Quantum Information and Computation

Sudebkumar Prasant Pal

Indian Institute of Technology Kharagpur

email: spp@cse.iitkgp.ac.in

Short CEP course on Quantum Information and Computation

July 24-28, 2023

July 30, 2023

1. Preliminaries
   - Shannon entropy
   - Density operators and von Neumann entropy
   - Operators on the state space
   - Logarithms of density operators
   - Klein's inequality
   - Projective measurements and entropy change
2. Holevo's bound
3. Qubit communication complexity results: Application of Holevo's bound [Cle+13]
4. Inner product
5. Quantum algorithms
   - Probabilistic Deustch's algorithm
   - Deterministic Deustch's algorithm
   - The Bernstein-Vazirani problem
   - Simon's problem
   - Quantum Fourier transforms and phase estimation
   - Order finding and factoring

Local computation and distant communication

- When quantum resources (apparatus) are distributed in two or more geographically separated locations, we may not be able to implement unitary operations without resorting to either quantum or classical communication.
- So, local operations and communication may be combined to achieve quantum computation.
- It is therefore necessary to determine the amount of communication necessary in such operations.
- Such costs may be deterministic worst case costs or even probabilistic costs, such as average communication costs over discrete distributions.
- We develop the necessary fundamentals and illustrate a few examples of analyses.

Shannon entropy

- For a random source of symbols from a certain discrete distribution, we know that (as small as) expected $H(X)$ bits of information (on the average), can be used for coding information coming out of X.
- If $p(x)$ is the probability of x in the source X, then this (Shannon Entropy) $H(X)$ or $H(p(x))$ is $\sum_x p(x) \log \frac{1}{p(x)}$.
- Once we have two such generators X and Y on the same set of symbols, we can define $H(X, Y)$ as $\sum_{(x,y)} p(x, y) \log \frac{1}{p(x,y)}$, where $p(x, y)$ is the probability that x comes out of X and y out of Y.
- If X and Y are independent (that is, $p(x, y) = p(x)p(y)$), then $H(X, Y) = H(X) + H(Y)$. Otherwise, $H(X, Y)$ is less than the sum of $H(X)$ and $H(Y)$.
- Naturally, joint entropy is less than sum of entropies if the processes are dependent.

Shannon entropy (cont.)

- We may view the joint entropy $H(X, Y)$ of X and Y as the sum of the entropy $H(Y)$ of Y and the conditional entropy $H(X|Y)$ of X given Y. In other words $H(X|Y)$, called the *conditional entropy of* X *given* Y, is the difference between the joint entropy and original entropy, i.e., $H(X|Y) = H(X, Y) - H(Y)$.
- $H(X|Y)$ defined as $H(X, Y) - H(Y)$ can now be written as $\sum_{(x,y)} p(x, y) \log \frac{p(y)}{p(x,y)} = \sum_{(x,y)} p(x, y) \log \frac{1}{p(x|y)}$.
- The conditional entropy is like the uncommon information between X and Y, because this information is needed for X conditional over Y.
- So, subtracting the conditional entropy $H(X|Y)$ from $H(X)$ gives the mutual or common information $H(X : Y)$ between the two sources X and Y, that is, $H(X) - H(X|Y) = H(Y) - H(Y|X)$, usually denoted as $H(X : Y)$ or $I(X : Y)$.

Shannon entropy (cont.)

- We may now view $H(X : Y)$ as $H(X) - H(X|Y) = H(X) - (H(X,Y) - H(Y)) = H(X) + H(Y) - H(X,Y)$, and the symmetry in its definition.

Density operators and von Neumann entropy

- Given the density operator ρ for a quantum state, determining the von Neumann entropy $S(\rho)$ amounts to determining the (real) eigenvalues λ_x of ρ and computing $\sum_x \lambda_x \log \frac{1}{\lambda_x}$.
- Indeed, the spectral decomposition of ρ is $\sum_x \lambda_x |\psi_x\rangle\langle\psi_x|$, where $|\psi_x\rangle$ are the eigenvectors defining an orthonormal basis for the Hilbert space.
- We will now see how these density operators can operate on individual states. If ρ operates on an eigenstate $|\psi_x\rangle$ then we get $\lambda_x \rho_x$.

Traces and postulates

- We already know that the expectation of a projective measurement with Hermitian observable M of a pure state $|\psi\rangle$ is $\langle\psi|M|\psi\rangle$.
- Writing the state as a density operator $\rho = |\psi\rangle\langle\psi|$, this expectation is $tr(M\rho) = tr(\rho M) = tr(|\psi\rangle\langle\psi|M) = \langle\psi|M|\psi\rangle$.
- For density operators of mixed states and measurements using POVM measurement operators M_m for results m, see section 2.4 in [NC00]. Here, the measurement elements are $E_m = M_m^+ M_m$, where (by definition, measurement postulate), E_m are positive, $\sum_m E_m = I$. $[M_m^+$ is the adjoint of $M_m.]$
- Further, for a pure state $|\psi\rangle$, $p(m) = \langle\psi|E_m|\psi\rangle$.
- Such measurements are called POVM and M_m is written as $\sqrt{E}_m$. For a mixed state denoted by a density operator ρ, a unitary operation would take it to a state represented by the density operator $\rho' = U\rho U^+$.

Traces and postulates (cont.)

- A measurement yields m with probability $p(m) = tr(M_m^+ M_m \rho)$. The state resulting due to measurement of m is $\frac{M_m \rho M_m^+}{tr(M_m^+ M_m \rho)}$.
- We use POVM measurements in applications where the Holevo bound is used to estimate upper bounds on the mutual information between a quantum information source at one end and a measured result at the other end.

Logarithms of the density operator

- A method for finding $\log A$ for a diagonalizable matrix A is as follows.
- Let V be the matrix of eigenvectors of A (each column of V is an eigenvector of A). Find the inverse V^{-1} of V.
- Consider AV; observe that $AV = VA'$, where A' is a diagonal matrix whose diagonal elements are eigenvalues of A.
- We get $\log A'$ by replacing each diagonal element of A' by its logarithm.
- Now, we can write $\log A$ as $V \log A' V^{-1}$. [It is now easy to check that the operator $e^{\log A}$ is identical to the operator A.
- In other words, verify that $e^{\log A}|\psi\rangle = A|\psi\rangle$, for all $\psi\rangle$.]
- So, for a density operator A, we can write $S(A) = -tr(A \log A) = -tr(AV \log A' V^{-1}) = -tr(V^{-1} AV \log A') = -tr(A' \log A') = \sum_x \lambda_x \log \frac{1}{\lambda_x}$. [Note also that $(\log A)^n = V(\log A')^n V^{-1}$.]

Klein's inequality

- This is from [NC00], Theorem 11.7, page 511. The relative entropy $S(\rho||\sigma)$ is defined as $-S(\rho) - tr(\rho\log\sigma)$.
- Using the orthonormal decomposition of $\rho = \sum_i p_i|i\rangle\langle i|$, the first term is $\sum_i p_i \log p_i$.
- Since unitary operators preserve trace, the second term can be written as $-\sum_i \langle i|\rho\log\sigma|i\rangle$.
- Also, $\langle i|\rho = p_i\langle i|$.
- Since we have the orthonormal decompostion of $\sigma = \sum_j q_j|j\rangle\langle j|$, we know that $\log\sigma$ is $V\log\sigma' V^{-1}$, where σ' is the diagonal matrix with $\log q_j$ as the jst diagonal element and V is the matrix with columns given by the eigenvectors $|j\rangle$ of σ.

Klein's inequality (cont.)

- So, the second term would be $-\sum_i p_i \langle i|V \log \sigma' V^{-1}|i\rangle = -\sum_i p_i \sum_j P_{ij} \log q_j$, where $P_{ij} = \langle i|j\rangle\langle j|i\rangle$.
- The rest of the proof that the relative entropy is non-negative is based on the double stochasticity of the matrix represented by $P'_{ij}s$, and the concavity of the log function.

Projective measurements and increase in entropy

- We know that entropy changes from $S(\rho)$ to $S(\rho')$ where $\rho' = \sum_i P_i \rho P_i$.
- Here, P_i are elements of the complete set of projectors of the Hermitian observable.
- We need to show that P_i commutes with $\log \rho' = V' \log \rho'' V'^{-1}$, where V is the matrix of eigenvectors of ρ' and ρ'' is the diagonal matrix of eigenvalues of ρ'.
- It is easy to show that $\rho' P_i = P_i \rho P_i = P_i \rho'$.
- Also, $P_i \log \rho' = \lambda_i' |v_i'\rangle\langle v_i'| = \log \rho' P_i$.
- That is, P_i commutes with ρ' as well as with $\log \rho'$.
- Here, λ_i' and $|v_i'\rangle$ are eigenvalues and eigenvectors of ρ'.
- We also use the facts (i) $\sum_i P_i = I$, and (ii) $P_i^2 = P_i$.

Projective measurements and increase in entropy (cont.)

- By Klein's inequality we know that $S(\rho||\rho') = -S(\rho) - tr(\rho\log\rho')$ is non-negative.
- We show that $S(\rho') = -tr(\rho\log\rho')$, thereby establishing $S(\rho') \geq S(\rho)$.
- We have

$$-tr(\rho\log\rho') = -tr((\sum_i P_i)\rho\log\rho')$$

$$= -tr(\sum_i P_i\rho\log\rho') = -tr(\sum_i P_i\rho\log\rho' P_i)$$

$$= -tr(\sum_i P_i\rho P_i\log\rho')$$

$$= -tr(\rho'\log\rho') = S(\rho')$$

Holevo's bound

- Bob is presented with a (mixed) state $\rho = \sum_{i=0}^{n} p_i \rho_i$ because Alice encodes $X = 0, 1, \cdots, n$ as states $\rho_0, \rho_1, \cdots, \rho_n$ (each of which could be mixed states) with probabilities $p_0, p_1, \cdots, p_n$, respectively, for state ρ.
- Bob performs a measurement described by POVM elements $\{E_y\} = \{E_0, E_1, \cdots, E_m\}$ on the (mixed) state provided by Alice and gets outcome Y.
- The Holevo (upper) bound on $H(X : Y)$ is $S(\rho) - \sum_i p_i S(\rho_i)$, often called the *Holevo Chi quantity*, $\chi(\rho_X)$.
- The superscipt X for ρ here is simply indicative of the probability distribution over the index set X of messages x (with probability p_x), from the classical generator X.

Holevo's bound (cont.)

- We consider the trio of the preparation system P, the quantum system Q, and the measuring device M and observe that initially the entire system may be viewed as represented by

$$\rho^{PQM} = \sum_x p_x |x\rangle\langle x| \otimes \rho_x \otimes |0\rangle\langle 0|$$

- This is like the system P with Alice, providing the state ρ_x to Bob for measurement into the system M through the set of POVM measurement elements $\{E_y\}$ in the quantum system Q.
- The subsystem QM realizes the POVM measurement operation defined by $\epsilon(\sigma \otimes |0\rangle\langle 0|)$ creating the state

$$\sum_y \sqrt{E_y}\sigma\sqrt{E_y} \otimes |y\rangle\langle y|$$

Holevo's bound (cont.)

- Observe that in the combined system QM, covering all the elements of the POVM measurement sets the result of the measurement in $M's$ register.
- Naturally, the mutual information between source X with Alice and measured Y with Bob, depends on the initial state ρ and POVM measurement.
- Now note that $S(P : Q) = S(P : Q, M)$ since M is initially isolated and therefore uncorrelated with P and Q.
- Applying the quantum operation ϵ to subsystem QM cannot increase mutual information between P and Q.
- So, $S(P : Q, M) \geq S(P' : Q', M')$.
- Finally, discarding Q' does not increase mutual information, i.e., $S(P' : Q', M') \geq S(P' : M')$.

Holevo's bound (cont.)

- So, we have
$$S(P' : M') \leq S(P : Q)$$
- The quantity $S(P : Q)$ is easily shown to be the expression of the Holevo χ quantity by using the definition of $S(\rho)$ and the Joint Entropy Thoerem, Theorem 11.10 in [NC00][1],
- whence $S(P : Q) = S(P) + S(Q) - S(P, Q) = H(p_x) + S(\rho) - (H(p_x) + \sum_x p_x S(\rho_x))$.
- So, all we need to do in order to establish Holevo's bound now is show that
$$H(X : Y) = S(P' : M')$$

Holevo's bound (cont.)

- This is done by tracing out Q' from $P'Q'M'$ and showing that

$$\rho^{P'M'} = \sum_{x,y} p(x,y)|x\rangle\langle x| \otimes |y\rangle\langle y|$$

where

$$\rho^{P'Q'M'} = \sum_{xy} p_x|x\rangle\langle x| \otimes \sqrt{E_y}\rho_x\sqrt{E_y} \otimes |y\rangle\langle y|$$

- To see this, recall the definition of POVM measurements and the expression for the probability of the result y as $tr(\rho_x E_y)$, so that $p(y|x) = tr(\rho_x E_y) = tr(\sqrt{E_y}\rho_x\sqrt{E_y}))$.
- So, tracing out Q' results in the above state $\rho^{P'M'}$, whose mutual information comes out directly from the joint entropy $S(\rho^{P'M'}) = S(P', M') = H(X, Y)$ and the two traced out systems' von Neumann entropies $S(P) = H(X)$ and $S(M') = H(Y)$.

Holevo's bound (cont.)

- These von Neumann entropies are identical to the Shannon entropies, where $H(X:Y) = H(X) + H(Y) - H(X,Y)$, and where $S(P':M') = S(P') + S(M') - S(P',M') = H(X) + H(Y) - H(X,Y) = H(X:Y)$.

[1] Observe that $\rho^{PQ} = \sum_x p_x |x\rangle\langle x| \otimes \rho_x$, and thus $S(P) = H(p_x)$, $S(Q) = S(\rho)$, and $S(P,Q) = H(p_x) + \sum_x p_x S(p_x)$

Qubit communication complexity

- Alice and Bob run a quantum protocol exchanging qubits. However, they do not exploit any pre-shared quantum entanglement resource.
- We show that at least $\lceil \frac{n}{2} \rceil$ qubits must be sent from Alice to Bob if Alice wishes to convey n bits of (classical) information to Bob [Cle+13].
- Bob wishes to extract n bits of information.
- What matters is the Holevo chi quantity at the end of the protocol in the quantum system with Bob.
- Let ρ_i be the density operator representing the state defined by the collection of qubits with Bob at the end of the ist step.
- Clearly, the information generator provides the states ρ_i^x, from the mixed state $\rho_i = \Sigma_x p_x \rho_i^x$.

Qubit communication complexity (cont.)

- The upper bound on the mutual information on measurements by Bob is the Holevo chi quantity $\chi(\rho_i^X) = S(\rho_i) - \Sigma_x p_x S(\rho_i^x)$.
- It is easy to see that Alice's unitary operations on its own qubits do not alter this χ quantity; the qubits in Bob's system are not tampered within such operations at Alice's end.
- That is, it does not alter ρ_i^X, and therefore does not alter either $S(\rho_i)$ or $\chi(\rho_i^X)$.
- Moreover, χ and S are invariant under unitary transformations at Bob's site.
- So, we consider only two non-trivial cases (i) when Alice sends a qubit to Bob, and (ii) when Bob sends a qubit to Alice.
- In case (i), let B denote the subsystem of qubits after i steps with Bob and Q the single new qubit obtained from Alice in the $(i+1)$st step.

Qubit communication complexity (cont.)

- We know that $S(Q) \leq 1$ (a single qubit !).
- Also, by subadditivity property,

$$S(BQ) \leq S(B) + S(Q) \leq S(B) + 1$$

- We can also show (Araki-Lieb inequality [NC00]) that

$$S(BQ) \geq S(B) - S(Q) \geq S(B) - 1$$

- Clearly therefore,

$$S(\rho_{i+1}) \leq S(\rho_i) + 1$$

(due to subadditivity as shown above), and

$$\chi(\rho_{i+1}^X) = S(\rho_{i+1}) - \Sigma_x p_x S(\rho_{i+1}^x)$$

$$\leq (S(\rho_i) + 1) - \Sigma_x p_x (S(\rho_i^x) - 1)$$

Qubit communication complexity (cont.)

$$= \chi(\rho_i^X) + 2$$

(due to the Araki-Lieb inequality as shown above)

- In case (ii), χ cannot increase [NC00]; we are tracing out a single qubit from Bob's site.
- So,

$$\chi(\rho_{i+1}^X) \leq \chi(\rho_i^X)$$

- Further, by the Araki-Lieb inequality, we have

$$S(\rho_{i+1}) \leq S(\rho_i) + 1$$

- We therefore conclude that the χ quantity goes up (by 2 units), only when a qubit is sent from Alice to Bob.

Qubit communication complexity (cont.)

- It is now clear that Alice would have to send at least $\lceil \frac{n}{2} \rceil$ qubits to Bob to raise the χ quantity at Bob's end to at least n, so that Bob may extract n bits of information.
- Further, observe that whenever a qubit is communicated (either way), the von Neumann entropy does not decrease at Bob's end.
- The entropy may rise by at most one unit.
- So, the total rise in entropy at Bob's end is less that the total number of qubits comunicated either way.
- Since the entropy was initially zero and second term in the χ quantity is also initially zero and finally non-zero, we can say that the rise in entropy exceeds the rise in the χ quantity, or equivalently, exceeds the final χ quantity.
- This χ quantity clearly must be larger than n, the number of classical bits conveyed from Alice to Bob.

Qubit communication complexity (cont.)

- Since the total number of qubits communicated exceeds the net rise in entropy, we can say that this also exceeds the total number of classical bits conveyed.
- So, although at least $\lceil \frac{n}{2} \rceil$ qubits need to be comminicated from Alice to Bob, a total of at least n qubits need to be communicated in order to transfer n bits of classical information from Alice to Bob.

Inner product lower bound

- We consider the computation of the boolean inner product of two n-bit vectors, X and Y, given to respectively, Alice and Bob.
- They, run a *quantum protocol* exchanging only qubits and finally come up with the inner product.
- We present the proof as in [Cle+13], that any such protocol where Bob first comes with the inner product (and then conveys it to Alice), must result in the communication of n classical bits of information from Alice to Bob.
- The main idea lies in the elegant use of the Hadamard operator simultaneously on all the qubits at Bob's end, after doing the highly parallel step of executing the quantum protocol on the balanced superposition of an exponential number of basis states at Bob's end.

Inner product lower bound (cont.)

- More precisely, Bob creates an equal superposition state of 2^n standard basis vectors and this quantum state interacts in the quantum protocol with Alice.
- Assuming that the protocol is *clean*, that is (i) all qubits used as *ancillaes* are reversed in transformations by (reversible) quantum operations, and (ii) all states except the ones storing/coding the bits of interaction in the inner product protocol are reset to initial conditions, we have the following scenario.
- We have the inner product of each of the basis vectors on Bob's side with X, stored as a superposition of 2^n results in a single answer register (a single qubit), at Bob's end.

Probabilistic Deustch's algorithm

- Deustch's problem is to decide whether a 1-bit input boolean function $f : \{0,1\} \Rightarrow \{0,1\}$, is *flat* or *uneven*, or in other words, whether $f(0) \oplus f(1)$ is 0 or 1.
- The problem is equivalent to guessing whether a given coin is genuine or fake.
- The question is how many times we need to look at the coin to find out which case it is. In the quantum world we show that only one look works, but we need to see the quantum superposition of both the sides!
- Note that any classical approach to solving this problem would require evaluating the function f two times.
- However, using quantum parallelism, only one quantum circuit for realizing a *quantum (unitary)* evalution of f suffices in solving this problem, as follows.

Probabilistic Deustch's algorithm (cont.)

- We create a superposition of the two basis states in the first qubit by doing a Hadamard operation[2], on the $|0\rangle$ state to create the $|X+\rangle = |0'\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle$ state, and then perform a *controlled-U_f* with this superposition on the $|0\rangle$ state in the second qubit. [3]
- If $f(0) = f(1)$ then we get,

$$
\begin{array}{rcl}
 & \frac{1}{\sqrt{2}}\left(|0, f(0)\rangle + |1, f(1)\rangle\right) & \\
= & \frac{1}{\sqrt{2}}\left(|0, 0\rangle + |1, 0\rangle\right) & [f(0) = 0] \\
(= & \frac{1}{\sqrt{2}}\left(|0, 1\rangle + |1, 1\rangle\right) & [f(0) = 1])
\end{array}
$$

Probabilistic Deustch's algorithm (cont.)

- For $f(0) = 0$:

$$\begin{aligned} & \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right)|0\rangle \\ = \quad & |0'\rangle \frac{(|0'\rangle + |1'\rangle)}{\sqrt{2}} \\ = \quad & \frac{|0'0'\rangle}{\sqrt{2}} + \frac{|0'1'\rangle}{\sqrt{2}} \end{aligned}$$

Here, $|1'\rangle = |X-\rangle$.

- For $f(0) = 1$:

$$\begin{aligned} & \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right)|1\rangle \\ = \quad & |0'\rangle \frac{(|0'\rangle - |1'\rangle)}{\sqrt{2}} \\ = \quad & \frac{|0'0'\rangle}{\sqrt{2}} - \frac{|0'1'\rangle}{\sqrt{2}} \end{aligned}$$

Probabilistic Deustch's algorithm (cont.)

- In both cases if the second qubit measures $|1'\rangle$, then the first qubit measures $|0'\rangle$ (see [Gru99]).
- On the other hand, if $f(0) \neq f(1)$ then,
  For $f(0) = 0$:

$$\begin{aligned} & \frac{1}{\sqrt{2}}\left(|00\rangle + |11\rangle\right) \\ = \quad & \frac{1}{\sqrt{2}}\left(|0'0'\rangle + |1'1'\rangle\right) \end{aligned}$$

For $f(0) = 1$:

$$\begin{aligned} & \frac{1}{\sqrt{2}}\left(|01\rangle + |10\rangle\right) \\ = \quad & \frac{1}{\sqrt{2}}\left(\frac{|0'\rangle+|1'\rangle}{\sqrt{2}} \otimes \frac{|0'\rangle-|1'\rangle}{\sqrt{2}}\right) + \frac{1}{\sqrt{2}}\left(\frac{|0'\rangle-|1'\rangle}{\sqrt{2}} \otimes \frac{|0'\rangle+|1'\rangle}{\sqrt{2}}\right) \\ = \quad & \frac{1}{\sqrt{2}}\left(\frac{|0'0'\rangle-|0'1'\rangle+|1'0'\rangle-|1'1'\rangle+|0'0'\rangle-|1'1'\rangle+|0'1'\rangle-|1'0'\rangle}{2}\right) \\ = \quad & \frac{1}{\sqrt{2}}\left(|0'0'\rangle - |1'1'\rangle\right) \end{aligned}$$

Probabilistic Deustch's algorithm (cont.)

- We summarize our observations and conclude that in either case, measuring state $|1'\rangle$ on the second qubit gives state $|1'\rangle$ on the first qubit too (see [Gru99]).
- On the other hand, observe that for the four cases above, the $|0'\rangle$ measured on qubit 2 gives no definite information in the first qubit !
- So, we see that this method has 50% success probability since the second qubit can settle into each of the two X-basis states $|0'\rangle$ and $|1'\rangle$ with equal probabilities on measurement.

[2] The Hadamard operation is defined as $H|0\rangle = \frac{|0\rangle+|1\rangle}{\sqrt{2}}$, and $H|1\rangle = \frac{|0\rangle-|1\rangle}{\sqrt{2}}$.

[3] The operation is $U_f(|x,y\rangle) \Rightarrow |x, y \oplus f(x)\rangle$, quite like the $CNOT$ operation $|x,y\rangle \Rightarrow |x, x \oplus y\rangle$.

Deterministic Deustch's algorithm

- Now consider the second approach. Instead of $|0\rangle$, we start with a $|1\rangle$ for the second qubit and do a Hadamard on both, the fisrt as well as the second qubit. The rest is explained below (see [NC00; Gru99]).

$$
\begin{aligned}
|0\rangle|1\rangle &\overset{H^{\otimes 2}}{\rightarrow} \frac{1}{2}\left(|0\rangle + |1\rangle\right)\left(|0\rangle - |1\rangle\right) \\
&\overset{U_f}{\rightarrow} \frac{1}{2}\left(\sum_{x=0}^{1}(-1)^{f(x)}|x\rangle\right)\left(|0\rangle - |1\rangle\right) \\
= \frac{1}{2}(-1)^{f(0)}&\left(|0\rangle + (-1)^{f(0)\oplus f(1)}|1\rangle\right)\left(|0\rangle - |1\rangle\right)
\end{aligned}
$$

Deterministic Deustch's algorithm (cont.)

The U_f step can be explained as follows:

$$\begin{aligned}
&\frac{1}{2}\left(|0\rangle\left(|0\rangle-|1\rangle\right)+|1\rangle\left(|0\rangle-|1\rangle\right)\right)\\
\xrightarrow{U_f}\ &\frac{1}{2}\left(|0\rangle\left(|0\oplus f(0)\rangle-|1\oplus f(0)\rangle\right)+|1\rangle\left(|0\oplus f(1)\rangle-|1\oplus f(1)\rangle\right)\right)\\
=\ &\frac{1}{2}\left[\left((-1)^{f(0)}|0\rangle\left(|0\rangle-|1\rangle\right)\right)+\left((-1)^{f(1)}|1\rangle\left(|0\rangle-|1\rangle\right)\right)\right]\\
=\ &\frac{(-1)^{f(0)}}{2}\left[|0\rangle+(-1)^{f(0)\oplus f(1)}|1\rangle\right]\left(|0\rangle-|1\rangle\right)
\end{aligned}$$

So, with 100% success we get $f(0)\oplus f(1)$, if we do a H on first qubit! This is explained as follows.

Deterministic Deustch's algorithm (cont.)

Let $R = f(0) \oplus f(1)$. Then, we have the following simplification of the first qubit of the above state.

$$\begin{aligned} & \tfrac{|0'\rangle + |1'\rangle}{2\sqrt{2}} + (-1)^R \tfrac{|0'\rangle - |1'\rangle}{2\sqrt{2}} \\ = \quad & \tfrac{|0'\rangle}{\sqrt{2}}(1 + (-1)^R) + \tfrac{|1'\rangle}{\sqrt{2}}(1 - (-1)^R) \end{aligned}$$

If $R = 1$, i.e.,$f(0) \neq f(1)$, then the above expression is $|1'\rangle$, measuring $|1\rangle$ in the standard basis after a Hadamard operation. Otherwise $R = 0$, i.e., $f(0) = f(1)$, and the above expression is $|0'\rangle$, measuring $|0\rangle$ in the standard basis after a Hadamard operation.

- The Hadamard operation is used to change the basis from standard to dual, and finally used again to revert back to the standard basis.

The Bernstein-Vazirani problem

- The Bernstein-Vazirani problem, like the previous ones, is another example of a problem of mathematical interest that is solvable efficiently on a quantum computer.
- The problem is as follows. Let a be an unknown positive integer, $0 \leq a < 2^n$. Let f be the evaluation function for this problem that takes another bit string in the range $0 \leq x < 2^n$ and outputs the modulo 2 sum of the bitwise product of a and x, denoted by $a.x$.
- What we want to determine here is, the number of invocations of the algorithm for evaluating or deciphering a.
- Surprisingly we need only one invocation quantum mechanically. Doing classically would have taken n invocations.

The Bernstein-Vazirani problem (cont.)

- We show this below. Consider U_f applied to $|x\rangle_n|y\rangle$ flipping y if and only if $f(x) = 1$. So, we have

$$U_f|x\rangle_n \frac{(|0\rangle - |1\rangle)}{\sqrt{2}}$$

$$= (-1)^{f(x)}|x\rangle_n \frac{(|0\rangle - |1\rangle)}{\sqrt{2}}$$

where we prepared $|y\rangle$ as $HX|0\rangle =$

$$H|1\rangle = \frac{(|0\rangle - |1\rangle)}{\sqrt{2}}$$

The Bernstein-Vazirani problem (cont.)

- To recap, we know that

$$H|x\rangle = (|0\rangle + (-1)^x|1\rangle)/\sqrt{2}$$

$$= (\sum_{y=0}^{1}(-1)^{xy}|y\rangle)/\sqrt{2}$$

- So, now consider

$$(H^{\otimes n} \otimes 1)U_f(H^{\otimes n} \otimes H)|0\rangle_n|1\rangle$$

$$= \frac{1}{2^{n/2}}(H^{\otimes n} \otimes 1)U_f(\sum_{x=0}^{2^n-1}|x\rangle)\frac{(|0\rangle - |1\rangle)}{\sqrt{2}}$$

$$= \frac{1}{2^{n/2}}(H^{\otimes n}\sum_{x=0}^{2^n-1}(-1)^{f(x)}|x\rangle)\frac{(|0\rangle - |1\rangle)}{\sqrt{2}}$$

The Bernstein-Vazirani problem (cont.)

$$= \frac{1}{2^n} \sum_{x=0}^{2^n-1} \sum_{y=0}^{2^n-1} (-1)^{f(x)+x.y} |y\rangle) \frac{(|0\rangle - |1\rangle)}{\sqrt{2}}$$

- Now consider the sum for a y over all x (we are given that $f(x) = a.x$):

$$\sum_{x=0}^{2^n-1} (-1)^{a.x} (-1)^{x.y}$$

$$= \sum_{x=0}^{2^n-1} (-1)^{a_0.x_0} (-1)^{a_1.x_1} ... (-1)^{a_{n-1}.x_{n-1}} (-1)^{y_0.x_0} (-1)^{y_1.x_1} ... (-1)^{y_{n-1}.x_{n-1}}$$

$$= \sum_{x=0}^{2^n-1} (-1)^{(a_0+y_0)x_0} \dots (-1)^{(a_{n-1}+y_{n-1})x_{n-1}}$$

The Bernstein-Vazirani problem (cont.)

$$= \Pi_{j=1}^{n} \sum_{x_j=0}^{1} (-1)^{(a_j+y_j)x_j}$$

- At least one sum in the product vanishes if $a_j \neq y_j$, i.e., the product equals 0, unless $y = a$.
- So, if we measure the input register finally, we get $|a\rangle$ because all $y \neq a$ will give a zero.
- That is,

$$H^{\otimes(n+1)} U_f H^{\otimes(n+1)} |0\rangle_n |1\rangle = |a\rangle_n |1\rangle$$

if $f(x) = a.x$.

Simon's problem

- In the previous sections we have discussed Deutsch's algorithm and the Deutsch-Jozsa algorithm. They deterministically answer whether (i) a given 1-bit Boolean function is constant or balanced, and (ii) a given n-bit Boolean function is constant or balanced (with a promise restriction on the n input bits), respectively.
- The promise in the second case is that the input string is either constant or balanced. The computation is possible with a single quantum gate for the given function realized as a control gate for that function.
- The Bernstein-Vazirani problem seeks a deterministic solution; finding an unknown bit string a by using a single quantum gate for the function realized as a control gate, in contrast to several evaluations in the classical case.

Simon's problem (cont.)

- In this discussion, we study another problem that has a probabilistic solution within quantum polynomial computation time, and whose best known classical probabilistic algorithm takes exponential time, about $2^{(\log n)^{\frac{1}{3}}}$ steps, for acheveing a certain lower bound of success probability.
- We may use Yao's lemma as in Lemma 3.1.15 in [Gru99] for this negative result.
- The problem is due to Daniel Simon and this exposition is based on [Gru99].
- The function f is defined as $f : \{0,1\}^n \rightarrow \{0,1\}^n$, a two-to-one periodic mapping.

Simon's problem (cont.)

- The problem deals with computing the period of this function. That is, for two distinct elements x, y (n-bit integers) from the domain, $f(x) = f(y)$ if and only if $y = x \oplus s$, where (i) $\oplus$ is the bitwise modulo 2 addition, (ii) x, y differ by an integral multiple, and (iii) s, the period, is an n-bit integer.
- The problem is to determine s, the period, given a quantum circuit, function control gate for f. So, we may find out two integers x and y, from the domain, that give rise to a match in $f(x) = f(y)$, giving period $s = y \oplus x$.
- Finding a match and thereby the period, requires an exponential number of trials in the classical case for finding the period using a probabilistic approach and a given lower bound of success probability.

Simon's problem (cont.)

- To formulate the above problem in mathematical form we state it as follows.
  Input: An integer $m \geq 1$ and a function $f : F_2^n \rightarrow R$, where R is finite set.
  Promise: Does there exists a nonzero element $s \in F_2^n$ such that for all $x, y \in F_2^n, f(x) = f(y)$ if and only if $x = y \oplus s$.
  Output: Element s.
- As the domain of the problem is of n bits, there will be a total of $N = 2^n$ elements.
- Below, we present an argument of the necessity of an exponential number of trials for solving the problem classically.
- We pick elements at random. The probability that we will fail to find a match after the first trial (or for the second trial) is $\frac{N-2}{N-1}$, as we have $N - 1$ remaining elements for selection in next trials and $N - 2$ unfavourable.

Simon's problem (cont.)

- After the second trial we have $N-2$ remaining elements to select from and $N-4$ unfavourable cases for in the third choice.
- So, the probability of failure after the second trial is $\frac{N-4}{N-2}$.
- Continuing in this manner, the $(m+1)$st has probability of failure $\frac{N-2m}{N-m}$.
- So, the failure probability for $(m+1)$ trials is,

$$\frac{N-2}{N-1}\times\frac{N-4}{N-2}\times\ldots\times\frac{N-2m}{N-m}=\frac{1-2/N}{1-1/N}\times\frac{1-4/N}{1-2/N}\times\ldots\times\frac{1-2m/N}{1-m/N}$$

- For sufficiently large N and small x we use $1-x\approx e^{-x}$ and simplify the above as

$$= e^{-2/N}\times e^{-4/N}\times\ldots e^{-2m/N}/e^{-1/N}\times e^{-2/N}\times\ldots e^{-m/N}$$

$$= e^{-(2/N+4/N+\ldots+2m/N)+(1/N+2/N+\ldots+m/N)}$$

$$= e^{-(1/N+2/N+\ldots+m/N)} = e^{-m(m+1)/2N}$$

Simon's problem (cont.)

- This failure probability can be made sufficiently small if $m(m+1)/2$ comparable to N.
- In other words, appreciable success probability results if m is be of the order of $\sqrt{N}$, and therefore exponential in n.
- If $n = 100$ bits then we need approximately $2^{n/2} = 2^{50} = 10^{15}$ trials to get an appreciable chance for finding s.
- Now we will see how we can solve this problem using a quantum algorithm which uses polynomial time probabilistic quantum computation.
- The procedure is as follows: The initial or input state is

$$|0\rangle_n|0\rangle_n$$

Simon's problem (cont.)

- The application of the Hadamard operator on the first $|0\rangle_n$ yields the superposition state

$$\frac{1}{2^{n/2}} \sum_{x=0}^{2^n-1} |x\rangle_n |0\rangle_n$$

- The unitary operator $(U_f : |x, y\rangle \to |x, y \oplus f(x)\rangle)$ converts the state to

$$\frac{1}{2^{n/2}} \sum_{x=0}^{2^n-1} |x\rangle_n |f(x)\rangle_n$$

- Now, applying $H^{\otimes n}$ on the first n-bit register we obtain,

$$\frac{1}{2^n} \sum_{x=0}^{2^n-1} \sum_{y=0}^{2^n-1} (-1)^{x.y} |y\rangle_n |f(x)\rangle_n$$

Simon's problem (cont.)

- We observe the resulting state on both vectors to get $|y, f(x)\rangle$.
- Now for two distinct x say x, x_1, as per the assumption in the problem definition, $f(x) = f(x_1) \Leftrightarrow x_1 = x \oplus s$ and $s \neq 0_n$.
- So, for both $x_1 = x \oplus s$ and x, $|y, f(x)\rangle$ and $|y, f(x \oplus s\rangle)$ are identical.
- So their total amplitude is

$$\frac{1}{2^n}((-1)^{x.y} + (-1)^{(x \oplus s).y})$$

- If $y.s = 0$ mod2 then $x.y = (x \oplus s).y$ mod 2. So the total amplitude becomes

$$(-1)^{x.y} * 2 * 2^{-n} = (\frac{1}{2})^{n-1}(-1)^{x.y}$$

- Suppose we have run the process repeatedly and obtained $n-1$ such linearly independent vectors $y^1, y^2, ..., y^{n-1}$ such that $y^1.s = 0, y^2.s = 0, y^3.s = 0, \ldots y^{n-1}.s = 0$.

Simon's problem (cont.)

- Solving this set of $n-1$ equations we determine the non-zero value of s, the required period.
- Since we have defined f to be a two-to-one mapping, we do not have to check for $f(0)$ or $f(s)$.
- However, if the definition of f is not restricted in the beginning, then we have to check whether f is one-one $(f(0) \neq f(s))$ or two-to-one $(f(0) = f(s))$.
- The total computation time is proportional to the number of repetitions and the time required for a single evaluation of f on an n-bit input.
- Let $t(n)$ be and the time required to execute the quantum circuit once. Let the time required to solve these $n-1$ linearly independent equations be $g(n)$. Then, the total required time is $O(nt(n) + g(n))$.

Simon's problem (cont.)

- Now we show that the $n-1$ observed vectors are linearly independent with probability at least $\frac{1}{4}$.
- Consider $y^1, y^2, \ldots, y^{n-1}$, the $n-1$ vectors measured in as many runs. For any set of $i-1$ vectors there are 2^{i-1} vectors resulting due to linear combination of these vectors. The probability that y^i is one of these (dependent vectors) is

$$\frac{2^{i-1}}{2^{n-1}} = \frac{1}{2^{n-i}}$$

because the total number of vectors satisfying $y.s = 0 mod 2$ is 2^{n-1} (the null space has dimension $n-1$). So, the probability that $y^1, y^2, \cdots, y^{n-2}$ are dependent is at most

$$\sum_{i=2}^{n-2} \frac{1}{2^{n-i}} \leq \frac{1}{2}$$

Simon's problem (cont.)

.

- So, the probability that the $n-2$ vectors are independent is at least $\frac{1}{2}$. Also, the probability that y^{n-1} is independent of the previous $n-2$ vectors is at least $1-\frac{1}{2^{n-(n-1)}} \geq \frac{1}{2}$. Whence, the result.

Quantum Fourier transforms and phase estimation

- We now consider more problems that lie in the class $BQP \subseteq PSPACE$, intuitively the class of problems solvable in probabilistic polynomial time on a QTM, such as order finding and integer factorization (see Definition 10.9 in [AB06]).
- Shor's factorization algorithm lies in BQP, and is not believed to be in $BPP \subseteq BQP$.
- Consider the *unitary* transformation of $\mathbf{x} = \sum_{j=0}^{N-1} x_j|j\rangle$ to $\mathbf{y} = \sum_{k=0}^{N-1} y_k|k\rangle$ for N-dimensional vectors, where x_j and y_k are complex numbers, $N = 2^n$, and n is the number of qubits.
- If y_1, y_2,. . ., y_k,. . .,y_{N-1} form the DFT of x_1, x_2,. . ., x_j,. . .,x_{N-1} then this transformation is called the QFT.

Quantum Fourier transforms and phase estimation (cont.)

- DFT is defined as follows:

$$y_k \equiv \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} x_j e^{2\pi ijk/N}$$

- Naturally QFT is as follows, depicting how basis vector $|j\rangle$ is mapped by the unitary operator:

$$|j\rangle = |j_1 j_2 \ldots j_n\rangle \rightarrow \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} e^{2\pi ijk/N} |k\rangle$$

- Viewing the unitary operator written in matrix form transforming **x** to **y**, we observe that $|j\rangle$ gets rotated to the vector whose components are the elements of the jth column in the transformation matrix.

Quantum Fourier transforms and phase estimation (cont.)

- Note that the inverse of this transformation would get back $|j\rangle$ from $\frac{1}{\sqrt{N}}\sum_{k=0}^{N-1} e^{2\pi ijk/N}|k\rangle$.
- So what goes into the phase of QFT viz., j, comes back by inverse QFT as $|j\rangle$.
- We see below a more complete construction using inverse QFT for estimating the phase in the eigenvalue of a unitary transformation U with eigenvalue $e^{2\pi i\phi}$ and eigenvector $|u\rangle$. That is,

$$U(|u\rangle) = e^{2\pi i\phi}|u\rangle$$

- Here, the phase ϕ need be only a fraction as any integral part turns the phase by four right angles. Let the binary representation of (the fraction) ϕ be $0.\phi_1\phi_2\ldots\phi_n$.

Quantum Fourier transforms and phase estimation (cont.)

- We initially assume that the binary representation of this phase is finite and well within n bits so that n qubits suffice in representing a standard basis vector $|\phi_1\phi_2\ldots\phi_n\rangle$.
- The first step in generating the QFT of the standard basis state $|\phi_1\phi_2\ldots\phi_n\rangle$ in the $2^n = N$ dimensional Hilbert space is the randomization step on the n qubits, where each of the qubits is initialized to $|0\rangle$.
- The ist qubit $(1 \leq i \leq n)$ then performs a controlled-$U^{2^{i-1}}$ operation on the eigen vector $|u\rangle$. The first such operation is for $i = 1$, accumulating a phase $.\phi_1\phi_2\ldots\phi_n \times 2^0 = \phi$; the last one for $i = n$, collects phase $.\phi_1\phi_2\ldots\phi_n \times 2^{n-1} = 0.\phi_n$.
- The ist qubit gathers phase $.\phi_1\phi_2\ldots\phi_n \times 2^{i-1} = 0.\phi_i\phi_{i+1}\ldots\phi_n$.

Quantum Fourier transforms and phase estimation (cont.)

- These phases lead to the computation of the tensor product of qubit states
$$\frac{1}{2^{\frac{n}{2}}}\Pi_{i=1}^{n}(|0\rangle + e^{2\pi i 0.\phi_i\phi_{i+1}\ldots\phi_n}|1\rangle)$$
- This is precisely the QFT of the basis vector $|\phi_1\phi_2\ldots\phi_n\rangle$! [See Nielsen and Chuang [NC00], Equations 5.2 through 5.10].
- So, an inverse QFT would take this state to the basis state $|\phi_1\phi_2\ldots\phi_n\rangle$.
- In this manner the *n* bits of the binary representation of the fractional phase can be determined.
- All we now need to do is to work out a quantum circuit for QFT followed by its counterpart, the quantum circuit for inverse-QFT.
- We know that the unitary transformation of QFT has its conjugate transpose operator as its inverse.

Quantum Fourier transforms and phase estimation (cont.)

- So, we can construct the circuit for inverse-QFT.
- In summary, we have now the mechanism for computing the phase exactly, for the eigenvector of an operator U given its eigenvector u.
- There are therefore two steps in the ensuing *order finding* algorithm for the integer multiplication operator (modulo a prime N) (see [NC00]).
- First, we determine an eigenvector for the operator, and then we determine the phase of the eigenvalue.
- Determining/creating the eigenvector is tricky; so, even if we do not have such an eigenstate, we can use any vector $|\psi\rangle = \sum_u c_u|u\rangle$ for eigenstates $|u\rangle$ of U.
- Use of such a vector would lower the lower bound of $1 - \epsilon$, on the success probability of estimating the right phase by a factor of $|c_u|^2$. See Exercise 5.8 [Nielsen and Chuang] [NC00].

Quantum Fourier transforms and phase estimation (cont.)

- To illustrate an example, we consider the unitary operator based on modulo N multiplication with a fixed number $x < N$ as

$$U|y\rangle = |xy(modN)\rangle, y \in \{0,1\}^L, L = \log N$$

[Show that U is unitary.]

- If r is the order of x i.e, $x^r = 1(modN)$, then there are r eigen vectors $|u_s\rangle$ of U, $0 \leq s \leq r-1$ as follows.
-

$$|u_s\rangle = \frac{1}{\sqrt{r}} \sum_{k=0}^{r-1} e^{\frac{-2\Pi isk}{r}} |x^k(modN)\rangle$$

Quantum Fourier transforms and phase estimation (cont.)

- We can see that

$$U|u_s\rangle = \frac{1}{\sqrt{r}} \sum_{k=0}^{r-1} e^{\frac{-2\Pi isk}{r}} |x^{k+1}(modN)\rangle$$

$$= e^{\frac{2\Pi is}{r}} |u_s\rangle$$

- This can be seen as (i) taking the eigenvalue exponent out increases the negative exponent in the term k by one unit, thereby matching x^{k+1} in the summation and (ii) rolling cyclically as r is the order of x.
- Now it is not hard to show that $\frac{1}{\sqrt{r}} \sum_{s=0}^{r-1} |u_s\rangle = |1\rangle$ (see Exercise 5.13, Nielsen and Chuang [NC00]).
- So, even in the absence of an eigenvector, we may proceed creating the state $|1\rangle$, which is easy to create.

Quantum Fourier transforms and phase estimation (cont.)

- This will however erode the success probability as mentioned above. Nevertheless, we can proceed with phase estimation, initially assuming we have a sufficient number of qubits to represent the binary form of the fractional phase $\phi = 0.\phi_1\phi_2\ldots\phi_n$ encoded/interpreted as the standard basis state $|\phi_1\phi_2\ldots\phi_n\rangle$.
- When we do not have enough (qu)bits for exactly representing an unknown ϕ, certain errors creep leading to approximations upto a limited number of bits with provably specified (high) probability as shown in the standard literature [NC00].
- Phase estimation has several applications, the most well-known being factoring the product of two large primes as in Peter Shor's seminal paper [**shor**].

Quantum Fourier transforms and phase estimation (cont.)

- In summary, phase estimation uses a sufficient number $t = L + \lceil \log_2(2 + \frac{1}{2\epsilon}) \rceil$ of qubits in the first register to compute an approximation of the phase ϕ_u of the eigenvalue of the eigenvector $|u\rangle$ of U with probability at least $1 - \epsilon$, running using $O(t^2)$ operations, including one call to each of the $L = n = \log_2 N$ controlled-U^j operations.

Order finding and factoring

- Phase estimation for the eigenvalue(s) of the unitary operator $U|y\rangle = |xy(modN)\rangle$ yields fractional phase $\frac{s}{r}$, $0 \leq s \leq r$, if there are a sufficient number of qubits to fully encode this fraction in as many binary digits.
- From the fraction $\frac{s}{r}$ computed correctly to at least $2L + 1$ bits, it is possible to determine r; this is done using the continued fraction expansion of the fractional estimate of the phase $\frac{s}{r}$ in $O(L^3)$ time (see Theorem 5.1 in [NC00]). Here $L = n = \log_2 N$.
- Now we detail an alternative scheme and its analysis, yielding the same results for the implementation of Shor's algorithm.
- The second (controlled) register $|y\rangle$ has $n = \log_2 N$ qubits, where N is the product of two large primes.

Order finding and factoring (cont.)

- The first register $|j\rangle$ requires $t = 2n + 1 + \lceil \log_2 (2 + \frac{1}{2\epsilon}) \rceil$ qubits; the additional n qubits are required for modular exponentiation (see Box 5.2, [NC00]).
- The phase $\frac{s}{r}$ is estimated accurately upto $2L + 1$ bits with probability exceeding $\frac{(1-\epsilon)}{r}$. The order of a random $x \leq N$ is $r \leq N (= 2^n)$, i.e., $x^r = 1 \bmod N$.
- We use the function $f(k) = x^k \bmod N$, $0 \leq k \leq M - 1$. Here, $t = \log_2 M$. The (qu)bit basis vector (integer) k, in the randomized superposition of the first register of $t = \log_2 M$ qubits, does a control operation over the second register of $n = \log_2 N$ qubits.

Order finding and factoring (cont.)

- The first step is the usual randomization step of t Hadamard operations, one on each of the t qubits of the first register. This gives the transition

$$|0\rangle^t|0\rangle^n \Rightarrow \frac{1}{\sqrt{M}}(\sum_{k=0}^{M-1}|k\rangle|0\rangle^n)$$

Using a controlled-U_f we further get

$$\frac{1}{\sqrt{M}}(\sum_{k=0}^{M-1}|k\rangle|x^k mod(N)\rangle^n)$$

which may be written (using periodicity of f) as follows.

$$\frac{1}{\sqrt{M}}\sum_{l=0}^{r-1}(\sum_{q=0}^{s_l}|qr+l\rangle|x^l mod(N)\rangle^n) \tag{1}$$

Order finding and factoring (cont.)

Here, l stands for each partition of the periodic function, q determines the starting index of each run of r elements, and s_l gives the number of full periods for the periods with offset l. Clearly, for $0 \leq l \leq r-1$, $(M-r) \leq s_l r + l + 1 \leq M$
In an initial and simplified analysis, we first measure the second register, yielding $y = x^l \bmod N$, for some l, and thereby the state

$$\frac{1}{\sqrt{s_l + 1}} \sum_{q=0}^{s_l} |qr + l\rangle \tag{2}$$

by ignoring the second register.

Order finding and factoring (cont.)

Assume for simplicity that $s_l + 1 = \frac{M}{r}$, that is, there is exact matching of full periods for index l. Now perform a QFT on the following state yielding the state

$$\sqrt{\frac{r}{M}} \sum_{q=0}^{s_l} |qr + l\rangle \tag{3}$$

yielding the state

$$\frac{1}{\sqrt{M}} \sum_{c=0}^{M-1} \sqrt{\frac{r}{M}} \sum_{q=0}^{s_l} e^{2\Pi i c(qr+l)/M} |c\rangle \tag{4}$$

Note that the probability amplitude of $|c\rangle$ is non-zero if and only if c is a multiple of $\frac{M}{r}$. Even when non-zero, the problematic item l

Order finding and factoring (cont.)

appears in the complex exponent, making its effect irrelevant to the final probability of the different periodic outputs !
A slightly complicated analysis is required when $s_l + 1$ is not the same as $\frac{M}{r}$. Continuing with the state in Equation 1, we perform a QFT on the first register yielding

$$\frac{1}{\sqrt{M}} \sum_{l=0}^{r-1} \sum_{q=0}^{s_l} \frac{1}{\sqrt{M}} \sum_{p=0}^{M-1} e^{\frac{2\Pi ip(qr+l)}{M}} |p\rangle |x^l mod(N)\rangle^n$$

The summation over $s_l + 1$ values of q determines the probability amplitude for each value of l. The multiplicative term $e^{\frac{2\Pi ipl}{M}}$ is inconsequential inspite of l, as it has unit modulus. So, we need only determine

$$b_{p,l} = \frac{1}{M} \sum_{q=0}^{s_l} e^{\frac{2\Pi iprq}{M}}$$

Order finding and factoring (cont.)

approximated as

$$\frac{1}{M}(\frac{1-e^{\frac{2\Pi ipr(s_l+1)}{M}}}{1-e^{\frac{2\Pi ipr}{M}}})$$

Now it is easy to show that $b_{p,l}b_{p,l}^*$ is

$$\frac{1}{M^2}\frac{sin^2(\frac{\Pi pr(s_l+1)}{M})}{sin^2(\frac{\Pi pr}{M})}$$

The outcome p is measured with this probability for a fixed l. This probability is roughly the same for all l since s_l is nearly $\frac{M}{r}$ for all $0 \leq l \leq r-1$. Therefore, the probability of measuring p is about $r|b_{p,l}|^2$, which can be shown to be at least $\frac{2}{5r}$ for p's such that p differs for an integral multiple of $\frac{M}{r}$ by at most $\frac{1}{2}$. Now these integral multiples can be from 0 through $r-1$, that is, r values, thereby

Order finding and factoring (cont.)

rendering the probability of measuring such a p to be at least $r \times \frac{2}{5r} = 0.40$, which is 40% guaranteed success ! The probability calculation goes as follows. We underestimate the numerator and overestimate the denominator in order to show the 40% lower bound on the probability of getting such outcomes p as

$$|p - \frac{dM}{r}| < \frac{1}{2}$$

where d is an integer. So, the overestimation is done by substituting $\frac{\Pi pr}{M}$ by $\Pi(d+e)$ where e is either positive or negative with $|e| = \frac{r}{2M}$. The denominator $M^2 sin^2(\frac{\Pi pr}{M})$ is therefore

$$M^2 sin^2(\Pi(d+e)) = M^2 sin^2(\Pi e) \leq M^2(\Pi e)^2$$

Order finding and factoring (cont.)

The numerator $sin^2(\frac{\Pi pr(s_l+1)}{M})$ is underestimated as

$sin^2(\Pi(s_l+1)(d+e)) = sin^2(\Pi(s_l+1)e) \geq (\Pi(s_l+1)e)^2 g^2(\Pi(s_l+1)e)$

where $g(x) = \frac{sin(x)}{x}$. But $|g(\Pi(s_l+1)e)| = |g(\frac{\Pi}{2}(1+\epsilon))|$ since $(s_l+1)r$ is nearly M and no lesser, whereas $|e| = \frac{r}{2M}$. The numerator's underestimate is therefore

$$(\Pi(s_l+1)e)^2 g^2(\frac{\Pi}{2}(1+\epsilon))$$

which is nearly

$$= \frac{(\Pi(s_l+1)e)^2}{(\frac{\Pi}{2})^2} = 4(s_l+1)^2 e^2$$

Order finding and factoring (cont.)

The probability estimate is therefore $\frac{4}{r^2\Pi^2}$, which when multiplied by r gives $\frac{4}{r\Pi^2}$, nearly $\frac{2}{5r}$. Furthermore, d can take r values as already mentioned above, thereby enhancing the probability to at least 0.4 or 40%.

Now we have observation outcome p satisfying

$$|\frac{p}{M} - \frac{d}{r}| < \frac{1}{2M} \leq \frac{1}{2N^2} < \frac{1}{2r^2}$$

This enables computing r from the continued fraction of $\frac{p}{M}$ because p (and therefore this fraction), has been computed correct to at least $2n+1$ binary bits (see itemize [NC00]).

Grover's search

- The problem we consider is to identify for a given function $f(x)$ ($f(x)$ is a function from $\{0,1\}^n$ to $\{0,1\}$), that x ($x \in \{0,1\}^n$) for which $f(x)$ is 1.
- In the classical sense this problem will take $N = 2^n$ evaluations of the function $f(x)$ in the worst case. There can be one or more values of x for which $f(x)$ is 1.
- Intuitively, however the problem of finding such a x in case $f(x)$ evaluates to 1 for just one x, is at least as hard as finding a x in the scenario where $f(x)$ evaluates to 1 for more that one values of x.
- We therefore consider the somewhat simpler case of where $f(x)$ is 1 for just one x, and will see that the method we develop can be generalized.

Grover's search (cont.)

- This method we consider uses quantum parallelism (and other clever techniques) to achieve quadratic speed up ($\sqrt{N}$) over classical methods.
- We are given the quantum implementation of the function f as below

$$O : |x\rangle_n |y\rangle \mapsto |x\rangle_n |y \bigoplus f(x)\rangle$$

- If we prepare $|y\rangle$ as $|X-\rangle$ $(H^{\otimes}|1\rangle)$, we can re-write above as

$$O : |x\rangle_n \frac{|0\rangle - |1\rangle}{\sqrt{2}} \mapsto (-1)^{f(x)} |x\rangle_n \frac{|0\rangle - |1\rangle}{\sqrt{2}}$$

Ignoring the state of the last qubit, the action of O on a general state of quantum register is

$$O : \sum_{x \in \{0,1\}^n} \alpha_x |x\rangle \mapsto \sum_{x \in \{0,1\}^n} (-1)^{f(x)} \alpha_x |x\rangle$$

Grover's search (cont.)

- The quantum register is prepared in the state $|0\rangle^{\otimes n}$, which is then put by applying the Hadamard transform $H^{\otimes n}$, in superposition state

$$|\psi\rangle_n = \frac{1}{\sqrt{N}} \sum_{x=0}^{N-1} |x\rangle_n$$

- We now apply the following sequence of operations calsled *Grover* operator:

$$G = H^{\otimes n} P_0 H^{\otimes n} O$$

- Where the conditional phase shift P_0 is given by

$$P_0|x\rangle \mapsto \begin{cases} |x\rangle & x = 0 \\ -|x\rangle & x > 0, \end{cases}$$

OR

Grover's search (cont.)

$$P_0 = 2|0\rangle\langle 0| - I$$

- One can easily verify that the following holds

$$H^{\otimes n}(2|0\rangle\langle 0| - I)H^{\otimes n} = 2|\psi\rangle\langle\psi| - I$$

- Now, $(2|\psi\rangle\langle\psi| - I)(\sum_x \alpha_x|x\rangle) = 2\sum_x |\psi\rangle\langle\psi|\alpha_x|x\rangle - \sum_x \alpha_x|x\rangle$

$$= 2|\psi\rangle\sum_x \alpha_x\langle\psi|x\rangle - \sum_x \alpha_x|x\rangle$$

$$= 2|\psi\rangle\sum_x \frac{\alpha_x}{\sqrt{N}} - \sum_x \alpha_x|x\rangle$$

$$= 2\sum_x |x\rangle\sum_x \frac{\alpha_x}{N} - \sum_x \alpha_x|x\rangle$$

Grover's search (cont.)

$$= \sum_x (-\alpha_x + 2\langle\alpha\rangle)|x\rangle \text{ where } \langle\alpha\rangle \text{ is } \sum_x \frac{\alpha_x}{N}$$

- It can be noted from the result of application of $(2|\psi\rangle\langle\psi| - I)$ on an arbitrary quantum state(in standard basis representation) that if we had negative amplitude(s), then they get boosted up (positively) at the cost of the remaining positive ones (remember that the square of probability amplitude is normalized to 1).
- In other words as much as the positive amplitudes of the arbitrary state before application were above the mean value, they will fall down the mean by the same amount after the application. And the negative ones will be boosted over the old mean by the amount they were negative with.

Grover's search (cont.)

- Application of O is precisely meant to negate the amplitude of those x for which $f(x)$ evaluates to 1 and therefore the above observation become applicable.
- Also, it is to be note that multiple application of G will keep on increasing the probability amplitude of those x in the original state $|\psi\rangle$ for which $f(x)$ is 1.
- After a "sufficient" number of applications of G we can expect to find such an x with high probability, completing our search.
- What exactly do we mean by this sufficient number of applications is discussed now.
- Let us define $T = x$ for which $f(x)$ is 1, and $S = \{0,1\}^n / T$

$$|\sigma\rangle = \frac{1}{\sqrt{N-1}} \sum_{x \in S} |x\rangle \quad \text{And} \quad |\tau\rangle = |x\rangle \quad ; x \in T$$

Grover's search (cont.)

- Now, any state in the hyper plane (of *n*-dimentional space) induced by $|\sigma\rangle$ and $|\tau\rangle$ can be written as $a|\sigma\rangle + b|\tau\rangle$ with $a^2 + b^2 = 1$
- We have $O(a|\sigma\rangle + b|\tau\rangle) = a|\sigma\rangle - b|\tau\rangle$, which shows that the action of O on the hyper plane induced by $|\sigma\rangle$ and $|\tau\rangle$ is a reflection about $|\sigma\rangle$.
- We can write $|\psi\rangle$ as $(\sqrt{\frac{N-1}{N}}|\sigma\rangle + \sqrt{\frac{1}{N}}|\tau\rangle)$.
- It can therefore easily be shown that the action of $(2|\psi\rangle\langle\psi| - I)$ in the $|\sigma\rangle$ and $|\tau\rangle$ plane is a reflection about $|\psi\rangle$.
- Since the composition of two reflections is a rotation 1, it follows that one application of G rotates state vectors in the $|\sigma\rangle$, $|\tau\rangle$ plane by θ towards $|\tau\rangle$, where $\frac{\theta}{2}$ is the angle between $|\psi\rangle$ and $|\sigma\rangle$, i.e.
  $|\psi\rangle = \cos(\frac{\theta}{2})|\sigma\rangle + \sin(\frac{\theta}{2})|\tau\rangle$
  Hence, after m iterations we have:
  $$G^m|\psi\rangle = cos(\frac{2m+1}{2}\theta)|\sigma\rangle + \sin(\frac{2m+1}{2}\theta)|\tau\rangle$$

Grover's search (cont.)

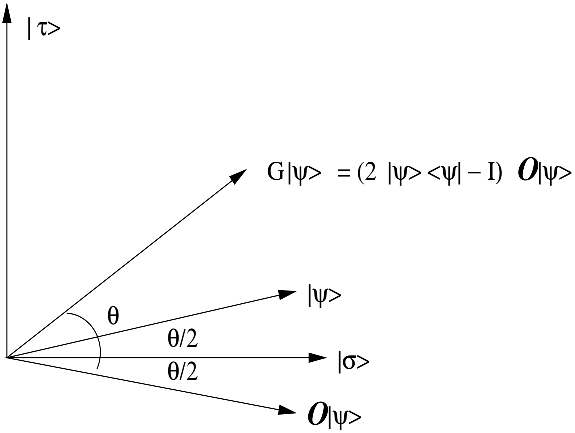

Figure: Action of Grover's operation.

Grover's search (cont.)

- It follows that when $\frac{2m+1}{2}\theta \approx \frac{\pi}{2}$, i.e. roughly after $(\frac{\pi}{2\theta} - \frac{1}{2})$ iterations, the state vector is within an angle $\leq \frac{\pi}{4}$ of $|\tau\rangle$.
- Measurement of the state vector now will give a solution with probability at least $\cos^2(\frac{\pi}{4}) = \frac{1}{2}$.
- Therefore in the light of the following

$$\frac{\theta}{2} \geq \sin(\frac{\theta}{2}) = \sqrt{\frac{1}{N}}$$

  we obtain an upper-bound for the number of iterations of G needed to find a solution: $m \leq \lfloor \frac{\pi}{4}\sqrt{N} \rfloor$.
- The intuition for the general case will be as follows. If $f(x)$ is 1 for more that one values of x, and because θ depends on the angle between $|\psi\rangle$ and $|\sigma\rangle$, θ will be larger than in the case discussed above.
- In other words the worst case (in terms of the number of iterations req.) will happen if $f(x)$ is 1 for just one value of x and θ will be very small.

Grover's search (cont.)

- In the general case the situation can only improve, which is also rational since it is easier to find something with more repetition in a set of fixed size than something which occurs just once.

References

S. Arora and B. Barak, "Computational complexity: A modern approach,", 2006. [Online]. Available: http://theory.cs.princeton.edu/complexity/.

R. Cleve, W. van Dam, M. Nielsen, and A. Tapp, "Quantum entanglement and the communication complexity of the inner product function," *Theor. Comput. Sci.*, vol. 486, pp. 11–19, 2013. DOI: 10.1016/j.tcs.2012.12.012. [Online]. Available: https://doi.org/10.1016/j.tcs.2012.12.012.

J. Gruska, *Quantum Computing*, ser. Advanced topics in computer science series. McGraw-Hill, 1999, ISBN: 9780077095031. [Online]. Available: https://books.google.co.in/books?id=Kb1QAAAAMAAJ.

References (cont.)

M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.