An exponential separation between quantum and classical one-way communication complexity

Ashley Montanaro

Centre for Quantum Information and Foundations, Department of Applied Mathematics and Theoretical Physics, University of Cambridge

arXiv:1007.3587
The field of communication complexity studies the amount of communication between parties required for them to compute some function of their joint inputs.
One-way communication complexity

- The field of communication complexity studies the amount of communication between parties required for them to compute some function of their joint inputs.

- One of the simplest models of communication complexity is the one-way model.

\[ x \]  
Alice

\[ y \]  
Bob
The field of communication complexity studies the amount of communication between parties required for them to compute some function of their joint inputs.

One of the simplest models of communication complexity is the one-way model.
The field of communication complexity studies the amount of communication between parties required for them to compute some function of their joint inputs.

One of the simplest models of communication complexity is the one-way model.
One-way communication complexity

- The field of communication complexity studies the amount of communication between parties required for them to compute some function of their joint inputs.

- One of the simplest models of communication complexity is the one-way model.

  ![Diagram](Alice -> m -> Bob -> f(x, y))

- The classical one-way communication complexity (1WCC) of a boolean function $f$ is the length of the shortest message $m$ sent from Alice to Bob that allows Bob to compute $f(x, y)$ with constant probability of success $> 1/2$. 
One-way quantum communication complexity

Can we do better by sending a quantum message?

\[ |\psi\rangle = f(x, y) \]

The quantum 1WCC of \( f \) is the smallest number of qubits sent from Alice to Bob that allows Bob to compute \( f(x, y) \) with constant probability of success greater than \( \frac{1}{2} \).

We don't allow Alice and Bob to share any prior entanglement or randomness.
One-way quantum communication complexity

Can we do better by sending a quantum message?

Alice \(x\) \(\ket{\psi}\) Bob \(y\)

The quantum 1WCC of \(f\) is the smallest number of qubits sent from Alice to Bob that allows Bob to compute \(f(x, y)\) with constant probability of success \(> \frac{1}{2}\).

We don't allow Alice and Bob to share any prior entanglement or randomness.
One-way quantum communication complexity

Can we do better by sending a quantum message?

\[ |\psi\rangle \]

The quantum 1WCC of \( f \) is the smallest number of qubits sent from Alice to Bob that allows Bob to compute \( f(x, y) \) with constant probability of success \( > \frac{1}{2} \).

We don't allow Alice and Bob to share any prior entanglement or randomness.
One-way quantum communication complexity

Can we do better by sending a quantum message?

\[ |\psi\rangle \]

The quantum 1WCC of \( f \) is the smallest number of qubits sent from Alice to Bob that allows Bob to compute \( f(x, y) \) with constant probability of success \( > \frac{1}{2} \).

- We don’t allow Alice and Bob to share any prior entanglement or randomness.
Quantum one-way communication complexity

The model of quantum one-way communication complexity is not well understood. The following results are known:
Quantum one-way communication complexity

The model of quantum one-way communication complexity is not well understood. The following results are known:

- If $f(x, y)$ is allowed to be a partial function (i.e. there is a promise on the inputs), there can be an exponential separation (qv) between quantum and classical 1WCC [Gavinsky et al ’08].
Quantum one-way communication complexity

The model of quantum one-way communication complexity is not well understood. The following results are known:

- If $f(x, y)$ is allowed to be a partial function (i.e. there is a promise on the inputs), there can be an exponential separation (qv) between quantum and classical 1WCC [Gavinsky et al '08].

- Recently it was shown that for partial functions, quantum one-way communication is exponentially stronger than even two-way classical communication [Klartag and Regev '10].
Quantum one-way communication complexity

The model of quantum one-way communication complexity is not well understood. The following results are known:

- If $f(x, y)$ is allowed to be a **partial** function (i.e. there is a promise on the inputs), there can be an **exponential** separation (qv) between quantum and classical 1WCC [Gavinsky et al ’08].

- Recently it was shown that for **partial** functions, quantum one-way communication is exponentially stronger than even **two-way** classical communication [Klartag and Regev ’10].

- If $f(x, y)$ is a **total** function, the best separation we have is a factor of 2 for equality testing [Winter ’04].
Why care about one-way communication complexity?

- One of the simplest interesting models of communication complexity, and still far from understood.
Why care about one-way communication complexity?

- One of the simplest interesting models of communication complexity, and still far from understood.
- Lower bounds on one-way communication complexity have many applications classically to lower bounds on data structures and streaming algorithms.
Why care about one-way communication complexity?

- One of the simplest interesting models of communication complexity, and still far from understood.

- Lower bounds on one-way communication complexity have many applications classically to lower bounds on data structures and streaming algorithms.

- Separating quantum and classical 1WCC is a first step to designing efficient quantum data structures.
Why care about one-way communication complexity?

- One of the simplest interesting models of communication complexity, and still far from understood.

- Lower bounds on one-way communication complexity have many applications classically to lower bounds on data structures and streaming algorithms.

- Separating quantum and classical 1WCC is a first step to designing efficient quantum data structures.

- On a more basic level: 1WCC allows us to address the question of how much information a quantum state “really” contains...
Why care about one-way communication complexity?

- One of the simplest interesting models of communication complexity, and still far from understood.

- Lower bounds on one-way communication complexity have many applications classically to lower bounds on data structures and streaming algorithms.

- Separating quantum and classical 1WCC is a first step to designing efficient quantum data structures.

- On a more basic level: 1WCC allows us to address the question of how much information a quantum state “really” contains...

Unfortunately, some of these applications only really make sense for total functions.
A potential separation for a total function?

It’s been conjectured for some time that there might be a \textbf{quadratic} separation between quantum and classical 1WCC for the following total function.
A potential separation for a total function?

It’s been conjectured for some time that there might be a **quadratic** separation between quantum and classical 1WCC for the following total function.

### Subgroup Membership

The **Subgroup Membership** problem is defined in terms of a group $G$, as follows.

- Alice gets a subgroup $H \leq G$.
- Bob gets an element $g \in G$.
- Bob has to output 1 if $g \in H$, and 0 otherwise.
A potential separation for a total function?

It’s been conjectured for some time that there might be a **quadratic** separation between quantum and classical 1WCC for the following total function.

**Subgroup Membership**

The **Subgroup Membership** problem is defined in terms of a group $G$, as follows.

- Alice gets a subgroup $H \leq G$.
- Bob gets an element $g \in G$.
- Bob has to output 1 if $g \in H$, and 0 otherwise.

For any group $G$, there’s an $O(\log^2 |G|)$ bit classical protocol: Alice just sends Bob the identity of her subgroup.
A potential separation for a total function?

However, for any group $G$, there is an $O(\log |G|)$ qubit quantum protocol...

Alice prepares two copies of the $O(\log |G|)$ qubit state $|H\rangle := \sum_{h \in H} |h\rangle$ and sends them to Bob. Bob applies the group operation $g$ to one copy of $|H\rangle$, to produce $|gH\rangle := \sum_{h \in H} |gh\rangle$. If $g \in H$, then $|H\rangle = |gH\rangle$. Otherwise, $\langle H | gH \rangle = 0$. Bob can distinguish these two cases with constant probability of success using the swap test.
A potential separation for a total function?

However, for any group $G$, there is an $O(\log |G|)$ qubit quantum protocol...

- Alice prepares two copies of the $O(\log |G|)$ qubit state $|H\rangle := \sum_{h \in H} |h\rangle$ and sends them to Bob.
A potential separation for a total function?

However, for any group $G$, there is an $O(\log |G|)$ qubit quantum protocol...

- Alice prepares two copies of the $O(\log |G|)$ qubit state $|H\rangle := \sum_{h \in H} |h\rangle$ and sends them to Bob.

- Bob applies the group operation $g$ to one copy of $|H\rangle$, to produce $|gH\rangle := \sum_{h \in H} |gh\rangle$. 
A potential separation for a total function?

However, for any group $G$, there is an $O(\log |G|)$ qubit quantum protocol...

- Alice prepares two copies of the $O(\log |G|)$ qubit state $|H\rangle := \sum_{h \in H} |h\rangle$ and sends them to Bob.

- Bob applies the group operation $g$ to one copy of $|H\rangle$, to produce $|gH\rangle := \sum_{h \in H} |gh\rangle$.

- If $g \in H$, then $|H\rangle = |gH\rangle$. Otherwise, $\langle H|gH\rangle = 0.$
A potential separation for a total function?

However, for any group $G$, there is an $O(\log |G|)$ qubit quantum protocol...

- Alice prepares two copies of the $O(\log |G|)$ qubit state $|H\rangle := \sum_{h \in H} |h\rangle$ and sends them to Bob.

- Bob applies the group operation $g$ to one copy of $|H\rangle$, to produce $|gH\rangle := \sum_{h \in H} |gh\rangle$.

- If $g \in H$, then $|H\rangle = |gH\rangle$. Otherwise, $\langle H | gH \rangle = 0$.

- Bob can distinguish these two cases with constant probability of success using the swap test.
A potential separation for a total function?

So have we obtained a quadratic separation between quantum and classical 1WCC?

Unfortunately not yet... for every group $G$ people have considered so far (e.g. abelian groups), there is in fact a more clever $O(\log |G|)$ bit classical protocol!

The complexity of the general problem has been an open problem for some time [Aaronson et al '09]... now it's even considered to be a "semi-grand challenge" for quantum computation: [http://scottaaronson.com/blog/?p=471]

Idea: can we prove any separation between quantum and classical 1WCC for a more general version of this problem?
A potential separation for a total function?

So have we obtained a quadratic separation between quantum and classical 1WCC?

- Unfortunately not yet... for every group $G$ people have considered so far (e.g. abelian groups), there is in fact a more clever $O(\log |G|)$ bit classical protocol!
A potential separation for a total function?

So have we obtained a quadratic separation between quantum and classical 1WCC?

- Unfortunately not yet... for every group $G$ people have considered so far (e.g. abelian groups), there is in fact a more clever $O(\log |G|)$ bit classical protocol!

- The complexity of the general problem has been an open problem for some time [Aaronson et al ’09]... now it’s even considered to be a “semi-grand challenge” for quantum computation: [http://scottaaronson.com/blog/?p=471]
A potential separation for a total function?

So have we obtained a quadratic separation between quantum and classical 1WCC?

- Unfortunately not yet... for every group $G$ people have considered so far (e.g. abelian groups), there is in fact a more clever $O(\log |G|)$ bit classical protocol!

- The complexity of the general problem has been an open problem for some time [Aaronson et al ’09]... now it’s even considered to be a “semi-grand challenge” for quantum computation: [http://scottaaronson.com/blog/?p=471]

- **Idea:** can we prove any separation between quantum and classical 1WCC for a more general version of this problem?
New results

In this talk, I will discuss an exponential separation between quantum and classical 1WCC for a partial function based on Subgroup Membership.

There are only one or two known functions showing a separation – more would be nice...
The known examples are arguably somewhat contrived – we'd like to find separations for problems we actually want to solve.
The new problem is a natural generalisation of a particular total function which people care about.
The techniques used seem a bit more applicable elsewhere.
New results

In this talk, I will discuss an exponential separation between quantum and classical 1WCC for a partial function based on Subgroup Membership.

Given that an exponential separation is already known for a partial function, why would we want to do this?
New results

In this talk, I will discuss an exponential separation between quantum and classical 1WCC for a partial function based on Subgroup Membership.

Given that an exponential separation is already known for a partial function, why would we want to do this?

- There are only one or two known functions showing a separation – more would be nice...
New results

In this talk, I will discuss an exponential separation between quantum and classical 1WCC for a partial function based on Subgroup Membership.

Given that an exponential separation is already known for a partial function, why would we want to do this?

- There are only one or two known functions showing a separation – more would be nice...
- The known examples are arguably somewhat contrived – we’d like to find separations for problems we actually want to solve.
New results

In this talk, I will discuss an exponential separation between quantum and classical 1WCC for a **partial** function based on **Subgroup Membership**.

Given that an exponential separation is already known for a partial function, why would we want to do this?

- There are only **one or two** known functions showing a separation – more would be nice...
- The known examples are arguably somewhat contrived – we’d like to find separations for problems we actually want to solve.
- The new problem is a natural generalisation of a particular total function which people care about.
New results

In this talk, I will discuss an exponential separation between quantum and classical 1WCC for a partial function based on Subgroup Membership.

Given that an exponential separation is already known for a partial function, why would we want to do this?

- There are only one or two known functions showing a separation – more would be nice...
- The known examples are arguably somewhat contrived – we’d like to find separations for problems we actually want to solve.
- The new problem is a natural generalisation of a particular total function which people care about.
- The techniques used seem a bit more applicable elsewhere.
The problem

**Perm-Invariance**

- Alice gets an $n$-bit string $x$.
- Bob gets an $n \times n$ permutation matrix $M$.
- Bob has to output

$$
\begin{cases}
1 & \text{if } Mx = x \\
0 & \text{if } d(Mx, x) \geq |x|/8 \\
\text{anything} & \text{otherwise},
\end{cases}
$$

where $|x|$ is the Hamming weight of $x$ and $d(x, y)$ is the Hamming distance between $x$ and $y$. 

Note that *Subgroup Membership* is the special case where $x$ is a $|G|$-bit string such that $x_i = 1 \iff i \in H$, and $M$ is the group action corresponding to $g$ (and we change $|x|/8$ to $2^{|x|}$).
The problem

**Perm-Invariance**

- Alice gets an $n$-bit string $x$.
- Bob gets an $n \times n$ permutation matrix $M$.
- Bob has to output

\[
\begin{cases}
1 & \text{if } Mx = x \\
0 & \text{if } d(Mx, x) \geq |x|/8 \\
\text{anything} & \text{otherwise},
\end{cases}
\]

where $|x|$ is the Hamming weight of $x$ and $d(x, y)$ is the Hamming distance between $x$ and $y$.

Note that **Subgroup Membership** is the special case where $x$ is a $|G|$ bit string such that $x_i = 1 \iff i \in H$, and $M$ is the group action corresponding to $g$ (and we change $|x|/8$ to $2|x|$).
Main result

Theorem

- There is a quantum protocol that solves PERM-INVARINACE with constant success probability and communicates $O(\log n)$ bits.

Any one-way classical protocol that solves PERM-INVARINACE with a constant success probability strictly greater than $\frac{1}{2}$ must communicate at least $\Omega\left(\frac{n^7}{16}\right)$ bits. Therefore, there is an exponential separation between quantum and classical one-way communication complexity for this problem.
Main result

Theorem

- There is a quantum protocol that solves \textsc{Perm-Invariance} with constant success probability and communicates $O(\log n)$ bits.

- Any one-way classical protocol that solves \textsc{Perm-Invariance} with a constant success probability strictly greater than $1/2$ must communicate at least $\Omega(n^{7/16})$ bits.
Theorem

- There is a quantum protocol that solves \textsc{Perm-Invariance} with constant success probability and communicates $O(\log n)$ bits.
- Any one-way classical protocol that solves \textsc{Perm-Invariance} with a constant success probability strictly greater than $1/2$ must communicate at least $\Omega(n^{7/16})$ bits.

Therefore, there is an exponential separation between quantum and classical one-way communication complexity for this problem.
The quantum protocol

The quantum protocol is a simple generalisation of the protocol used for Subgroup Membership:

\[ |\psi^x\rangle := \sum_{i, x_i = 1} |i\rangle \]

Bob performs the unitary operator corresponding to the permutation \(M\) on one of the states, to produce the state \(|\psi^M^x\rangle\), and then uses the swap test to check whether the states are equal. By the promise that either \(|\psi^M^x\rangle = |\psi^x\rangle\), or \(\langle \psi^M^x | \psi^x \rangle \leq 1/8\), these two cases can be distinguished with a constant number of repetitions.
The quantum protocol

The quantum protocol is a simple generalisation of the protocol used for Subgroup Membership:

- Alice prepares two copies of the $\log n$ qubit state $|\psi_x\rangle := \sum_{i,x_i=1}^n |i\rangle$ and sends them to Bob.
The quantum protocol

The quantum protocol is a simple generalisation of the protocol used for Subgroup Membership:

- Alice prepares two copies of the $\log n$ qubit state $|\psi_x\rangle := \sum_{i,x_i=1} |i\rangle$ and sends them to Bob.

- Bob performs the unitary operator corresponding to the permutation $M$ on one of the states, to produce the state $|\psi_{Mx}\rangle$, and then uses the swap test to check whether the states are equal.
The quantum protocol

The quantum protocol is a simple generalisation of the protocol used for Subgroup Membership:

- Alice prepares two copies of the $\log n$ qubit state $|\psi_x\rangle := \sum_{i, x_i=1} |i\rangle$ and sends them to Bob.

- Bob performs the unitary operator corresponding to the permutation $M$ on one of the states, to produce the state $|\psi_{Mx}\rangle$, and then uses the swap test to check whether the states are equal.

- By the promise that either $|\psi_{Mx}\rangle = |\psi_x\rangle$, or $\langle \psi_{Mx}|\psi_x \rangle \leq 1/8$, these two cases can be distinguished with a constant number of repetitions.
The classical lower bound

We prove a lower bound for a special case of Perm-Invariance.

**PM-Invariance**

- Alice gets a $2n$-bit string $x$ such that $|x| = n$.
- Bob gets a $2n \times 2n$ permutation matrix $M$, where the permutation entirely consists of disjoint transpositions (i.e. corresponds to a perfect matching on the complete graph on $2n$ vertices).
- Bob has to output
  \[
  \begin{cases} 
  1 & \text{if } Mx = x \\
  0 & \text{if } d(Mx, x) \geq n/8 \\
  \text{anything} & \text{otherwise}.
  \end{cases}
  \]
Relation to previous work

This is equivalent to the following problem.

PM-Invariance

- Alice gets a $2n$-bit string $x$.
- Bob gets an $n \times 2n$ matrix $M$ over $\mathbb{F}_2$, where each row contains exactly two 1s, and each column contains at most one 1.
- Bob has to output
  \[
  \begin{cases}
    0 & \text{if } Mx = 0 \\
    1 & \text{if } |Mx| \geq n/16 \\
    \text{anything} & \text{otherwise.}
  \end{cases}
  \]
Relation to previous work

A similar problem was used by [Gavinsky et al ’08] to separate quantum and classical 1WCC.

α-Partial Matching

- Alice gets an $n$-bit string $x$.
- Bob gets an $\alpha n \times n$ matrix $M$ over $\mathbb{F}_2$, where each row contains exactly two 1s, and each column contains at most one 1, and a string $w \in \{0, 1\}^{\alpha n}$.
- Bob has to output
  \[
  \begin{cases}
  0 & \text{if } Mx = w \\
  1 & \text{if } Mx = \bar{w} \\
  \text{anything} & \text{otherwise.}
  \end{cases}
  \]

So the main difference is the relaxation of the promise by removing this second string from Bob’s input.
Plan of attack

- Imagine Alice and Bob have a randomised protocol that uses a small amount of communication.
Plan of attack

- Imagine Alice and Bob have a randomised protocol that uses a small amount of communication.

- Fixing a distribution on the inputs, this corresponds to a partition of Alice’s inputs into large subsets, each corresponding to a short message.
Plan of attack

- Imagine Alice and Bob have a randomised protocol that uses a small amount of communication.

- Fixing a distribution on the inputs, this corresponds to a partition of Alice’s inputs into large subsets, each corresponding to a short message.

- Fix two “hard” distributions: one on Alice & Bob’s zero-valued inputs, and one on their one-valued inputs.
Plan of attack

- Imagine Alice and Bob have a randomised protocol that uses a small amount of communication.

- Fixing a distribution on the inputs, this corresponds to a partition of Alice’s inputs into large subsets, each corresponding to a short message.

- Fix two “hard” distributions: one on Alice & Bob’s zero-valued inputs, and one on their one-valued inputs.

- Show that the induced distributions on Bob’s inputs are close to uniform whenever Alice’s subset is large.
Plan of attack

- Imagine Alice and Bob have a randomised protocol that uses a small amount of communication.

- Fixing a distribution on the inputs, this corresponds to a partition of Alice’s inputs into large subsets, each corresponding to a short message.

- Fix two “hard” distributions: one on Alice & Bob’s zero-valued inputs, and one on their one-valued inputs.

- Show that the induced distributions on Bob’s inputs are close to uniform whenever Alice’s subset is large.

- This means they’re hard for Bob to distinguish.
More formally

For any distribution $\mathcal{D}$ on Alice and Bob’s inputs, let $\mathcal{D}^S$ be the induced distribution on Bob’s inputs, given that Alice’s input was in set $S$.

**Lemma**

- Let $f : \{0, 1\}^m \times \{0, 1\}^n \rightarrow \{0, 1\}$ be a function of Alice and Bob’s distributed inputs.
More formally

For any distribution $\mathcal{D}$ on Alice and Bob’s inputs, let $\mathcal{D}^S$ be the induced distribution on Bob’s inputs, given that Alice’s input was in set $S$.

**Lemma**

- Let $f : \{0, 1\}^m \times \{0, 1\}^n \rightarrow \{0, 1\}$ be a function of Alice and Bob’s distributed inputs.
- Let $\mathcal{D}_0, \mathcal{D}_1$ be distributions on the zero/one-valued inputs, respectively, that are each uniform over Alice’s inputs, when averaged over Bob’s inputs.
More formally

For any distribution $\mathcal{D}$ on Alice and Bob’s inputs, let $\mathcal{D}^S$ be the induced distribution on Bob’s inputs, given that Alice’s input was in set $S$.

**Lemma**

- Let $f : \{0, 1\}^m \times \{0, 1\}^n \to \{0, 1\}$ be a function of Alice and Bob’s distributed inputs.
- Let $\mathcal{D}_0$, $\mathcal{D}_1$ be distributions on the zero/one-valued inputs, respectively, that are each uniform over Alice’s inputs, when averaged over Bob’s inputs.
- Assume there is a one-way classical protocol that computes $f$ with success probability $1 - \epsilon$, for some $\epsilon < 1/3$, and uses $c$ bits of communication.
More formally

For any distribution $\mathcal{D}$ on Alice and Bob’s inputs, let $\mathcal{D}^S$ be the induced distribution on Bob’s inputs, given that Alice’s input was in set $S$.

**Lemma**

- Let $f : \{0, 1\}^m \times \{0, 1\}^n \rightarrow \{0, 1\}$ be a function of Alice and Bob’s distributed inputs.
- Let $\mathcal{D}_0, \mathcal{D}_1$ be distributions on the zero/one-valued inputs, respectively, that are each uniform over Alice’s inputs, when averaged over Bob’s inputs.
- Assume there is a one-way classical protocol that computes $f$ with success probability $1 - \epsilon$, for some $\epsilon < 1/3$, and uses $c$ bits of communication.
- Then there exists $S \subseteq \{0, 1\}^m$ such that $|S| \geq \epsilon 2^{m-c}$, and $\|\mathcal{D}_0^S - \mathcal{D}_1^S\|_1 \geq 2(1 - 3\epsilon)$. 
Proof idea: one-valued inputs

We want to show that Bob’s induced distribution on inputs such that $Mx = x$ is close to uniform (the argument for zero-valued inputs is similar but easier).
Proof idea: one-valued inputs

We want to show that Bob’s induced distribution on inputs such that $Mx = x$ is close to uniform (the argument for zero-valued inputs is similar but easier).

- Fix distribution $\mathcal{D}_1$ to be uniform over all pairs $(M, x)$ such that $Mx = x$. 
Proof idea: one-valued inputs

We want to show that Bob’s induced distribution on inputs such that $Mx = x$ is close to uniform (the argument for zero-valued inputs is similar but easier).

- Fix distribution $\mathcal{D}_1$ to be uniform over all pairs $(M, x)$ such that $Mx = x$.

- Let $p_M$ be the probability under $\mathcal{D}_1$ that Bob gets $M$, given that Alice’s input was in $A$, for an arbitrary set $A$. 
Proof idea: one-valued inputs

We want to show that Bob’s induced distribution on inputs such that $Mx = x$ is close to uniform (the argument for zero-valued inputs is similar but easier).

- Fix distribution $D_1$ to be uniform over all pairs $(M, x)$ such that $Mx = x$.

- Let $p_M$ be the probability under $D_1$ that Bob gets $M$, given that Alice’s input was in $A$, for an arbitrary set $A$.

- Let $N_{2n}$ be the number of partitions of $\{1, \ldots, 2n\}$ into pairs. Then

$$p_M = \frac{\binom{2n}{n}}{N_{2n}\binom{n}{n/2}} \Pr_{x \in A}[Mx = x].$$
Proof idea

We want to show that Bob’s induced distribution on inputs such that $Mx = x$ is close to uniform.
Proof idea

We want to show that Bob’s induced distribution on inputs such that $Mx = x$ is close to uniform.

- Upper bounding the 1-norm by the 2-norm, we have

$$\|\mathcal{D}_1^A - U\|_1 \leq \sqrt{N_{2n} \sum_M p_M^2 - 1}$$

where $U$ is the uniform distribution on Bob’s inputs.
Proof idea

We want to show that Bob's induced distribution on inputs such that $Mx = x$ is close to uniform.

- Upper bounding the 1-norm by the 2-norm, we have
  \[
  \|D_A^1 - U\|_1 \leq \sqrt{N_{2n} \sum_M p_M^2} - 1
  \]
  where $U$ is the uniform distribution on Bob's inputs.

- We can now calculate
  \[
  N_{2n} \sum_M p_M^2 = \frac{(\frac{2n}{n})^2}{N_{2n} \left(\frac{n}{n/2}\right)^2 |A|^2} \left( \sum_{x,y \in A} \sum_M [Mx = x, My = y] \right).
  \]
Proof idea

- It turns out that the sum over $M$ only depends on the Hamming distance $d(x, y)$:

\[
\sum_{M} [Mx = x, My = y] = h(x + y)
\]

where $h : \{0, 1\}^{2n} \to \mathbb{R}$ is a function such that $h(z)$ only depends on the Hamming weight $|z|$.
Proof idea

- It turns out that the sum over $M$ only depends on the Hamming distance $d(x, y)$:

$$
\sum_M [Mx = x, My = y] = h(x + y)
$$

where $h : \{0, 1\}^{2n} \to \mathbb{R}$ is a function such that $h(z)$ only depends on the Hamming weight $|z|$.

- So

$$
N_{2n} \sum_M p_M^2 = \frac{(2n)^2}{N_{2n} \binom{n}{n/2}^2 |A|^2} \left( \sum_{x,y} f(x)f(y)h(x + y) \right),
$$

where $f$ is the characteristic function of $A$. 
Proof idea

- It turns out that the sum over $M$ only depends on the Hamming distance $d(x, y)$:

$$
\sum_M [Mx = x, My = y] = h(x + y)
$$

where $h : \{0, 1\}^{2n} \rightarrow \mathbb{R}$ is a function such that $h(z)$ only depends on the Hamming weight $|z|$.

- So

$$
N_{2n} \sum_M p_M^2 = \frac{(2n)^2}{N_{2n}(n/2)^2|A|^2} \left( \sum_{x,y} f(x)f(y)h(x + y) \right),
$$

where $f$ is the characteristic function of $A$.

- This means that it’s convenient to upper bound $N_{2n} \sum_M p_M^2$ using Fourier analysis over the group $\mathbb{Z}_2^{2n}$. 
Fourier analysis in 2 lines

Informally:

- The Fourier transform of a function $f : \{0, 1\}^n \to \mathbb{R}$ is the function $\hat{f} : \{0, 1\}^n \to \mathbb{R}$ defined by

  $$\hat{f}(x) = \frac{1}{2^n} \sum_{y \in \{0,1\}^n} (-1)^{x \cdot y} f(y).$$
Fourier analysis in 2 lines

Informally:

- The Fourier transform of a function \( f : \{0, 1\}^n \rightarrow \mathbb{R} \) is the function \( \hat{f} : \{0, 1\}^n \rightarrow \mathbb{R} \) defined by
  \[
  \hat{f}(x) = \frac{1}{2^n} \sum_{y \in \{0, 1\}^n} (-1)^{x \cdot y} f(y).
  \]

- For any functions \( f, g : \{0, 1\}^n \rightarrow \mathbb{R} \),
  \[
  \sum_{x, y \in \{0, 1\}^n} f(x)f(y)g(x + y) = 2^{2n} \sum_{x \in \{0, 1\}^n} \hat{g}(x)\hat{f}(x)^2.
  \]
Fourier analysis in 2 lines

Informally:

- The Fourier transform of a function $f : \{0, 1\}^n \rightarrow \mathbb{R}$ is the function $\hat{f} : \{0, 1\}^n \rightarrow \mathbb{R}$ defined by

  $$\hat{f}(x) = \frac{1}{2^n} \sum_{y \in \{0, 1\}^n} (-1)^{x \cdot y} f(y).$$

- For any functions $f, g : \{0, 1\}^n \rightarrow \mathbb{R},$

  $$\sum_{x, y \in \{0, 1\}^n} f(x)f(y)g(x + y) = 2^{2n} \sum_{x \in \{0, 1\}^n} \hat{g}(x)\hat{f}(x)^2.$$

- This allows us to write

  $$N_{2n} \sum_{M} p_M^2 = \frac{(2n)^2 2^{4n}}{N_{2n} \binom{n}{n/2}^2 |A|^2} \sum_{x \in \{0, 1\}^{2n}} \hat{h}(x)\hat{f}(x)^2,$$

  where $f$ is the characteristic function of $A$, and $h$ is as on the previous slide.
We can upper bound this sum using the following crucial inequality.

**Lemma**

Let $A$ be a subset of $\{0, 1\}^n$, let $f$ be the characteristic function of $A$, and set $2^{-\alpha} = |A|/2^n$. Then, for any $1 \leq k \leq (\ln 2)\alpha$,

$$\sum_{x, |x|=k} \hat{f}(x)^2 \leq 2^{-2\alpha} \left( \frac{(2e \ln 2)\alpha}{k} \right)^k.$$
Upper bounding this sum

We can upper bound this sum using the following crucial inequality.

**Lemma**

Let $A$ be a subset of $\{0, 1\}^n$, let $f$ be the characteristic function of $A$, and set $2^{-\alpha} = |A|/2^n$. Then, for any $1 \leq k \leq (\ln 2)\alpha$,

$$\sum_{x, |x|=k} \hat{f}(x)^2 \leq 2^{-2\alpha} \left( \frac{(2e \ln 2)\alpha}{k} \right)^k.$$

- This inequality is based on a result of Kahn, Kalai and Linial (the **KKL Lemma**), which in turn is based on a “hypercontractive” inequality of Bonami, Gross and Beckner.
Upper bounding this sum

We can upper bound this sum using the following crucial inequality.

**Lemma**

Let $A$ be a subset of $\{0, 1\}^n$, let $f$ be the characteristic function of $A$, and set $2^{-\alpha} = |A|/2^n$. Then, for any $1 \leq k \leq (\ln 2)\alpha$,

$$\sum_{x, |x|=k} \hat{f}(x)^2 \leq 2^{-2\alpha} \left( \frac{(2e \ln 2)\alpha}{k} \right)^k.$$

- This inequality is based on a result of Kahn, Kalai and Linial (the KKL Lemma), which in turn is based on a “hypercontractive” inequality of Bonami, Gross and Beckner.

- Here $\alpha$ ends up (approximately) measuring the length of Alice’s message in bits.
Finishing up

To summarise:

- We calculate and upper bound the Fourier transform $\hat{h}(x)$, which turns out to be exponentially decreasing with $|x|$. 
Finishing up

To summarise:

- We calculate and upper bound the Fourier transform \( \hat{h}(x) \), which turns out to be exponentially decreasing with \( |x| \).

- We upper bound the “Fourier weight at the \( k \)'th level” of \( f \), \( \sum_{x, |x|=k} \hat{f}(x)^2 \), using the previous lemma.
Finishing up

To summarise:

- We calculate and upper bound the Fourier transform $\hat{h}(x)$, which turns out to be exponentially decreasing with $|x|$.

- We upper bound the “Fourier weight at the $k$’th level” of $f$, $\sum_{x, |x|=k} \hat{f}(x)^2$, using the previous lemma.

- Combining the two upper bounds, we end up with something that’s smaller than a constant unless $|A| \leq 2^{2n-\Omega(n^{7/16})}$. 

Thus, unless Alice sends at least $\Omega(n^{7/16})$ bits to Bob, he can’t distinguish the distribution $D_A$ from uniform with probability better than a fixed constant.

So the classical 1WCC of PM-I variance is $\Omega(n^{7/16})$. 
Finishing up

To summarise:

- We calculate and upper bound the Fourier transform $\hat{h}(x)$, which turns out to be exponentially decreasing with $|x|$.

- We upper bound the “Fourier weight at the $k$’th level” of $f$, $\sum_{x, |x|=k} \hat{f}(x)^2$, using the previous lemma.

- Combining the two upper bounds, we end up with something that’s smaller than a constant unless $|A| \leq 2^{2n-\Omega(n^{7/16})}$.

- Thus, unless Alice sends at least $\Omega(n^{7/16})$ bits to Bob, he can’t distinguish the distribution $D_1^A$ from uniform with probability better than a fixed constant.
Finishing up

To summarise:

- We calculate and upper bound the Fourier transform $\hat{h}(x)$, which turns out to be exponentially decreasing with $|x|$.

- We upper bound the “Fourier weight at the $k$’th level” of $f$, $\sum_{x, |x|=k} f(x)^2$, using the previous lemma.

- Combining the two upper bounds, we end up with something that’s smaller than a constant unless $|A| \leq 2^{2n-\Omega(n^{7/16})}$.

- Thus, unless Alice sends at least $\Omega(n^{7/16})$ bits to Bob, he can’t distinguish the distribution $D^A_1$ from uniform with probability better than a fixed constant.

- So the classical 1WCC of PM-Invariance is $\Omega(n^{7/16})$. 
Conclusions

- We’ve found an \( \Omega(n^{7/16}) \) lower bound on the classical 1WCC of the PM-INVARINANCE problem, implying an exponential separation between quantum and classical 1WCC.
Conclusions

We’ve found an $\Omega(n^{7/16})$ lower bound on the classical 1WCC of the PM-INvariance problem, implying an exponential separation between quantum and classical 1WCC.

How far is this from optimal? There’s an $O(n^{1/2})$ upper bound on the classical 1WCC of PM-INvariance, which is probably actually the right answer.
Conclusions

- We’ve found an $\Omega(n^{7/16})$ lower bound on the classical 1WCC of the PM-INвариантность problem, implying an exponential separation between quantum and classical 1WCC.

- How far is this from optimal? There’s an $O(n^{1/2})$ upper bound on the classical 1WCC of PM-INвариантность, which is probably actually the right answer.

- The original question still remains: can we get a quadratic separation between quantum and classical 1WCC for Subgroup Membership?
Conclusions

- We’ve found an $\Omega(n^{7/16})$ lower bound on the classical 1WCC of the PM-INvariance problem, implying an exponential separation between quantum and classical 1WCC.

- How far is this from optimal? There’s an $O(n^{1/2})$ upper bound on the classical 1WCC of PM-INvariance, which is probably actually the right answer.

- The original question still remains: can we get a quadratic separation between quantum and classical 1WCC for Subgroup Membership?

- Or indeed any asymptotic separation for any total function?
Thanks!

arXiv:1007.3587