Quantum algorithms **Monsoon 2007**

Lecture 5

Lecturer: Pranab Sen *Date: August 13*

1 Motivating density matrices

In this section, we start off with the postulate that the state of a large system is pure since it is assumed to be isolated. We then show how density matrices give us the right definition to work with if we want to study subsystems of the large system.

Let A, B be two finite dimensional quantum systems. We also use A, B to denote their respective Hilbert spaces. The Hilbert space of the joint system is $A \otimes B$. Suppose AB is an isolated system with state vector $|\psi\rangle_{AB}$, where the subscript denotes the quantum system whose state is being described. In general, $|\psi\rangle_{AB}$ is entangled i.e. it cannot be written as a tensor product of a state vector in A and a state vector in B. Nevertheless, we still want a mathematical description of the *reduced state* of A when the global state is $|\psi\rangle_{AB}$.

It is easy to see that $|\psi\rangle_{AB}$ can be written as $|\psi\rangle_{AB} = \sum_j |\alpha_j\rangle_A \otimes |b_j\rangle_B$, where $\{|b_j\rangle\}_j$ form an orthonormal basis of B and $\{|\alpha_j\rangle\}_j$ are vectors in A. Such a decomposition can be obtained, for example, by expressing $|\psi\rangle_{AB}$ in terms of an orthonormal tensor basis of $A \otimes B$ and then collecting terms in the first multiplicand for each basis vector $|b_j\rangle_B$ of B to form the vectors $|\alpha_j\rangle_A$ in A. In general, $|\alpha_j\rangle_A$ are neither of unit length, nor are they orthogonal.

Now, suppose we apply a unitary U on system A only. To understand the action on the system AB, let us consider the case when the system AB is in a separable initial state $|\alpha\rangle_A|\beta\rangle_B$. We would then expect the final state to be $(U|\alpha\rangle)_A \otimes |\beta\rangle_B$. Thus, the action on AB is $U_A \otimes \mathbb{1}_B$. Now for an, in general entangled, state vector $|\psi\rangle_{AB}$ as the initial state, the final state vector will be

$$(U_A \otimes \mathbb{1}_B)|\psi\rangle_{AB} = \sum_j (U|\alpha_j\rangle)_A \otimes |b_j\rangle_B.$$

This seems to suggest that the ensemble $\{|\alpha_j\rangle_A\}_j$ is a first candidate for mathematically describing the reduced state of A; the actual identity of the vectors $\{|b_j\rangle_B\}_j$ is unimportant as long as they form an orthonormal basis of B. Note that $\sum_j \||\alpha\rangle_j\|^2 = 1$; thus the ensemble $\{|\alpha_j\rangle\}$ can be viewed as a probability distribution of unit length state vectors $|\hat{\alpha}_j\rangle := \frac{|\alpha_j\rangle}{\||\alpha_j\rangle\|}$ with probabilities $p_j := \||\alpha_j\rangle\|^2$. We call such probability distributions of unit length state vectors as *mixed states*; a single state vector is sometimes called a *pure state*. We shall use the notations $\{(p_j, |\hat{\alpha}_j\rangle)\}_j$ and $\{|\alpha_j\rangle_A\}_j$ interchangeably to denote mixed states. Under action by unitary U on system A, the state evolves to the ensemble $\{U|\alpha_j\rangle_A\}_j$, in line with our intution about the mixed state being a *probabilistic mixture* of pure states.

Now let us study what happens when we apply a projective measurement on A alone. We shall use only the projective measurement version of the measurement postulate of quantum mechanics, but the ensuing mathematical arguments also work for more general versions of the measurement postulate. As will be seen later on in the *Quantum information and error correction* course, the most general quantum measurement can be implemented by tensoring to the given system an ancillary system initialised to a fixed state and then performing a projective measurement on the joint system. Thus, it suffices to check whether our notion of mixed states as the mathematical description of the reduced state of A stands its ground against projective measurements on A. Suppose Π_A is an orthogonal projection operator on A corresponding to a particular outcome of a projective measurement. By a similar reasoning as above, the projection operator for this outcome viewed over the joint system AB is $\Pi_A \otimes \mathbb{1}_B$. Now,

$$(\Pi_A \otimes \mathbb{1}_B)|\psi\rangle_{AB} = \sum_j (\Pi|\alpha_j\rangle)_A \otimes |b_j\rangle_B.$$

The probability of this projection succeeding is

$$\Pr[\Pi \text{ succeeds}] = \|(\Pi_A \otimes \mathbb{1}_B)|\psi\rangle_{AB}\|^2 = \sum_j \|\Pi|\alpha_j\rangle\|^2 = \sum_j p_j \|\Pi|\hat{\alpha}_j\rangle\|^2,$$

where p_j, $|\hat{\alpha}_j\rangle$ are defined above. This is exactly what one would expect from a mixed state being a probabilistic mixture of pure states. Moreover, the collapsed state of AB if $\Pi_A \otimes \mathbb{1}_B$ succeeds is

$$\frac{1}{\sqrt{\Pr[\Pi \text{ succeeds}]}} \sum_j (\Pi|\alpha_j\rangle)_A \otimes |b_j\rangle_B,$$

which leads us to consider the mixture $\left\{\frac{\Pi|\alpha_j\rangle}{\sqrt{\Pr[\Pi \text{ succeeds}]}}\right\}_j$ as the reduced state of A if Π succeeds. Again, this is what one would expect if we view a mixed state as a probabilistic mixture of pure states. The explanation follows. The probability p'_j of the jth state of the mixture after Π succeeds is

$$\begin{aligned} p'_j &:= \Pr[j \mid \Pi \text{ succeeds}] = \frac{\Pr[j \wedge \Pi \text{ succeeds}]}{\Pr[\Pi \text{ succeeds}]} = \frac{p_j \Pr[\Pi \text{ succeeds} \mid j]}{\Pr[\Pi \text{ succeeds}]} = \frac{p_j \|\Pi|\hat{\alpha}_j\rangle\|^2}{\Pr[\Pi \text{ succeeds}]} \\ &= \frac{\|\Pi|\alpha\rangle_j\|^2}{\Pr[\Pi \text{ succeeds}]}. \end{aligned}$$

If Π succeeds on the jth state, then the collapsed state is $\frac{\Pi|\hat{\alpha}\rangle_j}{\|\Pi|\hat{\alpha}\rangle_j\|} = \frac{\Pi|\alpha\rangle_j}{\|\Pi|\alpha\rangle_j\|}$. Thus, the collapsed mixed state of A if Π succeeds is

$$\left\{\left(p'_j, \frac{\Pi|\alpha\rangle_j}{\|\Pi|\alpha\rangle_j\|}\right)\right\}_j = \left\{\frac{\Pi|\alpha\rangle_j}{\sqrt{\Pr[\Pi \text{ succeeds}]}}\right\}_j.$$

The above arguments strengthen our idea of using mixed states to mathematically describe the reduced state of system A when it is part of an isolated joint system AB. A careful second look at the above arguments leads us to consider the operator

$$\rho_A := \sum_j (|\alpha_j\rangle\langle\alpha_j|)_A = \sum_j p_j (|\hat{\alpha}_j\rangle\langle\hat{\alpha}_j|)_A$$

as a more refined notion of the reduced state of A. The subscript indicates that the operators in question act on the space A. The operator ρ_A is known as the *density operator* or *density matrix* of A. Now observe that

$$\Pr[\Pi \text{ succeeds}] = \sum_j \|\Pi|\alpha_j\rangle\|^2 = \sum_j \mathrm{Tr}\,(\Pi|\alpha_j\rangle\langle\alpha_j|) = \mathrm{Tr}\,\left(\Pi\left(\sum_j |\alpha_j\rangle\langle\alpha_j|\right)\right) = \mathrm{Tr}\,(\Pi_A\rho_A),$$

where we use the subscript in the last term to emphasise the fact that the operators considered act only on the subsystem A. Note that if P, Q are two Hermitian positive semidefinite operators, then $\mathrm{Tr}\,(PQ) \geq 0$ even though PQ, in general, is not even Hermitian (this happens if P, Q do not commute). Thus, $\mathrm{Tr}\,(\Pi\rho) \geq 0$ for any projection Π and any density matrix ρ, which agrees with our intuition that whatever be the state of system A, the probability of a measurement outcome can never be negative. If Π succeeds, the resulting state of A under the density matrix formalism is

$$\sum_j \frac{\Pi|\alpha_j\rangle\langle\alpha_j|\Pi}{\Pr[\Pi \text{ succeeds}]} = \frac{\Pi_A\rho_A\Pi_A}{\mathrm{Tr}\,(\Pi_A\rho_A\Pi_A)}.$$

Under evolution by a unitary U on system A, the new state of A becomes

$$\sum_j U|\alpha_j\rangle\langle\alpha_j|U^\dagger = U_A\rho_A U_A^\dagger$$

in the density matrix formalism. Thus, we have seen how two of the fundamental kinds of dynamics of the quantum system A can be modelled as appropriate functions acting on density matrices.

We now study the tensor product postulate that describes the bringing together two hitherto independent quantum systems. Thus, if we bring in a new system C with density matrix σ_C into consideration, the joint state of system AC will be $\rho_A \otimes \sigma_C$ as can be verified from the tensor product postulate when two isolated quantum systems are brought together for consideration, treating ρ to be part of a pure state on AB and σ to be part of a pure state on CD, where CD is independent of AB.

Now suppose the joint state of system AB is described by a density matrix ω_{AB}. We now describe a mathematical operation called *partial trace* to obtain the density matrix for the reduced state of A from ω_{AB}. In general, ω_{AB} can be a mixed state. This can happen if AB is entangled with another quantum system E, for example, E can be the ‘environment’ or the rest of the universe. Let system ABE be in a pure state $|\psi\rangle_{ABE}$. We can express $|\psi\rangle_{ABE}$ as

$$|\psi\rangle_{ABE} = \sum_{j,k} |\alpha_{jk}\rangle_A \otimes |b_j\rangle_B \otimes |e_k\rangle_E,$$

where $\{|b_j\rangle\}_j$, $\{|e_k\rangle\}_k$ form orthonormal bases of B, E and $\{|\alpha_{jk}\rangle\}_{jk}$ are vectors in A. Then, the reduced state of A is the probabilistic mixture $\{|\alpha_{jk}\rangle\}_{jk}$ with density matrix

$$\rho_A := \sum_{jk}(|\alpha_{jk}\rangle\langle\alpha_{jk}|)_A.$$

Also, the reduced state of AB is the probabilistic mixture $\{\sum_j |\alpha_{jk}\rangle|b_j\rangle\}_k$ with density matrix

$$\omega_{AB} = \sum_k \left(\sum_j |\alpha_{jk}\rangle \otimes |b_j\rangle\right)\left(\sum_j \langle\alpha_{jk}| \otimes \langle b_j|\right) = \sum_{k,j,j'} (|\alpha_{jk}\rangle\langle\alpha_{j'k}|)_A \otimes (|b_j\rangle\langle b_{j'}|)_B.$$

Now, note that the density matrix corresponding to $|\psi\rangle_{ABE}$ is

$$|\psi\rangle\langle\psi| = \sum_{j,k,j',k'} (|\alpha_{jk}\rangle\langle\alpha_{j'k'}|)_A \otimes (|b_j\rangle\langle b_{j'}|)_B \otimes (|e_k\rangle\langle b_{k'}|)_E.$$

So if we define the *partial trace over* B, Tr_B , also known as *tracing out* B, as a $\mathbb{C}$-linear map from operators on AB to operators on A by

$$\mathrm{Tr}_E\,(|a_i\rangle\langle a_{i'}|)_A \otimes (|b_j\rangle\langle b_{j'}|)_B := \delta_{j,j'}(|a_i\rangle\langle a_{i'}|)_A,$$

where $\{|a_i\rangle\}_i$, $\{|b_j\rangle\}_j$ are orthonormal bases for A, B, then $\mathrm{Tr}_E\,(|\psi\rangle\langle\psi|)_{ABE} = \omega_{AB}$ and

$$\mathrm{Tr}_{BE}\,(|\psi\rangle\langle\psi|)_{ABE} = \rho_A = \mathrm{Tr}_B\,\omega_{AB}.$$

Observe that the definition of Tr_B is independent of the choice of orthonormal bases of A, B that are used to denote operators on AB and A; thus, Tr_B is indeed a map from operators on AB to operators on A. It is easy to check that the partial trace is the unique $\mathbb{C}$-linear map from operators on AB to operators on A satisfying the property

$$\mathrm{Tr}_B\,(M_A \otimes M'_B) = \mathrm{Tr}\,(M'_B) \cdot M_A,$$

where M_A, M_B are linear operators on A, B; this justifies the name partial trace. It is also easy to check that the above definition of partial trace agrees with the definition given in Lecture 4, namely,

$$\mathrm{Tr}_B\,(M_{AB}) := \sum_j (\mathbb{1}_A \otimes |b_j\rangle_B) M_{AB} (\mathbb{1}_A \otimes \langle b_j|_B),$$

where M_{AB} is a linear operator on AB. The $\mathbb{C}$-linearity of Tr_B implies that Tr_B respects the interpretation of a mixed state as a probability distribution over pure states. Since by the spectral theorem any Hermitian positive semidefinite operator is a non-negative linear combination of projections onto its eigenvectors, and partial trace applied to a rank one projector gives a density matrix (which is easy to check), we see that partial trace maps positive operators to positive operators and preserves their traces. Thus, density matrices and partial trace together give us a mathematical formalism to describe the relation between the state of a joint system and the reduced state of a component.

Thus, it appears that ρ_A is an accurate notion of the reduced state of A in the following senses:

- ρ_A suffices to describe the dynamics of system A when operations are applied to A alone, viz. unitary evolution in A, projective measurement in A, attaching new quantum systems to A and taking reduced states of subsystems of A;
- Two mixed states having the same density matrix are indistinguishable by any quantum process since the above operations characterise all quantum processes as per the postulates;
- For any two different density matrices ρ, ρ' of A, there exists a projective measurement on A that gives different probability distributions for ρ and ρ' (to be seen later during the *Quantum information and error correction* course);

In view of this, we propose the density matrix to be the correct mathematical notion of the state of a quantum system A, even when A is part of a larger joint system. Note that if system A is in a pure state $|\alpha\rangle$, its density matrix $|\alpha\rangle\langle\alpha|$ is independent of the global phase of $|\alpha\rangle$. This shows that global phase does not matter when studying the dynamics of a system in a pure state, and hence, a pure state is quite correctly modelled by a one-dimensional subspace of the Hilbert space.

2 Embedding classical computaton into quantum

Let $f : \{0,1\}^n \rightarrow \{0,1\}^m$ be a (deterministic) function. In general, the mapping $x \mapsto f(x)$ is not reversible; so it cannot be directly thought of as a unitary transformation. Hence, we use the following convention to 'embed' f into a deterministic reversible function $F(f) : \{0,1\}^{n+m} \rightarrow \{0,1\}^{n+m}$ defined by $(x,b) \mapsto (x, b \oplus f(x))$, where the first and second registers are n bits and m bits long respectively, and $b \oplus f(x)$ denotes the string of length m obtained by the bitwise XOR of the strings b and $f(x)$. We are now free to think of $F(f)$ as a unitary quantum map, extend by $\mathbb{C}$-linearity to arbitrary superpositions. Henceforth, when we talk of implementing f by a quantum computer, we will actually mean designing a circuit for the unitary $F(f)$ using gates from our (finite) basic set of unitaries.

We now need to replace the universal gates for classical, in general irreversible, computation by universal gates for classical reversible computation. It turns out that we only need a single reversible 3-bit gate for universal classical reversible computation called the *Toffoli* gate. It can be shown that one and two-bit reversible gates are not universal for reversible classical computation, unlike the situation with classical, in general irreversible, computation. The circuit diagram of the Toffoli gate is given below (note that this is just a schematic diagram and one should not think of the vertical lines as additional wires nor the junctions as fanouts).

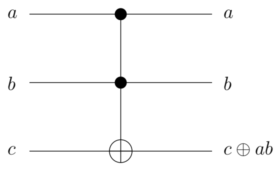

Figure 1. The Toffoli gate

A few words on the above schematic diagram are in order. In general, if U is a unitary on n qubits, then *controlled-*U is a unitary on $n+1$ qubits acting as follows, where the first register is a single qubit called the *control* and the second register contains n qubits:

$$\begin{array}{rcl} |0\rangle \otimes |x\rangle & \mapsto & |0\rangle \otimes |x\rangle \\ |1\rangle \otimes |x\rangle & \mapsto & |1\rangle \otimes (U|x\rangle), \end{array}$$

extended by $\mathbb{C}$-linearity to superpositions. Thus, controlled-U applies U on the second register iff the control qubit is $|1\rangle$.

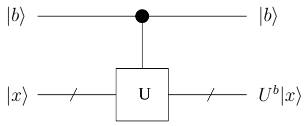

Figure 2. A controlled-U gate

This justifies the name controlled-NOT for the CNOT gate, also called controlled-X since the single qubit

Pauli operator X is nothing but the NOT gate. Viewed in this light, the Toffoli gate is a controlled-controlled-NOT.

Since NOT, NAND and FANOUT gates are universal for classical, in general irreversible, computation, it suffices to show how to simulate the last two by the Toffoli gate (since NOT is already reversible).

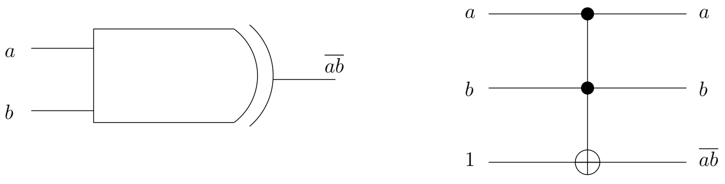

Figure 3. Implementing NAND using Toffoli

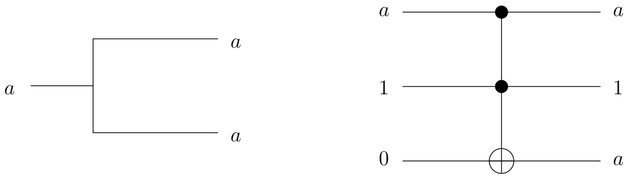

Figure 4. Implementing FANOUT using Toffoli

Thus, if a boolean circuit C computes $x \mapsto f(x)$ using possibly extra work bits (initialised to zero by convention), then we can embed C into a classical reversible circuit C' using possibly at most three times as many work bits. Now C' can be thought of as a quantum circuit also using $\mathbb{C}$-linearity to define its action on superpositions. The circuit C' implements a unitary transformation that happens to be a permutation of computational basis states. In general on input x with work bits initialised to zero, the output of C' will contain, besides $f(x)$, a string correlated to x on the remaining bits which we called garbage in our later discussion (the reason for this nomenclature will become clear very soon).

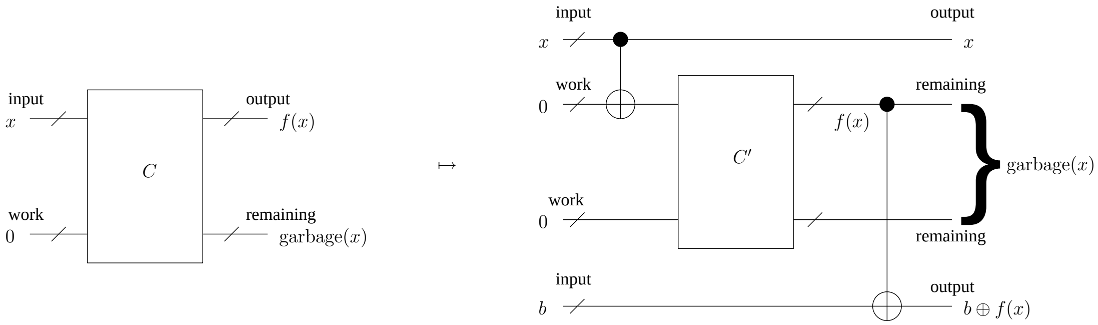

Figure 5. Making a classical circuit reversible naively

The two CNOTs depicted in the circuit as operating on several qubits are actually a *transversal system* of two-qubit CNOTs, as detailed in the next figure.

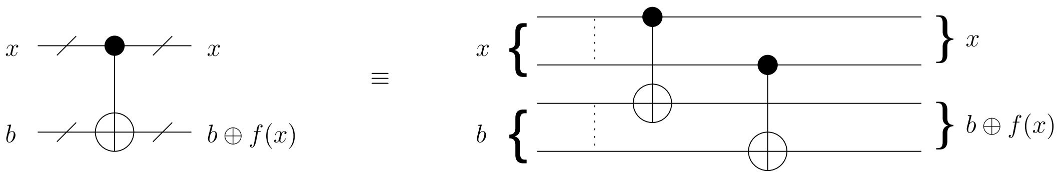

Figure 6. A transversal system of two-qubit CNOTs

Such transversal systems are very useful in fault tolerant quantum computation.

In a classical randomised circuit C, we have an additional input besides x and the work bits called the random string r where uniform independent random bits are fed. We can transform C to a classical reversible circuit C' as before, with x, b, r and work bits as its input. We can then transform C' into a quantum circuit C'' where the random input to r is generated not by a randomness source but by feeding zeroes to Hadamard gates and measuring their outputs in the computational basis.

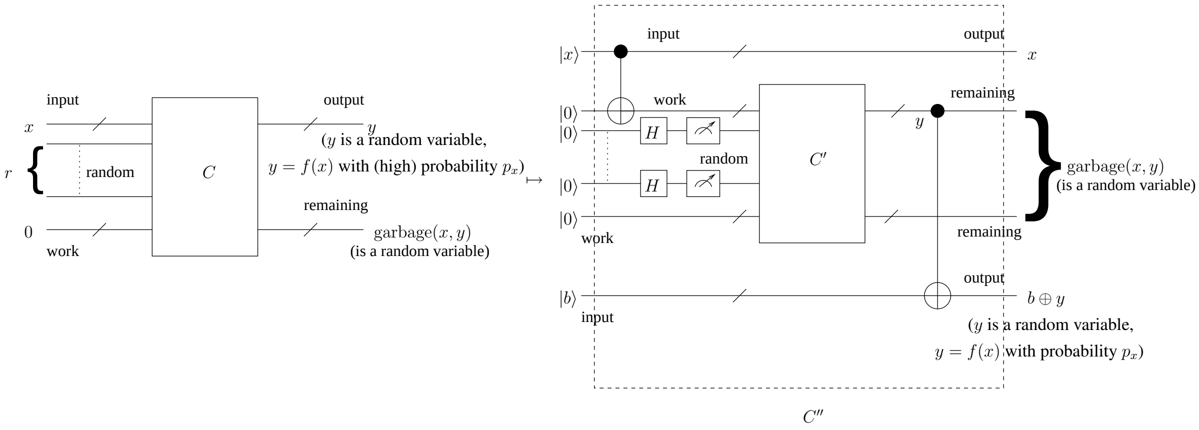

Figure 7. Making a classical randomised circuit quantum naively with intermediate measurements

We can avoid the measurement just after the Hadamards by using a general principle of replacing measurement in the computational basis by bringing in fresh ancilla qubits initialised to zero, applying CNOT and then ignoring these ancillas for the rest of the computation. It is easy to see that this principle mimics measurement in the computational basis by doing a reduced state calculation for the original qubits of the circuit.

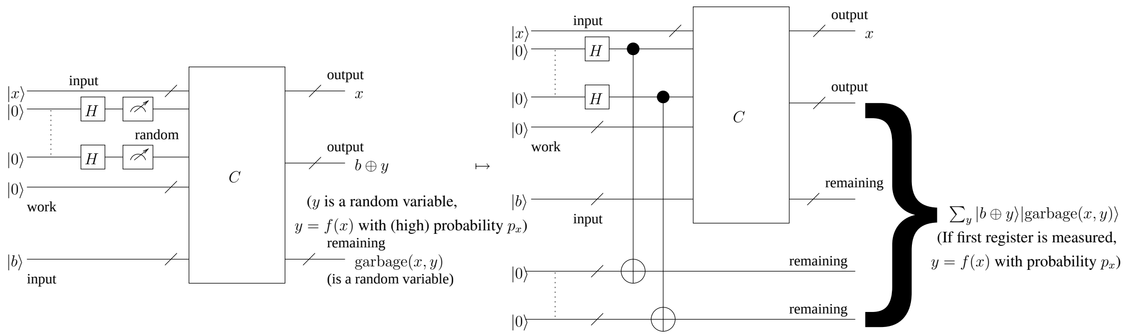

Figure 8. Removing intermediate measurements in a quantum circuit

From the above discussion, we can conclude that $\mathbf{P} \subseteq \mathbf{BPP} \subseteq \mathbf{BQP}$, where $\mathbf{P}$, $\mathbf{BPP}$ and $\mathbf{BQP}$ are the class of problems that can be solved using polynomial time classical deterministic, classical randomised and quantum algorithms respectively. There is an issue about uniformity of the circuit families out here, which we resolve by requiring that all circuit families, classical or quantum, be generated by classical deterministic Turing machines running in polynomial time.

We shall see in the next couple of lectures that there is a finite universal set of basic quantum gates consisting of only one and two-qubit unitaries. Now, given a uniformly generated polynomial size quantum circuit composed of gates from this universal set, we can write down the $M \times M$ unitary matrix for the overall circuit in classical deterministic exponential time ($\mathbf{EXP}$), where M is the dimension of the Hilbert space of the overall circuit. Note that M is single exponential in the input size. This is because we can multiply out the unitaries corresponding to the individual gates of the circuit in deterministic time polyomial in M. This shows that $\mathbf{BQP} \subseteq \mathbf{EXP}$. We can prove the tighter containment $\mathbf{BQP} \subseteq \mathbf{PSPACE}$, where $\mathbf{PSPACE}$ is the class of problems that can be solved in polynomial space by a classical deterministic algorithm, by observing that the multiplication described above can be done in deterministic space polynomial in $\log M$, which is polynomial in the input size. We have thus shown

$$\mathbf{P} \subseteq \mathbf{BPP} \subseteq \mathbf{BQP} \subseteq \mathbf{PSPACE}.$$

Since $\mathbf{P} \overset{?}{=} \mathbf{PSPACE}$ is still open (resolving it will be a big breakthrough for computer science), we do not know for sure if quantum computation could ever offer a superpolynomial speedup over classical deterministic computation. We have evidence of superpolynomial speedups for some problems like integer factoring which is not known to be in $\mathbf{BPP}$ but which lies in $\mathbf{BQP}$ due to Shor's algorithm. But at the present moment, that's all we have, evidence but no proof.

Getting rid of garbage: The embedding of a classical circuit C into a quantum circuit C' described above generates garbage, which is bad if C' is to be used as a unitary subroutine in a larger quantum algorithm. Suppose we would like to ideally use the reversible function $F(f) : |x\rangle|b\rangle \mapsto |x\rangle|b \oplus f(x)\rangle$ as a unitary subroutine in a quantum algorithm. Suppose however, what we actually have is a 'dirty' deterministic reversible implementation of $F(f)$ using possibly some work qubits, of the form

$$|x\rangle|b\rangle|0\rangle \mapsto |x\rangle|b \oplus f(x)\rangle|\mathrm{garbage}(x, b)\rangle.$$

Above, $|\text{garbage}(x,b)\rangle$ indicates that the work qubits at the end of the 'dirty' implementation may contain a string that is correlated to the input (x,b). Now suppose that the outer quantum algorithm prepares a superposition $\sum_{x,b}\alpha_{x,b}|x\rangle|b\rangle$ and feeds it to the unitary $F(f)$. It would ideally expect the resulting state to be $\sum_{x,b}\alpha_{x,b}|x\rangle|b\oplus f(x)\rangle$. However, because of our 'dirty' implementation of $F(f)$, the actual resulting state will be

$$\sum_{x,b}\alpha_{x,b}|x\rangle|b\oplus f(x)\rangle|\text{garbage}(x,b)\rangle.$$

In general, the reduced state of the first two registers after the application of the 'dirty' $F(f)$ will be nowhere close to the state of the first two registers after the application of the ideal $F(f)$.

Thankfully, we can get get rid of garbage in a deterministic reversible implementation of $F(f)$ by the following procedure, which gives a 'clean' reversible circuit for $F(f)$ with complexity at most a constant times the original deterministic, in general irreversible, circuit complexity of f. By a 'clean' reversible circuit for $F(f)$, we mean a map of the form $|x\rangle|b\rangle|0\rangle\mapsto|x\rangle|b\oplus f(x)\rangle|0\rangle$ implemented by reversible gates.

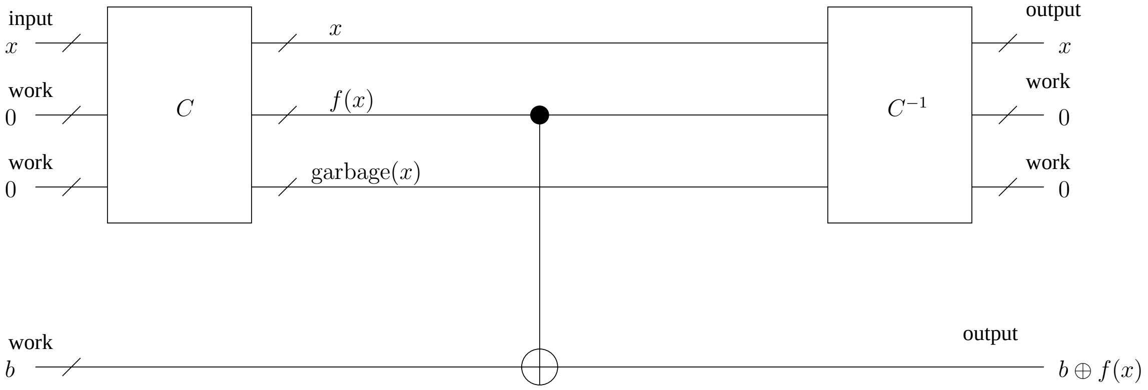

Figure 9. Cleaning garbage

The CNOT gate in the above figure is actually transversal. Note that if the reversible circuit C is made up of NOT, Hadamard, CNOT and Toffoli gates only, then $C^{-1}=C$ as these gates are their own inverses.

The above procedure can also be applied to 'dirty' quantum circuits that evaluate $F(f)$ with bounded error, e.g. as in Figure 2, but the resulting circuit will, in general, only be an approximation to a clean implementation $|x\rangle|b\rangle|0\rangle\mapsto|x\rangle|b\oplus f(x)\rangle|0\rangle$. The approximation error can be reduced by reducing the error in evaluting $F(f)$ by the 'dirty' circuit, which can be done by standard techniques. Suppose on input $|x\rangle|0\rangle|0\rangle$, the second output of the quantum circuit C in Figure 2 would give $f(x)$ with probability $1-\epsilon$, if measured, that is, the output of C is the superposition $|x\rangle\otimes(\sum_y|y\rangle|\text{garbage}(x,y)\rangle)$, where $\||\text{garbage}(x,f(x))\rangle\|^2\geq 1-\epsilon$. The global state of the four registers is $|x\rangle\otimes(\sum_y|y\rangle|\text{garbage}(x,y)\rangle)\otimes|b\rangle$. After applying the CNOT gate, it becomes $|x\rangle\otimes(\sum_y|y\rangle|\text{garbage}(x,y)\rangle|b\oplus y\rangle)$, which is within an ℓ_2-distance of $\sqrt{\epsilon}$ from $|x\rangle\otimes(\sum_y|y\rangle|\text{garbage}(x,y)\rangle)\otimes|b\oplus f(x)\rangle$. This shows that the final state of the circuit of Figure 2 is within an ℓ_2-distance of $\sqrt{\epsilon}$ from the ideal state $|x\rangle|0\rangle|0\rangle|b\oplus f(x)\rangle$. Also by construction, the reduced state of the first register in the final state is $|x\rangle$. This implies by the Cauchy-Schwarz inequality that for any superposition $\sum_{x,b} \alpha_{x,b} |x\rangle|0\rangle|0\rangle|b\rangle$ as input, the output state of the circuit is within an ℓ_2-distance of $\sqrt{\epsilon 2^m}$ from the ideal state $\sum_{x,b} \alpha_{x,b} |x\rangle|0\rangle|0\rangle|b \oplus f(x)\rangle$, where the range of f is $\{0,1\}^m$. The important point to note is that the above bound on the approximation error is independent of the number of work qubits, which enables us to reduce ϵ by standard techniques (which in general increase the number of work qubits), decreasing the approximation error to an exponentially small quantity at the expense of a polynomial blowup in the circuit size. Now, it can be seen via Equation 1 that if the circuit is used in place of $F(f)$ in a larger quantum algorithm, the larger algorithm will produce an output state within ℓ_2-distance of $\sqrt{\epsilon 2^m}$ of its ideal output state. This implies, as shall be seen later on in the *Quantum information and error correction* course, that the probability distribution given by any measurement of the actual output state of the larger quantum algorithm is within ℓ_1-distance of $2\sqrt{\epsilon 2^m}$ from the probability distribution given by the same measurement on the ideal output state.

Sometimes, when $x \mapsto f(x)$ is a permutation of bit strings of the same length, we may want to use f itself as a unitary subroutine of a larger quantum algorithm. Hence, we want a 'clean' implementation $|x\rangle|0\rangle \mapsto |f(x)\rangle|0\rangle$ of f. It turns out that we can do this if we have, possibly irreversible, classical circuits for f and f^{-1}. Given classical circuits for f and f^{-1}, we can construct 'clean' reversible circuits computing $F(f)$ and $F(f^{-1})$ exactly. Using them, we can get a 'clean' implementation of f as follows:

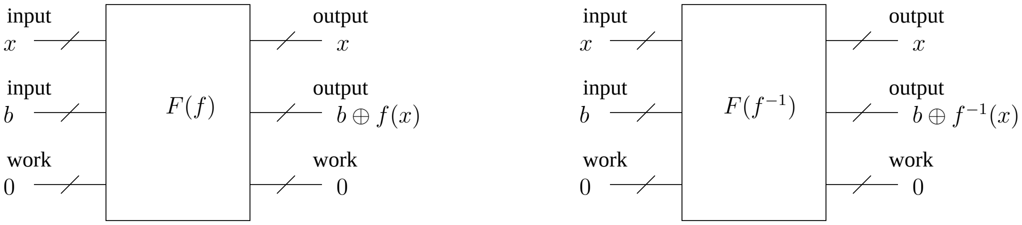

Figure 10. Clean reversible implementations of $F(f)$ and $F(f^{-1})$

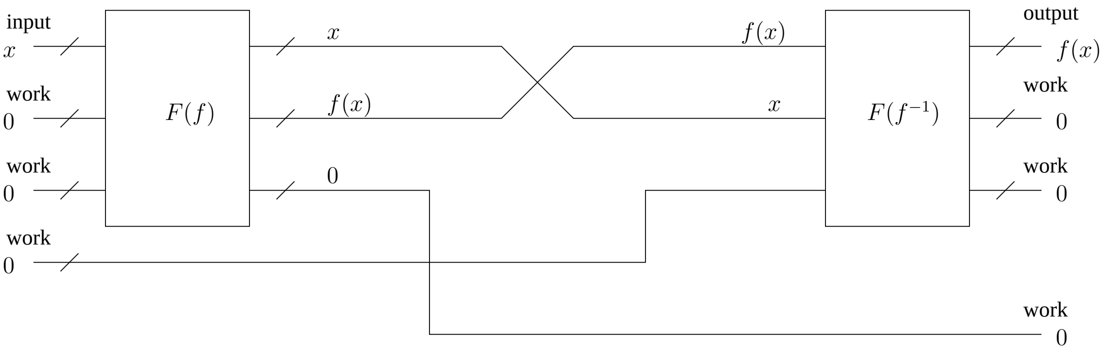

Figure 11. Clean reversible implementation of an invertible function

A similar comment as above can be made if we start off with 'dirty' reversible circuits evaluating $F(f)$ and $F(f^{-1})$ with bounded error.