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: \(x, y \subseteq \{1,2, ..., n\}\), 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
- \(\text{QCC}(\text{Disj}) = \Theta(\sqrt{n})\) [BCW98, Razb03, AA03]
- \(\text{CC}(\text{Disj}) = \Omega(n)\) [KS92]

Input: \(x\)  
Initialize: \(\frac{1}{n} \sum_i |i\rangle\)
Oracle call: \(\frac{1}{n} \sum_i |i\rangle x_i\rangle\)
\(\frac{1}{n} \sum_i (-1)^{x_i \land y_i} |i\rangle\)
Inversion about the mean
Repeat \(\approx \sqrt{n}\) times
Measure to get desired \(i\) if intersection

Input: \(y\)

[Buhrman, Cleve and Wigderson 1998; Razborov 2003; Aaronson and Ambainis 2003; Kalyanasundaram and Schnitger 1992]
Quantum Advantage for Disjointness

- Disjointness: $x, y \subseteq \{1, 2, ..., n\}$, 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
- $\text{QCC(Disj)} = \Theta(\sqrt{n})$ [BCW98, Razb03, AA03]
- $\text{CC(Disj)} = \Omega(n)$ [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 $(x, y) \sim \mu$
  - Take information-theoretic view: Information Complexity
    - How much information to compute $f$ on $(x, y) \sim \mu$
    - 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) ∈ Ω(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) ∈ \Omega(N^{\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) ∈ \Omega\left(\frac{n}{t^2}\right) or QIC_{B\rightarrow A}(\Pi, \mu_0) ∈ \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 $\Pi$: 

\begin{align*}
\mu \quad |\psi\rangle \\
\downarrow \\
A_0 \quad U_1 \\
\downarrow \\
X A_1 \quad X A_2 \quad X A_3 \\
\downarrow \\
C_1 \quad C_2 \quad C_3 \\
\downarrow \\
C_{M-1} \quad C_M \\
\downarrow \\
B_0 \quad U_2 \\
\downarrow \\
Y B_2 \quad Y B_{M-1} \\
\downarrow \\
Y \quad B_f \\
\downarrow \\
U_f \\
\downarrow \\
A_f \\
\downarrow \\
\text{Output: } f(X,Y)
\end{align*}
Quantum Communication Complexity

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

Protocol $\Pi$: [Diagram showing quantum states and operations]
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 $T_{B \rightarrow BC}$ such that $-\log F(\rho_{RBC}, T_{B \rightarrow BC}(\rho_{RB})) \leq I(R: C | B)_{\rho}$

[Lieb and Ruskai 1973; Devetak and Yard 2008; Yard and Devetak 2009; Fawzi and Renner 2015]
Quantum Information Complexity (QIC)

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

![Quantum Circuit Diagram]

Referee

$|\psi\rangle_{ABCR}$
Quantum Information Complexity (QIC)

- $\text{QIC}(f, \mu) = \inf_{\Pi} \text{QIC}(\Pi, \mu)$
- Optimization over all $\Pi$ computing $f$
- $\text{QIC}(\Pi, \mu) = \sum_{i \text{ odd}} I( R_X R_Y : C_i | Y B_i ) + \sum_{i \text{ even}} I( R_X R_Y : C_i | X A_i )$
- Properties [T15]:
  - Information equals amortized communication
  - Additivity
  - $\text{QIC} \leq \text{QCC}$
  - 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?
  - \( \text{HIC}(\Pi, \mu) = I(X: B_f | Y) + I(Y: A_f | X) \)
  - Compare to classical \( \text{IC}(\Pi_C, \mu) = I(X: \Pi_C | Y) + I(Y: \Pi_C | X) \), with \( \Pi_C = M_1 M_2 \ldots \) the transcript of messages
  - But reversible computing makes \( \text{HIC}(f) \) 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?
  - $\text{HIC}(\Pi, \mu) = I(X:B_f | Y) + I(Y:A_f | X)$
  - Compare to classical $\text{IC}(\Pi_C, \mu) = I(X: \Pi_C | Y) + I(Y: \Pi_C | X)$, with $\Pi_C = M_1 M_2 \cdots$ the transcript of messages
  - But reversible computing makes HIC(f) trivial...
- Can we measure only new classical information?
  - $\text{CIC}(\Pi, \mu) = \sum_{l \ odd} I(X:C_l | Y B_l) + \sum_{l \ even} I(Y:C_l | X A_l)$ [$\text{KLLGR16}$]
  - Compare to classical $\text{IC}(\Pi_C, \mu) = \sum_{l \ odd} I(X: M_l | Y M_{<l}) + \sum_{l \ even} I(Y: M_l | X M_{<l})$
  - 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?
  - \( \text{HIC}(\Pi, \mu) = I(X:B_f|Y) + I(Y:A_f|X) \)
  - Compare to classical \( \text{IC}(\Pi_C, \mu) = I(X: \Pi_C|Y) + I(Y: \Pi_C|X) \), with \( \Pi_C = M_1M_2 \cdots \) the transcript of messages
  - But reversible computing makes \( \text{HIC}(f) \) trivial...

- Can we simply measure new classical information?
  - \( \text{CIC}(\Pi, \mu) = \sum_{l \text{ odd}} I(X:C_l|YB_l) + \sum_{l \text{ even}} I(Y:C_l|XA_l) \) [KLLGR16]
  - Compare to classical \( \text{IC}(\Pi_C, \mu) = \sum_{l \text{ odd}} I(X:M_l|YM_{<l}) + \sum_{l \text{ even}} I(Y:M_l|XM_{<l}) \)
  - Motivated by privacy concerns
  - \( \text{HIC}(\Pi, \mu) \leq \text{CIC}(\Pi, \mu) \leq \text{QIC}(\Pi, \mu) \)
  - Is there a deeper relationship?

[Kerenidis, Lauriere, Le Gall and Rennela 2016]
Tool 1: Information Flow Lemma

- Lemma: $I(X: YB_f) - I(X: Y) = I(X: B_f | Y) = \sum_{i \text{ odd}} I(X: C_i | YB_i) - \sum_{i \text{ even}} I(X: C_i | YB_i)$
- Can also handle fully quantum processes and arbitrary extension of inputs
Th. 1: Cost of Forgetting

- Rewrite QIC(\(\Pi, \mu\)) = \(\sum_i I(X; C_i | YB_i) + I(Y; C_i | XA_i)\)
- What are those extra terms compared to CIC?
- CRIC(\(\Pi, \mu\)) = \(\sum_{i \text{ even}} I(X; C_i | YB_i) + \sum_{i \text{ odd}} I(Y; C_i | XA_i)\)
Th. 1: Cost of Forgetting

- Rewrite $\text{QIC}(\Pi, \mu) = \sum_i I(X:C_i|YB_i) + I(Y:C_i|XA_i)$
  - What are those extra terms compared to CIC?
  - CRIC$(\Pi, \mu) = \sum_{\text{even}} I(X:C_i|YB_i) + \sum_{\text{odd}} I(Y:C_i|XA_i)$
- Using Info. Flow Lemma, rewrite
  - Th. 1.1: HIC$(\Pi, \mu) = \text{CIC}(\Pi, \mu) - \text{CRIC}(\Pi, \mu)$
  - $\text{QIC}(\Pi, \mu) = \text{CIC}(\Pi, \mu) + \text{CRIC}(\Pi, \mu)$
- CRIC corresponds to cost of forgetting
  - Exactly assess back-flow of information
  - No need to introduce purification registers $R_XR_Y$ to define QIC (for classical tasks)
Tool 2: Superposition-Average Encoding Th.

- Average encoding theorem [KNTZ07]: \( E_X [h^2(\rho_B^X, \rho_B)] \leq I(X:B)_\rho \)
  - \( \rho_{XB} = \sum_x p_X(x) |x\rangle\langle x| \otimes \rho_B^x \)
  - \( \rho_B = E_X [\rho_B^X] \), average state
  - \( h^2(\sigma, \theta) = 1 - F(\sigma, \theta) \), 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]: $E_X[h^2(\rho_B^X, \rho_B)] \leq I(X:B)_\rho$
- What about superposition over (part of) $X$?
- Recall F-R theorem (stated in terms of $h$)
  - There exists a recovery map $T_{B\rightarrow BC}$ such that $h^2(\rho_{RBC}, T_{B\rightarrow BC}(\rho_{RB})) \leq I(R:C|B)_\rho$
- Theorem: If for odd $i$ then $h^2(\rho_{RXR_YYB_f}^f, \sigma_{RXR_YYB_f}^f) \leq M \sum_i \varepsilon_i$

\[ I(R_XR_Y:C_i|YB_i) = \varepsilon_i \]

[R]: F-R maps

[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
  \[
  h \left( V_{B_t}^{y_1 \rightarrow y_2} (\rho_{A_t B_t C_t}^{t, x_2 y_1} ) \rho_{A_t B_t C_t}^{t, x_2 y_2} \right) \leq 2 \sum_{j \leq i} \delta_j
  \]
  \[
  h \left( \rho_{A_i C_i}^i, \rho_{A_i C_i}^i, x_1 y_2 \right) = \delta_i
  \]
Applications
Th. 2: Disjointness

- Recall Disjointness: $x, y \subseteq [n], Disj_n(x, y) = ? [x \cap y = \emptyset]$
- $CC(Disj_n) \in \Omega(n), QCC(Disj_n) \in \Omega(\sqrt{n})$
- For $r$ rounds, $QCC^r(Disj_n) \in \tilde{\Omega}(\frac{n}{r})$ [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), $QCC^{NF}(Disj_n) \in \Omega(n)$

[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 $\Pi_C$ 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 $\Pi_C$ in canonical form first
  - Then QIC looks classical... almost!

- Known: $QCC(IP_n) = n$ [CDNT99], $QCC(f_{rdm}) = n(1 - o(1))$ [MW07]
  - $IP_n(x, y) = \oplus_i x_i \land y_i$, $f_{rdm}$ 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)

- DYCK(2) = ε + [DYCK(2)] + (DYCK(2)) + DYCK(2) · 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

[Mag niez, Mathieu and Nayak 2014]
Th. 4: Streaming Algorithms for DYCK(2)

- \( \text{DYCK}(2) = \epsilon + [\text{DYCK}(2)] + (\text{DYCK}(2)) + \text{DYCK}(2) \cdot \text{DYCK}(2) \)
- 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

[Magniez, Mathieu and Nayak 2014]
Th. 5: Augmented Index

- Index\( (x_1 \ldots x_i \ldots x_n, i) = x_i \)
- Augmented Index: \( AI_n(x_1 \ldots x_n, (i, x_1 \ldots x_{<i}, b)) = x_i \oplus b \)
- Th. 2.2: For any \( r \)-round protocol \( \Pi \) for \( AI_n \), either
  - \( QIC_{A\rightarrow B}(\Pi, \mu_0) \in \Omega\left( \frac{n}{r^2} \right) \) or
  - \( QIC_{B\rightarrow A}(\Pi, \mu_0) \in \Omega\left( \frac{1}{r^2} \right) \)
  - \( \mu_0 \) 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)

[Jain and Nayak 2014]
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

\[ I(E_A:B_f|E_B) - I(E_A:B_0|E_B) = \sum_i I(E_A:C_i|E_BB_i) - \sum_i I(E_A:D_i|E_BB_i) \]
ASCENSION

[MMN14]

[Magniez, Mathieu and Nayak 2014]