

Quantum Shannon Theory

Aram Harrow (MIT)

QIP 2016 tutorial 9-10 January, 2016

the prehistory of quantum information

ideas present in disconnected form

- 1927 Heisenberg uncertainty principle
- 1935 EPR paper / 1964 Bell's theorem
- 1932 von Neumann entropy subadditivity (Araki-Lieb 1970) strong subadditivity (Lieb-Ruskai 1973)
- measurement theory (Helstrom, Holevo, Uhlmann, etc., 1970s)

relativity: a close relative

- Before Einstein, Maxwell's equations were known to be incompatible with Galilean relativity.
- Lorentz proposed a mathematical fix, but without the right physical interpretation.
- Einstein's solution redefined space/time, mass/momentum/ energy, etc.
- Space and time had solid mathematical foundations (Descartes, etc.), unlike information and computing.

theory of information and computing

- 1948 Shannon created modern information theory (and to some extent cryptography) and justified entropy as a measure of information independent of physics. units of bits.
- Turing, Church, von Neumann, ..., Djikstra described a theory of computation, algorithms, complexity, etc.
- This made it possible to formulate questions such as: how do "quantum effects" change the capacity?
 (→ Holevo bound)

what is the thermodynamic cost of computing? (Landauer principle, Bennett reversible computing)

what is the computational complexity of simulating QM? $(\rightarrow \text{DMRG/QMC}, \text{ and also Feynman})$

some wacky ideas



Feynman '82: "Simulating Physics with Computers"

- Classical computers require exponential overhead to simulate quantum mechanics.
- But quantum systems obviously don't need exp overhead to simulate *themselves*.
- Therefore they are doing something more computationally powerful than our existing computers.
- (Implicitly requires the idea of a universal Turing machine, and the strong Church-Turing thesis.)

Wiesner '70: "Conjugate Coding"

- The uncertainty principle restricts possible measurements.
- In experiments, this is a disadvantage, but in crypto, limiting information is an advantage.
- (Requires crypto framework, notion of "adversary.")
- Paper initially rejected by IEEE Trans. Inf. Th. ca. 1970

towards modern QIT

- Deutsch, Jozsa, Bernstein, Vazirani, Simon, etc. impractical speedups required oracle model, precursors to Shor's algorithm, following Feynman.
- quantum key distribution (BB84, B90, E91) following Weisner.
- ca. 1995
 - Shor and Grover algorithms
 - quantum error-correcting codes
 - fault-tolerant quantum computing
 - teleportation, super-dense coding
 - Schumacher–Jozsa data compression
 - HSW coding theorem
 - resource theory of entanglement

modern QIT

semiclassical

- compression: $S(\rho) = -tr [\rho log(\rho)]$
- CQ or QC channels: $\chi(\{p_x, \rho_x\}) = S(\Sigma_x p_x \rho_x) \Sigma_x p_x S(\rho_x)$
- hypothesis testing: $D(\rho || \sigma) = tr[\rho(log(\rho) log(\sigma)]$

"fully quantum"

- complementary channel: $N(\rho) = tr_2 V \rho V^+$, $N^c(\rho) := tr_1 V \rho V^+$
- quantum capacity: $Q^{(1)}(N) = \max_{\rho} [S(N(\rho)) S(N^{c}(\rho))]$ $Q(N) = \lim_{n \to \infty} Q^{(1)}(N^{\otimes n})/n$
- tools: purifications (Stinespring), decoupling

recent

- one-shot: $S_{\alpha}(\rho) := \log(tr \rho^{\alpha})/(1-\alpha)$
- applications to optimization, condensed matter, stat mech.

Relevant talks

- Wed 9. Omar Fawzi and Renato Renner. Quantum conditional mutual information and approximate Markov chains.
- Wed 9:50. Omar Fawzi, Marius Junge, Renato Renner, David Sutter, Mark Wilde and Andreas Winter. Universal recoverability in quantum information theory.
- Thurs 11. David Sutter, Volkher Scholz, Andreas Winter and Renato Renner. Approximate degradable quantum channels
- Thurs 4:15. Mario Berta, Joseph M. Renes, Marco Tomamichel, Mark Wilde and Andreas Winter.
 Strong Converse and Finite Resource Tradeoffs for Quantum Channels.

semi-relevant talks

- Tues 11:50. Ryan O'Donnell and John Wright. Efficient quantum tomography merged with Jeongwan Haah, Aram Harrow, Zhengfeng Ji, Xiaodi Wu and Nengkun Yu. Sample-optimal tomography of quantum states
- Tues 3:35. Ke Li. Discriminating quantum states: the multiple Chernoff distance
- Thurs 10. Mark Braverman, Ankit Garg, Young Kun Ko, Jieming Mao and Dave Touchette. Near optimal bounds on bounded-round quantum communication complexity of disjointness
- Thurs 3:35. Fernando Brandao and Aram Harrow. Estimating operator norms using covering nets with applications to quantum information theory
- Thurs 4:15. Michael Beverland, Gorjan Alagic, Jeongwan Haah, Gretchen Campbell, Ana Maria Rey and Alexey Gorshkov. Implementing a quantum algorithm for spectrum estimation with alkaline earth atoms.

outline

- metrics
- compressing quantum ensembles (Schumacher coding)
- sending classical messages over q channels (HSW)
- remote state preparation (RSP)
- Schur duality
- RSP and the strong converse
- hypothesis testing
- merging
- quantum conditional mutual information and q Markov states

metrics

Trace distance $T(\rho, \sigma) := \frac{1}{2} \| \rho - \sigma \|_{1}$

- Is a metric.
- monotone: $T(\rho, \sigma) \ge T(N(\rho), N(\sigma))$
- and this is achieved by a measurement
 → T = max m'mt bias

Fidelity
$$F(
ho,\sigma):=\|\sqrt{
ho}\sqrt{\sigma}\|_1=\mathrm{tr}\,\sqrt{\sqrt{\sigma}
ho}\sqrt{\sigma}$$

- F=1 iff $\rho = \sigma$ and F=0 iff $\rho \perp \sigma$
- monotone $F(\rho, \sigma) \leq F(N(\rho), N(\sigma))$
- and this is achieved by a measurement!

Relation: 1-F ≤ T ≤ (1-F²)^{1/2} Pure states with angle θ : F = cos(θ) and T = sin(θ). (exercise: which m'mts saturate?)

the case for fidelity

Uhlmann's theorem:

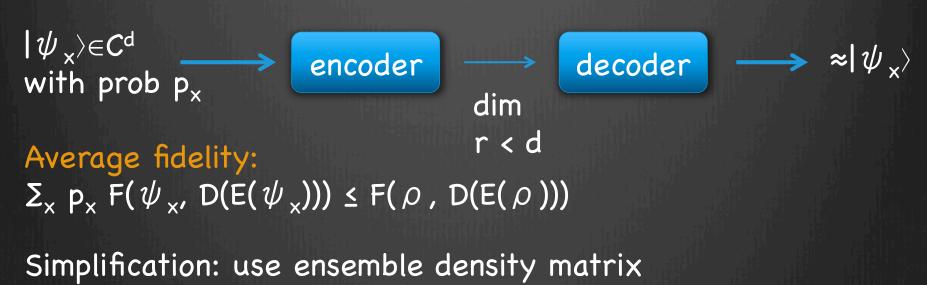
 $F(\rho_{A}, \sigma_{A}) = \max_{\psi, \phi} F(\psi_{AB}, \phi_{AB}) \text{ s.t.}$ $\psi = |\psi\rangle \langle \psi|, \phi = |\phi\rangle \langle \phi|, \psi_{A} = \rho_{A}, \phi_{A} = \sigma_{A}.$ Note:

≥ from monotonicity.
 = requires sweat

Church of the Larger Hilbert Space

- 2. Can fix either ψ or ϕ and max over the other.
- 3. $F(\psi, \phi) = |\langle \psi | \phi \rangle|$. (Some use different convention.)
- 4. Implies that $(1-F)^{1/2}$ is a metric.
- Also F is multiplicative.

Compression



 $\rho = \Sigma_x p_x \psi_x$ with eigenvalues $\lambda_1 \ge \lambda_2 \ge \dots \ge \lambda_d \ge 0$

 $\begin{aligned} \operatorname{rank}(\sigma) = r \Rightarrow F(\rho, \sigma)^2 \leq \operatorname{tr} [P_r \rho] &= \lambda_1 + \dots + \lambda_r \\ P_r \text{ projects onto top } r \text{ eigenvectors} \\ \sigma &= \frac{P_r \rho P_r}{\operatorname{tr} [P_r \rho]} \end{aligned}$

Suggests optimal fidelity = $(\lambda_1 + ... + \lambda_r)^{1/2}$.

Too good to be true!

Ensemble density matrix: $\rho = \Sigma_x p_x \psi_x$ Yes compression depends only on ρ . But reproducing ρ is not enough! consider:

E(·)=|0><0| D(·)= *ρ*

Gets the average right but not the correlations.

Reference system

Average fidelity: $\Sigma_{x} p_{x} F(\psi_{x}, E(D(\psi_{x}))))$ $= F(\Sigma_{x} p_{x} |x\rangle\langle x| \otimes \psi_{x}, \Sigma_{x} p_{x} |x\rangle\langle x| \otimes E(D(\psi_{x})))$

Not so easy to analyze. Instead follow the Church of the Larger Hilbert Space.

$$\left|\varphi\right\rangle_{RQ} := \sum \sqrt{p_x} \left|x\right\rangle_R \left|\psi_x\right\rangle_Q$$

 \boldsymbol{x}

Avg fidelity $\geq F(\varphi, (id_R \otimes D \circ E_Q)(\varphi))$ (pf: monotonicity under map that measures R.)

Protocol: $E(\omega) = P_r \ \omega \ P_r$. D = id. achieves $F = \langle \varphi | (I \otimes P_r) | \varphi \rangle = tr [\rho P_r] = \lambda_1 + ... + \lambda_r$



Optimality

Complication: E, D might be noisy.

Solution: purify!

1. Write $D(E(\omega)) = tr_G V \omega V^+$ where V is an isometry from Q -> Q \otimes G.

2. Uhlmann \rightarrow F(φ , tr_G V φ V⁺) = $|\langle \varphi |_{RQ} \langle 0 |_{G} V | \varphi \rangle_{RQ}|$

3. a little linear algebra →
F ≤ tr[ρP] for P rank-r and ||P||≤1
≤ λ₁ + ... + λ_r

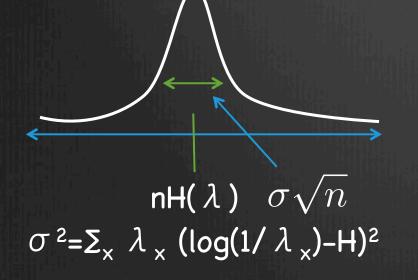


compressing i.i.d. sources

Quantum story ≈ classical story

 $\rho^{\otimes n}$ has eigenvalues $\lambda_{x_1} \lambda_{x_2} \cdots \lambda_{x_n}$ for X=(x₁,...,x_n) \in [d]ⁿ. Typically this is $\approx \lambda_1^{n\lambda_1} \cdots \lambda_d^{n\lambda_d} = \exp(-nH(\lambda))$ H(λ) = - $\Sigma_x \lambda_x \log(\lambda_x)$ = S(ρ) = -tr[$\rho \log(\rho)$]

distribution of $-\log(\lambda_{x_1} \lambda_{x_2} \cdots \lambda_{x_n})$



qubits	fidelity	
nH(λ) + 2 σ n ^{1/2}	0.98	
nH(λ) – 2 σ n ^{1/2}	0.02	
n(H(λ)+ δ)	1-exp(-n $\delta^2/2\sigma^2$)	
n(H(λ)– δ)	exp(-n δ ²/2 σ ²)	

typicality

Definitions:

An eigenvector of $\rho^{\otimes n}$ is k-typical if its eigenvalue is in the range exp(-nS(ρ) ± k σ n^{1/2}).

Typical subspace V = span of typical eigenvectors Typical projector P = projector onto V

Structure theorem for iid states: "asymptotic equipartition"

- tr [P ρ^{⊗n}] ≥ 1 − k⁻²
- $\exp(-nS(\rho) k\sigma n^{1/2}) P \leq P\rho^{\otimes n} P \leq \exp(-nS(\rho) + k\sigma n^{1/2}) P$
- likewise tr[P] $\approx \exp(nS(\rho) + k\sigma n^{1/2})$

Almost flat spectrum. Plausible because of permutation symmetry.

Quantum Shannon Theory

Aram Harrow (MIT)

QIP 2016 tutorial day 2 10 January, 2016

entropy S(A|B) $S(\rho) = -tr [\rho \log \rho]$ range: $0 \leq S(\rho) \leq \log(d)$ • B A symmetry: $S(\rho) = S(U \rho U^{\dagger})$ • multiplicative: $S(\rho \otimes \sigma) = S(\rho) + S(\sigma)$ • • continuity (Fannes-Audenaert): $|S(\rho) - S(\sigma)| \leq \varepsilon \log(d) + H(\varepsilon, 1 - \varepsilon)$ I(A:B) $\varepsilon := || \rho - \sigma ||_1 / 2$ B A multipartite systems: ρ_{AB} $S(A) = S(\rho_A), S(B) = S(\rho_B), etc.$ conditional entropy: S(A|B) := S(AB) - S(B), can be < 0 • mutual information: I(A:B) = S(A) + S(B) - S(AB)= $S(A) - S(A|B) = S(B) - S(B|A) \ge 0$ "subadditivity"

 $\begin{array}{c} CQ \ channel \ coding\\ CQ = Classical \ input, \ Quantum \ output\\ |x > \langle x| \longrightarrow N \ & \rho_x = N(|x > \langle x|) \end{array}$

Given n uses of N, how many bits can we send?

Allow error that $\rightarrow 0$ as $n \rightarrow \infty$.

HSW theorem: Capacity = max χ $\chi (\{p_x, \rho_x\}) = S(\Sigma_x p_x \rho_x) - \Sigma_x p_x S(\rho_x)$



HSW coding

 $\rho = \Sigma_{x} p_{x} \rho_{x}$ $\chi = S(\rho) - \Sigma_{x} p_{x} S(\rho_{x})$ = S(Q) - S(Q|X)

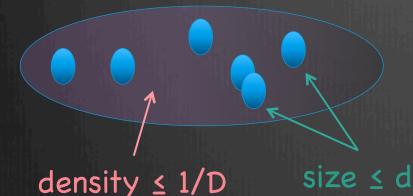
total information ambiguity in each message

typical subspace of $\rho \otimes^n$ has dim $\approx \exp(n S(\rho))$ If $x=(x_1,...,x_n)$ is p-typical then $\rho_{x_1} \otimes \rho_{x_2} \otimes ... \otimes \rho_{x_n}$ has typical subspace of dim $\approx \exp(n\Sigma_x p_x S(\rho_x))$

> "Packing lemma" Can fit $\approx \exp(n \chi)$ messages.

Packing lemma

Classically: random coding and maximum-likelihood decoding Quantumly: messages do not commute with each other



Suppose $\sigma = \sum_{x} p_{x} \sigma_{x}$ and there exist Π , { Π_{x} } s.t. 1. $tr[\Pi \sigma_{x}] \ge 1 - \varepsilon$ 2. $tr[\Pi_{x}\sigma_{x}] \ge 1 - \varepsilon$ 3. $tr[\Pi_{x}] \le d$ 4. $\Pi \sigma \Pi \le \Pi / D$

Packing lemma: We can send M messages with error $O(\varepsilon^{1/2} + Md/D)$

For HSW: $\sigma = \rho^{\otimes n}$ with typ proj Π . $\sigma_{x} = \rho_{x_{1}}^{\otimes} \dots \otimes \rho_{x_{n}}^{\otimes}$ with typ proj Π_{x} . $d \approx \exp(n S(Q|X))$.

Upper bound

 $\rho_{\mathrm{X}} = \rho_{\mathrm{X}_{1}} \otimes \ldots \otimes \rho_{\mathrm{X}_{n}}$

Q

proof: $n \chi_{z} \ge I(X;Q) \ge I(X;Y) \ge (1-O(\varepsilon)) \log(M)$

additivity Wed 10:50 Cross-Li-Smith. also Shannon 1948

D

Q

N⊗n

 $X \in \{X^1, \dots, X^M\}$

data-processing inequality

 $Q \longrightarrow \bigvee_{D} \longrightarrow Y$ $V_{D} \longrightarrow Q'$ isometry

I(X:Q) = I(X:YQ') ≥ I(X:Y)

 $\Sigma_{\rm Y} D_{\rm Y} = I$

 $\Pr[Y|X] = tr[\rho_X D_Y]$

continuity

D

conditional mutual information Claim that $I(A:BC) - I(A:B) \ge 0$. =: I(A:C|B) conditional mutual information = S(A|B) + S(C|B) - S(AC|B)= S(AB) + S(BC) - S(ABC) - S(B)

CMI C

A

If B is classical, $\rho = \Sigma_b p(b) |b\rangle\langle b| \otimes \sigma(b)_{AC}$ then I(A:C|B) = $\Sigma_b p(b) I(A:C)_{\sigma(b)} \ge 0$ from subadditivity

I(A:C|B) ≥ 0 is strong subadditivity [Lieb-Ruskai '73].

I(A:C|B) = 0 for "quantum Markov states" Wed morning you will hear I(A:C|B) ≥ "non-Markovianity"

capacity of QQ channels

Additional degree of freedom: channel inputs $|\psi_{x}\rangle$.

$$C^{(1)}(N) = \max_{\{p_x, \psi_x\}} \chi(\{p_x, \psi_x\})$$

NP-hard optimization problem [Beigi-Shor, H.-Montanaro]

Worse: $C(N) = \lim_{n \to \infty} C^{(1)}(N^{\otimes n})/n$. and \exists channels where $C(N) > C^{(1)}(N)$.

Open questions: Non-trivial upper bounds on capacity. Strong converse ($p_{succ} \rightarrow 0$ when sending n(C+ δ) bits.) (see Berta et al, Thurs 4:15pm).

quantum capacity

How many qubits can be sent through a noisy channel?

R

Q⁽¹⁾(N) := max S(B) - S(E) = max S(B) - S(RB) = max -S(R|B) "coherent information"

→ B

N

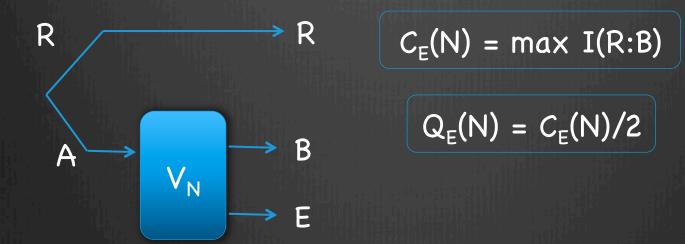
A

isometry $Q(N) = \lim_{n \to \infty} Q^{(1)}(N^{\otimes n})/n$ not known when > 0. sometimes $Q^{(1)}(N) = 0 < Q(N)$.

▶ R

entanglement-assisted capacity

Alice and Bob share unlimited free EPR pairs.

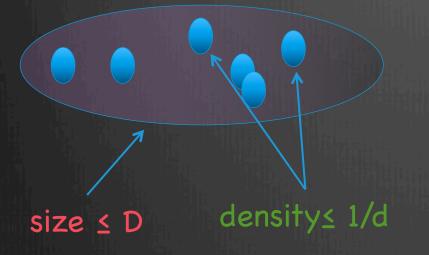


Bennett Shor Smolin Thapliyal q-ph/0106052

additive
 concave in input

 \rightarrow efficiently computable

covering lemma

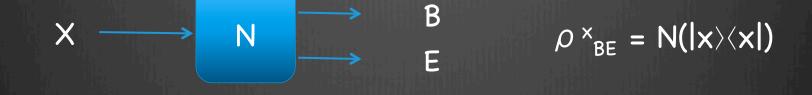


Suppose $\sigma = \sum_{x} p_{x} \sigma_{x}$ and there exist Π , { Π_{x} } s.t. 1. $tr[\Pi \sigma_{x}] \ge 1 - \varepsilon$ 2. $tr[\Pi_{x}\sigma_{x}] \ge 1 - \varepsilon$ 3. $tr[\Pi] \le D$ 4. $\Pi_{x}\sigma_{x}\Pi_{x} \le \Pi_{x} / d$

Covering lemma:

If $x_1, ..., x_M$ are sampled randomly from p and M >> (D/d) log(D)/ ε^3 then with high probability $\sigma \approx_{O(\epsilon^{1/4})} \frac{\sigma_{x_1} + \cdots + \sigma_{x_M}}{M}$

wiretap (CQQ) channel



<u>Thm</u>: Alice can send secret bits to Bob at rate I(X:B) - I(X:E).

Proof: packing lemma -> coding ≈nI(X:B) bits for Bob covering lemma -> sacrifice ≈nI(X:E) bits to decouple Eve

remote state preparation (RSP)

Q: Cost to transmit n qubits?

A: 2n cbits, n ebits using teleportation.

Cost is optimal given super-dense coding and entanglement distribution.

visible coding: What if the sender knows the state?

We want to simulate the map: " $\psi'' \rightarrow |\psi\rangle$. Requires $\geq n$ cbits, but above optimal arguments break.

RSP via covering

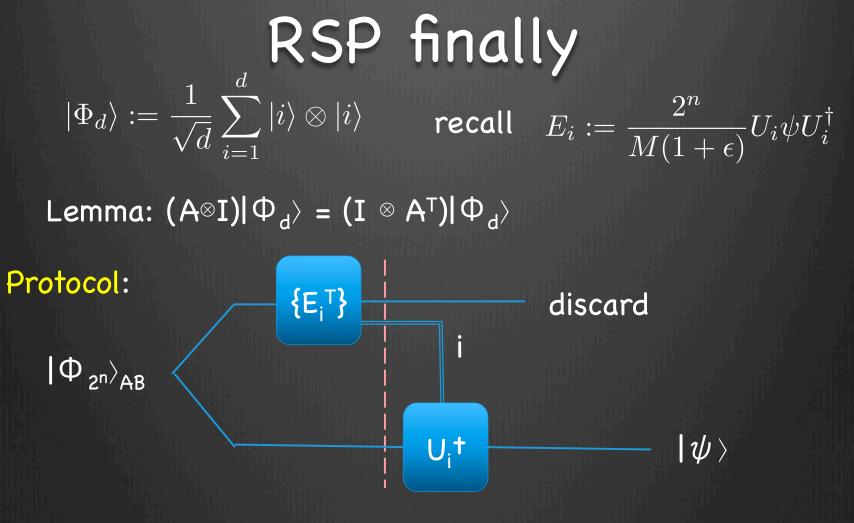
Consider the ensemble $\{U\psi U^{\dagger}\}\$ for random U. Average state is $I/2^{n}$.

Covering-type arguments [Aubrun arXiv:0805.2900] \rightarrow If we choose U₁, ..., U_M randomly with M >> 2ⁿ / ε ² then

with high probability,
$$orall \psi = \left\|rac{1}{M}\sum_{i=1}^M U_i\psi U_i^\dagger - rac{I}{2^n}
ight\| \leq rac{\epsilon}{2^n}$$

Set $E_i := \frac{2^n}{M(1+\epsilon)} U_i \psi U_i^{\dagger}$ Then (1- ε)I $\leq \Sigma_i E_i \leq I$

So $\{E_i\}$ is \approx a valid measurement. So what?



 $\underline{\mathbf{x}} \mathbf{E}_{i}^{\mathsf{T}} \otimes \mathbf{E}_{i} \mathbf{x} (\mathbf{U}_{i} \psi \mathbf{U}_{i}^{\dagger})^{\mathsf{T}} \otimes (\mathbf{U}_{i} \psi \mathbf{U}_{i}^{\dagger})$

 $cost \approx n cbits + n ebits.$

RSP of ensembles

can simulate x -> ρ_x with cost χ

 $\approx n \chi$ cbits + some ebits $\geq N^{\otimes n} \geq \approx n \chi$ cbits

Lemma: Converting $n(C-\delta)$ cbits $+\infty$ ebits into nC cbits will have success probability $\leq \exp(-n\delta)$.

implies strong converse: sending $n(\chi + \delta)$ bits through $N^{\otimes n}$ has $exp(-n\delta')$ success prob

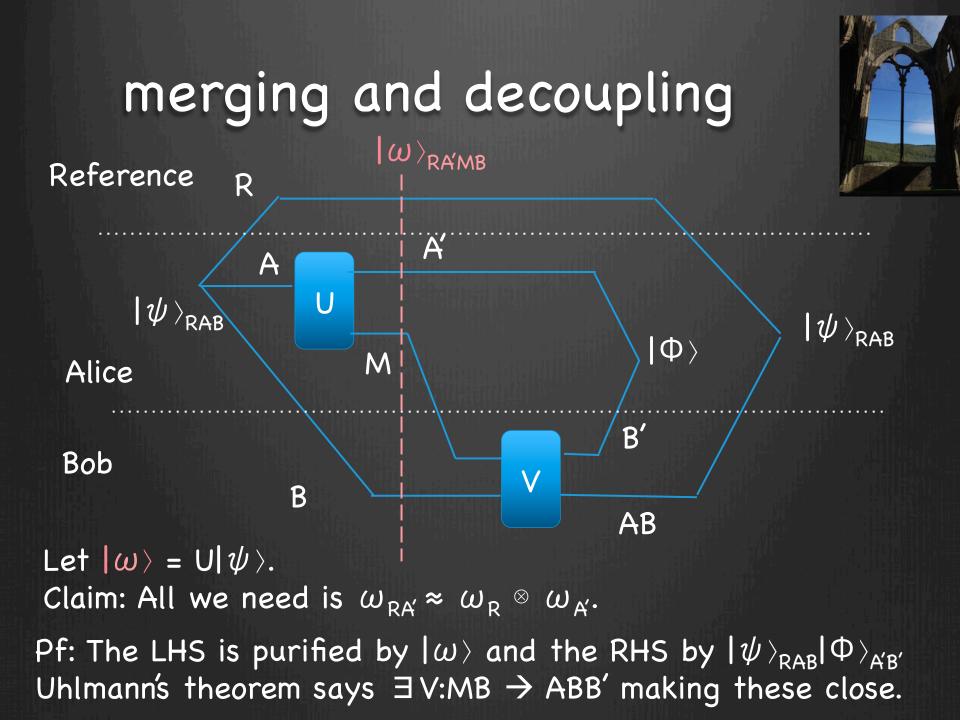
simulation and strong converses

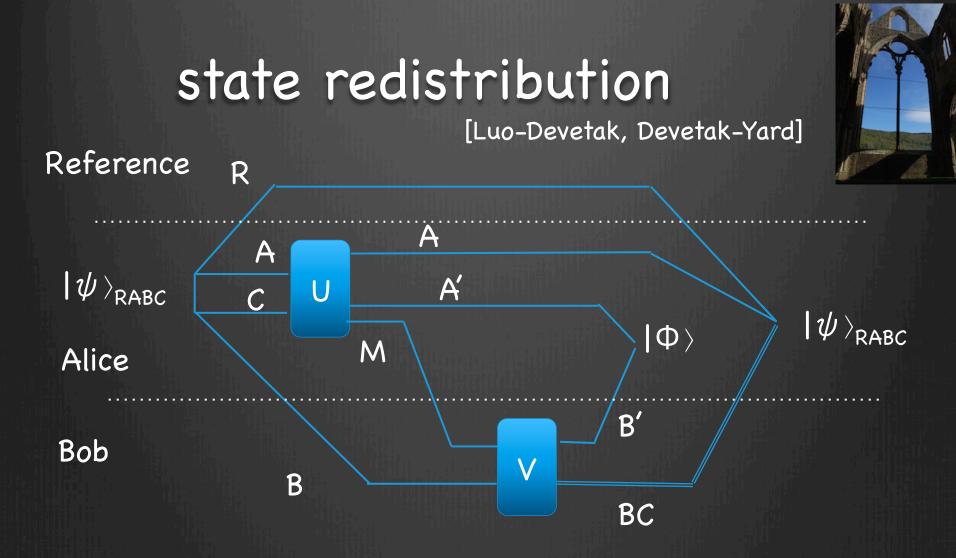
Let N be a general q channel.

Type of simulation	cbit simulation cost	also needs
visible product input	X	EPR
visible arbitrary input	R	EPR
arbitrary quantum input	C _E	embezzling

R is "strong converse rate"; i.e. min s.t. sending $n(R+\delta)$ bits has success prob $\leq exp(-n\delta')$

 $\chi \leq C \leq R \leq C_{\rm E}$





 $|M| = \frac{1}{2} I(C:R|B) = \frac{1}{2} I(C:R|A)$ qubits communicated entanglement consumed/created = H(C|RB)

quantum Markov states

B

С

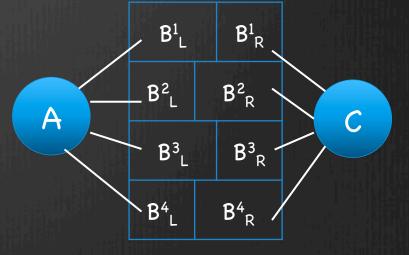
Bob can "redistribute" C to E with ½ I(A:C|B) qubits. If I(A:C|B)=0 then this is reversible! Implies recovery map R : B -> BC such that $(id_A \otimes R_{B->BC})(\rho_{AB}) = \rho_{ABC}$

A

structure theorem: I(A:C|B)=0 iff

$$B = \bigoplus_{i} B_{i}^{L} \otimes B_{i}^{R}$$

$$\rho_{ABC} = \bigoplus_{i} p_{i} \rho_{AB_{i}}^{i} \otimes \rho_{B_{i}}^{i} \otimes \rho_{$$

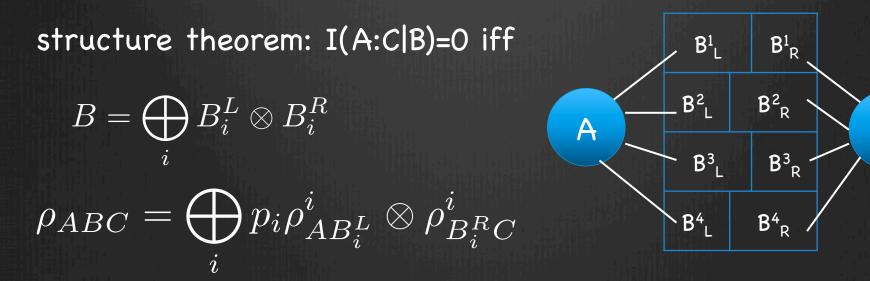


Ε

approximate Markov states

towards a structure thm: [Fawzi-Renner 1410.0664, others] If I(A:C|B) \approx 0 then \exists approximate recovery map R, i.e. (id_A \otimes R_{B->BC})(ρ _{AB}) \approx ρ _{ABC}

states with low CMI appear in condensed matter, optimization, communication complexity, ...



С

Relevant talks

- Wed 9. Omar Fawzi and Renato Renner. Quantum conditional mutual information and approximate Markov chains.
- Wed 9:50. Omar Fawzi, Marius Junge, Renato Renner, David Sutter, Mark Wilde and Andreas Winter. Universal recoverability in quantum information theory.
- Thurs 11. David Sutter, Volkher Scholz, Andreas Winter and Renato Renner. Approximate degradable quantum channels
- Thurs 4:15. Mario Berta, Joseph M. Renes, Marco Tomamichel, Mark Wilde and Andreas Winter. Strong Converse and Finite Resource Tradeoffs for Quantum Channels.

QCMI

channel capacities

semi-relevant talks

Tues 11:50. Ryan O'Donnell and John Wright. Efficient quantum tomography merged with Jeongwan Haah, Aram Harrow, Zhengfeng Ji, Xiaodi Wu and Nengkun Yu. Sample-optimal tomography of quantum states

- Tues 3:35. Ke Li. Discriminating quantum states: the multiple Chernoff distance
- Thurs 10. Mark Braverman, Ankit Garg, Young Kun Ko, Jieming Mao and Dave Touchette. Near optimal bounds on bounded-round quantum communication complexity of disjointness
- Thurs 3:35. Fernando Brandao and Aram Harrow. Estimating operator norms using covering nets with applications to quantum information theory
- Thurs 4:15. Michael Beverland, Gorjan Alagic, Jeongwan Haah, Gretchen Campbell, Ana Maria Rey and Alexey Gorshkov. Implementing a quantum algorithm for spectrum estimation with alkaline earth atoms.

HSW metrics

QCMI

covering

entropy

reference

Mark Wilde. arXiv:1106.1445. "From Classical to Quantum Shannon Theory" Last update Dec 2, 2015. 768 pages.