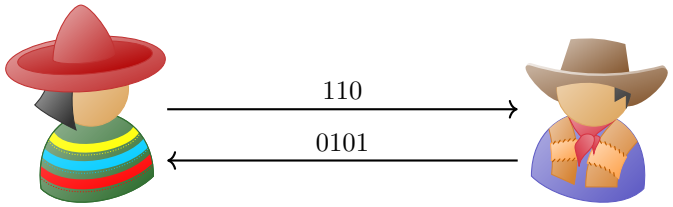
Quantum Communication Complexity

Seyed Sajjad Nezhadi, Dmitry Paramonov, Akash Rakheja

December 7, 2018

Communication Complexity



Communication complexity is an extensively studied concrete measure of complexity for boolean functions, with rich connections to many areas of Mathematics and Computer science including Game theory, Data structures, Proof complexity, Differential Privacy, and Circuit Lower Bounds.

Let Alice and Bob be two players who have access to sets X, Y respectively (typically $X = Y = \{0,1\}^n$) and a boolean function $f : X \times Y \to \{0,1\}$. They would like to evaluate the expression $f(x,y)$ where $x \in X$ and $y \in Y$. Bob may simply call up Alice and tell her y and Alice then proceeds to compute $f(x,y)$. Let us consider though a situation where Alice and Bob have access to supercomputers but they live very far away from each other (separate galaxies?) and would like to minimize the messages they send between them. The trivial method requires Bob to communicate all n bits of his input(maximum communication). Could we do better? Luckily for many functions, we can!

A communication protocol π is an algorithm previously agreed upon by the two players, instructing them to send information based on previous messages. After some number of bits have been communicated using the protocol we would like the function to have been evaluated by one of the players.

In the randomized setting, we allow Alice and Bob to have access to public(shared between players) or private(individual access) coin flips. Therefore a randomized protocol π depends both on previous messages and coin flips and must evaluate $f(x,y)$ with probability $\geq 2/3$.

$C(f,\pi) = max_{x \in X, y \in Y}($ bits communicated to evaluate $f(x,y)$ using $\pi)$ is the maximum communication cost over all possible inputs evaluated on $f : X \times Y \to \{0,1\}$ using the protocol π. Then the deterministic communication complexity of the boolean function f is denoted $D(f) = min_\pi C(f,\pi)$ the minimum communication over all deterministic protocols. And the Public and Private Randomized Communication complexity are denoted $R_{pub}(f) = min_{\pi_{pub}} C(f, \pi_{pub})$, $R(f) = min_\pi C(f,\pi)$ the minimum communication over all randomized public and private protocols respectively.

Newman's $Theorem : R(f) \leq R_{pub}(f) + O(log(n))$.

Equality

The equality problem, $EQ(x, y) = 1$ if and only if $x = y$, is a classic example of the power of randomness in communication complexity. It can be shown that the deterministic communication complexity for equality is maximal, $D(EQ) = \Omega(n)$. But we can do much better with randomness.

Public randomness protocol:

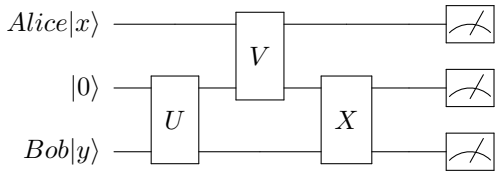
1. Alice and Bob generate a uniform random string $z \in \{0,1\}^n$
2. Alice sends the bit $(\sum_i x_i z_i) \bmod 2$
3. Bob compares the bit with $(\sum_i y_i z_i) \bmod 2$, and decides if $x = y$

   If $x = y$ then with probability 1 they decide correctly.

   if $x \neq y$ with probability $1/2$ they decide correctly.

We can decide the equality problem with probability $\geq 2/3$ by running the above protocol twice. Therefore $R_{pub}(EQ) = O(1)$, and by applying Newman's Theorem we get $R(EQ) = O(log(n))$. Therefore we have an exponential separation between randomized and deterministic communication complexity!

Quantum Communication Models



Equality is one of many functions that induce separations between randomized and deterministic communication complexity, one can ask the same question of quantum information. Can Alice and Bob use less communication to evaluate their function if the had access to a quantum communication channel and/or entanglement?

The model introduced by Yao[4] allows Alice and Bob to have access to the quantum states $|x\rangle$ and $|y\rangle$ respectively. They will also share access to an intermediary quantum state (communication channel) initialized to $|0\rangle$. A protocol is a quantum circuit, where Alice and Bob apply unitary transformations to their state and the channel, computing $f(x, y)$ w.h.p (we may also restrict it to be deterministic) by measuring the result of the circuit. The communication cost of the protocol π for f, $C(f, \pi) =$ 'number of channel qubits affected by each of the gates in the circuit'. The Quantum communication complexity of the boolean function f is calculated as $Q(f) = min_\pi C(f, \pi)$.

Combining Yao's model and the model introduced by Cleve and Buhrman[5] we get the more powerful setting where Alice and Bob also have access to an unlimited supply of entanglement(shared EPR pairs). In the next two sections we will showcase some separations between the quantum and randomized communication complexities.

Disjointness

The Disjointness function $DISJ : \{0,1\}^{2n} \to \{0,1\}$ is defined as $DISJ(x, y) = 1$ if and only if $\exists i \in [n] : x_i = y_i = 1$. It was shown that $R(DISJ) \in \Theta(n)$ by Razbarov[6], and $Q(DISJ) \in \Theta(\sqrt{n})$ by Aaronson and

Ambainis[7], Razbarov[8]. Therefore, DISJ gives us a quadratic separation between quantum and randomized communication complexity!

We will present the quantum protocol by Buhrman, Cleve and Widgerson[9]. If Alice had access to y she could solve $DISJ(x, y)$ by running Grover's algorithm on $Z = x \wedge y$ (bit-wise). We will take this observation and have Alice independently run Grover's algorithm, calling upon Bobs assistance only when we query Z during a Grover Iterate.

Protocol ($k = log(n)$, R Phase Shift, U XOR Oracle, $CCNOT$ Toffoli Gate):

1. Alice would like to apply the oracle on Z to her state $|\psi\rangle = \frac{1}{\sqrt{n}} \sum_{i\in\{0,1\}^k} |i\rangle$
2. Alice sends $|\psi_2\rangle = U_{x_{(1,2)}}|\psi\rangle_1|0\rangle_2|0\rangle_3|0\rangle_4 = \frac{1}{\sqrt{n}} \sum_{i\in\{0,1\}^k} |i\rangle_1|x_i\rangle_2|0\rangle_3|0\rangle_4$
3. Bob constructs $|\psi_3\rangle = U_{y_{(1,3)}}|\psi_2\rangle = \frac{1}{\sqrt{n}} \sum_{i\in\{0,1\}^k} |i\rangle_1|x_i\rangle_2|y_i\rangle_3|0\rangle_4$
4. Bob constructs $|\psi_4\rangle = CCNOT_{(2,3,4)}|\psi_3\rangle = \frac{1}{\sqrt{n}} \sum_{i\in\{0,1\}^k} |i\rangle_1 CCNOT|x_i\rangle_2|y_i\rangle_3|0\rangle_4$
   $= \frac{1}{\sqrt{n}} \sum_{i\in\{0,1\}^k} |i\rangle_1|x_i\rangle_2|y_i\rangle_3|x_i y_i\rangle_4$
5. Bob constructs $|\psi_5\rangle = R_{\pi_{(4)}}|\psi_4\rangle = \frac{1}{\sqrt{n}} \sum_{i\in\{0,1\}^k} |i\rangle_1|x_i\rangle_2|y_i\rangle_3 R_\pi|x_i y_i\rangle_4 = \frac{1}{\sqrt{n}} \sum_{i\in\{0,1\}^k} (-1)^{x_i y_i}|i\rangle_1|x_i\rangle_2|y_i\rangle_3|x_i y_i\rangle_4$
6. Bob sends back $|\psi_6\rangle = U_{y_{(1,3)}} \cdot CCNOT_{(2,3,4)}|\psi_5\rangle = \frac{1}{\sqrt{n}} \sum_{i\in\{0,1\}^k} (-1)^{x_i y_i}|i\rangle_1|x_i\rangle_2|y_i \oplus y_i\rangle_3|x_i y_i \oplus x_i y_i\rangle_4$
   $= \frac{1}{\sqrt{n}} \sum_{i\in\{0,1\}^k} (-1)^{x_i y_i}|i\rangle_1|x_i\rangle_2|0\rangle_3|0\rangle_4$
7. Alice constructs $|\psi_7\rangle = U_{x_{(1,2)}}|\psi_6\rangle = \frac{1}{\sqrt{n}} \sum_{i\in\{0,1\}^k} (-1)^{Z_i}|i\rangle_1|x_i \oplus x_i\rangle_2|0\rangle_3|0\rangle_4 = \frac{1}{\sqrt{n}} \sum_{i\in\{0,1\}^k} (-1)^{Z_i}|i\rangle_1|0\rangle_2|0\rangle_3|0\rangle_4$
8. Alice has therefore applied the oracle to her state!

Since Grover's algorithm has $O(\sqrt{n})$ Grover iterates and each time we do Grover iterate we communicate $O(log(n))$ qubits $\implies$ $Q(DISJ) \in O(log(n)\sqrt{n})$. While this is not the optimal protocol it still separates the quantum complexity from the randomized complexity.

Vector in Subspace

Now, we describe a problem for which the gap between quantum and classical communication complexity is exponential.

The problem is described as follows:

- Alice is given an m-dimensional vector $v \in \mathbb{R}^m$
- Bob is given two projection operators P_0, P_1 from $\mathbb{R}^m \to \mathbb{R}^m$ such that $P_0 + P_1 = I$
- It is given that either $P_0 v = v$ or $P_1 v = v$
- The task is to find out whether $P_0 v = v$ or $P_1 v = v$

The problem involves continuous inputs as any entry can be any real number but it can be discretized by approximating each entry using $O(\log m)$ bits. Thus, the total input size to the problem is $n = O(m^2 \log m)$ bits. A simple quantum protocol that solves the problem is such that Alice views her input v as a $\log m$-qubits quantum state and sends it to Bob; Bob then measures with operators P_0 and P_1, and outputs the result. Therefore we have $Q(VSP) \in O(log(m)) = O(log(n))$.

A similar technique is not possible in the classical case. In fact it has been shown that $R(VSP) \in \Omega(n^{1/3})$ [14]. Thus, there is an exponential separation between the quantum and randomized communication complexity for this problem.

Quantum Communication and Classical Information

Information Theory Preliminaries

Definition 1. *The Shannon Entropy H(X) of a random variable X is defined as* $H(X) = E_x[log(\frac{1}{Pr(x)})] = \sum_x [Pr(x) log(\frac{1}{Pr(x)})]$

Definition 2. *the Conditional Entropy of X given Y (X, Y random variables) is defined as* $H(X|Y) = E_y[H(X|Y = y)]$

Chain rule : $H(X, Y) = H(X) + H(Y|X)$

Definition 3. *The mutual information of random variables X, Y is defined as* $I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X)$ *with conditional mutual information* $I(X;Y|Z) = H(X|Z) - H(X|Z,Y)$

Information Complexity

Now we may define the notion of Information Complexity of a boolean function with domain $X \times Y$ and bounded-error ϵ, due to Braverman[12].

Definition 4. *Information Complexity* $IC_\epsilon(F) = Min_\pi Max_\mu IC_\mu(\pi)$, *where* π *is a protocol with error* $\leq \epsilon$, *and* μ *is a distribution on the domain* $X \times Y$.

Definition 5. *The information cost of a protocol* π, $IC_\mu(\pi) = I(\pi; X|Y) + I(\pi; Y|X)$

Information cost allows us to measure the information Alice and Bob learn about each other's input during the protocol π.

Information Complexity IC is one of the most important tools for proving lower bounds for classical and randomized communication complexity. A natural question is how this complexity measure compares to quantum communication? Kerenidis, Laplante, Lerays, Roland and Xiao[10] showed that the Information complexity for the Vector in subspace problem $IC(VSP) \in \Omega(n^{1/3})$. Therefore the Information Complexity of the Vector in subspace problem is exponentially larger than its Quantum communication complexity.

Also, Anshu, Touchette, Yao and Yu[13] showed that the Symmetric k-ary Pointer Jumping function by Rao and Sinha[11] has Quantum communication complexity exponentially larger than its Information complexity.

Theorem 1. *([13]) There exists a (family of) Boolean function f and a distribution* μ *on the domain* $X \times Y$ *such that* $Q(f, \mu, 1/3) \geq 2^{\Omega(IC(f,\mu,1/3))} \geq 2^{\Omega(QIC(f,\mu,1/3))}$

Here QIC is the quantum analogue of information complexity, presented in the next section. Therefore we have that the two complexity measures of Information Complexity and Quantum Communication Complexity are incomparable!

In the same paper they showed that there exists a (tight) communication trade off for any quantum protocol computing the Greater-Then function, as a consequence of their proofs.

Theorem 2. *([13]) Any quantum protocol computing Greater-Than with error* $\leq 1/3$, *requires Alice to communicate* $\frac{n}{2^{O(b)}}$ *bits to Bob, where b is number of bits sent by Bob.*

Here $Greater - Than(x, y) = 1$ if and only if $x \geq y$.

Quantum Information and Amortized Communication

We consider the model for quantum communication in which we are given quantum registers A and B having associated set of mixed states as $\mathcal{D}(A)$ and $\mathcal{D}(B)$ respectively, and a channel from A to B is denoted by $\mathcal{N}^{A\to B}$ which takes a mixed state in A to a mixed state in B. $\mathcal{C}(A, B)$ is the set of all channels from A to B and $\mathcal{U}(A, B)$ is the set of all unitary channels. $Tr_{\neg A}(.) = Tr_B(.) \in C(A \otimes B, B)$ is an operator from

$A \otimes B$ to A that tells mixed state in $A \otimes B$ seen as mixed state in A. For a state $\rho^A \in \mathcal{D}(A)$, the purification is a pure state ρ^{AR} for some reference system R satisfying $Tr_R(\rho^{AR}) = \rho^A$. The distance between two states ρ_1 and ρ_2 in $\mathcal{D}(A)$ is defined to be $||\rho_1 - \rho_2||_A = Tr(|\rho_1 - \rho_2|)$.

Uncertainty or information in state $\rho \in \mathcal{D}(A)$ is defined as $H(A)_\rho = Tr(\rho \log_2 \rho)$. If $\rho = \Sigma_j \eta_j |j\rangle\langle j|$, then $H(A)_\rho = \Sigma_j \eta_j \log_2 \eta_j$. We define $0 \log_2 0$ as 0.

For pure bipartite state ρ^{AB} we have that $H(A) = H(B)$ as it should be since the amount of information is same as seen in A or B. For isomorphic A and B, and ρ^{AB} maximally entangled; $H(A) = \log_2 dim(A)$.

Now, the problem that we are trying to solve is that we are given a communication channel $\mathcal{N} \in C(A_{in} \otimes B_{in}, A_{out} \otimes B_{out})$, an input state $\rho \in \mathcal{D}(A_{in} \otimes B_{in})$, Alice has registers A_{in}, A_{out} and Bob has registers B_{in}, B_{out} and they have to compute state $\mathcal{N}(\rho)$.

A protocol π for implementing $\mathcal{N}$ is given by a set of unitaries $U_i{}_{i=1}^{M+1}$ Alice and Bob compute by themselves along with a pure state they share, $\psi \in \mathcal{D}(T_A \otimes T_B)$ where T_A and T_B are of arbitrary size belonging to Alice and Bob respectively. For appropriate finite dimensional registers $A_1, A_3, \ldots, A_{M-1}, A'$ belonging to Alice and $B_2, B_4, \ldots, B_{M-2}, B'$ belonging to Bob and communication registers $C_{i}{}_{i=1}^{M}$; we have $U_1 \in \mathcal{U}(A_{in} \otimes T_A, A_1 \otimes C_1)$, $U_2 \in \mathcal{U}(B_{in} \otimes T_B \otimes C_1, B_2 \otimes C_2)$, $U_3 \in \mathcal{U}(A_1 \otimes C_2, A_3 \otimes C_3)$, $U_4 \in \mathcal{U}(B_2 \otimes C_3, B_4 \otimes C_4)$,..., $U_M \in \mathcal{U}(B_{M-2} \otimes C_{M-1}, B_{out} \otimes B' \otimes C_M)$, $U_{M+1} \in \mathcal{U}(A_{M-1} \otimes C_M, A_{out} \otimes A')$. We also denote the channel implemented by the protocol as π.

$$\pi(\rho) = Tr_{A'B'}(U_{M+1} U_M \cdots U_2 U_1 (\rho \otimes \psi))$$

If $\rho^{A_{in}B_{in}}$ has purification $\rho_{A_{in}B_{in}R}$, we say that protocol π has error $\epsilon \in [0, 2]$ if

$$||\pi(\rho) - \mathcal{N}(\rho)||_{A_{out}B_{out}R} \leq \epsilon$$

The set of all protocols implementing $\mathcal{N}, \rho$ having error at most ϵ is denoted by $\mathcal{T}(\mathcal{N}, \rho, \epsilon)$. If such protocols are restricted in number of communication registers to M, the set is denoted by $\mathcal{T}^M(\mathcal{N}, \rho, \epsilon)$.

Quantum Communication Cost of protocol π is calculated as $Q(\pi) = \Sigma_i \log_2 dim(C_i)$ since $\log_2 dim(C_i)$ is the cost of communicating C_i. ϵ-error quantum communication complexity of $\mathcal{N}$ on input ρ is defined as $Q(\mathcal{N}, \rho, \epsilon) = \min_{\pi \in \mathcal{T}(\mathcal{N}, \rho, \epsilon)} Q(\pi)$.

Quantum Information Cost for protocol π, input state ρ is calculated as $QIC(\pi, \rho) = \Sigma_{i>0, odd} \frac{1}{2} I(C_i; R|B_{i-1}) + \Sigma_{i>0, even} \frac{1}{2} I(C_i; R|A_{i-1})$ where $B_0 = B_{in} \otimes T_B$. The quantum information cost roughly denotes the amount of information Alice and Bob gain about each other's input using the qubits transferred between them while the communication cost is the amount of qubits transferred. We take the information cost as this since it was proved in [1, 2] that the amortized cost of communicating C when qubits C are transferred from Alice to Bob where A is the feedback to Alice, B is the side information held by Bob and R is the reference system is $\frac{1}{2} I(C; R|B)$. Quantum Information Complexity is calculated as $QIC(\mathcal{N}, \rho, \epsilon) = \inf_{\pi \in \mathcal{T}(\mathcal{N}, \rho, \epsilon)} QIC(\pi, \rho)$.

Protocol π_n computes n-fold product channel $\mathcal{N}^{\otimes n}$ on $\rho^{\otimes n}$ with error ϵ if $\forall i \in [n]$, $||Tr_{\neg(A^i_{in}, B^i_{in}, R^i)} \circ \pi_n(\rho^{\otimes n}) - \mathcal{N}(\rho)||_{A^i_{out}B^i_{out}R^i} \leq \epsilon$. n-fold quantum communication complexity is denoted as $Q_n(\mathcal{N}^{\otimes n}, \rho^{\otimes n}, \epsilon)$. Amortized quantum communication complexity is calculated as

$$AQCC(\mathcal{N}, \rho, \epsilon) = \limsup_{n \to \infty} \frac{1}{n} Q_n(\mathcal{N}^{\otimes n}, \rho^{\otimes n}, \epsilon)$$

The following two inequalities follows from the fact that $\frac{1}{2} I(C; R|B) \leq \log_2 dim(C)$.

$$0 \leq QIC(\pi, \rho) \leq Q(\pi)$$
$$QIC(\mathcal{N}, \rho, \epsilon) \leq Q(\mathcal{N}, \rho, \epsilon)$$

Our goal here is to establish the fact that $QIC(\mathcal{N}, \rho, \epsilon) = AQCC(\mathcal{N}, \rho, \epsilon)$. This is intutively true since the information transferred is same as data transferred in amortized sense. Intutively, if we have 0.1 bits of information to be sent, we have to send 1 bit but if we send 0.1 bits of information ten times, we can send

it in nearly one bit by compression. This is only an intuition behind why the above equation is true but it needs a formal proof. Due to lack of space, we won't give a formal proof but would refer to [3]. However, the proof follows from the following main lemmas.

Lemma 1 ([3]). *For any M-message protocol π, any input state ρ and any $\epsilon \in (0, 2]$, $\delta > 0$, there exists a large enough n_0 such that for any $n \geq n_0$, there exists a protocol $\pi_n \in \mathcal{T}(\pi^{\otimes n}, \rho^{\otimes n}, \epsilon)$ satisfying*

$$\frac{1}{n} Q(\pi_n) \leq QIC(\pi, \rho) + \delta$$

Taking limit of n to infinity, δ to 0 and infimum over both sides gives us intutively that $AQCC(\mathcal{N}, \rho, \epsilon) \leq QIC(\mathcal{N}, \rho, \epsilon)$.

The following two lemmas are useful in proving the opposite direction.

Lemma 2 ([3]). *For any two protocols π^1, π^2 with M_1, M_2 messages, respectively, there exists a M-message protocol π_2, satisfying, $M = \max(M_1, M_2)$, such that the following holds for any corresponding input states ρ^1, ρ^2:*

$$QIC(\pi_2, \rho^1 \otimes \rho^2) = QIC(\pi^1, \rho^1) + QIC(\pi^2, \rho^2)$$

Lemma 3 ([3]). *For any M-message protocol π_2 and any input states $\rho^1 \in \mathcal{D}(A^1_{in} \otimes B^1_{in}), \rho^2 \in \mathcal{D}(A^2_{in} \otimes B^2_{in})$, there exists protocols π^1, π^2 satisfying $\pi^1(.) = Tr_{A^2_{out} B^2_{out}} \circ \pi_2(. \otimes \rho^2), \pi^2(.) = Tr_{A^1_{out} B^1_{out}} \circ \pi_2(\rho^1 \otimes .)$, and the following holds:*

$$QIC(\pi^1, \rho^1) + QIC(\pi^2, \rho^2) = QIC(\pi_2, \rho^1 \otimes \rho^2)$$

By using additivity proved in the previous two lemmas, we get that $nQIC(\mathcal{N}, \rho, \epsilon) = QIC_n(\mathcal{N}^{\otimes n}, \rho^{\otimes n}, \epsilon) \leq Q_n(\mathcal{N}^{\otimes n}, \rho^{\otimes n}, \epsilon)$. By dividing this inequality by n and taking limits as n goes to infinity, we get that $AQCC(\mathcal{N}, \rho, \epsilon) \geq QIC(\mathcal{N}, \rho, \epsilon)$.

These proofs presented are not meant to be formal and for the interested reader, we refer to [3].

Simultaneous Message Passing

Now, we wish to consider another area of communication. In the standard model, two players can continuously exchange information, in either direction. A more restrictive model is the *simultaneous message passing* (SMP) model. Here, Alice receives an input $x \in X$ and Bob receives an input $y \in Y$. Then, they each get to send one message to a third player, the referee. Following this, the referee must output $f(x, y)$ with high probability. In other words, Alice and Bob must both send their messages without knowing any information about the other's input, and they can only send messages in one batch, unable to get any responses back.

The deterministic communication complexity of f over a distribution μ on $X \times Y$ is denoted by $D^{\|}_{\mu,\varepsilon}$ and is the number of bits Alice and Bob need to send to the referee for the referee to output $f(x, y)$ with probability at least $1 - \varepsilon$ when (x, y) are sampled from μ. In this case, all algorithms and protocols are deterministic, with all randomness coming from sampling from μ.

The randomized communication complexity of a function $f : X \times Y \to \{0, 1\}$ is denoted by $R^{\|}(f)$ and is the number of bits Alice and Bob need to send to the referee for the referee to output $f(x, y)$ with probability $\frac{2}{3}$. Here, Alice and Bob may use randomness in their algorithms, as may the referee.

The quantum communication complexity $Q^{\|}(f)$ is defined analogously in terms of the qubits that Alice and Bob need to send.

Note that the above classes assume that Alice and Bob act independently. However, we can also assume that Alice and Bob might have access to shared randomness. This thus leads to variations in the communication complexity, leading to $R^{\|,pub}(f)$ and $Q^{\|,pub}(f)$.

We now wish to show an exponential separation between $R^{\|,pub}(f)$ and $Q^{\|}(f)$, for a certain function f.

To do this, we will construct variations on the basic equality problem. The most important is presented below. Given some set $T \subseteq \{0,1\}^n$, we define $\widetilde{cEq}\text{-}neg_T$ as the approximate equality problem under cyclic shifts and negations in T. This problem was originally defined in [15].

$$\widetilde{cEq}\text{-}neg_T(x,y) = \begin{cases} 1, & |\sigma_j(x) \oplus y \oplus \tau| \leq \frac{6n}{15} \text{ for some } \tau \in T \text{ and } j \in [n] \\ & \quad \text{and } |\sigma_j(x) \oplus y \oplus \tau| \notin \left(\frac{6n}{15}, \frac{7n}{15}\right) \forall \tau \in T, j \in [n] \\ 0, & |\sigma_j(x) \oplus y \oplus \tau| \geq \frac{7n}{15} \forall \tau \in T, j \in [n] \\ \text{undefined}, & \text{otherwise} \end{cases}$$

Here, $|x|$ denotes the Hamming weight of a vector x, and σ_j refers to the j'th cyclic shift.

We also define a distribution $\mu_{\widetilde{cEq}\text{-}neg_T}$ over $X \times Y$. This is defined as the distribution uniform over $j \in [n]$, uniform over $\tau \in T$, uniform over subsets $u \in \binom{[n]}{\frac{n}{3}}$ (the subsets of $[n]$ with cardinality $\frac{n}{3}$) such that $(\sigma_j(x) \oplus \tau)$ has equal probability of being equal to y on the indices in u.

Of particular importance will be T which are small-bias spaces.

Definition 6 (Small-Bias Space)**.** *A set $T \subseteq \{0,1\}^n$ is an ε-bias space if for every set of indices $S \subseteq [n]$ such that $S \neq \emptyset$*

$$\left| \mathop{\mathbb{E}}_{\tau \in T} \left[(-1)^{|\tau_S|} \right] \right| \leq \varepsilon$$

where τ_S is the vector obtained by taking only the indices of τ which lie in S.

It is possible to construct such small bias spaces efficiently in time poly $\left(\frac{n}{\varepsilon}\right)$. The resulting small bias space additionally satisfies the property that for every pair $\tau_1 \neq \tau_2 \in T$, $|\tau_1 \oplus \tau_2| = \frac{n}{2} + o(n)$ [16].

Now, let us demonstrate the communication complexity of $\widetilde{cEq}\text{-}neg_T$, in both regimes we care about.

Consider first $Q^{\|}\left(\widetilde{cEq}\text{-}neg_T\right)$. For now, fix some $j \in [n]$ and some $\tau \in T$. Consider the following protocol, specified in [15]. If Alice sends $\frac{1}{\sqrt{n}} \sum_{i=1}^n |i\rangle|x_i\rangle$ and Bob sends $\frac{1}{\sqrt{n}} \sum_{i=1}^n |i\rangle|y_i\rangle$, then the referee can transform Alice's state into $\frac{1}{\sqrt{n}} \sum_{i=1}^n |\sigma_j(i)\rangle|x_i \oplus \tau_{\sigma_j(i)}\rangle$ using only unitary operations. After this, the referee can apply the swap test to the received messages, and output the result. Thus, they will output 1 with probability $\frac{1}{2} + \frac{|\sigma_j(x) \oplus y \oplus \tau|}{2n}$. Thus, if Alice and Bob send $O(\log n)$ copies of the above superpositions, then the referee can estimate $|\sigma_j(x) \oplus y \oplus \tau|$ with $\frac{1}{\text{poly}(n)}$ accuracy.

Note also that the referee can "reuse" the messages for different τ and j. Thus, they can actually estimate $|\sigma_j(x) \oplus y \oplus \tau|$ for all $j \in [n]$ and all $\tau \in T$ simultaneously, using only $O\left((\log n)^2 + \log n \log |T|\right)$ messages. Thus, we see that $Q^{\|}\left(\widetilde{cEq}\text{-}neg_T\right) = O\left((\log n)^2\right)$, as long as $|T| = \text{poly}(n)$.

Meanwhile, in the classical case, there exists a family of small bias spaces $\mathcal{T} = T_1, T_2, \ldots$ such that for each i, $T_i \subseteq \{0,1\}^i$, T_i can be constructed in time poly (i), and $R^{\|,pub}\left(\widetilde{cEq}\text{-}neg_{T_i}\right) = \Omega\left(\frac{\sqrt{n}}{\log n}\right)$.

To see this, let n be sufficiently large and let $\delta = \Theta\left(\frac{1}{n}\right)$ be sufficiently small. Let T be a δ-biased space of size poly $\left(\frac{n}{\delta}\right)$. This can be constructed by our earlier claim.

Now, for any given protocol P, let $Al : \{0,1\}^n \to \{0,1\}^r$ be Alice's message function, mapping from inputs x to the message that Alice will send. Let $Bo : \{0,1\}^n \to \{0,1\}^q$ be defined likewise.

Thus, consider the following lemma.

Lemma 4 ([15])**.** *For sufficiently large n, there exists some $\delta = \Theta\left(\frac{1}{n}\right)$ such that for all δ-biased spaces T of size $2^{o(n)}$, there exists some $\epsilon = \Theta\left(\frac{1}{|T|n^2}\right)$ such that for all protocols P which solve $\widetilde{cEq}\text{-}neg_T$ in $D^{\|}_{\mu_{\widetilde{cEq}\text{-}neg_T},\epsilon}$ satisfies that*

$$\mathop{\mathbb{E}}_{i_1,u_1}\left[I(X_{i_1}|X_{u_1}; Al(X))\right] \mathop{\mathbb{E}}_{i_2,u_2}\left[I(Y_{i_2}|Y_{u_2}; Bo(Y))\right] \geq \frac{1}{2n}$$

where i_1, i_2 are sampled from $[n]$, u_1, u_2 are sampled from $\binom{[n]\setminus\{i_1\}}{\frac{2n}{3}}$ and $\binom{[n]\setminus\{i_2\}}{\frac{2n}{3}}$, respectively, and X and Y are sampled from $\{0,1\}^n$, with all samplings being uniform over their domains.

This lemma, when used alongside the construction mentioned above, gives us that there exist T such that $D^{\|}_{\mu_{\widetilde{cEq}\text{-}neg_T},\frac{1}{3}}\left(\widetilde{cEq}\text{-}neg_T\right)$ takes $\Omega\left(\frac{\sqrt{n}}{\log n}\right)$ messages to compute. Therefore, the same results hold for $R^{\|,pub}\left(\widetilde{cEq}\text{-}neg_T\right)$.

Putting these results together, we thus see that there exists a problem where $Q^{\|}$ has an exponential advantage over $R^{\|,pub}$. At the same time, there exist other problems such that $R^{\|,pub}$ has an exponential advantage over $Q^{\|}$ ([17]), so these two classes are incomparable.

Sampling Complexity

In addition to the communication complexity to compute functions, as covered in the earlier parts of the paper, we can also measure the communication complexity of other tasks. In particular, we can consider the sampling complexity.

Definition 7 (Sampling Complexity)**.** *The* classical sampling complexity $\mathring{R}_{\varepsilon}(f,\mathcal{D})$ *of a function* $f: X\times Y\to\{0,1\}$ *with a given distribution* $\mathcal{D}$ *on* $X\times Y$ *is the amount of bits needed to be exchanged between two parties to sample* (x,y,z) *such that the resulting distribution is at most* ε *away from the distribution* $\mathcal{D}\times f(\mathcal{D})$ *in* 1*-norm.*

The quantum sampling complexity $\mathring{Q}_{\varepsilon}(f,\mathcal{D})$ *is defined analogously with regards to the number of qubits needed.*

If no distribution $\mathcal{D}$ *is provided, it is to be taken to be uniform.*

We also have, in the quantum case, the idea of generating a superposition which encodes the function into the phase.

Definition 8 (q-Generation Complexity)**.** *The* q-generation complexity $\dot{Q}_{\varepsilon}(f,\mu)$ *of a function* $f: X\times Y\to\{0,1\}$*, with a given* ℓ_2 *distribution* μ *on* $X\times Y$*, is the amount of communication qubits needed to be exchanged by two players for them to generate a superposition* $\sum_{x,y}\mu_{x,y}(-1)^{f(x,y)}|x\rangle|y\rangle$ *to within* ε *error.*

If no distribution μ *is provided, it is to be taken to be uniform.*

Note that we can equivalently talk about the q-generation complexity of an arbitrary state $|\psi\rangle=\sum_{x,y}a_{x,y}|x\rangle|y\rangle$, by asking how many communication qubits are needed to be exchanged in order to generate the state $|\psi\rangle$.

Given the above definitions, [18] then showed that we can strictly bound $\dot{Q}_{\varepsilon}(\psi)$. Given a superposition $|\psi\rangle=\sum_{x,y}a_{x,y}|x\rangle|y\rangle$, we define a matrix M_{ψ} with entries $a_{x,y}$.

Theorem 3 ([18])**.** *If* $|\psi\rangle=\sum_{x,y}a_{x,y}|x\rangle|y\rangle$ *is a pure state and* M_{ψ} *is defined above, we then bound* $\dot{Q}_{\varepsilon}(\psi)$ *via*

$$\lceil\log K_{2\varepsilon}(M_{\psi})\rceil\leq\dot{Q}_{\varepsilon}(\psi)\leq\lceil\log K_{\varepsilon}(M_{\psi})\rceil$$

where $K_{\varepsilon}(B)=\min_{A:\|A-B\|_2^2\leq\varepsilon}\operatorname{rank}A$.

Proof. The upper bound can be demonstrated by construction. Alice can begin by finding the singular value decomposition $U_1\Sigma U_2$ of M_{ψ}. Then, Alice can take the first $k=K_{\varepsilon}(M_{\psi})$ entries $\sigma_{i,i}$ of Σ and construct a superposition $c\sum_{i=1}^{k}\sigma_{i,i}|i\rangle|i\rangle$ where c is a normalizing factor. This uses $2\lceil\log k\rceil$ qubits. This superposition represents a renormalized version of the first k columns of Σ. Alice can then send the latter $2\lceil\log k\rceil$ qubits to Bob. Both players can then pad their qubits to the appropriate lengths and apply U_1 and $U_2^{\top}$, respectively, to their qubits. The resulting state $|\phi\rangle$ approximates $|\psi\rangle$ to within ε error.

The lower bound is shown by considering the rank of any approximating M_{ϕ}, as seen in [19]. If $|\phi\rangle$ approximates $|\psi\rangle$ to within ε error and requires l qubits to generate, then we can show that $\operatorname{rank}M_{\phi}\leq 2^l$. Meanwhile, because it approximates $|\psi\rangle$ with ε error, we know that $\|M_{\phi}-M_{\psi}\|_2^2\leq 2\varepsilon$. Thus, we see that $K_{2\varepsilon}(M_{\psi})\leq\operatorname{rank}M_{\phi}\leq 2^l$, which completes the proof. □

We can also see some relations between the different kinds of complexity in the quantum case. In particular, we define a *product function* as a function $g : X \times Y \to M$ such that $g(x,y) = g_1(x)\, g_2(y)$ for some x and y.

Theorem 4 ([18]). *Given a function $f : X \times Y \to \{0,1\}$ and a product ℓ_2 distribution μ, let $\mathcal{D}$ be the classical distribution generated from μ, given by $\mathcal{D}(x,y) = |\mu(x,y)|^2$. Then,*

$$\mathring{Q}_{4\sqrt{\varepsilon}}(f,\mathcal{D}) \leq \dot{Q}_{\varepsilon}(f,\mu) + O(1)$$

Theorem 5 ([18]). *Given a function $f : X \times Y \to \{0,1\}$ and a product ℓ_2 distribution μ, then*

$$\dot{Q}_{2\varepsilon}(f,\mu) \leq 2Q_{\varepsilon}(f)$$

Now, we wish to apply this to the disjointness problem.

Definition 9 (The $DISJ_k$ Problem). *Given as input two sets $S, T \subseteq \{1,\ldots,n\}$, such that each is promised to be of cardinality k, $DISJ_j(S,T) = 1$ if and only if $S \cap T = \emptyset$.*

Now, we wish to consider the sampling cost of $DISJ_k$ under a uniform distribution.

Theorem 6 ([18]). *If $k = \Theta(\sqrt{n})$, then $\mathring{Q}_{\varepsilon}(DISJ_k) = O\left(\log n \log \varepsilon^{-1}\right)$.*

Proof. To show this, we will bound $\dot{Q}_{\varepsilon}(DISJ_k)$. By analysing the eigenspaces of M_{DISJ_k}, we can see that $K_{\varepsilon}(M_{DISJ_k}) = O\left(\log n \frac{\log \varepsilon^{-1}}{\log\log \varepsilon^{-1}}\right)$. Thus, we see that $\mathring{Q}_{4\sqrt{\varepsilon}}(DISJ_k) \leq \dot{Q}_{\varepsilon}(DISJ_k) = O\left(\log n \log \varepsilon^{-1}\right)$, so $\mathring{Q}_{\varepsilon}(DISJ_k) = O\left(\log n \log \varepsilon^{-1}\right)$. □

This result can be improved further. Alice and Bob can actually sample $DISJ_k$ over the distribution $\mathcal{D}$ such that $\mathcal{D}$ is a uniform distribution over all disjoint sets S and T. In other words, Alice and Bob can sample two sets which are guaranteed to be disjoint, with these two sets being completely distinct. This also takes time $O\left(\log n \log \varepsilon^{-1}\right)$. This also holds if Alice starts with some set S, and Bob wants to sample some T such that $S \cap T = \emptyset$.

At the same time, there exists some $\varepsilon > 0$ such that for $k = \sqrt{n}$, $\mathring{R}_{\varepsilon}(DISJ_k) = \Omega(\sqrt{n})$. Thus, we see an exponential separation between quantum and classical sampling complexity.

Finally, we close with some words on 0-error sampling. We have a strict lower bound on 0-error quantum sampling complexity, that $\mathring{Q}_0(f,\mathcal{D}) \geq \frac{\log \operatorname{rank}_{f,\mathcal{D}}}{2} - 1$, as shown in [18]. This can be seen by analysing the rank of any matrix generated by l qubits of communication, and finding it to be 2^{2l}. This means that generating $M_{f,\mathcal{D}}$ requires the requisite amount of qubits. This combines with the earlier upper bound to show that in the uniformly distributed case, $\mathring{Q}_0(f) = \Theta(\log \operatorname{rank}_f)$.

At the same time, classically, $\sqrt{D(f)} \leq \mathring{R}_0(f) \leq D(f)$, where $D(f)$ is the deterministic communication complexity to compute f. As such, to prove the log-rank hypothesis, it suffices to show that $\mathring{R}_0(f,\mathcal{D}) = \operatorname{poly}\left(\mathring{Q}_0(f,\mathcal{D})\right)$.

References

[1] Jon T. Yard, and Igor Devetak. Optimal Quantum Source Coding With Quantum Side Information at the Encoder and Decoder. IEEE Transactions on Information Theory 55.11 (2009): 5339-5351.

[2] Igor Devetak, and Jon Yard. Exact cost of redistributing multipartite quantum states. Physical Review Letters 100.23 (2008): 230501.

[3] Touchette, Dave. "Quantum information complexity and amortized communication. April 2014." arXiv preprint arXiv:1404.3733.

[4] A. C-C. Yao. Quantum circuit complexity. InProceedings of 34th IEEE FOCS, pages 352–360,1993.

[5] R. Cleve and H. Buhrman. Substituting quantum entanglement for communication.PhysicalReview A, 56(2):1201–1204, 1997. quant-ph/9704026.

[6] A. A. Razborov. On the distributional complexity of disjointness.Theoretical ComputerScience, 106(2):385–390, 1992.

[7] S. Aaronson and A. Ambainis. Quantum search of spatial regions. InProceedings of 44th IEEEFOCS, pages 200–209, 2003. quant-ph/0303041.

[8] A. Razborov. Quantum communication complexity of symmetric predicates.Izvestiya of theRussian Academy of Sciences, mathematics, 67(1):159–176, 2003. quant-ph/0204025.

[9] H. Buhrman, R. Cleve, and A. Wigderson. Quantum vs. classical communication and compu-tation. InProceedings of 30th ACM STOC, pages 63–68, 1998. quant-ph/9802040.

[10] Iordanis Kerenidis, Sophie Laplante, Virginie Lerays, Jérémie Roland, and DavidXiao. Lower bounds on information complexity via zero-communication protocols andapplications.SIAM Journal on Computing, 44(5):1550–1572, 2015.

[11] Anup Rao and Makrand Sinha. Simplified separation of information and communication.Electronic Colloquium on Computational Complexity (ECCC), 15:057, 2015.

[12] Mark Braverman. Interactive information complexity. Electronic Colloquium on Computational Complexity (ECCC), Report No.123, 2011.

[13] Anurag Anshu, Dave Touchette, Penghui Yao and Nengkun Yu . Exponential Separation of Quantum Communicationand Classical Information. arXiv:1611.08946 [quant-ph]. Nov 2016.

[14] B. Klartag and O. Regev. Quantum one-way communication is exponentially stronger than classical communication. In Proceedings of 43rd ACM STOC, 2011. arXiv:1009.3640.

[15] D. Gavinsky. Quantum versus classical simultaneity in communication complexity. 2017. arXiv:1705.07211 [cs.CC]

[16] J. Naor and M. Naor. Small-Bias Probability Spaces: Efficient Constructions and Applications. SIAM Journal on Computing 22(4), pages 838–856, 1993.

[17] D. Gavinsky, J. Kempe, O. Regev, and R. de Wolf. Bounded-Error Quantum State Identification and Exponential Separations in Communication Complexity. Proceedings of the 38th Symposium on Theory of Computing, pages 594–603, 2006.

[18] A. Ambainis, L. Schulman, A. Ta-Shma, U. Vazirani, A. Wigderson. The Quantum Communication Complexity of Sampling. Society for Industrial and Applied Mathematics (Siam) Journal on Computing, vol. 32, pp. 1570-1585, 2003.

[19] I. Kremer. Quantum communication. Master's thesis, The Hebrew University of Jerusalem, 1995.