Exponential Separation of Quantum Communication and Classical Information

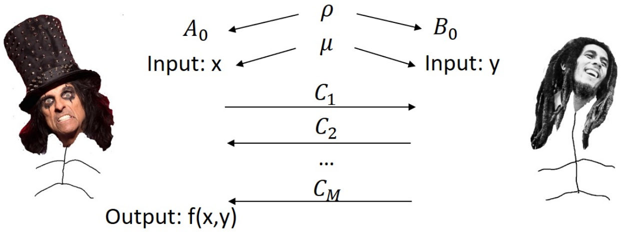
Dave Touchette
IQC and C&O,
University of Waterloo,
and Perimeter Institute for Theoretical Physics

jt. work with
Anurag Anshu (CQT, NUS), Penghui Yao (QuICS, UMD) and
Nengkun Yu (QSI, UTS)
arXiv: 1611.08946

QIP 2017,
Seattle, 20 January 2017

Interactive Communication

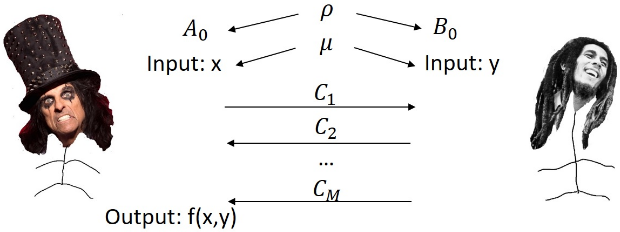
- Communication complexity setting:



- How much **communication/information** to compute f on $(x, y) \sim \mu$

Interactive Communication

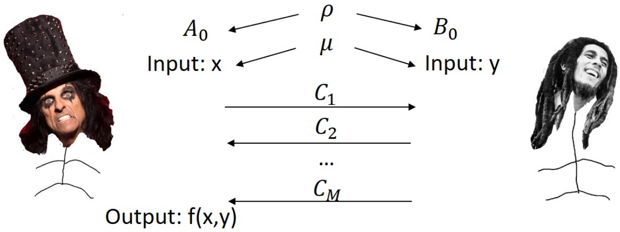
- Communication complexity setting:



- How much **communication/information** to compute f on $(x, y) \sim \mu$
- Information content of interactive protocols?

Interactive Communication

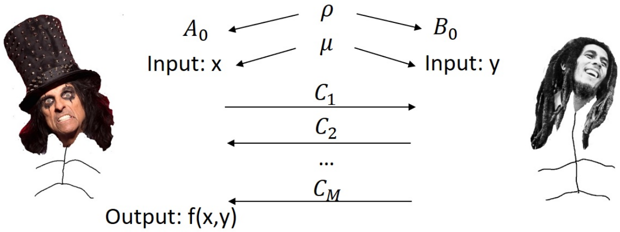
- Communication complexity setting:



- How much **communication/information** to compute f on $(x, y) \sim \mu$
- Information content of interactive protocols?
- Information vs. Communication: Compression?

Interactive Communication

- Communication complexity setting:



- How much **communication/information** to compute f on $(x, y) \sim \mu$
- Information content of interactive protocols?
- Information vs. Communication: Compression?
- Classical vs. Quantum ?

Main result

- Th.: $\exists$ classical task (f, μ, ϵ) s.t. $QCC(f, \mu, \epsilon) \geq 2^{\Omega(IC(f,\mu,\epsilon))}$
  - $f(x, y)$ Boolean function, $(x, y) \sim \mu$, ϵ constant error, say $1/10$
  - QCC: quantum communication complexity
  - IC: classical information complexity

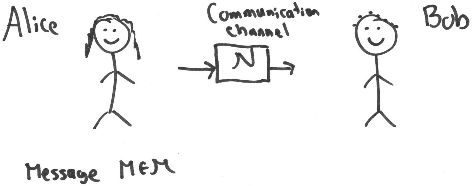
Main result

- Th.: $\exists$ classical task (f, μ, ϵ) s.t. $QCC(f, \mu, \epsilon) \geq 2^{\Omega(IC(f, \mu, \epsilon))}$
  - $f(x, y)$ Boolean function, $(x, y) \sim \mu$, ϵ constant error, say $1/10$
  - QCC: quantum communication complexity
  - IC: classical information complexity
- Cor.: Limit on Direct sum theorems:
  - Amortized CC: $ACC(f, \mu, \epsilon) = \lim_{n \to \infty} \frac{1}{n} CC((f, \mu, \epsilon)^{\otimes n})$
  - $IC(f, \mu, \epsilon) = ACC(f, \mu, \epsilon) \geq AQCC(f, \mu, \epsilon)$
  - $QCC((f, \mu, \epsilon)^{\otimes n}) \not\geq \Omega(n \cdot QCC(f, \mu, \epsilon))$

Main result

- Th.: $\exists$ classical task (f, μ, ϵ) s.t. $QCC(f, \mu, \epsilon) \geq 2^{\Omega(IC(f,\mu,\epsilon))}$
  - $f(x, y)$ Boolean function, $(x, y) \sim \mu$, ϵ constant error, say $1/10$
  - QCC: quantum communication complexity
  - IC: classical information complexity
- Cor.: Limit on Direct sum theorems:
  - Amortized CC: $ACC(f, \mu, \epsilon) = \lim_{n\to\infty} \frac{1}{n} CC((f, \mu, \epsilon)^{\otimes n})$
  - $IC(f, \mu, \epsilon) = ACC(f, \mu, \epsilon) \geq AQCC(f, \mu, \epsilon)$
  - $QCC((f, \mu, \epsilon)^{\otimes n}) \not\geq \Omega(n \cdot QCC(f, \mu, \epsilon))$
- Cor.: Limit on interactive compression:
  - QIC: quantum information complexity
  - $IC(f, \mu, \epsilon) \geq QIC(f, \mu, \epsilon)$
  - $QCC(f, \mu, \epsilon) \not\leq O(QIC(f, \mu, \epsilon))$
- In more details . . .

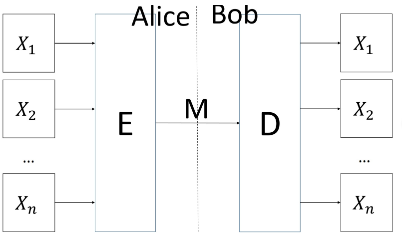
Unidirectional Classical Communication



- Compress messages with "low information content"
- Today, interested in **noiseless** communication channel

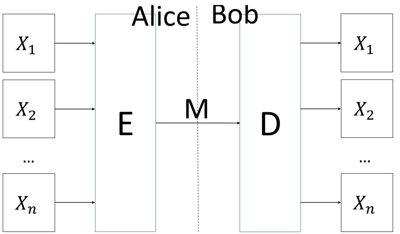
Information Theory I

- How to quantify classical information?
- Shannon's entropy!
- (Finite) Random Variable X of distribution p_X has entropy $H(X)$
- Operational significance: optimal asymptotic rate of compression for i.i.d. copies of X : $\frac{1}{n}|M| \rightarrow H(X)$ bits



Information Theory I

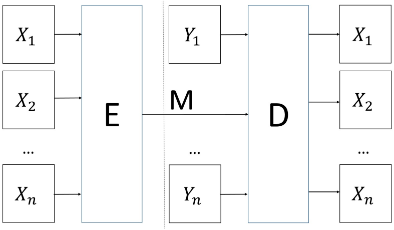
- How to quantify classical information?
- Shannon's entropy!
- (Finite) Random Variable X of distribution p_X has entropy $H(X)$
- Operational significance: optimal asymptotic rate of compression for i.i.d. copies of X : $\frac{1}{n}|M| \rightarrow H(X)$ bits



- Single-copy, optimal variable length encoding, e.g. Huffman code: $H(X) \leq \mathbb{E}(|M|) \leq H(X) + 1$

Information Theory II

- Many Derived Quantities
- Conditional Entropy $H(X|Y) = \mathbb{E}_y H(X|Y = y)$
  - Chain rule for entropy: $H(XY) = H(Y) + H(X|Y)$
  - Operational interpretation: Source X, side information Y, $\lim_{n\to\infty} \frac{1}{n}|M| = H(X|Y)$: Alice does not know Y!

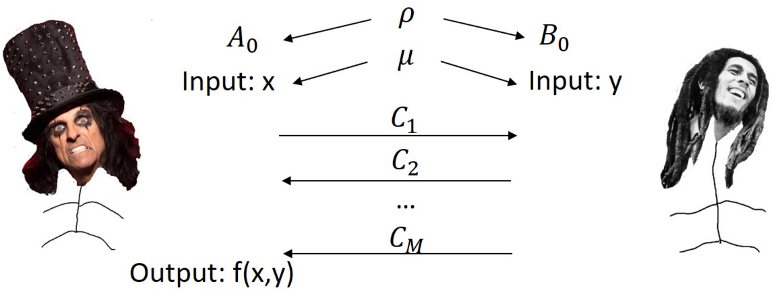


Information Theory II

- Many Derived Quantities
- Conditional Entropy $H(X|Y) = \mathbb{E}_y H(X|Y=y)$
  - Chain rule for entropy: $H(XY) = H(Y) + H(X|Y)$
  - Operational interpretation: Source X, side information Y, $\lim_{n\to\infty} \frac{1}{n}|M| = H(X|Y)$: Alice does not know Y!
- Mutual Information $I(X;C) = H(X) - H(X|C) = I(C;X)$
  - Data Processing $I(X;C) \geq I(X;N(C))$, with N a stochastic map
- Conditional Mutual Information $I(X:Y|Z) = \mathbb{E}_z I(X;Y|Z=z)$
  - **Chain rule**: $I(X_1 X_2 \cdots X_n; C|B) = \sum_{i\leq n} I(X_i; C|BX_1X_2\cdots X_{<i})$
  - $I(X_1X_2\cdots X_n; C|B) \leq H(C)$: at most bit length

Interactive Classical Communication

- Communication complexity of bipartite functions



- $c_1 = f_1(x, r_A), c_2 = f_2(y, c_1, r_B), c_3 = f_3(x, c_1, c_2, r_A), \cdots$
- **Protocol transcript** $\Pi(x, y, r_A, r_B) = c_1 c_2 \cdots c_M$
- Classical protocols: Π memorizes whole history
- $CC(f, \mu, \epsilon) = \min_\Pi CC(\Pi)$
- $CC(\Pi) = |c_1| + |c_2| + \cdots + |c_M|$

Information Cost of Interactive Protocols

- Can we compress protocols that "do not convey much information"
  - For many copies run in parallel?
  - For a single copy?

Information Cost of Interactive Protocols

- Can we compress protocols that "do not convey much information"
  - For many copies run in parallel?
  - For a single copy?
- What is the amount of information conveyed by a protocol?
  - Total amount of information leaked at end of protocol?
  - Sum of information content of each transmitted message?
  - Optimal asymptotic compression rate?

Classical Information Complexity

- Information cost: $IC(\Pi, \mu) = I(X : \Pi | Y) + I(Y : \Pi | X)$ [Barak, Braverman, Chen, Rao 2010]
  - Amount of information each party learns about the other's input from the final transcript

Classical Information Complexity

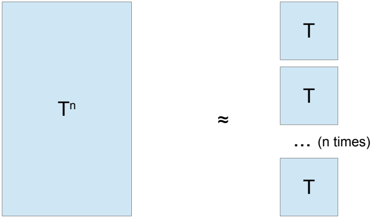
- Information cost: $IC(\Pi, \mu) = I(X : \Pi | Y) + I(Y : \Pi | X)$ [Barak, Braverman, Chen, Rao 2010]
  - Amount of information each party learns about the other's input from the final transcript
- Information complexity: $IC(f, \mu, \epsilon) = \inf_{\Pi} IC(\Pi, \mu)$
  - Least amount of info Alice and Bob must reveal to compute (f, μ, ϵ)

Classical Information Complexity

- Information cost: $IC(\Pi, \mu) = I(X : \Pi | Y) + I(Y : \Pi | X)$ [Barak, Braverman, Chen, Rao 2010]
  - Amount of information each party learns about the other's input from the final transcript
- Information complexity: $IC(f, \mu, \epsilon) = \inf_{\Pi} IC(\Pi, \mu)$
  - Least amount of info Alice and Bob must reveal to compute (f, μ, ϵ)
- Important properties:
  - $T = (f, \mu, \epsilon)$: Task of computing f with average error ϵ w.r.t. μ
  - $T_1 \otimes T_2$: Product task
  - **Additivity**: $IC(T_1 \otimes T_2) = IC(T_1) + IC(T_2)$
  - Lower bounds communication: $IC(T) \leq CC(T)$
  - Operational interpretation: $IC(T) = ACC(T) = \lim_{n \to \infty} \frac{1}{n} CC(T^{\otimes n})$ [Braverman, Rao 2011]
  - Continuity, etc.

Direct Sum

- Direct sum: $CC((f,\epsilon)^{\otimes n}) \geq \Omega(n \cdot CC(f,\epsilon))$?
- Remember $IC(f,\epsilon) = \lim_{n\to\infty} \frac{1}{n} CC\left((f,\epsilon)^{\otimes n}\right)$
  - Direct sum related to one-shot compression down to IC



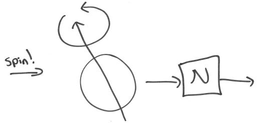
Direct Sum

- Direct sum: $CC((f,\epsilon)^{\otimes n}) \geq \Omega(n \cdot CC(f,\epsilon))$?
- Remember $IC(f,\epsilon) = \lim_{n\to\infty} \frac{1}{n} CC\left((f,\epsilon)^{\otimes n}\right)$
  - Direct sum related to one-shot compression down to IC
- Partial results ...
  - Classical: [Chakrabarti, Shi, Wirth, Yao 2001;
    Jain Radhakrishnan, Sen 2003;
    Harsha, Jain, McAllister, Radhakrishnan 2007;
    Jain, Klauck, Nayak 2008; Barak, Braverman, Chen, Rao 2010;
    Braveman, Rao 2011; Braverman 2012; Kol 2015; Sherstov 2016; ...]
  - Quantum: [Jain, Radhakrishnan, Sen 2005;
    Ambainis, Spalek, De Wolf 2006; Klauck, Jain 2009; Sherstov 2012;
    Anshu, Jain, Mukhopadhyay, Shayeghi, Yao 2014; T. 2015; ...]

Direct Sum

- Direct sum: $CC((f,\epsilon)^{\otimes n}) \geq \Omega(n \cdot CC(f,\epsilon))$?
- Remember $IC(f,\epsilon) = \lim_{n\to\infty} \frac{1}{n} CC\left((f,\epsilon)^{\otimes n}\right)$
  - Direct sum related to one-shot compression down to IC
- Partial results ...
- Fails in general [Ganor, Kol, Raz 2014, 2015, 2016; Rao, Sinha 2015]:
  - $\exists (f,\mu,\epsilon)$ s.t. $CC(f,\mu,\epsilon) \geq 2^{IC(f,\mu,\epsilon)}$

Quantum Information Theory I



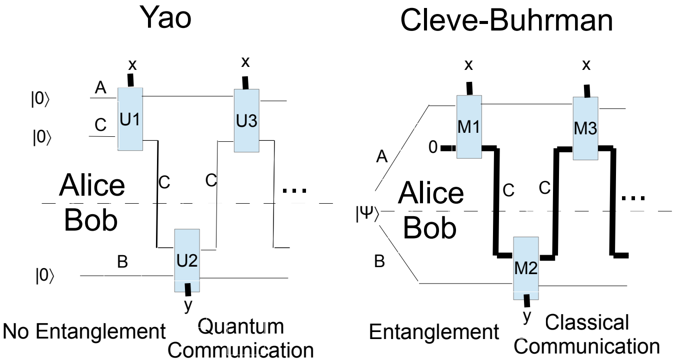
- von Neumann's quantum entropy: $H(A)_\rho$
- Characterizes optimal rate for quantum source compression [Schumacher]

Quantum Information Theory II

- Derived quantities defined in formal analogy to classical quantities
- $H(A|B) \neq \mathbb{E}_b H(A|B=b)$
  - Use $H(A|B) = H(AB) - H(B)$
  - Conditional entropy can be negative!
- $I(A;B) = H(A) - H(A|B) = I(B;A)$
- $I(A;B|C) \neq \mathbb{E}_c I(A;B|C=c)$
  - Use $I(A;B|C) = I(A;BC) - I(A;C)$
  - Get Chain Rule
  - Non negativity holds [Lieb, Ruskai 73]
  - Data Processing also holds

Quantum Communication Complexity

- 2 Models for computing classical $f : \mathcal{X} \times \mathcal{Y} \rightarrow \mathcal{Z}$



- Exponential separations in communication complexity
  - Classical vs. quantum
  - N-rounds vs. N+1-rounds

Quantum Information Complexity

- $QIC(f, \mu, \epsilon) = \inf_\Pi QIC(\Pi, \mu)$
- $QIC(\Pi, \mu)$: based on $I(X; C|YB)$
- Properties:
  - Additivity: $QIC(T_1 \otimes T_2) = QIC(T_1) + QIC(T_2)$
  - Lower bounds communication: $QIC(T) \leq QCC(T)$
  - Operational interpretation [T.]:
    $QIC(T) = AQCC(T) = \lim_{n\to\infty} \frac{1}{n} QCC(T^{\otimes n})$

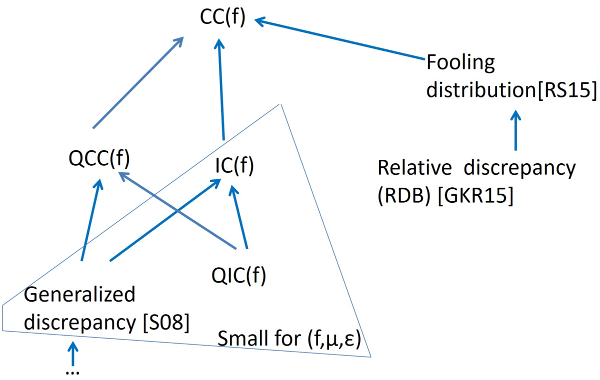
Implications of Main Result

- Recall: $\exists(f, \mu, \epsilon)$ s.t. $QCC(f, \mu, \epsilon) \geq 2^{\Omega(IC(f,\mu,\epsilon))}$
  - f Boolean-valued function, ϵ constant, say $1/10$
- Implications...

Implications of Main Result

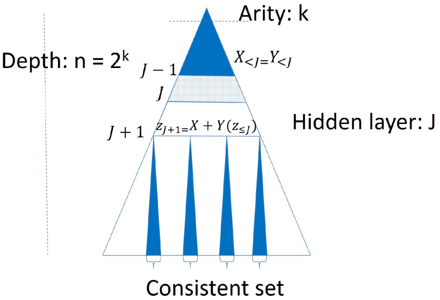
- Recall: $\exists(f,\mu,\epsilon)$ s.t. $QCC(f,\mu,\epsilon) \geq 2^{\Omega(IC(f,\mu,\epsilon))}$
  - f Boolean-valued function, ϵ constant, say $1/10$
- Implications...
- No strong Direct sum theorem for Quantum Communication Complexity: $QCC((f,\mu,\epsilon)^{\otimes n}) \not\geq \Omega(nQCC(f,\mu,\epsilon))$
- Even stronger: $ACC(f,\mu,\epsilon) \leq O(\lg QCC(f,\mu,\epsilon))$!
- IC and QCC incomparable: $\exists(g,\eta,\delta)$ s.t. $IC(g,\eta,\delta) \geq 2^{\Omega(QCC(g,\eta,\delta))}$ [Kerenidis, Laplante, Lerays, Roland, Xiao 2012]
- GDM bound is not (poly) tight for QCC

Need for a New Lower Bound Method on QCC



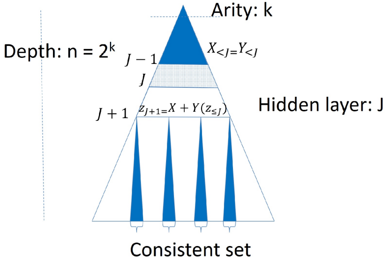
- Is RDB $\leq$ QCC (poly)? No [Klartag, Regev 2011]: QCC (VSP)$\leq$ lg (RDB (VSP))!
- We want new method M s.t. M $\leq$ QCC and IC $\not\geq$ M : Quantum Fooling Distribution!

(f, μ, ϵ): Rao-Sinha's k-ary pointer jumping function



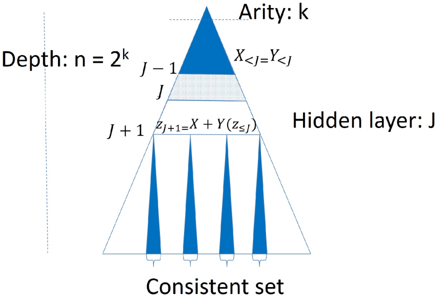
- Two parameters: arity k, depth n. Fix $n = 2^k$.
- Input: (x, h_A) with Alice, (y, h_B) with Bob
- $x, y : [k]^{\leq n} \to [k]$, $x + y \mod k$ defines "good" path
- $h_A, h_B : [k]^n \to \{0, 1\}$, $h_A \oplus h_B$ on a "good" leaf defines output

(Quantum) Fooling Distribution



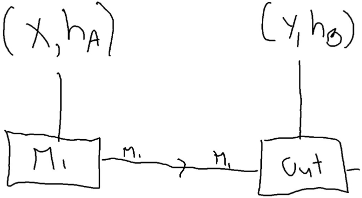
- Two distributions: fooling dist. p, hard dist. μ, with $\mu = \frac{1}{2}\mu_0 + \frac{1}{2}\mu_1$
- Hidden layer $j \in [n]$
- $x_{<j} = y_{<j}$
- Fix $j, x_{<j} = y_{<j}$
- Let G set of "good" leaves/paths: determined by $x_j + y_j$ only
- p: $(x, h_A) \otimes (y, h_B)$ product distribution
- μ_b: $x_G = y_G$, $h_A^G \oplus h_B^G = b$
- For low QCC: $\Pr_{\mu_0}[Out = 1] \approx \Pr_p[Out = 1] \approx \Pr_{\mu_1}[Out = 1]$

Low Information?



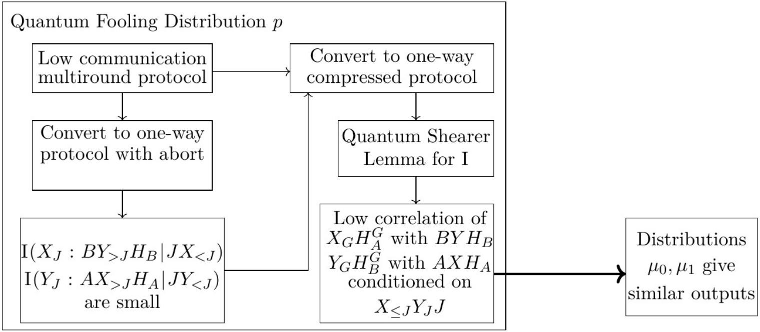
- On "good" path, $x = y$ except at level j ...
- If can "hide" $j \in [n]$, then information $\approx \lg k$, values of $x(z_j), y(z_j), |z_j| = j$
- *Must* hide j: *CC* to find $j \approx \lg n = H(J) = k = 2^{O(IC)}$
- Hide j by adding noise [Rao, Sinha 2015]: IC $\leq O(\lg k)$,
- We show QCC is at least poly(k)
- For one round, QCC is $\Omega(k)$ .. then poly (k) from **round elimination**.

Baby Case for QCC Lower Bound: One-way protocols

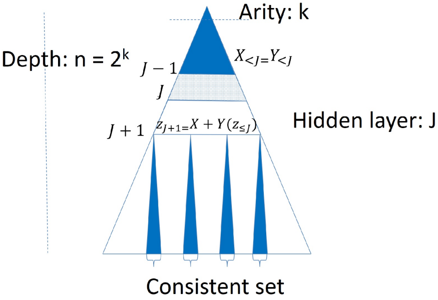


- One Message M depends only on (x, h_A)
- Output depends only on M and (y, h_B)
- p vs. μ_0: massage to $M \otimes (X^G, H_A^G)$ vs. non-$\otimes$ $M(X^G H_A^G)$
- Can relate $|\Pr_{\mu_0}[Out = 1] - \Pr_p[Out = 1]|$ to $I(M; X^G H_A^G)$
- $\approx$ Shearer: $\Pr_p[Leaf\ \ell \in G] \leq \frac{1}{k} \rightarrow I(M; X^G H_A^G) \leq \frac{|M|}{k}$

Structure of Proof for Multi-Round Protocols

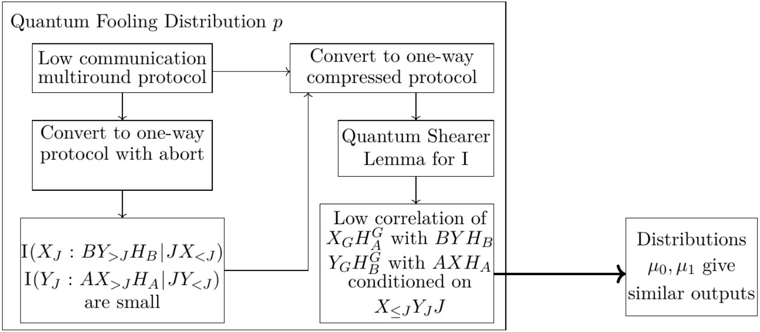


First Conversion to One-Way

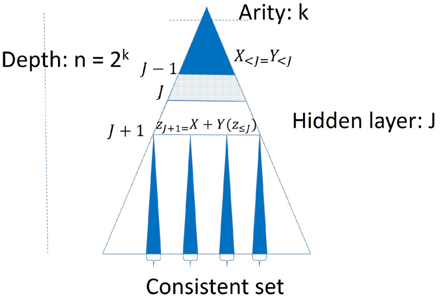


- One-round: Alice does not know J
- Need to send information about all X_J, $J \in [n]$
- Multi-round to one-round: guess teleportation transcript, parallel-repeat 2^{QCC} times
- Need $2^{QCC} \geq n = 2^k \leftrightarrow QCC \geq k$
- Technical issue: repeat once, abort if guess wrong

Structure of Proof for Multi-Round Protocols

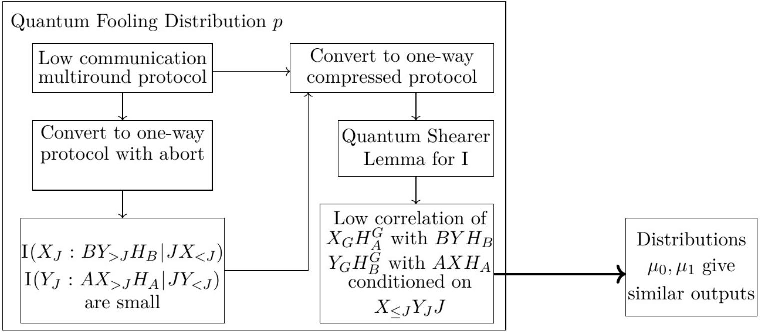


Second Conversion to One-Way

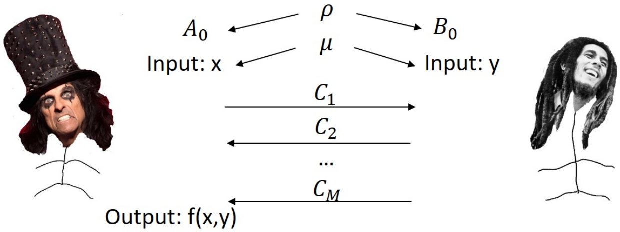


- Want to perform compression to $QIC_{B\to A}2^{QIC_{A\to B}}$ [Jain, Radhakrishnan, Sen 2005]
- Use product structure of distribution p
- Need to fix protocol, J, embed input $X_J \ldots$
- Need $QCC \geq k$ to send information about all k possible paths

Structure of Proof for Multi-Round Protocols



Going from p to μ_0



- Distributional Cut-and-Paste
- If local state is independent of other party's input under $p = \mu_X \otimes \mu_Y$
- Then local state is independent of other party's input under μ_{XY}

Outlook

- Summary:
  - Exponential separation between QCC and IC
  - Strong Direct Sum fails for QCC
  - New QCC lower bound tools
- Open Questions:
  - *External* classical Information complexity?
  - What is power of quantum fooling distribution method? Quantum Relative Discrepancy?
  - 2^{QIC} compression? BBCR-type interactive compression?
  - Partial Direct Sum?

Outlook

- Summary:
  - Exponential separation between QCC and IC
  - Strong Direct Sum fails for QCC
  - New QCC lower bound tools
- Open Questions:
  - *External* classical Information complexity?
  - What is power of quantum fooling distribution method? Quantum Relative Discrepancy?
  - 2^{QIC} compression? BBCR-type interactive compression?
  - Partial Direct Sum?
- Thank you!

Outlook

- Summary:
  - Exponential separation between QCC and IC
  - Strong Direct Sum fails for QCC
  - New QCC lower bound tools
- Open Questions:
  - *External* classical Information complexity?
  - What is power of quantum fooling distribution method? Quantum Relative Discrepancy?
  - 2^{QIC} compression? BBCR-type interactive compression?
  - Partial Direct Sum?
- Thank you!
- See you next year!