

# Lecture Topic: Grover Search and Dynamic Programming

# Overview

So far, we have seen examples of quantum algorithms with an exponential speed-up, but only for problems that are *not* NP-Hard. For NP-Hard problems, we know only algorithms with quadratic speed-up so far, and even that is disputed.

# Grover Problem

In the problem of Grover, we have

- a dimension  $n$
- a black-box function  $f(x) : \{0, 1\}^n \rightarrow \{0, 1\}$  parametrized by a secret  $n$ -bit string  $j$ , which returns 1 if  $x = j$  and 0 otherwise.

The functional version of the problem asks what is the unknown  $w$ . It is clear that classically, one may need to perform  $2^n$  oracle calls in the worst-case, and that randomized algorithms would not help much. Notice that  $N = 2^n$  is sometimes referred to as the “library size” we are searching.

# The Oracle

The black-box function is usually thought of as an *oracle operator*  $U_w$  such that for states  $|j\rangle$  in the computational basis

$$U_w |j\rangle = (-1)^{f(j)} |j\rangle = \mathbf{I} - 2 |w\rangle \langle w| = \begin{cases} -|j\rangle, & \text{if } j = w \\ +|j\rangle, & \text{if } j \neq w \end{cases} \quad (1.1)$$

This is sometimes known as the  $\pm$ -oracle or phase oracle. One can generalize this to the situation where there are multiple secrets.

## Another Operator

We will also use a generalization, which is known as the *diffusion operator* or inversion for an arbitrary state  $|s\rangle$ :

$$U_s = 2 |s\rangle \langle s| - \mathbf{I} \quad (1.2)$$

Let us have a bit of a geometric detour: any state  $|\phi\rangle$  can be uniquely expressed as  $|\phi\rangle = \alpha |\psi\rangle + \beta |\psi^\perp\rangle$ , where  $|\psi^\perp\rangle$  is orthogonal to  $|\psi\rangle$ . Then:

$$U_s |\phi\rangle = -\alpha |\psi\rangle + \beta |\psi^\perp\rangle \quad (1.3)$$

that is, amplitudes of basis states orthogonal to  $|\psi\rangle$  are left unchanged, while signs of amplitudes of the basis state  $|\psi\rangle$  are flipped. Furthermore, for any state  $\phi$ ,  $U_\psi$  preserves the subspace spanned by  $|\phi\rangle$  and  $|\psi\rangle$ .

# Grover's Algorithm

Grover's algorithm performs the following steps:

- 1 creates an initial,  $n$ -qubit state  $|0\rangle^{\otimes n}$
- 2 apply Hadamard transform on it to obtain the uniform superposition  $\frac{1}{\sqrt{n}} \sum_{k=0}^{n-1} |x\rangle$
- 3 apply the function oracle operator  $U_w$  and the diffusion operator  $U_s$  repeatedly,  $q$  times.
- 4 obtain  $\hat{w}$  by measuring the  $n$ -qubit register. With probability  $\sin^2((q + \frac{1}{2})\theta)$  for some  $\theta$  depending on  $\frac{1}{\sqrt{N}}$ , estimate  $\hat{w}$  will be the correct  $f(\hat{w}) = 1$ . Otherwise, we repeat.

Ideally, one considers  $q \approx \frac{\pi}{4} 2^{n/2}$ . If the Grover iteration  $U_s U_w$  could be implemented in unit time (a big if!), this would correspond to  $O(2^{n/2}) = O(\sqrt{2^n}) = O(\sqrt{N})$  algorithm and quadratic speed-up compared to the linear search in time  $O(2^n) = O(N)$ .

# The Compound-Pendulum View

The Grover iteration has a number of appealing interpretations: Perhaps the most physical is due to Grover and Sengupta, which will surely please the Kybernetika and Robotika students.

Recall the discussion of the oscillators from the second lecture. The oscillator could describe a weight (or bob) suspended from a pivot on a (massless) cord such that the bob can swing freely. Now, consider  $N$  oscillators, one of which has a slightly shorter cord, and hence a different frequency. We seek to find the one with the shorter cord. We could check the frequency of the  $N$  oscillators one by one. Alternatively, we can consider a compound pendulum.

# The Compound-Pendulum View

The Grover iteration has a number of appealing interpretations: Perhaps the most physical is due to Grover and Sengupta, which will surely please the Kybernetika and Robotika students.

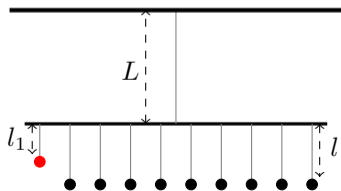
Recall the discussion of the oscillators from the second lecture. The oscillator could describe a weight (or bob) suspended from a pivot on a (massless) cord such that the bob can swing freely. Now, consider  $N$  oscillators, one of which has a slightly shorter cord, and hence a different frequency. We seek to find the one with the shorter cord. We could check the frequency of the  $N$  oscillators one by one. Alternatively, we can consider a compound pendulum.



# The Compound-Pendulum View

To this end, we consider a system where the  $N$  oscillators are suspended from a *support pendulum*. We use the following notation:

- The length, mass and displacement coordinate for the support pendulum are denoted  $L, M, X$ ;
- the pendulum we aim to identify has length, mass and displacement  $l_1, \frac{m_1}{N}, x_1$ ;
- the remaining  $N - 1$  oscillators have length, mass and displacements  $l, \frac{m}{N}, x_j$  for  $j = 2, \dots, N$ .



## The Compound-Pendulum View

The Lagrangian (kinetic energy *minus* potential energy) is then:

$$\frac{1}{2}[M\dot{X}^2 - KX^2 + \frac{1}{N}(m_1\dot{x}_1^2 - k_1(x_1 - X)^2) + \frac{1}{N}\sum_{j=2}^N(m_j\dot{x}_j^2 - k(x_j - X)^2)] \quad (1.4)$$
$$K \equiv (M + \frac{m}{N})\frac{g}{L}, \quad k_j \equiv m_j\frac{g}{l_j},$$

where

- $g$  is the acceleration due to gravity;
- $K$ ,  $k_1$  and  $k$  are the spring, or stiffness, constants, of the corresponding oscillators. For a simple, uncoupled, harmonic oscillator with mass  $m$ , this is related to the frequencies  $\omega$  through  $\omega = \sqrt{\frac{k}{m}}$ .

## The Compound-Pendulum View

Through a simple change of variables, one obtains:

$$L_{red} \approx \frac{1}{2} [M\dot{X}^2 - KX^2 + m_1\dot{\xi}^2 - k_1(\xi - \frac{1}{\sqrt{N}}X)^2 + m\dot{\bar{x}}^2 - k(\bar{x} - X)^2]. \quad (1.5)$$

Note that this has 3 degrees of freedom, two that are strongly coupled  $X$  and  $\bar{x}$ , while the third,  $\xi$ , is weakly coupled due to the  $1/\sqrt{N}$  factor. Solving first the  $X, \bar{x}$  system gives us two modes with frequencies  $\omega_a$  and  $\omega_b$ . The natural frequency of the  $\xi$  degree of freedom that corresponds to the special pendulum is approximately  $\omega_1 = \sqrt{\frac{k_1}{m_1}}$ . If  $\omega_1$  is close to either  $\omega_a$  or  $\omega_b$ , there will be resonant transfer of energy between the two weakly coupled systems. In  $O(\sqrt{N})$  cycles, one should be able to identify the correct pendulum by having amplified its energy. If we instead had  $n$  shorter cords, it would take  $O(\sqrt{N/n})$  cycles.

# The Compound-Pendulum View

Imagine that one starts by a single push to the support pendulum and can change parameters of any pendulum and then observe their frequency with a finite precision that is independent of  $N$ .

By bisection, we can adjust the cords of  $1/2$  of the pendula,  $1/4$  of the pendula, etc., until we identify the one pendulum. This would have a runtime of  $O(\sqrt{N} \log N)$ .

# The Compound-Pendulum View

Imagine that one starts by a single push to the support pendulum and can change parameters of any pendulum and then observe their frequency with a finite precision that is independent of  $N$ .

By bisection, we can adjust the cords of  $1/2$  of the pendula,  $1/4$  of the pendula, etc., until we identify the one pendulum. This would have a runtime of  $O(\sqrt{N} \log N)$ .

## A Probability-Theory Point of View

The diffusion operator should be viewed as a quantum amplitude amplification procedure, with the aim to increase the probability amplitude of the target state. Following Gruska and Brassard, one could consider  $|\phi\rangle = A|0^n\rangle = \sum_i \alpha_i |i\rangle$  and some partition:

$$|\phi\rangle = \sum_{i \in \text{good}} \alpha_i |i\rangle + \sum_{i \in \text{bad}} \alpha_i |i\rangle,$$

with  $\mathbb{P}(\text{good}) = \sum_{i \in \text{good}} |\alpha_i|^2$ . Then,

$$|\phi\rangle = \sqrt{\mathbb{P}(\text{good})} |\phi_{\text{good}}\rangle + \sqrt{1 - \mathbb{P}(\text{good})} |\phi_{\text{bad}}\rangle = \sin(\theta) |\phi_{\text{good}}\rangle + \cos(\theta) |\phi_{\text{bad}}\rangle$$

with  $\sin^2(\theta) = \mathbb{P}(\text{good})$  and  $\sin^2(\theta) + \cos^2(\theta) = 1$ . The state  $|\phi\rangle$  is thus orthogonal to  $|\phi^\perp\rangle = \cos(\theta) |\phi_{\text{good}}\rangle - \sin(\theta) |\phi_{\text{bad}}\rangle$ .  $\{|\phi_{\text{good}}\rangle, |\phi_{\text{bad}}\rangle\}$  and  $\{|\phi\rangle, |\phi^\perp\rangle\}$  are thus two orthonormal bases in a 2-dimensional subspace.

# A Probability-Theory Point of View

One obtains

$$U_w (\sin(\theta) |\phi_{\text{good}}\rangle + \cos(\theta) |\phi_{\text{bad}}\rangle) = -\sin(\theta) |\phi_{\text{good}}\rangle + \cos(\theta) |\phi_{\text{bad}}\rangle \quad (1.6)$$

$$U_{\phi^\perp} (\sin(\theta) |\phi\rangle + \cos(\theta) |\phi^\perp\rangle) = \sin(\theta) |\phi\rangle - \cos(\theta) |\phi^\perp\rangle \quad (1.7)$$

$$U_{\phi^\perp} U_w |\phi\rangle = \cos(2\theta) |\phi\rangle + \sin(2\theta) |\phi^\perp\rangle \quad (1.8)$$

$$= \sin(3\theta) |\phi_{\text{good}}\rangle + \cos(3\theta) |\phi_{\text{bad}}\rangle. \quad (1.9)$$

We will see this view in the following lecture.

## A Geometric View

This amplitude amplification also has a geometric interpretation: one should see  $U_w$  and  $U_s$  as Householder reflections. Grover's algorithm stays in a subspace spanned by  $(|s\rangle, |w\rangle)$ . The two operators are reflections with respect to the hyper-planes perpendicular to  $w$  and  $s$ . It is an elementary fact of Euclidean geometry that when  $M_1$  and  $M_2$  are two lines in the plane intersecting at point  $O$  with intersection angle  $\alpha$ , the operation of reflection with respect to  $M_1$ , followed by reflection with respect to  $M_2$ , is rotation by angle  $2\alpha$  around  $O$ . Then, the product  $U_s U_w$  is a rotation in the  $(|s\rangle, |w\rangle)$  plane (for the first  $\approx \pi\sqrt{N}/4$  iterations from  $|s\rangle$  to  $|w\rangle$ ) by  $\theta = 2 \arcsin \frac{1}{\sqrt{N}}$ . This view is beautifully elaborated by Gruska.



# A Geometric View

Without giving a complete derivation here, let us consider  $|x_0^\perp\rangle$  and  $|\phi_0\rangle$  at an angle  $\beta$ .

Then  $U_{|\phi_0^\perp\rangle} U_{|x_0\rangle}$  is a rotation around the origin by angle  $2\beta$ .

Starting with a state  $|\phi_0\rangle = \sin(\beta) |x_0\rangle + \cos(\beta) |x_0^\perp\rangle$ ,

after  $q$  Grover iterations, we obtain:

$$|\phi_k\rangle = \sin((2q + 1)\beta) |x_0\rangle + \cos((2q + 1)\beta) |x_0^\perp\rangle.$$

We thus wish to pick  $q$  such that  $\sin((2q + 1)\beta)$  is as close as possible to 1.

## A Geometric View

Without giving a complete derivation here, let us consider  $|x_0^\perp\rangle$  and  $|\phi_0\rangle$  at an angle  $\beta$ .

Then  $U_{|\phi_0^\perp\rangle}U_{|x_0\rangle}$  is a rotation around the origin by angle  $2\beta$ .

Starting with a state  $|\phi_0\rangle = \sin(\beta)|x_0\rangle + \cos(\beta)|x_0^\perp\rangle$ ,

after  $q$  Grover iterations, we obtain:

$$|\phi_k\rangle = \sin((2q+1)\beta)|x_0\rangle + \cos((2q+1)\beta)|x_0^\perp\rangle.$$

We thus wish to pick  $q$  such that  $\sin((2q+1)\beta)$  is as close as possible to 1.

# A Geometric View

Without giving a complete derivation here, let us consider  $|x_0^\perp\rangle$  and  $|\phi_0\rangle$  at an angle  $\beta$ .

Then  $U_{|\phi_0^\perp\rangle} U_{|x_0\rangle}$  is a rotation around the origin by angle  $2\beta$ .

Starting with a state  $|\phi_0\rangle = \sin(\beta) |x_0\rangle + \cos(\beta) |x_0^\perp\rangle$ ,

after  $q$  Grover iterations, we obtain:

$$|\phi_k\rangle = \sin((2q + 1)\beta) |x_0\rangle + \cos((2q + 1)\beta) |x_0^\perp\rangle.$$

We thus wish to pick  $q$  such that  $\sin((2q + 1)\beta)$  is as close as possible to 1.

## A Geometric View

Without giving a complete derivation here, let us consider  $|x_0^\perp\rangle$  and  $|\phi_0\rangle$  at an angle  $\beta$ .

Then  $U_{|\phi_0^\perp\rangle} U_{|x_0\rangle}$  is a rotation around the origin by angle  $2\beta$ .

Starting with a state  $|\phi_0\rangle = \sin(\beta) |x_0\rangle + \cos(\beta) |x_0^\perp\rangle$ ,

after  $q$  Grover iterations, we obtain:

$$|\phi_k\rangle = \sin((2q + 1)\beta) |x_0\rangle + \cos((2q + 1)\beta) |x_0^\perp\rangle.$$

We thus wish to pick  $q$  such that  $\sin((2q + 1)\beta)$  is as close as possible to 1.

## A Geometric View

Without giving a complete derivation here, let us consider  $|x_0^\perp\rangle$  and  $|\phi_0\rangle$  at an angle  $\beta$ .

Then  $U_{|\phi_0^\perp\rangle} U_{|x_0\rangle}$  is a rotation around the origin by angle  $2\beta$ .

Starting with a state  $|\phi_0\rangle = \sin(\beta) |x_0\rangle + \cos(\beta) |x_0^\perp\rangle$ ,

after  $q$  Grover iterations, we obtain:

$$|\phi_k\rangle = \sin((2q + 1)\beta) |x_0\rangle + \cos((2q + 1)\beta) |x_0^\perp\rangle.$$

We thus wish to pick  $q$  such that  $\sin((2q + 1)\beta)$  is as close as possible to 1.

## Tensor-Analytical Point of View

Stoudenmire and Waintal suggests that  $U_w$  should be seen as:

$$\mathbf{U}_w = \begin{bmatrix} 1 & 1 \end{bmatrix} \left( \prod_{i=1}^n \mathbf{M}_i \right) \begin{bmatrix} 1 \\ -2 \end{bmatrix} \quad (1.10)$$

with

$$\mathbf{M}_i = \begin{bmatrix} \mathbf{I}_i & 0 & 0 & \dots \\ 0 & |w_i^1\rangle\langle w_i^1| & 0 & \dots \\ 0 & 0 & |w_i^2\rangle\langle w_i^2| & \dots \\ \dots & \dots & \dots & \dots \\ \dots & 0 & 0 & |w_i^S\rangle\langle w_i^S| \end{bmatrix} \quad (1.11)$$

where  $\mathbf{I}_i$  is the  $2 \times 2$  identity matrix acting on qubit  $i$  and  $|w_i^\alpha\rangle\langle w_i^\alpha|$  projects  $\alpha$  on the bitstring  $i$ .

## Tensor-Analytical Point of View

The diffusion operator  $U_S$  is similar, except for the replacement of  $M_i$  by

$$\mathbf{M}'_i = \begin{bmatrix} \mathbf{I}_i & 0 \\ 0 & |+\rangle\langle +| \end{bmatrix} \quad (1.12)$$

# The Controversy

As we have mentioned at the beginning, there is also a fair amount of controversy, which centers around three issues:

- one needs to be able to run the oracle with an error that scales with  $N^{-1/4} = 1/2^{n-4}$ . This is a very exacting standard which may be difficult to obtain for non-trivial  $n$ .
- quantumly, one needs to be able to implement the oracle in unit amount of time, but not to be able to implement the product of the Grover iteration  $U_S U_W$  in unit amount of time, and not to be able to implement *many* things classically.
- the tensor-analytic view suggests that one uses rank-2 matrix product operation, which are classically simulable in polytime. Then, one efficiently simulates the product of the Grover iterations as well.



# The NP-Hard Problems

Let us now consider two NP-Hard functional (optimization) problems. In the TRAVELLING SALESMAN PROBLEM (TSP), we seek the shortest simple cycle that visits each vertex in a weighted graph  $G$  once (*Hamiltonian circuit*). In the MINIMUM SET COVER, we seek the minimum cardinality subset  $S' \subseteq \mathcal{S}$  such that

$$\bigcup_{S \in S'} S = \mathcal{U}$$

for some given  $\mathcal{S} \subset \mathcal{U}$ , with the cardinality of the ground set  $|\mathcal{U}| = n$  and  $|\mathcal{S}| = m$ .

# Dynamic Programming for the TSP

A naive classical approach to either problem would construct a dynamic programming tableau, where in each row  $r$  in the tableau, we would have the lengths of Hamiltonian circuits in  $r$ -vertex subgraphs. Following ?, let  $f(S, u, v)$  denote the length of the shortest path in the graph induced by a subset of vertices  $S$  that starts in  $u \in S$ , ends in  $v \in S$  and visits all vertices in  $S$  exactly once.

Then:

$$f(S, u, v) = \min_{\substack{t \in N(u) \cap S \\ t \neq v}} \{w(u, t) + f(S \setminus \{u\}, t, v)\}, \quad f(\{v\}, v, v) = 0. \quad (2.1)$$

where  $N(u)$  is the neighbourhood of  $u$  in  $G$ . For  $k \in [2, |S| - 1]$  fixed,

$$f(S, u, v) = \min_{\substack{X \subset S, |X|=k \\ u \in X, v \notin X}} \min_{\substack{t \in X \\ t \neq u}} \{f(X, u, t) + f((S \setminus X) \cup \{t\}, t, v)\}. \quad (2.2)$$

# Dynamic Programming for the TSP

The algorithm of Ambainis et al. picks some  $\alpha \in (0, 1/2]$  and classically precomputes  $f(S, u, v)$  for all  $|S| \leq (1 - \alpha)n/4$  using dynamic programming. That is, it computes the bottom rows of the tableau classically, in time exponential in  $n$ . Quantumly, it obtains

$$\min_{\substack{S \subset V \\ |S|=n/2}} \min_{\substack{u, v \in S \\ u \neq v}} \{f(S, u, v) + f((V \setminus S) \cup \{u, v\}, v, u)\}$$

over all subsets  $S \subset V$  such that  $|S| = n/2$  by taking the following steps:

- 1 Run Grover with  $k = \alpha n/4$  to calculate  $f(S, u, v)$  for  $|S| = n/4$  starting with the rows of the tableau obtained classically.
- 2 Run Grover with  $k = n/4$  to calculate  $f(S, u, v)$  for  $|S| = n/2$ .

# Dynamic Programming for the TSP

Under very strong assumptions about storing the data in quantum RAM, Ambainis et al. claim a speed-up. Notice that much of the controversy surrounding the original Grover applies to this setting as well, compounded by the QRAM assumptions.

	Classical (best known)	Ambainis et al.
Vertex Ordering Problems	$O^*(2^n)$	$O^*(1.817^n)$
Travelling Salesman Problem	$O(n^2 2^n)$	$O^*(1.728^n)$
Minimum Set Cover	$O(nm 2^n)$	$O(\text{poly}(m, n) 1.728^n)$