Quantum algorithms **Monsoon 2007**

Lecture 6

Lecturer: Pranab Sen *Date: August 16*

In this and the next lecture, we shall learn how to approximate a unitary operation on n qubits by a circuit composed of CNOT and single qubit gates. Moreover, we shall show that only a finite set of single qubit gates are required, thus proving the existence of a finite universal set of basic gates for quantum computation. Our approach will be different from the approach detailed in Nielsen and Chuang's textbook; in fact, our approach is the solution to a problem at the end of Chapter 4 of the textbook.

The advantage of the approach in the textbook is that it shows by a direct decomposition that any unitary on n qubits can be expressed exactly by a circuit consisting of CNOTs and single qubit gates, provided arbitrary single qubit gates are allowed. The textbook then indicates how an arbitrary single qubit gate can be approximated by a circuit made up of a finite universal set of basic single qubit gates. Our approach will incur an approximation error in decomposing a unitary on n qubits into a circuit of CNOTs and single qubit gates, but it will use a restricted, though still infinite, set of single qubit gates. The attraction in our approach is that it uses some nice and important mathematics on the way, and also, it is easier to prove that the single qubit gates arising in our approach can be approximated by circuits of gates from a finite universal set of basic single qubit gates. At the end of the day, the textbook approach appeals to a deep result called *Solovay-Kitaev theorem* in order to improve the approximation error of the circuit for a generic n-qubit unitary at the expense of a mild blowup in the circuit size; it uses the theorem for 2×2 unitaries. Similarly, we can use the Solovay-Kitaev theorem for $2^n \times 2^n$ unitaries to achieve a similar approximation errors and circuit sizes for generic n-qubit unitaries, however, it will be less practical than the textbook approach.

Before we can dive into our decomposition technique, we need to develop a couple of mathematical concepts, namely, exponentials of linear operators and generalised Pauli matrices on n qubits.

1 Exponential of a linear operator

Let A be a complex $n \times n$ matrix. The exponential of A is defined as

$$e^A := \mathbb{1} + A + \frac{A^2}{2!} + \cdots$$

It is fairly easy to show that the series converges entry-by-entry to some $n \times n$ matrix; thus, the matrix exponential is well-defined. For an invertible $n \times n$ matrix B, it is easy to see that $e^{BAB^{-1}} = Be^AB^{-1}$, which proves that the exponential is actually a map from linear operators on $\mathbb{C}^n$ to linear operators on $\mathbb{C}^n$.

If $n \times n$ matrices A, B commute, then it is easy to see that $e^{A+B} = e^A e^B$. In general however, $e^{A+B} \neq e^A e^B$. The *Trotter formula* proved below quantifies the difference between the left and right hand sides of the above expression in terms of the *spectral norms* of A and B. Recall that the *spectral norm* of an operator A on $\mathbb{C}^n$, also known as the $\ell - 2$*-induced operator norm*, is defined by $\|A\| := \max_{v:\|v\|=1} \|Av\|$, where the norm for vectors is assumed to be the usual ℓ_2 (Euclidean) norm. It is easy to verify the norm properties of $\|\cdot\|$ and also its submultiplicativity viz. $\|AB\| \leq \|A\| \cdot \|B\|$. Also, it is easy to check that $\|A\|$ is equal to the largest singular value of A.

Theorem 1.1 (Trotter formula). *For complex $n \times n$ matrices A, B such that $\|A\| + \|B\| \le 1$,*

$$\|e^{A+B} - e^A e^B\| \le (\|A\| + \|B\|)^2.$$

Proof: It is easy to see that

$$\begin{aligned} e^{A+B} - e^A e^B &= \frac{BA - AB}{2} + \left(\frac{(A+B)^3}{3!} + \cdots\right) - \left(\frac{A^3}{3!} + \cdots\right) - \left(\frac{B^3}{3!} + \cdots\right) - \\ & \quad A\left(\frac{B^2}{2!} + \cdots\right) - \left(\frac{A^2}{2!} + \cdots\right)B + \left(\frac{A^2}{2!} + \cdots\right)\left(\frac{B^2}{2!} + \cdots\right). \end{aligned}$$

Using the triangle inequality (which also holds for the infinite summation here since the exponential series converges) and submultiplicativity of the spectral norm, we get

$$\begin{aligned} \|e^{A+B} - e^A e^B\| &\le \frac{\|AB - BA\|}{2} + \left(\frac{(\|A\| + \|B\|)^3}{3!} + \cdots\right) + \left(\frac{\|A\|^3}{3!} + \cdots\right) + \left(\frac{\|B\|^3}{3!} + \cdots\right) + \\ & \quad \|A\|\left(\frac{\|B\|^2}{2!} + \cdots\right) + \left(\frac{\|A\|^2}{2!} + \cdots\right)\|B\| + \left(\frac{\|A\|^2}{2!} + \cdots\right)\left(\frac{\|B\|^2}{2!} + \cdots\right). \end{aligned}$$

Using elementary calculus, it can be checked that $\frac{x^3}{3!} + \cdots \le \frac{x^3}{3}$ and $\frac{x^2}{2!} + \cdots \le x^2$ for $0 \le x \le 1$. Hence,

$$\begin{aligned} \|e^{A+B} - e^A e^B\| &\le \frac{\|AB - BA\|}{2} + \frac{(\|A\| + \|B\|)^3}{3} + \left(\frac{\|A\|^3}{3} + \frac{\|B\|^3}{3} + \|A\|\|B\|^2 + \|A\|^2\|B\|\right) + \\ & \quad \|A\|^2\|B\|^2 \\ &\le \|A\| \cdot \|B\| + \frac{2(\|A\| + \|B\|)^3}{3} + \|A\|^2\|B\|^2 \\ &\le \frac{(\|A\| + \|B\|)^2}{4} + \frac{2(\|A\| + \|B\|)^3}{3} + \frac{(\|A\| + \|B\|)^4}{16} \\ &\le (\|A\| + \|B\|)^2, \end{aligned}$$

where the last two inequalities are easy to verify. This completes the proof of the theorem. □

We conclude this section with a digression about how to compute e^A for an arbitrary complex $n \times n$ matrix A in time polynomial in n, assuming that e^z, $z \in \mathbb{C}$ can be computed in unit time. The *Jordan canonical form* expresses A as $A = B(D + N)B^{-1}$, where B is invertible, D is diagonal, N is upper triangular with zeroes on the diagonal, and $DN = ND$. Then, $e^A = Be^D e^N B^{-1}$. Since D is diagonal, e^D is easy to compute (just take exponentials of the diagonal elements). It is easy to verify that $N^n = 0$; thus, e^N is a finite series and can be computed in time polynomial in n. Since the Jordan canonical form can be computed in time polynomial in n, we get a polynomial time algorithm for computing e^A. The Jordan canonical form also enables us to prove the following interesting result:

Proposition 1.1. *For any complex $n \times n$ matrix A, $\det A = e^{\mathrm{Tr}\, A}$.*

Proof: Consider the Jordan canonical form $A = B(D + N)B^{-1}$ described above. It is obvious that $\mathrm{Tr}\, A = \mathrm{Tr}\,(D + N) = \mathrm{Tr}\, D$. Observe that for a diagonal matrix D, $\det e^D = e^{\mathrm{Tr}\, D}$. Also, e^N is an upper triangular matrix with ones on the diagonal; hence $\det e^N = 1$. Thus, $\det e^A = (\det e^D)(\det e^N) = e^{\mathrm{Tr}\, D} = e^{\mathrm{Tr}\, A}$. □

2 Generalised Pauli matrices on n qubits

The linear operators on $\mathbb{C}^n$ form an inner product space over $\mathbb{C}$ under the *Hilbert-Schmidt inner product* $\langle A, B\rangle := \mathrm{Tr}\, A^\dagger B$. Fix an orthonormal basis $|p\rangle$, $p = 1, \ldots, n$ for $\mathbb{C}^n$. If we think of linear operators A, B as matrices with respect to this orthonormal basis, the Hilbert-Schmidt inner product is nothing but the usual dot product applied to the vectors obtained by stretching out the matrices A, B, that is, $\langle A, B\rangle = \sum_{p,q} A^*_{pq} B_{pq}$. The norm on operators defined by the Hilbert-Schmidt inner product is called the *Frobenius norm* $\|A\|_F := \sqrt{\mathrm{Tr}\, A^\dagger A} = \sqrt{\sum_{p,q} |A_{ij}|^2}$. Observe that $\|A\|_F = \|UAV\|_F$ for any unitary operators U, V on $\mathbb{C}^n$. This implies that $\|A\|_F = \sqrt{\sum_{p=1}^n s_p^2(A)}$, where $\{s_p(A)\}_{p=1}^n$ are the singular values of A. The n^2 linear operators $|p\rangle\langle q|$, $p, q = 1, \ldots, n$ form an orthonormal basis under the Hilbert-Schmidt inner product for the space of linear operators on $\mathbb{C}^n$.

Now, observe that the Hermitian linear operators on $\mathbb{C}^n$ form an inner product space over $\mathbb{R}$ under the Hilbert-Schmidt inner product, that is, $\mathbb{R}$-linear combinations of Hermitian operators are Hermitian, and $\langle A, B\rangle \in \mathbb{R}$ for Hermitian A, B as can be easily checked. The dimension of this $\mathbb{R}$-linear space is also n^2; the linear operators $|p\rangle\langle p|$, $\frac{|p\rangle\langle q| + |q\rangle\langle p|}{\sqrt{2}}$ and $\frac{-\sqrt{-1}|p\rangle\langle q| + \sqrt{-1}|q\rangle\langle p|}{\sqrt{2}}$, $p, q = 1, \ldots, n, p \neq q$ form an orthonormal basis under the Hilbert-Schmidt inner product. For $n = 2$, the 2×2 Pauli matrices

$$\mathbb{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},\ X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},\ Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix},\ Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

where $i := \sqrt{-1}$ form an orthogonal basis for the $\mathbb{R}$-linear space of 2×2 complex Hermitian matrices. Morevoer, $\|P\|_F = \sqrt{2}$ for all $P \in \{\mathbb{1}, X, Y, Z\}$.

The n-fold tensor products of the 2×2 Pauli matrices are called *generalised Pauli matrices on n qubits*. By standard properties of tensor product, the 4^n generalised Pauli matrices on n qubits are orthogonal under Hilbert-Schmidt inner product and each of them has Frobenius norm $\sqrt{2^n}$. Thus, the generalised Pauli matrices on n qubits form an orthogonal basis for the $\mathbb{R}$-linear space of $2^n \times 2^n$ complex Hermitian matrices. In the rest of this lecture and the next, we use $\vec{P}$ to denote a generalised Pauli matrix on n qubits. Hence, any $2^n \times 2^n$ Hermitian matrix H can be expressed as $H = \sum_{\vec{P}} h_{\vec{P}} \vec{P}$, where $h_{\vec{P}} \in \mathbb{R}$. Moreover for any $\vec{P}$, $\langle \vec{P}, H\rangle = h_{\vec{P}} \langle \vec{P}, \vec{P}\rangle$ implying that

$$|h_{\vec{P}}| \leq \frac{\|\vec{P}\|_F \cdot \|H\|_F}{\|\vec{P}\|_F^2} \leq \frac{\|H\|\sqrt{2^n}}{\|P\|_F} = \|H\|.$$

The two inequalities above follow from Cauchy-Schwarz inequality and the facts that the spectral and Frobenius norms of H are the ℓ_∞ and ℓ_2-norms of the vector of singular values of H.

3 Simplifying arbitrary n-qubit unitary matrices

In this section, we take the first step towards approximating an arbitrary n-qubit unitary U by a circuit composed of CNOTs and single qubit gates. We will show in this section that U can be approximated by a product of unitaries of a special kind viz. those unitaries whose Hamiltonians are scaled generalised Pauli matrices on n qubits.

By working in the eigenbasis of U and using the spectral theorem, it is easy to see that U can be written as $U = e^{-iH}$, where $i = \sqrt{-1}$ and H is an n-qubit Hermitian matrix with eigenvalues in the range $(-\pi, \pi]$, that is, $\|H\| \leq \pi$. Thus, H can be expressed as $H = \sum_{\vec{P}} h_{\vec{P}} \vec{P}$, where $h_{\vec{P}} \in \mathbb{R}$ and $|h_{\vec{P}}| \leq \|H\| \leq \pi$.

Lemma 3.1. *Let H be a $2^n \times 2^n$ Hermitian matrix and $0 \leq \delta \leq (3\pi)^{-1}$. Let $H = \sum_{\vec{P}} h_{\vec{P}} \vec{P}$, $h_{\vec{P}} \in \mathbb{R}$. Then,*

$$\left\| e^{-iH\delta} - \prod_{\vec{P}} e^{-i\delta h_{\vec{P}} \vec{P}} \right\| \leq (3\pi\delta)^2 4^n.$$

Proof: The proof is by induction on the number of non-zero terms m in the decomposition of H into a $\mathbb{R}$-linear combination of generalised Pauli matrices on n qubits. The base case $m = 1$ is trivial. Now suppose $m > 1$. Let $T(m)$ be the left hand side of the above inequality when the decomposition of H has m non-zero terms, that is, $H = \sum_{l=1}^{m} h_l \vec{P}_l$. Then, by the triangle inequality and submultiplicativity of the spectral norm,

$$\begin{aligned} T(m) &:= \left\| e^{-i\delta \sum_{l=1}^{m} h_l \vec{P}_l} - \prod_{l=1}^{m} e^{-i\delta h_l \vec{P}_l} \right\| \\ &\leq \left\| e^{-i\delta \sum_{l=1}^{m} h_l \vec{P}_l} - e^{-i\delta \sum_{l=1}^{m-1} h_l \vec{P}_l} \cdot e^{-i\delta h_m \vec{P}_m} \right\| + \\ &\quad \left\| e^{-i\delta \sum_{l=1}^{m-1} h_l \vec{P}_l} - \prod_{l=1}^{m-1} e^{-i\delta h_l \vec{P}_l} \right\| \cdot \|e^{-i\delta h_m \vec{P}_m}\| \\ &\leq \left(\left\| \delta \sum_{l=1}^{m-1} h_l \vec{P}_l \right\| + \|\delta h_m \vec{P}_m\| \right)^2 + T(m-1) \cdot 1. \end{aligned}$$

The last inequality follows because $e^{-i\delta h_m \vec{P}_m}$ is a unitary matrix, and by using Trotter's formula (Theorem 1.1). Note that we can indeed use Trotter's formula because

$$\begin{aligned} \left\| \delta \sum_{l=1}^{m-1} h_l \vec{P}_l \right\| + \|\delta h_m \vec{P}_m\| &= \delta(\|H - h_m \vec{P}_m\| + \|h_m \vec{P}_m\|) \\ &\leq \delta(\|H\| + 2h_m \|\vec{P}_m\|) \\ &\leq \delta(\pi + 2\pi \cdot 1) \\ &= 3\pi\delta \leq 1. \end{aligned}$$

Above, we used the triangle inequality and the facts that $\vec{P}_m$ is unitary, $|H| \leq \pi$ and $|h_m| \leq |H| \leq \pi$.

We thus get a recurrence $T(m) \leq (3\pi\delta)^2 + T(m-1)$, $T(1) = 0$. Solving the recurrence, we get $T(m) \leq (3\pi\delta)^2(m-1)$. Putting $m \leq 4^n$ gives us the right hand side of the desired inequality. □

Lemma 3.2. *Let H be a $2^n \times 2^n$ Hermitian matrix, $0 \leq \epsilon \leq 1$ and $k := (3\pi)^2 4^n \epsilon^{-1}$. Let $H = \sum_{\vec{P}} h_{\vec{P}} \vec{P}$, $h_{\vec{P}} \in \mathbb{R}$. Then,*

$$\left\| e^{-iH} - \left(\prod_{\vec{P}} e^{-ih_{\vec{P}} \vec{P}/k} \right)^k \right\| \leq \epsilon.$$

Proof: Let $1 \leq m \leq k$. By induction on m, we shall show that

$$T(m) := \left\| e^{-imH/k} - \left(\prod_{\vec{P}} e^{-ih_{\vec{P}}\vec{P}/k} \right)^m \right\| \leq \frac{m\epsilon^2}{(3\pi)^2 4^n}.$$

The base case $m = 1$ holds because we can set $\delta := k^{-1}$ in the previous lemma (Lemma 3.1) to get

$$\left\| e^{-iH/k} - \prod_{\vec{P}} e^{-ih_{\vec{P}}\vec{P}/k} \right\| \leq \frac{\epsilon^2}{(3\pi)^2 4^n}.$$

Now suppose $m > 1$. We use the triangle inequality and submultiplicativity of $\|\cdot\|$ to prove the induction step.

$$\begin{aligned} T(m) &\leq \left\| e^{-imH/k} - e^{-i(m-1)H/k} \prod_{\vec{P}} e^{-ih_{\vec{P}}\vec{P}/k} \right\| + \\ &\quad \left\| e^{-i(m-1)H/k} \prod_{\vec{P}} e^{-ih_{\vec{P}}\vec{P}/k} - \left(\prod_{\vec{P}} e^{-ih_{\vec{P}}\vec{P}/k} \right)^m \right\| \\ &= \left\| e^{-i(m-1)H/k} \right\| \cdot \left\| e^{-iH/k} - \prod_{\vec{P}} e^{-ih_{\vec{P}}\vec{P}/k} \right\| + \\ &\quad \left\| e^{-i(m-1)H/k} - \left(\prod_{\vec{P}} e^{-ih_{\vec{P}}\vec{P}/k} \right)^{m-1} \right\| \cdot \left\| \prod_{\vec{P}} e^{-ih_{\vec{P}}\vec{P}/k} \right\| \\ &\leq 1 \cdot \frac{\epsilon^2}{(3\pi)^2 4^n} + T(m-1) \cdot 1 \\ &\leq \frac{m\epsilon^2}{(3\pi)^2 4^n}. \end{aligned}$$

Above, we use the fact that $e^{-i(m-1)H/k}$ and $\prod_{\vec{P}} e^{-ih_{\vec{P}}\vec{P}/k}$ are unitary. Setting $m = k$ finishes the proof of the lemma. □

So far we have seen how to approximate a $2^n \times 2^n$ unitary U to within spectral distance ϵ by a product of $(3\pi)^2 4^{2n} \epsilon^{-1}$ unitaries of the form $e^{-i\alpha\vec{P}}$, $\alpha \in \mathbb{R}$ and $\vec{P}$ a generalised Pauli matrix on n qubits. We performed the approximation by breaking up the time evolution of the Hamiltonian of U into small intervals, and then approximating the evolution in each interval by a product of unitaries of the above special kind. The crucial reason why we could approximate U to an arbitrary accuracy was that the approximation error in a single interval was $O(k^{-2})$, where k is the number of intervals. The overall approximation error became $k \cdot O(k^{-2}) = O(k^{-1})$. Thus, increasing the number of intervals k increased the accuracy of approximating U. The same idea extends to discretely integrating the Schrödinger equation; increasing the number of intervals increases the accuracy of the numerical solution.

In the next lecture, we shall see how to implement a unitary of the form $e^{-i\alpha\vec{P}}$ using $O(n)$ CNOT gates and single qubit gates of the form $e^{-i\beta P}$, $\beta \in \mathbb{R}$ and P a Pauli matrix. Note that in general, $e^{-i\alpha\vec{P}}$ is an entangled operation over n qubits. After that, we shall indicate how we can approximate any gate of the form $e^{-i\beta P}$ by repeated applications of a basic gate $e^{-i\beta_0 P}$, where β_0 is a fixed irrational multiple of π.

Finally, we motivate the reason behind using the spectral distance to approximate a desired unitary U. Suppose U' is an approximant to U within spectral distance ϵ. Then, if the initial state of the quantum computer is $|\psi_0\rangle$, the final states of the computer if the ideal unitary versus the approximating unitary were used will be $|\psi\rangle := U|\psi_0\rangle$ versus $|\psi'\rangle := U'|\psi_0\rangle$. The difference between the two final states will be

$$\||\psi\rangle - |\psi'\rangle\| = \|U|\psi_0\rangle - U'|\psi_0\rangle\| \leq \|U - U'\| \leq \epsilon.$$

Now let P, P' be the probability distributions got by performing a measurement in the computational basis on states $|\psi\rangle$, $|\psi'\rangle$. It will be proved later on in the *Quantum information and error correction* course that $\|P - P'\|_1 \leq 2\||\psi\rangle - |\psi'\rangle\| \leq 2\epsilon$, where $\|P - P'\|_1 := \sum_m |P(m) - P'(m)|$ is the ℓ_1-distance, also known as the *total variation distance*, between P, P' (here m ranges over all possible computational basis states). In fact, the above bound on the total variation distance also holds for generalised measurements. This shows that P, P' are nearly indistinguishable information-theoretically. In short, approximating a unitary in the spectral distance ensures that the actual output probability distribution is information-theoretically nearly indistinguishable from the ideal output probability distribution.