatorname{Tr}_3 \mathcal{A}_{123}^p)^{1/p} \qquad \begin{matrix} p \in (0,1) \\ p \in (1,2) \end{matrix}$$

Compare Renyi $\frac{1}{1-p} \log \operatorname{Tr} A^p$ and Tsallis $\frac{1}{p-1}(1 - \operatorname{Tr} A^p)$ entropy

Aside on (Alfred) Horn's Lemma

Def: Assume sequence $\{a_k\}, \{b_k\}$ of length d in non-increasing order. Then a_k majorizes b_k, written $a_k \succ b_k$ means

$$a_1 \geq b_1, \quad \sum_{k=1}^{m} a_k \geq \sum_{k=1}^{m} b_k \ (m = 1, \ldots d-1) \qquad \sum_{k=1}^{n} a_k = \sum_{k=1}^{n} b_k$$

Plays important role in quantum info, e.g., entanglement tests

Horn's Lemma: Given positive sequences $\{\lambda_k\}, \{d_k\}$ of length n, there exists a positive semi-definite $n \times n$ matrix A with e-vals λ_k and diagonal elements d_k if and only if $\lambda_k \succ d_k$.

Remark: Any sequence with $\lambda_k \geq 0$ and $\sum_{k=1}^{n} \lambda_k = \sum_{k=1}^{n} d_k$ majorizes $d_k = \frac{1}{n}$

For simple inductive proof see Carlen & .Lieb arXiv:0904.0734

Corollary to A. Horn's Lemma

Cor: Let A be a $n \times n$ pos semi-def matrix with $\mathrm{Tr}\, A = 1$. Then $\exists$ n normalized (not nec orthog) vectors $\mathbf{x}_m$ s. t. $A = \sum_{m=1}^{n} \frac{1}{n} \mathbf{x}_m \mathbf{x}_m^*$

Pf: By Horn $\exists$ $B = B^*$ s.t. $A = UB^2U^*$ and $(B^2)_{kk} = \frac{1}{d}$ $\forall k$

$U = \sum_k \mathbf{u}_k \mathbf{e}_k^* = \sum_k |\mathbf{u}_k\rangle\langle\mathbf{e}_k|$ unitary with els u_{jk} and cols $\mathbf{u}_k$.

Let $\mathbf{x}_m = \sqrt{d} \sum_j \mathbf{u}_j b_{jm}$. Then

$$\begin{aligned} A &= \sum_{jk} |\mathbf{u}_j\rangle\langle\mathbf{e}_j|\, B^2\, |\mathbf{e}_k\rangle\langle\mathbf{u}_k| = \sum_{jk} \mathbf{u}_j \langle \mathbf{e}_j, B^2 \mathbf{e}_k\rangle \mathbf{u}_k^* \\ &= \sum_{jk} \sum_m \mathbf{u}_j \langle \mathbf{e}_j, B\, \mathbf{e}_m\rangle \langle \mathbf{e}_m, B\, \mathbf{e}_k\rangle \mathbf{u}_k^* = \sum_m \tfrac{1}{d} \mathbf{x}_m \mathbf{x}_m^* \end{aligned}$$

$$\begin{aligned} \|\mathbf{x}_m\|^2 &= d \sum_{jk} \mathbf{u}_j^* \overline{b}_{jm} b_{km} \mathbf{u}_k = d \sum_{jk} \overline{b}_{jm} b_{km} \mathbf{u}_j^* \mathbf{u}_k \\ &= \sum_{jk} \delta_{jk} \overline{b}_{jm} b_{km} = d(B^2)_{mm} = d \tfrac{1}{d} = 1. \qquad \square \end{aligned}$$

Conjectured generalization of Horn's Lemma to CPT maps

Conj 1: Let $\Phi : M_{d_1} \mapsto M_{d_2}$ be a CPT map. Then $\exists$ d_2 CPT maps Φ_m with Choi rank $\leq d_1$ such that $\Phi = \sum_{m=1}^{d_2} \frac{1}{d_2} \Phi_m$.

Conj 2: Let $\Phi : M_{d_2} \mapsto M_{d_1}$ be a CP map with $\widehat{\Phi}(I_2) = I_1$. Then $\exists$ d_2 unital CP maps $\widehat{\Phi}_m$ with Choi rank $\leq d_1$ s.t. $\widehat{\Phi} = \sum_{m=1}^{d_2} \frac{1}{d_2} \widehat{\Phi}_m$

Conjectures of K.M.R. Audenaert and M.B. Ruskai
strongly supported by numerical work of Audenaert

Can prove for $d_1 = 1$ or $d_2 = 2$ using block matix version.

Using only true extreme points may need $d_A d_B$ maps in general.

Block Matrix forms of Audenaert-Ruskai conjecture

Conj 3: Let $\mathbf{A}$ be a $d_1 d_2 \times d_1 d_2$ pos semi-def. matrix with $d_2 \times d_2$ blocks A_{jk} each $d_1 \times d_1$, with $\sum_j A_{jj} = M$. $\exists$ d_2 block matrices $\mathbf{B}_m$, each of rank $\leq d_1$, s.t. $\sum_j B_{jj} = M$, and $\mathbf{A} = \sum_{m=1}^{d_2} \frac{1}{d_2} \mathbf{B}_m$.

Restate using vectors of matrices $\mathbf{X}_m^* = \begin{pmatrix} X_{1m}^* & X_{2m}^* & \ldots & X_{d_2 m}^* \end{pmatrix}$ with each block X_{jm} $d_1 \times d_1$.

Conj 4: Let $\mathbf{A}$ be a $d_1 d_2 \times d_1 d_2$ pos semi-def. matrix with $d_2 \times d_2$ blocks A_{jk} each $d_1 \times d_1$, with $\sum_j A_{jj} = M$. Then $\exists$ d_2 vectors $\mathbf{X}_m$ composed of d_2 blocks X_{jm} of size $d_1 \times d_1$ such that

$$\mathbf{A} = \sum_{m=1}^{d_2} \frac{1}{d_2} \mathbf{X}_m \mathbf{X}_m^*, \qquad \text{and} \qquad \sum_k X_{km} X_{km}^* = M \qquad \forall\ m$$

References

A. Wehrl "General Properties of Entropy"
Rev. Mod. Phys. **50**, 221–260 (1978).

M.B. Ruskai, "Inequalities for Quantum Entropy: A Review with Conditions for Equality" *J. Math. Phys.* **43**, 4358–4375 (2002); erratum 46, 019901 (2005). (quant-ph/0205064).

M.B. Ruskai, "Lieb's simple proof of concavity of $\text{Tr}\, A^p K^* B^{1-p} K$ and remarks on related inequalities"
Int. Jour. Quant. Info. **3**, 570–590 (2005). quant-ph/0404126

M.B. Ruskai, "Another short and elementary proof of strong subadditivity of quantum entropy"
Reports on Math. Physics **60**, 1–12 (2007). quant-ph/0604206

A. Jencova and M.B. Ruskai, arXiv:0903.2895
"A Unified Treatment of Convexity of Relative Entropy and Related Trace Functions, with Conditions for Equality"

M.B. Ruskai, arXiv:0708.1902
"Some Open Problems in Quantum Information Theory"

Eric A. Carlen, Elliott H. Lieb
"A Minkowski Type Trace Inequality and Strong Subadditivity of Quantum Entropy II: Convexity and Concavity"
Lett. Math. Phys. **83**, 107–126 (2008). arXiv:0710.4167

E. Carlen "Trace Inequalities and Quantum Entropy" lecture notes available at http://www.ueltschi.org/AZschool/program/

T. Ando, *Topics on Operator Inequalities* Sapporo Lecture Notes (1978).

T. Ando, "Concavity of certain maps on pos. definite matrices and appl. to Hadamard prods" *Lin, Alg. Appl.* **26**, 203-241 (1979)

A. Lesniewski and M.B. Ruskai, "Relative Entropy and Monotone Riemannian Metrics on Non-Commutative Probability Space"
J. Math. Phys. **40**, 5702-5724 (1999).