Group representations and quantum information theory

Aram Harrow (Bristol)
NII, Tokyo
5 Feb, 2007

outline

- Types in classical information theory
- A quantum analogue: Schur duality
- Joint types
- Applications

The method of types

Given a string $x^n = (x_1,...,x_n) \in [d]^n$ $\quad$ $[d] := \{1, ..., d\}$

define the type of x^n to be the letter frequency distribution:
$t = T(x^n) = (t_1, ..., t_d)$
where $t_c := |\{ j : x_j = c \}|$.

For a type t, the type class of t is the set of strings with type t:
$\mathcal{T}_t = \{x^n : T(x^n) = t\}$

Example:
$T(babcba) = (2,1,3)$
$\mathcal{T}_{(2,3,1)} = \{aabbbc, ababbc, abcaab, cbbaaa, ...\}$

Total of $\frac{6!}{2!1!3!}$ strings

Properties of types

1. Size of type classes is given by entropic expression

$$|\mathcal{T}_t| = \binom{n}{t} = \frac{n!}{t_1! \ldots t_d!}$$

$$\bar{t} := t/n$$

$$(n+1)^{-d} \exp(nH(\bar{t})) \leq |\mathcal{T}_t| \leq \exp(nH(\bar{t}))$$

2. Number of types is polynomial

$$\binom{n+d-1}{d-1} \leq (n+1)^d \sim \text{poly}(n)$$

$$D(\bar{t}\|p) := \sum_{i=1}^{d} \bar{t}_i \log \frac{\bar{t}_i}{p_i}$$

3. i.i.d. sources

$$\begin{aligned} p^{\otimes n}(x^n) \quad &:= \quad p(x_1) \cdots p(x_n) = p(1)^{t_1} \cdots p(d)^{t_d} \\ &= \quad \exp\left(\sum_{i=1}^{d} t_i \log p_i\right) = \exp\left(-n\left[\left(H(\bar{t}) + D(\bar{t}\|p)\right)\right]\right) \end{aligned}$$

types and i.i.d. sources

Further note that:

1. $p^{\otimes n}(x^n)$ depends only on the type of x^n

(i.e. conditional on the type, x^n is uniformly distributed)

2. The observed type t is closely concentrated near p.

$$p^{\otimes n}(\mathcal{T}_t) := \sum_{x^n \in \mathcal{T}_t} p^{\otimes n}(x^n) \leq \exp\left(-nD(\bar{t}\|p)\right)$$

Application: randomness concentration

Given x^n distributed according to $p^{\otimes n}$,

we would like to extract a uniformly distributed random variable.

Algorithm:

Condition on the type of x^n.

This yields $\approx nH(p)$ random bits w.h.p.

Applications of types

Data compression:

Given a string x^n drawn from an i.i.d. source $p^{\otimes n}$,

represent it using $\approx nH(p)$ bits.

Algorithm:

1. Transmit the type t using $O(\log n)$ bits.
2. Transmit the index of the string within $\mathcal{T}_t$ using

$$\log |\mathcal{T}_t| \leq nH(\bar{t}) \approx nH(p)$$ bits.

Alternate perspective on types

Divide a type t (e.g. t = (2,5,2,3)) into
a list of frequences λ (e.g. λ = (5,3,2,2))
and a list of corresponding letters q (e.g. q = (b, d, a, c)).

Note that:

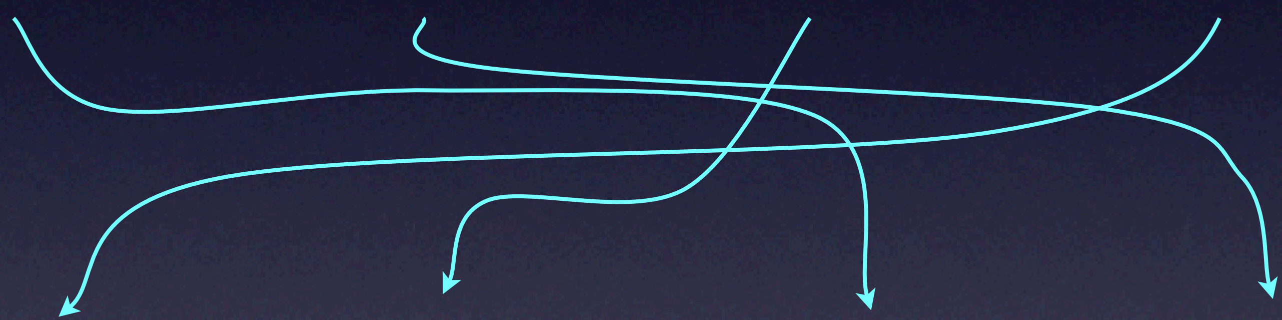
1. Each x^n can be uniquely written as (λ, q, p), where $1 \leq p \leq \binom{n}{\lambda}$
2. The range of λ and q is ≤ poly(n).
3. S_d acts on q, while S_n acts on p. Both leave λ unchanged.

symmetries of $(\mathbb{C}^d)^{\otimes n}$

$U \in \mathcal{U}_d \to U \otimes U \otimes U \otimes U$

$(\mathbb{C}^d)^{\otimes 4} = \mathbb{C}^d \otimes \mathbb{C}^d \otimes \mathbb{C}^d \otimes \mathbb{C}^d$

$(1324) \in \mathcal{S}_4 \to$



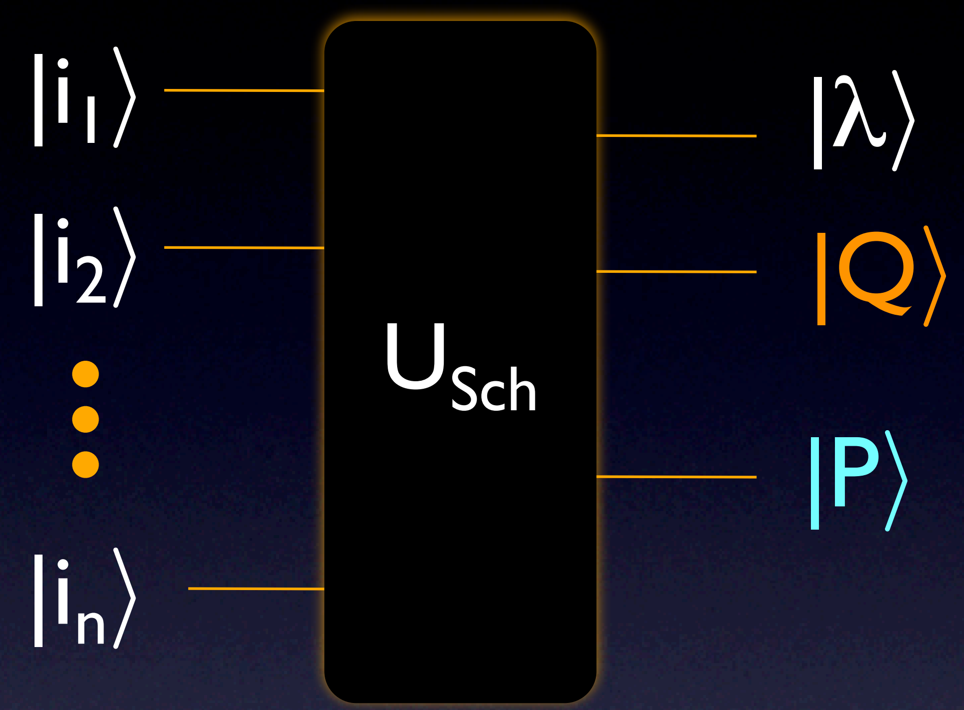
Schur duality

$$(\mathbb{C}^d)^{\otimes n} \cong \bigoplus_{\lambda \in \mathrm{Par}(n,d)} Q_\lambda^d \otimes \mathcal{P}_\lambda$$

The Schur Transform

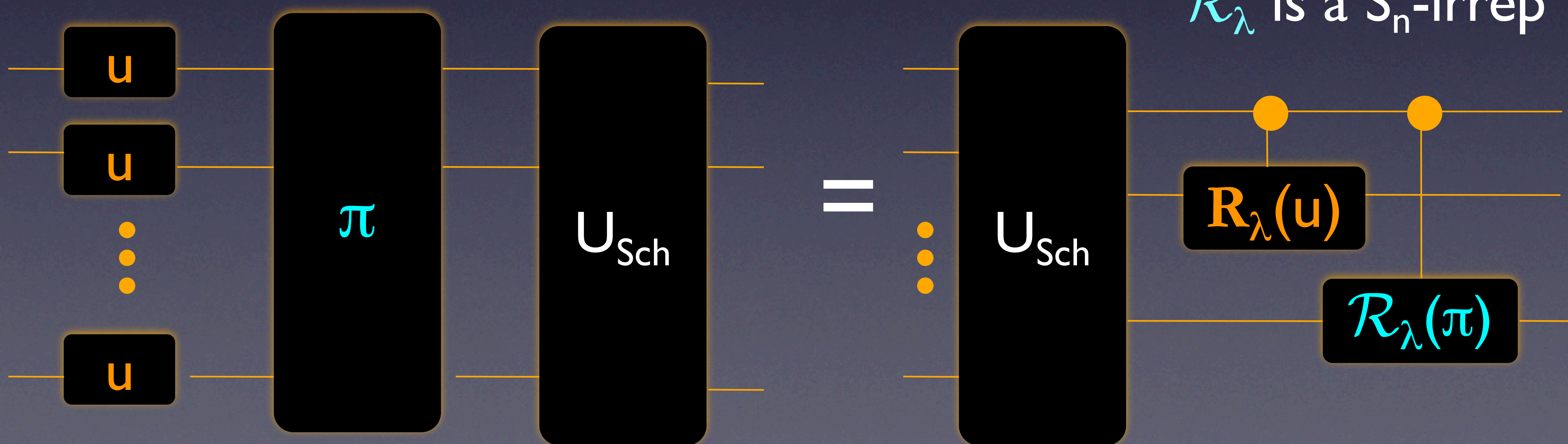
$u \in \mathcal{U}_d$

$\pi \in \mathcal{S}_n$



$\mathbf{R}_\lambda$ is a U_d-irrep

$\mathcal{R}_\lambda$ is a S_n-irrep



Properties of the Schur basis

1. $|\mathrm{Par}(n,d)| \leq (n+1)^d \sim \mathrm{poly}(n)$

2. $\dim \mathcal{Q}_\lambda^d \leq (n+1)^{d^2}$

3. $$\frac{1}{\mathrm{poly}(n)} \exp(nH(\bar{\lambda})) \leq \dim \mathcal{P}_\lambda \leq \exp(nH(\bar{\lambda}))$$

4. i.i.d. sources:
   (a) The P_λ register of $\rho^{\otimes n}$ is maximally mixed.
   (b) $\mathrm{tr}\, \Pi_\lambda \rho^{\otimes n} \leq \exp(-nD(\bar{\lambda} \| \mathrm{spec}\ \rho))$

Summary:
Most of the dimensions are in the P_λ register.
There the spectrum is flat for i.i.d. sources and the dimension is controlled by λ, which is tightly concentrated.

Schur duality applications

Entanglement concentration:

Given $|\psi_{AB}\rangle^{\otimes n}$, Alice and Bob both perform the Schur transform, measure λ, discard $\mathcal{Q}_\lambda$ and are left with a maximally entangled state in $\mathcal{P}_\lambda$ (by Schur's Lemma) equivalent to $\log \dim \mathcal{P}_\lambda \approx nS(\psi^A)$ EPR pairs.

Hayashi and Matsumoto, Universal entanglement concentration. quant-ph/0509140

Universal data compression:

Given $\rho^{\otimes n}$, perform the Schur transform, weakly measure λ, and we obtain $\mathcal{P}_\lambda$ with dimension $\approx \exp(nS(\rho))$. (λ and $\mathcal{Q}_\lambda$ can be sent uncompressed.)

Hayashi and Matsumoto, quant-ph/0209124 and references therein.

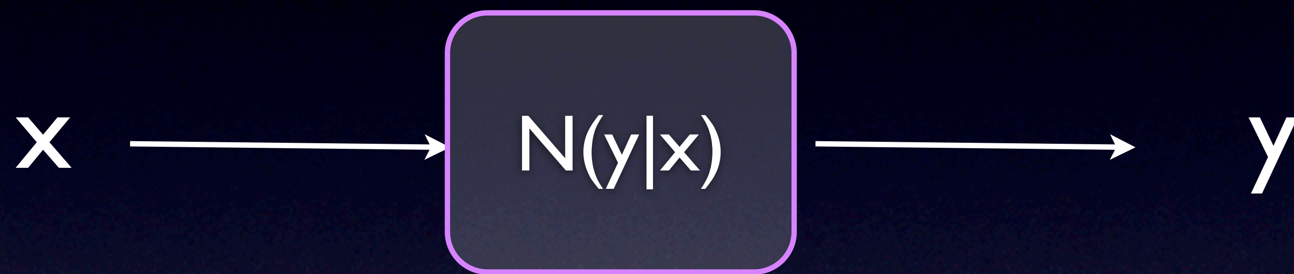
State estimation:

Given $\rho^{\otimes n}$, estimate the spectrum of ρ, or estimate ρ, or test to see whether the state is $\sigma^{\otimes n}$. λ is related to the spectrum, $\mathcal{Q}_\lambda$ to the eigenbasis, and $\mathcal{P}_\lambda$ is maximally mixed.

Keyl, quant-ph/0412053 and references therein.

Joint types

Classical noisy channel:



$t_x = T(x^n)$ is the type of the input
$t_y = T(y^n)$ is the type of the output
$t_{xy} = T(x_1y_1, \ldots, x_ny_n)$ is the joint type

Properties of joint types:

1. For each N, n, $\epsilon > 0$, there is a set of feasible joint types, which can occur with probability $\geq \epsilon$ on some inputs. These correspond roughly to the feasible pairs $(p, N(p))$.

2. $N^{\otimes n}(y^n|x^n)$ depends only on t_{xy}.

Classical Reverse Shannon Theorem

Goal:
Simulate n uses of a noisy channel N using shared randomness and an optimal ($\approx n \max_{p(X)} I(X;Y)$) rate of communication.

Approach:

1. On input x_n, Alice first simulates $N^{\otimes n}$ to obtain a joint type t_{xy}.
2. She sends t_{xy} to Bob using $O(\log n)$ bits.
3. Conditioned on t_{xy}, the action of $N^{\otimes n}$ is easy to describe and to simulate, using the appropriate amount of communication.

Bennett et al., quant-ph/0106052. Winter, quant-ph/0208131.

quantum channels

Church of the Larger Hilbert space:
Purify the input and output of a channel to obtain a tripartite pure state.

Two definitions of feasible joint types:

1. $(\mathrm{Spec}\ \psi^A, \mathrm{Spec}\ \psi^B, \mathrm{Spec}\ \psi^E)$ if $|\psi\rangle^{ABE}$ is the output of one use of $\mathcal{N}$.
2. $(\lambda_A, \lambda_B, \lambda_E)$ such that $\langle\psi| \Pi_{\lambda_A} \otimes \Pi_{\lambda_B} \otimes \Pi_{\lambda_E} |\psi\rangle \geq \epsilon$ for some $|\psi\rangle^{ABE}$ resulting from *n* uses of $\mathcal{N}$.

quantum joint types

Two definitions of feasible joint types:

1. $(\mathrm{Spec}\ \psi^A, \mathrm{Spec}\ \psi^B, \mathrm{Spec}\ \psi^E)$ if $|\psi\rangle^{ABE}$ is the output of one use of $\mathcal{N}$.

2. $(\lambda_A, \lambda_B, \lambda_E)$ such that $\langle\psi|\ \Pi_{\lambda_A} \otimes \Pi_{\lambda_B} \otimes \Pi_{\lambda_E} |\psi\rangle \geq \epsilon$ for some $|\psi\rangle^{ABE}$ resulting from n uses of $\mathcal{N}$.

Results:

1. These two definitions are approximately equivalent.

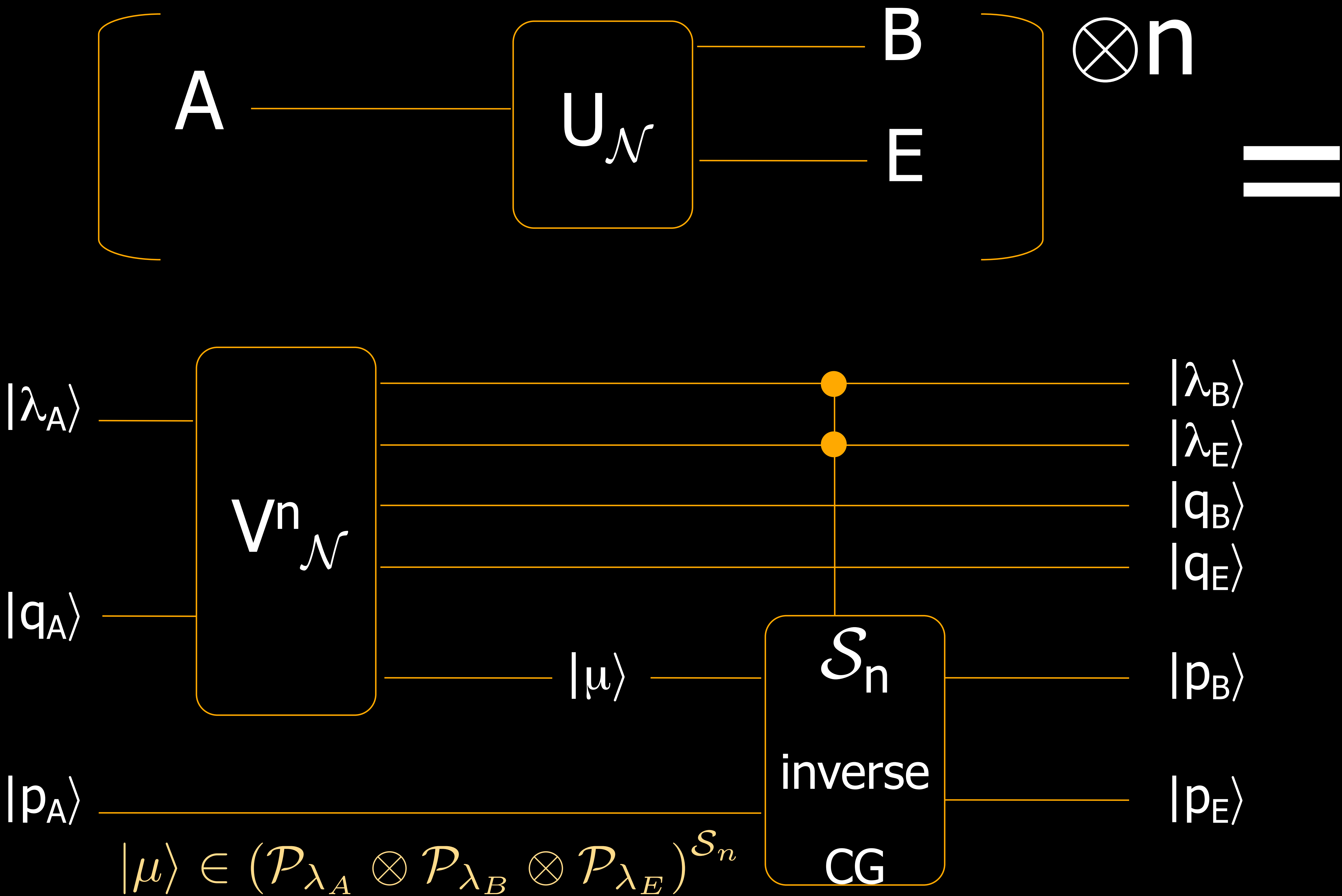
2. There is a sense in which conditioning on a joint type makes all transition probabilities equal.

References:

H., PhD thesis, 2005; quant-ph/0512255.

Christandl, H., Mitchison, "On nonzero Kronecker coefficients and their consequences for spectra." CMP 2007; quant-ph/0511029.

normal form of i.i.d. channels



Tripartite S_n-invariant pure states

Obtained e.g. by purifying $(\rho^{AB})^{\otimes n}$.

$$|\psi\rangle^{ABE} = \sum_{\substack{\lambda_A \in \mathrm{Par}(n,d_A) \\ \lambda_B \in \mathrm{Par}(n,d_B) \\ \lambda_E \in \mathrm{Par}(n,d_B)}} \sum_{\substack{q_A \in \mathcal{Q}^{d_A}_{\lambda_A} \\ q_B \in \mathcal{Q}^{d_B}_{\lambda_B} \\ q_E \in \mathcal{Q}^{d_E}_{\lambda_E}}} \sum_{\mu \in (\mathcal{P}_{\lambda_A} \otimes \mathcal{P}_{\lambda_B} \otimes \mathcal{P}_{\lambda_E})^{\mathcal{S}_n}} C^{q_A, q_B, q_E}_{\lambda_A, \lambda_B, \lambda_E; \mu} |\lambda_A, q_A\rangle^A |\lambda_B, q_B\rangle^B |\lambda_E, q_E\rangle^E |\mu\rangle^{ABE}$$

Interpretation:
This is almost completely general!
Except that μ^A, μ^B and μ^E are each maximally mixed (by Schur's Lemma).

Application: additivity of minimum output entropy

$S_{min}(\mathcal{N}) := \min_\psi S(\mathcal{N}(\psi))$

Additivity question: Does $S_{min}(\mathcal{N}_1 \otimes \mathcal{N}_2) = S_{min}(\mathcal{N}_1) + S_{min}(\mathcal{N}_2)$?

Equivalently: Does $\lim_{n\to\infty} S_{min}(\mathcal{N}^{\otimes n}) / n = S_{min}(\mathcal{N})$?

Our result: $\min_{|\psi\rangle \in \mathrm{Sym}^n \mathbb{C}^d} \geq n\, S_{min}(\mathcal{N}) - o(n)$

where $\mathrm{Sym}^n\mathbb{C}^d = \{|\psi\rangle : \pi|\psi\rangle = |\psi\rangle\ \forall \pi \in S_n\}$

Proof: Most of the entropy is in the $|\mu\rangle$ register. If λ_A is trivial then P_{λ_B} and P_{λ_E} are maximally entangled, so Bob's entropy $\approx \log \dim P_{\lambda_B} \approx nH(\lambda_B)$. Finally, λ_B is ϵ-feasible $\Leftrightarrow \exists$ a nearby feasible single-copy state.

joint work with P. Hayden and A. Winter.

Application: Quantum Reverse Shannon Theorem

Goal: Simulate $\mathcal{N}^{\otimes n}$ using an optimal rate of communication.

Establish qualitative equivalence of all channels.

Idea: Previously constructions were known for i.i.d. input, or for inputs restricted to a single value of λ_A and q_A.
To generalize, Alice splits her input according to λ_A and q_A and simulates $V^n_{\mathcal{N}}$ locally to generate λ_B, q_B, λ_E, q_E and μ.

- μ is simple, and easily compressible.
- λ_B, q_B are small, and can be sent uncompressed.

Subtlety: Different values of λ_B require different amounts of entanglement.

joint work with C. Bennett, I. Devetak, P. Shor and A. Winter.

Entropy-increasing quantum channels

Result: If $S(\mathcal{N}(\rho))>S(\rho)$ for all ρ then $\mathcal{N}^{\otimes n} \approx \sum_\nu p_\nu V_\nu$, where $\{p_\nu\}$ is a probability distribution and $\{V_\nu\}$ are isometries.

Related to quantum Birkhoff conjecture: If $\mathcal{N}$ is unital (i.e. $\mathcal{N}(I/d)=I/d$, or equivalently, $d_A=d_B$ and $S(\mathcal{N}(\rho))\geq S(\rho)$ for all ρ) then $\mathcal{N}^{\otimes n} \approx \sum_\nu p_\nu V_\nu$, where $\{p_\nu\}$ is a probability distribution and $\{V_\nu\}$ are unitaries.

Noisy state analogue: For any state ρ_{AB}, one can decompose $\rho_{AB}^{\otimes n}$ as a mixture of pure states with average entanglement $\approx n \min(S(\rho_A), S(\rho_B))$.

Proof idea for states: If $\dim P_{\lambda_B} \gg \dim P_{\lambda_A}$, then a random measurement on P_{λ_C} will leave P_{λ_A} nearly maximally mixed w.h.p.

Proof idea for channels: Consider the Jamiolkowski state obtained from inputting $\sum_{\lambda_A} |\lambda_A\rangle\langle\lambda_A| / |\mathrm{Par}(n,d_A)| \otimes I/\dim Q_{\lambda_A} \otimes I/\dim P_{\lambda_A}$ to $\mathcal{N}^{\otimes n}$.

[Smolin-Verstraete-Winter, quant-ph/0505038]

joint work with A. Childs

Future research directions

1. The quantum Birkhoff conjecture.

2. Applying the Schur basis to core questions of quantum information theory: additivity, strong converse of HSW theorem, coding.

3. Analyzing product states, e.g. in HSW coding.

4. Performing more protocols efficiently.

References

H., Ph.D thesis, 2005. quant-ph/0512255
Christandl, Ph.D thesis, 2006. quant-ph/0604183
Christandl, H., Mitchison, "On nonzero Kronecker..." q-ph/0511029
Mitchison, "A dual de Finetti theorem." q-ph/0701064
Hayashi and Matsumoto. q-ph/0202001, q-ph/0209030, q-ph/0209124, q-ph/0509140
Hayashi. q-ph/9704040, q-ph/0107004, q-ph/0202002, q-ph/0208020.
Keyl and Werner. q-ph/0102027. Keyl. q-ph/0412053