Information-Theoretic Tools for Interactive Quantum Protocols, and Applications: Flow of Information, Augmented Index, and DYCK(2)

MATHIEU LAURIÈRE, ASHWIN NAYAK, AND DAVE TOUCHETTE

QIP 2017, Seattle

16 January 2017

Interactive Quantum Protocols,

MATHIEU LAURIÈRE, ASHWIN NAYAK, AND DAVE TOUCHETTE

QIP 2017, Seattle

16 January 2017

Information-Theoretic Tools for Interactive Quantum Protocols,

MATHIEU LAURIÈRE, ASHWIN NAYAK, AND DAVE TOUCHETTE

QIP 2017, Seattle

16 January 2017

Information-Theoretic Tools for Interactive Quantum Protocols, and Applications: Flow of Information, Augmented Index, and DYCK(2)

MATHIEU LAURIÈRE, ASHWIN NAYAK, AND DAVE TOUCHETTE

QIP 2017, Seattle

16 January 2017

Quantum Advantage for Disjointness

- Disjointness: $x, y \subseteq \{1,2,\dots,n\}$, is $x \cap y = \emptyset$?
- $x = x_1 \cdots x_n, y = y_1 \cdots y_n \in \{0,1\}^n$, looking for i such that $x_i = y_i = 1$
- Quantum Protocol [BCW98]: distributed version of Grover search
- QCC(Disj) = $\Theta(\sqrt{n})$ [BCW98, Razb03, AA03]
- CC(Disj) = $\Omega(n)$ [KS92]

Input: x

Initialize: $\frac{1}{n}\sum_i |\, i \,\rangle$

Oracle call: $\frac{1}{n}\sum_i |\, i \,\rangle |\, x_i \rangle \longrightarrow \frac{1}{n}\sum_i (-1)^{x_i \wedge y_i} |\, i \,\rangle |\, x_i \rangle$

$\frac{1}{n}\sum_i (-1)^{x_i \wedge y_i} |\, i \,\rangle \longleftarrow$

Inversion about the mean
Repeat $\approx \sqrt{n}$ times
Measure to get desired i if intersection

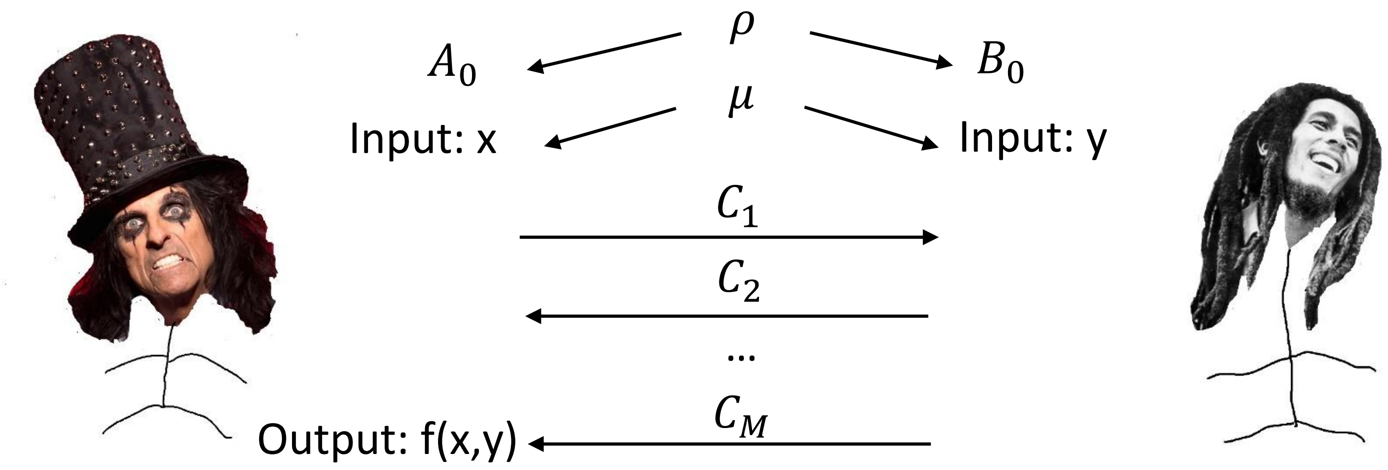
Input: y

Quantum Advantage for Disjointness

- Disjointness: $x, y \subseteq \{1,2,\ldots,n\}$, is $x \cap y = \emptyset$?
- $x = x_1 \cdots x_n, y = y_1 \cdots y_n \in \{0,1\}^n$, looking for i such that $x_i = y_i = 1$
- Quantum Protocol [BCW98]: distributed version of Grover search
- QCC(Disj) = $\Theta(\sqrt{n})$ [BCW98, Razb03, AA03]
- CC(Disj) = $\Omega(n)$ [KS92]
- How does information flow in this protocol?
- Can we avoid transmitting back/forgetting information?

Interactive Communication

- Communication Complexity setting:



- How much **communication** to compute f on $(\mathrm{x},\mathrm{y}) \sim \mu$
- Take information-theoretic view: Information Complexity
  - How much **information** to compute f on $(\mathrm{x},\mathrm{y}) \sim \mu$
- Information content of interactive protocols?
- Classical vs. Quantum?

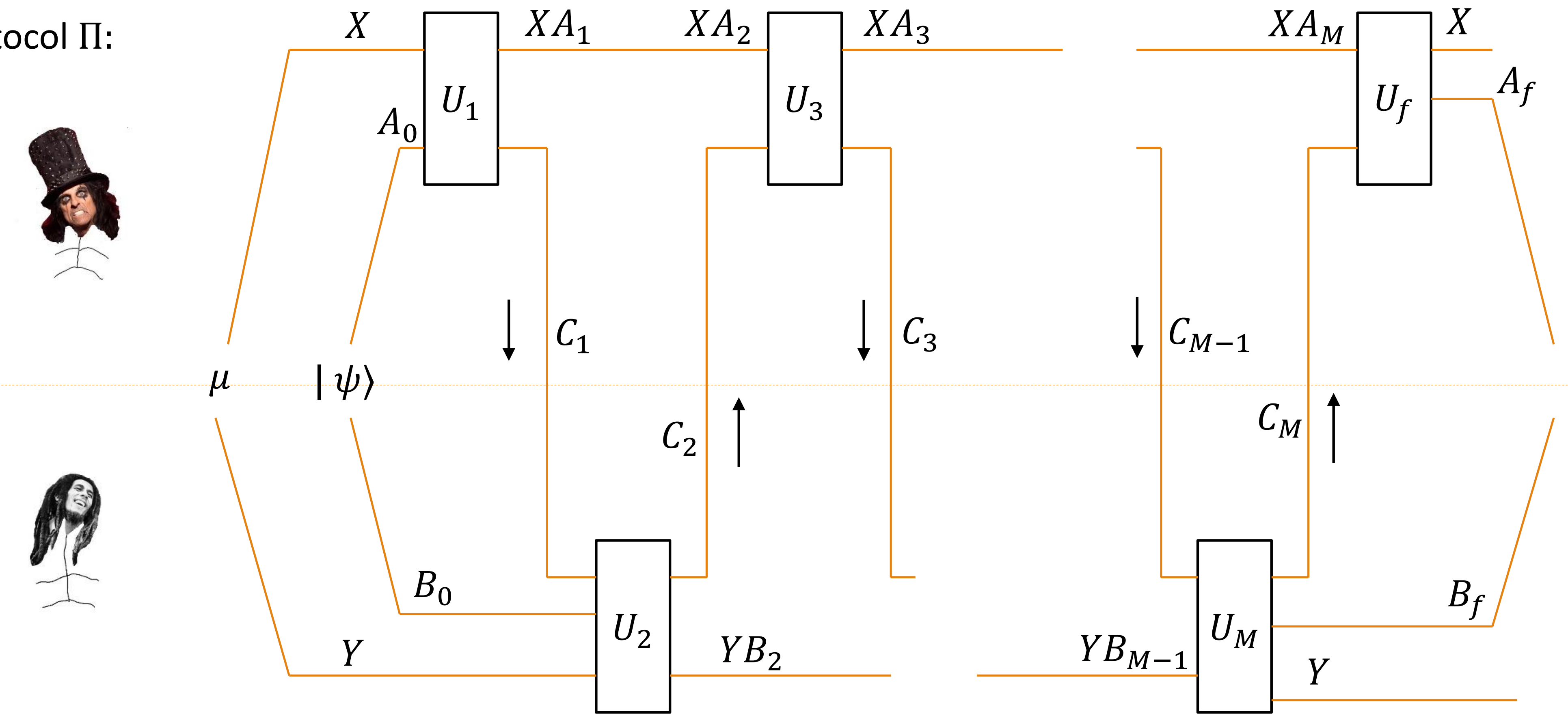
Overview

Based on 2 papers

- 1701.02062: ML & DT, Info. Flow & Cost of Forgetting
  - Th 1: HIC = CIC – CRIC, QIC = CIC + CRIC
    - Tool 1 : Information Flow Lemma
  - Th 2: Π not forgetting for Disjointness => QCC(Π) $\in \Omega(n)$
  - Th 3: Can maintain IC for quantum simulation of classical protocols, and then IC(f_{rdm}) = n (1 - o(1))
- 1610.04937: AN & DT, Aug. Index & Streaming algo. for DYCK(2)
  - Th 4: Any T-pass one-way qu. Streaming algorithm for DYCK(2) requires space $s(N) \in \Omega(\frac{\sqrt{N}}{T^3})$ on length N inputs
  - Th 5: Any t-round protocol for Augmented Index satisfies a QIC trade-off $QIC_{A\rightarrow B}(\Pi, \mu_0) \in \Omega\left(\frac{n}{t^2}\right)$ or $QIC_{B\rightarrow A}(\Pi, \mu_0) \in \Omega\left(\frac{1}{t^2}\right)$
    - Tool 2: Superposition-Average Encoding Theorem
    - Tool 3: Quantum Cut-and-Paste
    - Application of Tool 1

Quantum Communication Complexity

Protocol Π:

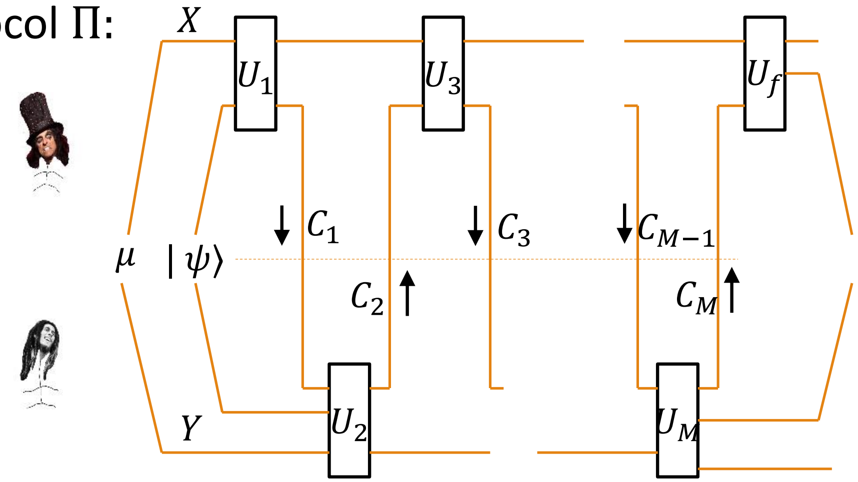


Output: f(X,Y)

Quantum Communication Complexity

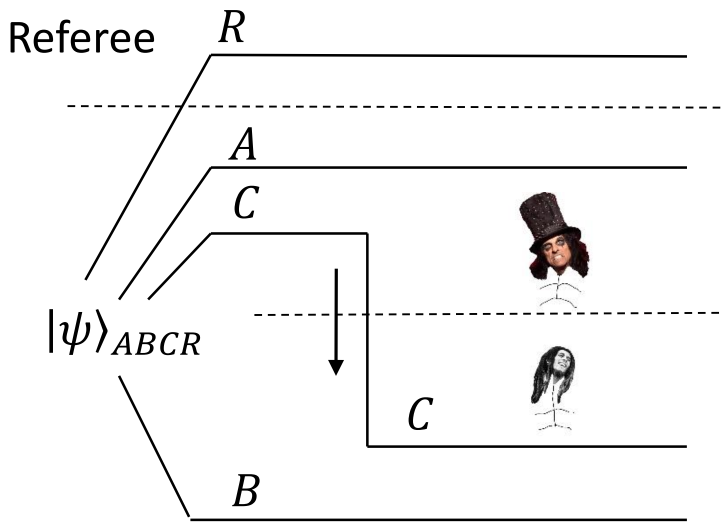
- QCC(f) = $\min_{\Pi}$ QCC(Π)
- Minimization over all Π computing f
- QCC(Π) =$\sum_i \log(\dim(C_i))$; total number of qubits exchanged

Protocol Π:



Quantum Information Theory

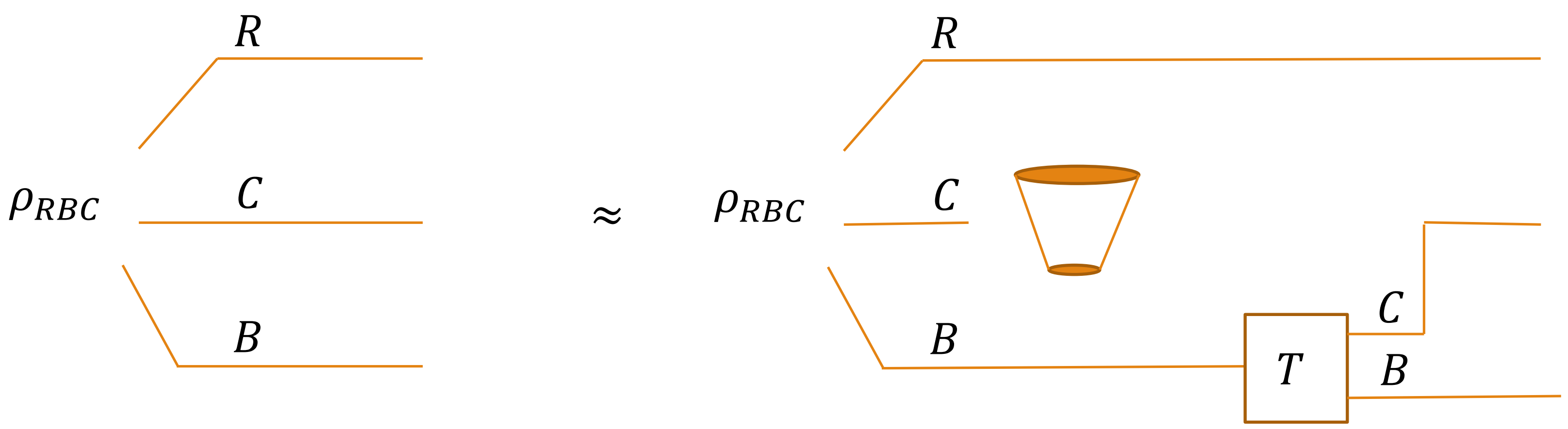
- Conditional Quantum Mutual Information
  - $I(R:C|B) = I(R:BC) - I(R:B) = H(R|B) - H(R|BC) = H(RB) + H(BC) - H(B) - H(RBC)$
  - Non-negativity: $I(R:C|B) \geq 0$ [LR73]
  - Chain rule: $I(A:BD|C) = I(A:B|C) + I(A:D|BC)$
  - Invariance under local isometry, satisfies a data processing inequality…
  - Operational interpretation [DY08, YD09]: Quantum state redistribution, optimal communication rate $I(R:C|B) = I(R:C|A)$



[Lieb and Ruskai 1973; Devetak and Yard 2008; Yard and Devetak 2009]

Quantum Information Theory

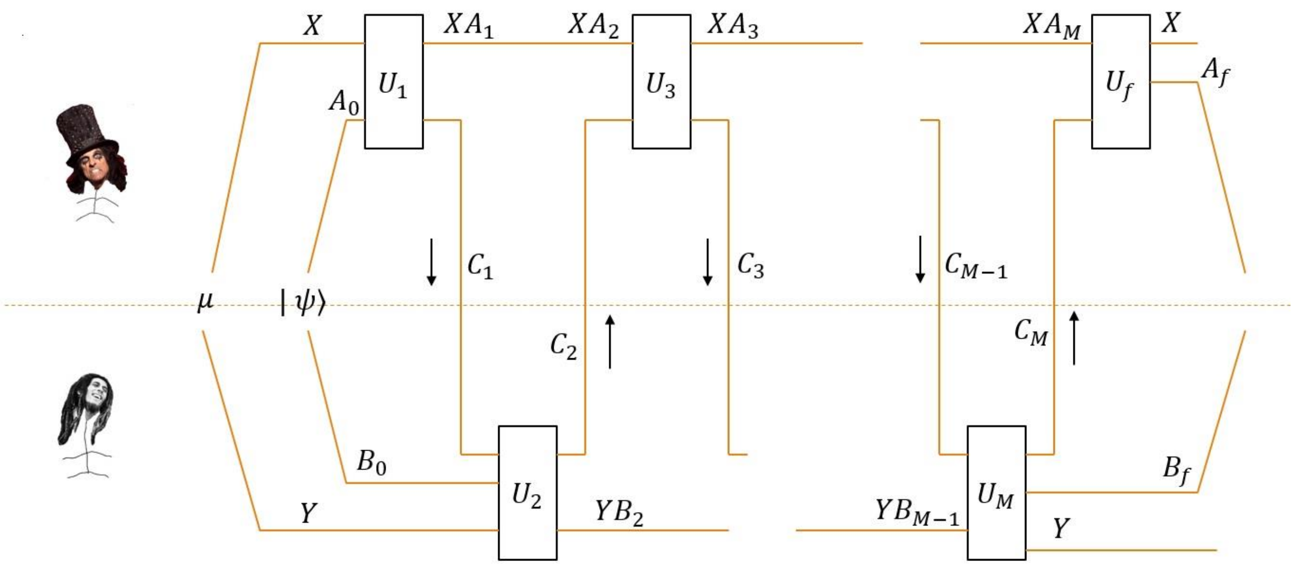
- Conditional Quantum Mutual Information
  - $I(R:C|B) = I(R:BC) - I(R:B) = H(R|B) - H(R|BC) = H(RB) + H(BC) - H(B) - H(RBC)$
  - Non-negativity: $I(R:C|B) \geq 0$ [LR73]
  - Chain rule: $I(A:BD|C) = I(A:B|C) + I(A:D|BC)$
  - Invariance under local isometry, satisfies a data processing inequality…
  - Operational interpretation [DY08, YD09]: Quantum state redistribution, optimal communication rate $I(R:C|B) = I(R:C|A)$
  - Recoverability [FR15]
    - There exists a recovery map $T_{B\to BC}$ such that $-\lg F\left(\rho_{RBC}, T_{B\to BC}(\rho_{RB})\right) \leq I(R:C|B)_\rho$

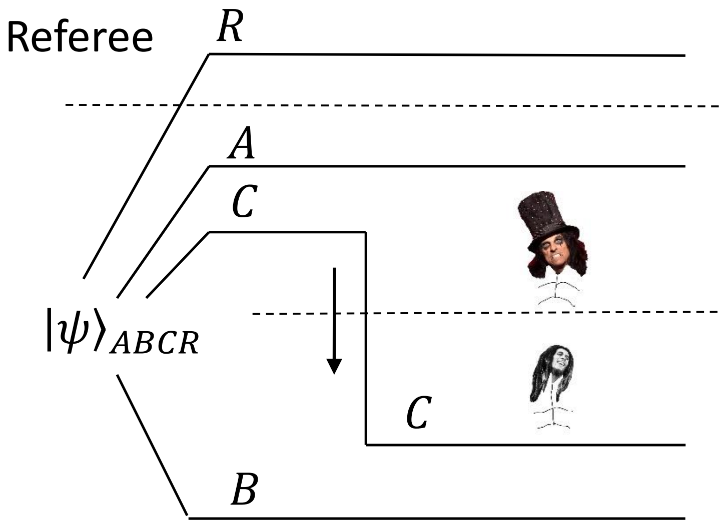


[Lieb and Ruskai 1973; Devetak and Yard 2008; Yard and Devetak 2009; Fawzi and Renner 2015]

Quantum Information Complexity (QIC)

- QIC(f,μ) = $\inf_{\Pi}$ QIC(Π, μ)
- Optimization over all Π computing f
- QIC(Π, μ) = $\sum_{i\ odd} I(R_X R_Y : C_i | Y B_i) + \sum_{i\ even} I(R_X R_Y : C_i | X A_i)$
  - Motivated by quantum state redistribution, with $R_X R_Y$ purifying the XY input registers: $|\rho_\mu\rangle_{R_X X R_Y Y} = \sum_{x,y} \sqrt{\mu(x,y)}\, |xxyy\rangle_{R_X X R_Y Y}$





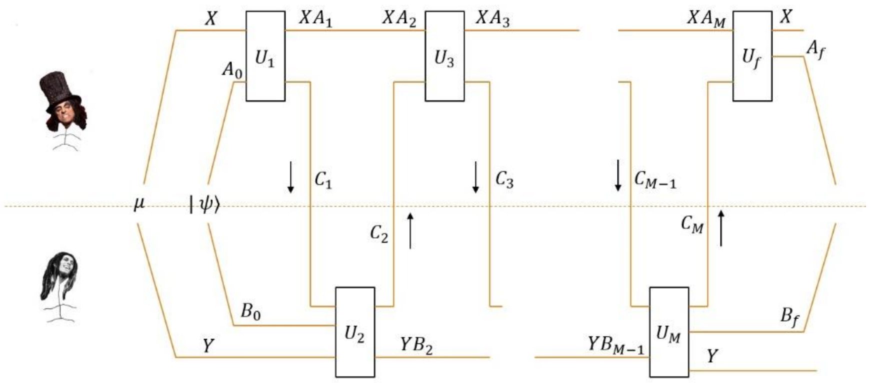
Quantum Information Complexity (QIC)

- QIC(f,μ) = $\inf_{\Pi}$ QIC(Π, μ)
- Optimization over all Π computing f
- QIC(Π, μ) = $\sum_{i\ odd} I(R_X R_Y : C_i | Y B_i) + \sum_{i\ even} I(R_X R_Y : C_i | X A_i)$
- Properties [T15]:
  - Information equals amortized communication
  - Additivity
  - QIC $\leq$ QCC
  - Continuity, …

[T. 2015]

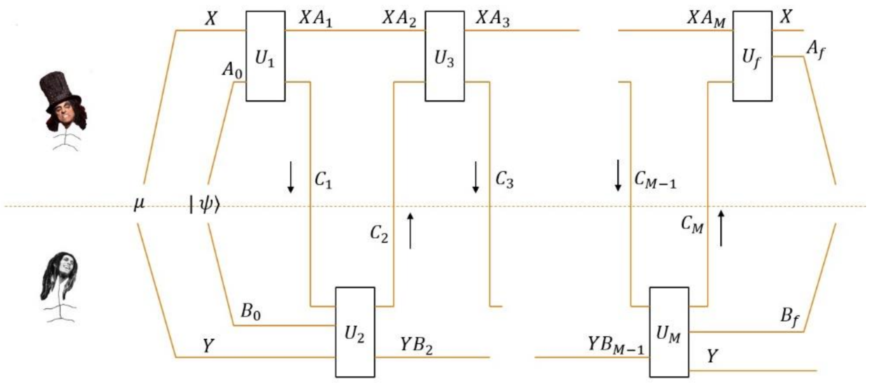
Alternative Notions of QIC

- QIC measures information about what?
  - Satisfies Information equals amortized communication
  - What about these purification registers for classical inputs?



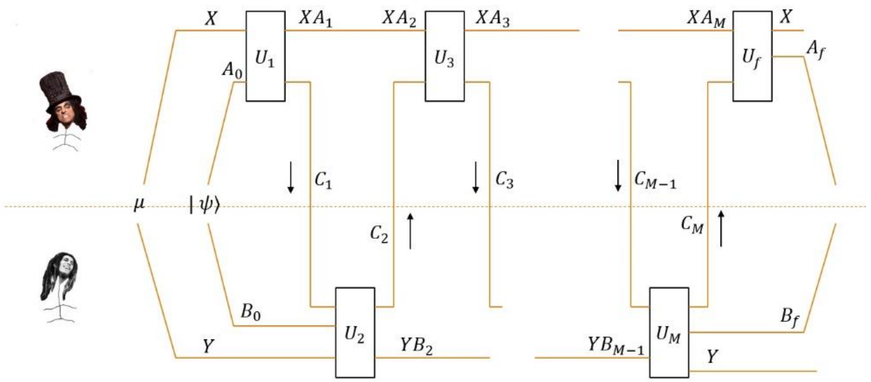
Alternative Notions of QIC

- QIC measures information about what?
  - Satisfies Information equals amortized communication
  - What about these purification registers for classical inputs?
- Can we simply measure the final information?
  - HIC(Π, μ) = $I\left(X\colon B_f \middle| Y\right) + I(Y\colon A_f|X)$
  - Compare to classical IC(Π_C, μ) = $I(X\colon \Pi_C|Y) + I(Y\colon \Pi_C|X)$, with $\Pi_\mathrm{C} = \mathrm{M}_1\mathrm{M}_2 \cdots$ the transcript of messages
  - But reversible computing makes HIC(f) trivial…



Alternative Notions of QIC

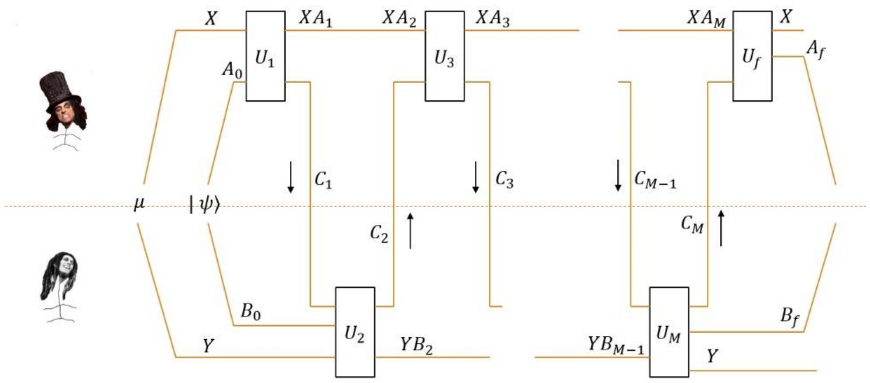
- QIC measures information about what?
  - Satisfies Information equals amortized communication
  - What about these purification registers for classical inputs?
- Can we simply measure the final information?
  - HIC(Π, μ) = $I\left(X : B_f \middle| Y\right) + I(Y : A_f | X)$
  - Compare to classical IC(Π_C, μ) = $I(X : \Pi_C | Y) + I(Y : \Pi_C | X)$, with $\Pi_\mathrm{C} = \mathrm{M}_1 \mathrm{M}_2 \cdots$ the transcript of messages
  - But reversible computing makes HIC(f) trivial…
- Can we measure only new classical information?
  - CIC(Π, μ) = $\sum_{i\ odd} I(X : C_i | Y B_i) + \sum_{i\ even} I(Y : C_i | X A_i)$ [KLLGR16]
  - Compare to classical IC(Π_C, μ) = $\sum_{i\ odd} I(X : M_i | Y M_{<i}) + \sum_{i\ even} I(Y : M_i | X M_{<i})$
  - Motivated by privacy concerns



[Kerenidis, Lauriere, Le Gall and Rennela 2016]

Alternative Notions of QIC

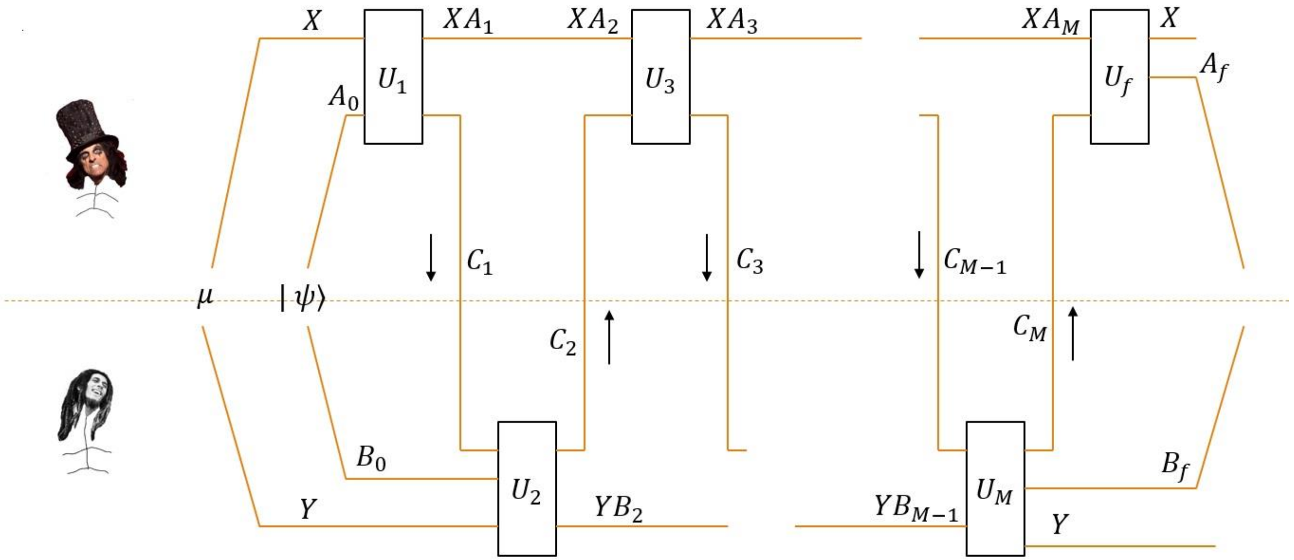
- QIC measures information about what?
  - Satisfies Information equals amortized communication
  - What about these purification registers for classical inputs?
- Can we simply measure the final information?
  - HIC(Π, μ) = $I\big(X\colon B_f \big| Y\big) + I(Y\colon A_f | X)$
  - Compare to classical IC(Π_C, μ) = $I(X\colon \Pi_C | Y) + I(Y\colon \Pi_C | X)$, with $\Pi_C = M_1 M_2 \cdots$ the transcript of messages
  - But reversible computing makes HIC(f) trivial…
- Can we simply measure new classical information?
  - CIC(Π, μ) = $\sum_{i\ odd} I(X\colon C_i | Y B_i) + \sum_{i\ even} I(Y\colon C_i | X A_i)$ [KLLGR16]
  - Compare to classical IC(Π_C, μ) = $\sum_{i\ odd} I(X\colon M_i | Y M_{<i}) + \sum_{i\ even} I(Y\colon M_i | X M_{<i})$
  - Motivated by privacy concerns
- HIC$(\Pi, \mu) \leq$ CIC$(\Pi, \mu) \leq$ QIC(Π, μ)
  - Is there a deeper relationship?



[Kerenidis, Lauriere, Le Gall and Rennela 2016]

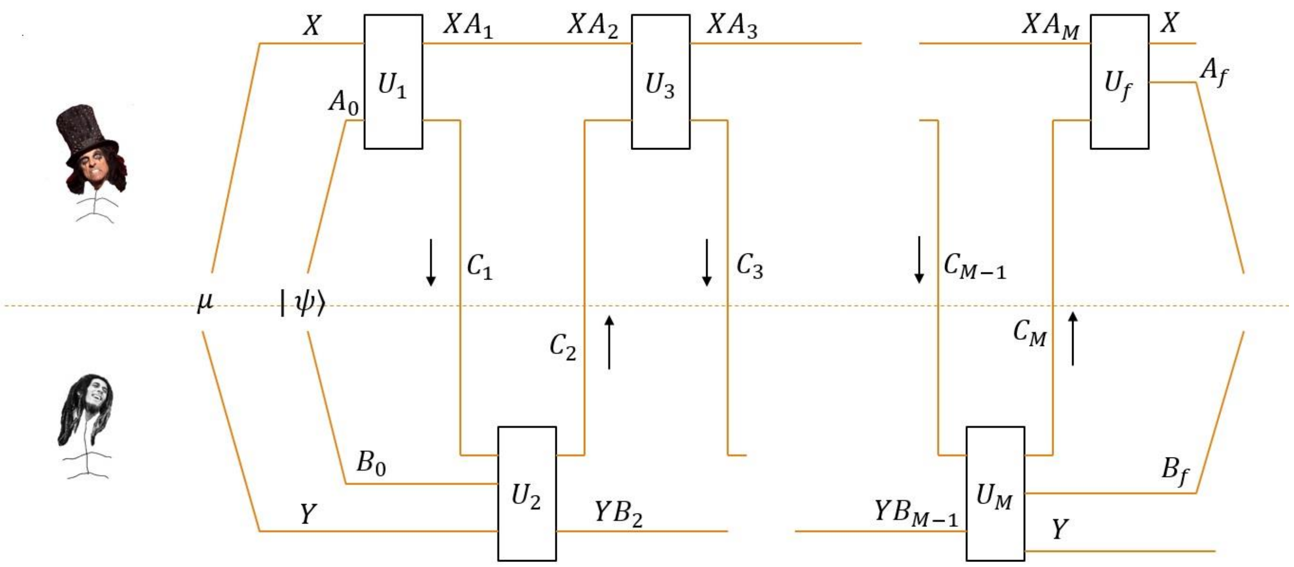
Tool 1: Information Flow Lemma

- Lemma: $I(X\!:\!YB_f) - I(X\!:\!Y) = I(X\!:\!B_f|Y) = \sum_{i\ odd} I(X\!:\!C_i|YB_i) \ - \sum_{i\ even} I(X\!:\!C_i|YB_i)$
  - Can also handle fully quantum processes and arbitrary extension of inputs



Th. 1: Cost of Forgetting

- Rewrite QIC(Π, μ) = $\sum_i I(X: C_i|YB_i) + I(Y: C_i|XA_i)$
  - What are those extra terms compared to CIC?
  - CRIC(Π, μ) = $\sum_{i\ even} I(X: C_i|YB_i) + \sum_{i\ odd} I(Y: C_i|XA_i)$



Th. 1: Cost of Forgetting

- Rewrite QIC(Π, μ) = $\sum_i I(X: C_i | YB_i) + I(Y: C_i | XA_i)$
  - What are those extra terms compared to CIC?
  - CRIC(Π, μ) = $\sum_{i\ even} I(X: C_i | YB_i) + \sum_{i\ odd} I(Y: C_i | XA_i)$
- Using Info. Flow Lemma, rewrite
  - Th. 1.1: HIC(Π, μ) = CIC(Π, μ) - CRIC(Π, μ)
  - QIC(Π, μ) = CIC(Π, μ) + CRIC(Π, μ)
- CRIC corresponds to cost of forgetting
  - Exactly assess back-flow of information
  - No need to introduce purification registers $R_X R_Y$ to define QIC (for classical tasks)

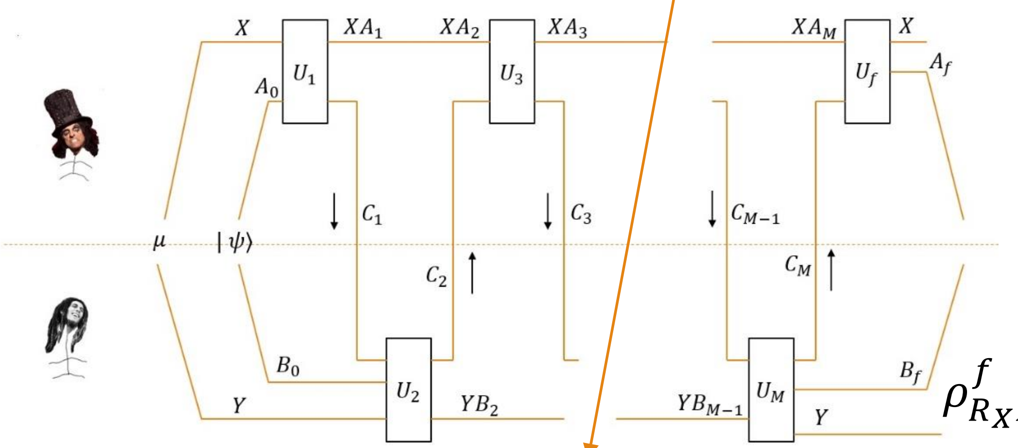
Tool 2: Superposition-Average Encoding Th.

- Average encoding theorem [KNTZ07]: $\mathbb{E}_X[h^2(\rho_B^X, \rho_B)] \leq I(X:B)_\rho$
  - $\rho_{XB} = \sum_x p_X(x)|x\rangle\langle x|_X \otimes \rho_B^x$
  - $\rho_B = \mathbb{E}_X[\rho_B^X]$, average state
  - $\mathrm{h}^2(\sigma, \theta) = 1 - F(\sigma, \theta)$, Bures distance, with $F(\sigma, \theta) = ||\sqrt{\sigma}\sqrt{\theta}||_1$
  - Follows from Pinsker's inequality
  - Many applications, e.g. together with a round-by-round variant of HIC [JRS03]

[Klauck, Nayak, Ta-Shma and Zuckerman 2007; Jain, Radhakrishnan and Sen 2003]

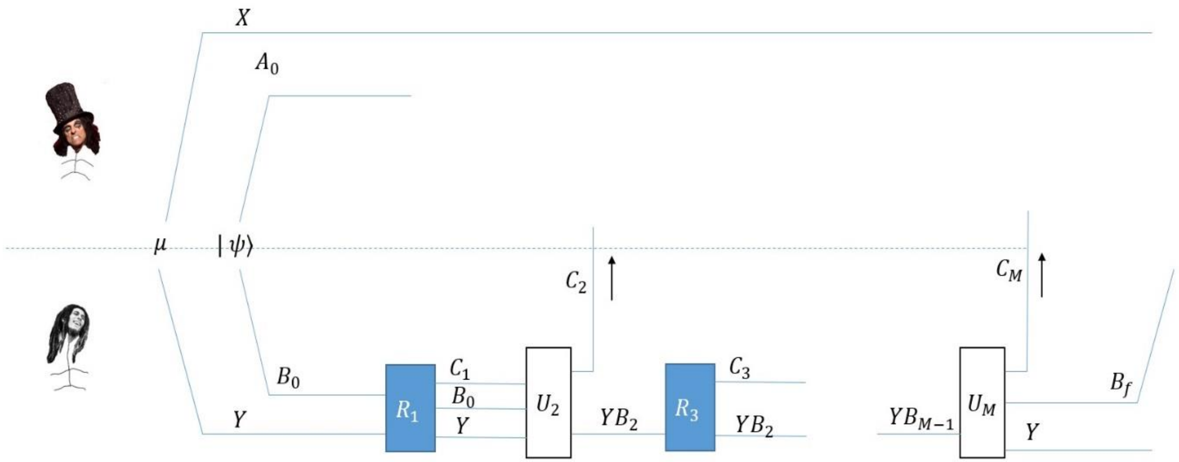
Tool 2: Superposition-Average Encoding Th.

- Average encoding theorem [KNTZ07]: $\mathbb{E}_X[h^2(\rho_B^X, \rho_B)] \leq I(X\!:B)_\rho$
- What about superposition over (part of) X?
- Recall F-R theorem (stated in terms of h)
  - There exists a recovery map $T_{B\to BC}$ such that $h^2(\rho_{RBC}, T_{B\to BC}(\rho_{RB})) \leq I(R\!:C|B)_\rho$
- Theorem: If for odd i then $h^2\left(\rho^f_{R_X R_Y Y B_f}, \sigma^f_{R_X R_Y Y B_f}\right) \leq M\ \sum_i \varepsilon_i$



$\rho^f_{R_X R_Y Y B_f}$

$I(R_X R_Y\!: C_i | Y\ B_i) = \varepsilon_i$



R_i: F-R maps

$\sigma^f_{R_X R_Y Y B_f}$

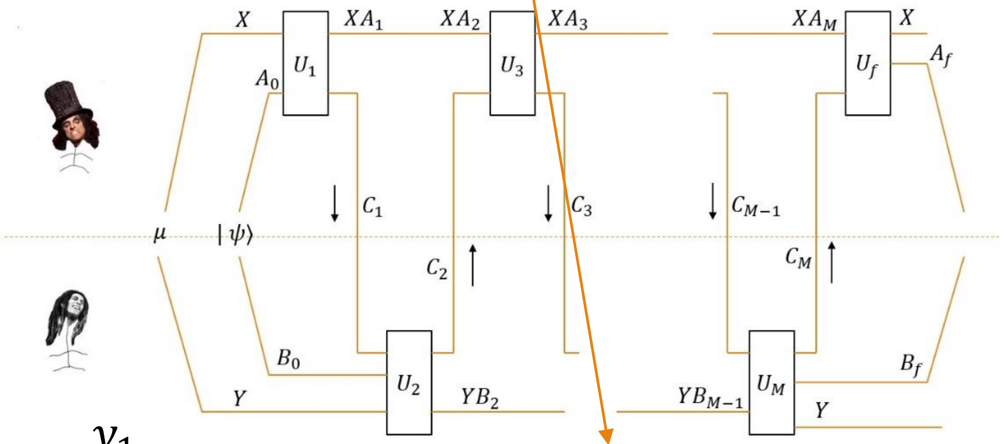
[Klauck, Nayak, Ta-Shma and Zuckerman 2007]

Tool 3: Quantum Cut-and-Paste Lemma

- Variant of a tool developed in [JRS03, JN14]
- Consider input subset $\{x_1, x_2\} \times \{y_1, y_2\}$
- Lemma: If for odd i and for even i, then $h\left(V_{B_t}^{y_1 \to y_2}(\rho^{t\ \ x_2 y_1}_{A_t B_t C_t})\rho^{t,x_2 y_2}_{A_t B_t C_t}\right) \leq 2\sum_{j \leq i} \delta_j$

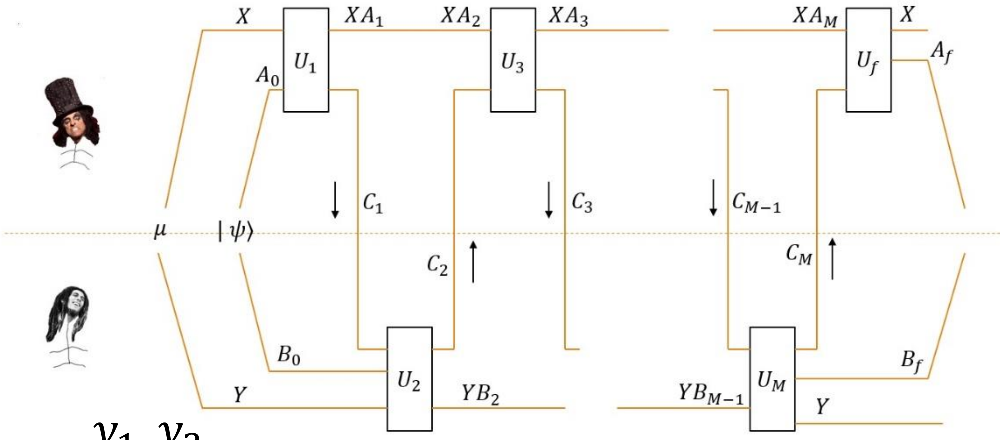
$$h\left(\rho^{i,\,x_1 y_1}_{A_i C_i,}\,\rho^{i,\,x_1 y_2}_{A_i C_i}\right) = \delta_i$$

x_1, x_2



y_1

x_1



y_1, y_2

$$h\left(\rho^{i,\,x_1 y_1}_{B_i C_i,}\,\rho^{i,\,x_2 y_1}_{B_i C_i}\right) = \delta_i$$

[Jain, Radhakrishnan and Sen 2003; Jain and Nayak 2014]

Applications

Th. 2: Disjointness

- Recall Disjointness: $x, y \subseteq [n]$, $Disj_n(x, y) =^? [x \cap y = \emptyset]$
- $\mathrm{CC}(Disj_n) \in \Omega(n)$, $\mathrm{QCC}(Disj_n) \in \Omega(\sqrt{n})$
- For r rounds, $\mathrm{QCC}^r(Disj_n) \in \widetilde{\Omega}(\frac{n}{r})$ [BGKMT15]
- Number of rounds r appears only through a continuity argument
  - Not there for classical protocols
  - Due to possibility of forgetting and retransmitting in quantum protocols
- With no-forgetting (NF), $\mathrm{QCC}^{NF}(Disj_n) \in \Omega\,(n)$

[Braverman, Garg, Kun Ko, Mao and T. 2015]

Th. 3: QIC and IC of Random functions

- Can we simulate classical protocols with quantum ones?
  - Of course!
  - What about maintaining IC?
  - Must be careful with private randomness
  - Bring Π_C in canonical form first
  - Then QIC looks classical… almost!

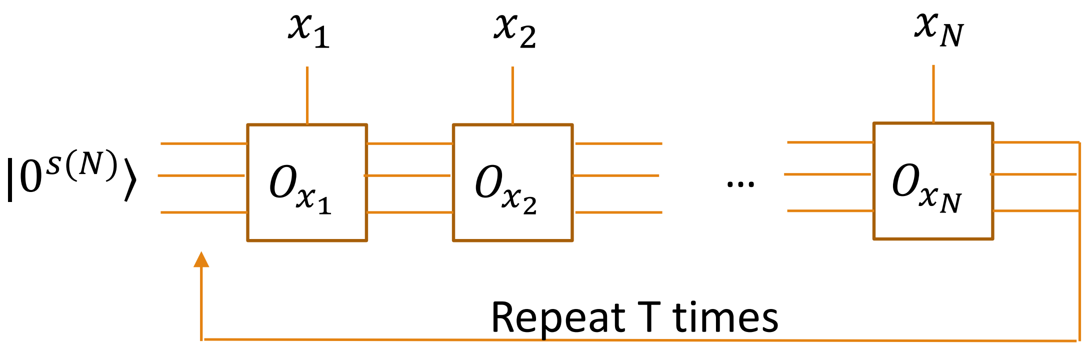
Th. 3: QIC and IC of Random functions

- Can we simulate classical protocols with quantum ones?
  - Of course!
  - What about maintaining IC?
  - Must be careful with private randomness
  - Bring Π_C in canonical form first
  - Then QIC looks classical… almost!
- Known: $QCC(IP_n) = n$ [CDNT99], $QCC(f_{rdm}) = n(1 - o(1))$ [MW07]
  - $IP_n(x, y) = \bigoplus_i x_i \wedge y_i$, f_{rdm} random function on n + n bits
  - Using Info. Flow Lemma, QCC lower bound transfers to QIC lower bound (at zero error)
  - Already known: $IC(IP_n) = n$ [BGPW], $IC(f_{rdm}) = \Omega(n)$ [BW]
- By above simulation, $IC(f_{rdm}) = n(1 - o(1))$

[Cleve, van Dam, Nielsen and Tapp 1999; Montanaro and Winter 2007;
Braverman, Garg, Pankratov and Weinstein 2013; Braverman and Weinstein 2012]

Th. 4: Streaming Algorithms for DYCK(2)

- $\mathrm{DYCK}(2) = \epsilon + [DYCK(2)] + \left(DYCK(2)\right) + DYCK(2) \cdot DYCK(2)$
- Reduction from multi-party QCC to streaming algorithm to DYCK(2) [MMN14]
  - Consider T-pass, one-way quantum streaming algorithms
  - Space s(N) in algorithm corresponds to communication between parties
  - Multi-party problem consists of OR of multiple instances of two-party problem



[Magniez, Mathieu and Nayak 2014]

Th. 4: Streaming Algorithms for DYCK(2)

- $\mathrm{DYCK}(2) = \epsilon + [DYCK(2)] + \left(DYCK(2)\right) + DYCK(2) \cdot DYCK(2)$
- Reduction from multi-party QCC to streaming algorithm to DYCK(2) [MMN14]
  - Consider T-pass, one-way quantum streaming algorithms
  - Space s(N) in algorithm corresponds to communication between parties
  - Multi-party problem consists of OR of multiple instances of two-party problem
- Direct sum argument allows to reduce from a two-party problem
  - Multi-party QCC lower bounds requires two-party QIC lower bound on "easy distribution"
- Th. 2.1: Any T-pass 1-way qu. streaming algo. for DYCK(2) needs space $s(N) \in \Omega(\frac{\sqrt{N}}{T^3})$ on length N inputs

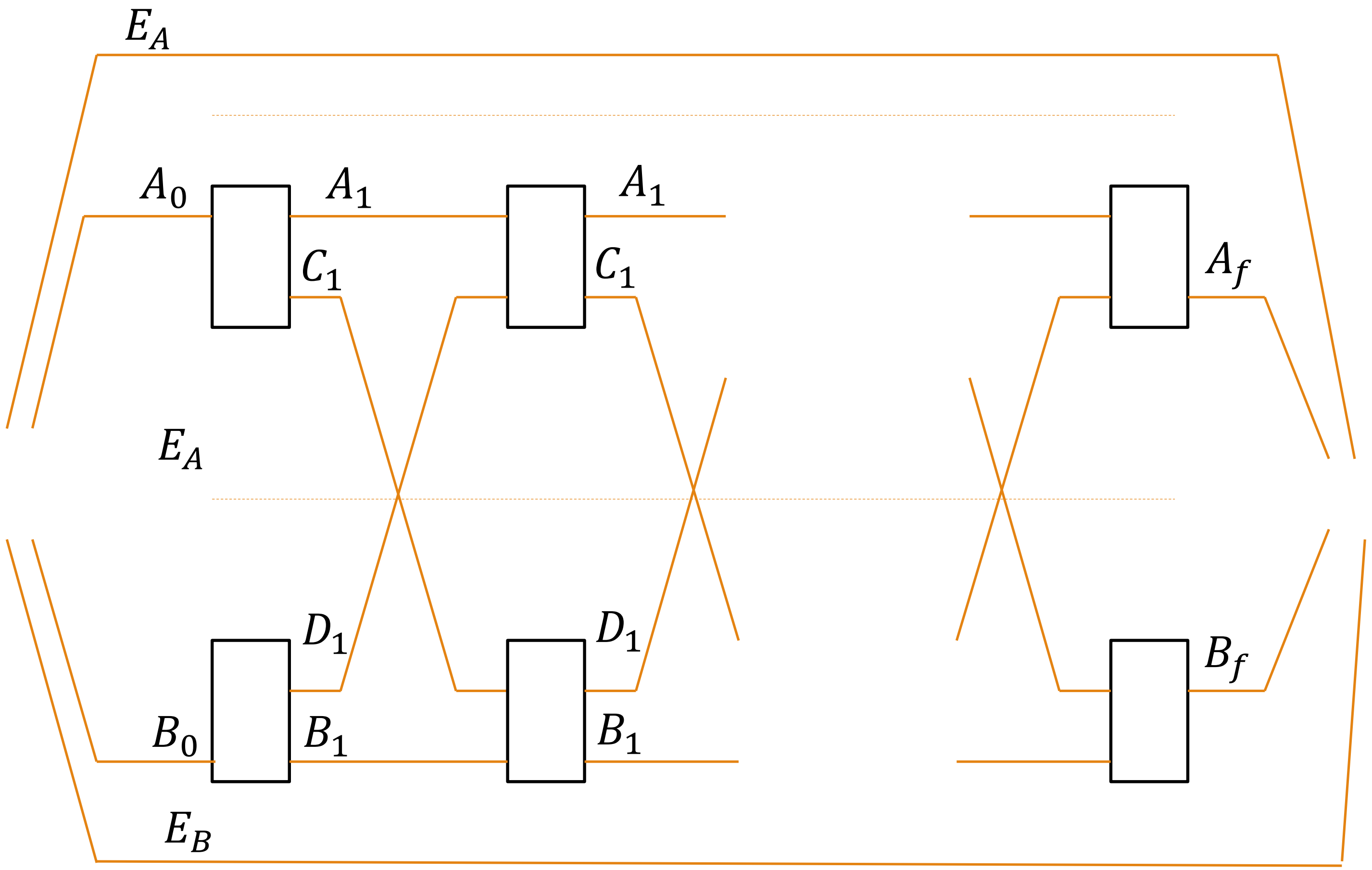
Th. 5: Augmented Index

- Index$(x_1 \dots x_i \dots x_n, i) = x_i$
- Augmented Index: $AI_n\big(x_1 \dots x_n, (i, x_1 \dots x_{<i}, b)\big) = x_i \oplus b$
- Th. 2.2: For any r-round protocol Π for AI_n, either
  - $QIC_{A \to B}(\Pi, \mu_0) \in \Omega\left(\frac{n}{r^2}\right)$ or
  - $QIC_{B \to A}(\Pi, \mu_0) \in \Omega\left(\frac{1}{r^2}\right)$ with
  - μ_0 the uniform distribution on zeros of AI_n ("easy distribution")
- Builds on direct sum approach of [JN14]
- General approach uses Tools 2, 3 (Sup.-Average Encoding Th., Qu. Cut-and-Paste)
- More specialized approach uses Tool 1 (Info. Flow Lemma)

Outlook

- Information-Theoretic Tools for Interactive Quantum Protocols
  - Information Flow Lemma
  - Superposition-average encoding theorem
  - Quantum Cut-and-Paste Lemma
- Applications
  - Intuitive interpretation of QIC, links with CIC, HIC (and other notions)
  - Forgetting an essential feature of quantum protocols for Disjointness
  - Quantum simulation of classical protocols leads to n(1-o(1)) lower bound on IC of random functions
  - Space lower bound on quantum streaming algorithms for DYCK(2)
  - Quantum information trade-off for Augmented Index
  - Further applications..?

V2: Information Flow Lemma

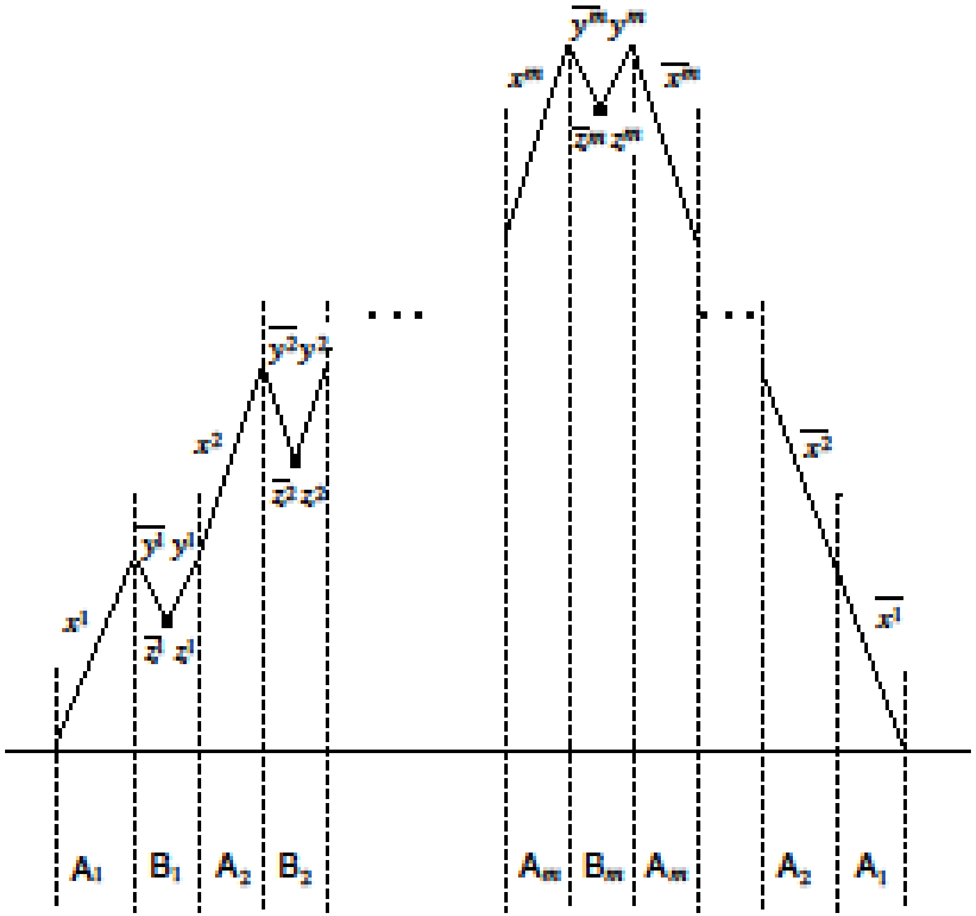


$$I\left(E_A{:}\,B_f|E_B\right) - I(E_A{:}\,B_0|E_B) =$$

$$\sum_i I(E_A{:}\,C_i|E_B B_i)$$

$$-\sum_i I(E_A{:}\,D_i|E_B B_i)$$

ASCENSION



[MMN14]

[Magniez, Mathieu and Nayak 2014]