CS 880: Quantum Algorithms 3/2/2021

Lecture 10: Oblivious Amplitude Amplification

Instructor: Dieter van Melkebeek Scribe: Sam Kottler

In this lecture we will be extending the amplitude amplification algorithm we have been studying for the last several lectures. We will begin by reviewing the amplitude amplification problem, as well as the two-domain view by going over the solution to last lecture's exercise. Then we will extend the amplitude amplification problem by allowing unknown, but still limited, initial superpositions. Finally, we will see how the oblivious amplitude amplification problem relates to block encoding, which we will see more of later in the semester.

1 Review of the Two-Domain view

First recall the problem. We assume we have blackbox access to a function $f : \{0,1\}^n \to \{0,1\}$ and a unitary circuit A on n qubits such that $A\left|0^n\right\rangle = \sum_x \alpha_x \left|x\right\rangle$ has nonzero amplitude α_x for some x with $f(x) = 1$. Define $p = \sum_{x:f(x)=1} |\alpha_x|^2$, which may or may not be known. Then denote the desirable states as $\left|G\right\rangle = \frac{1}{\sqrt{p}} \sum_{x:f(x)=1} \alpha_x \left|x\right\rangle$ and the undesirable states as $\left|B\right\rangle = \frac{1}{\sqrt{1-p}} \sum_{x:f(x)=1} \alpha_x \left|x\right\rangle$ so that $A\left|0^n\right\rangle = \sqrt{1-p}\left|B\right\rangle + \sqrt{p}\left|G\right\rangle$. The goal is to return $\left|G\right\rangle$ with probability at least 1/2 and with a success indicator. The algorithm for amplitude amplification is to start from $A\left|0^n\right\rangle$ and repeatedly apply

1. Reflect over undesirable states to flip the phase on desirable states using U_f
2. Reflect over the initial superposition by applying $AR_{\left|0^n\right\rangle}A^{-1}$

We can interpret this algorithm as a series of reflections and domain switches in the following way. First, we start in the time domain at $\left|0^n\right\rangle$ and switch the the frequency domain by applying A. Then for one iteration of the algorithm, we first do the reflection R_{bad} in the frequency domain and switch back to the time domain by applying A^{-1}. Note that A (and also A^{-1}) are unitary, so they preserve angles. Finally we reflect over the initial superposition ($\left|0^n\right\rangle$) in the time domain and switch back to frequency domain ending the first iteration. The diagrams for this view are in Figure 1.

Now, using this view, we can consider the solution to the previous lecture's exercise.

Exercise 1. *Consider algorithm for amplitude amplification but evaluate the success indicator initially and after every iteration*

1. *Determine the probability of no success within the first k iterations as a function of p.*
2. *Determine the expected number of iterations until the first success as a function of p.*

First, let us consider this through the two-domain view as in figure 2. The initial state is $\left|\psi_0\right\rangle = A\left|0^n\right\rangle$. As we have seen previously, the probability of failure on a measurement is $1-p$. In this case the state becomes $\left|B\right\rangle$. Now the reflection R_{bad} is trivial so we switch back to the time domain. Now we reflect over $\left|0^n\right\rangle$ in the time domain and switch back to the frequency domain.

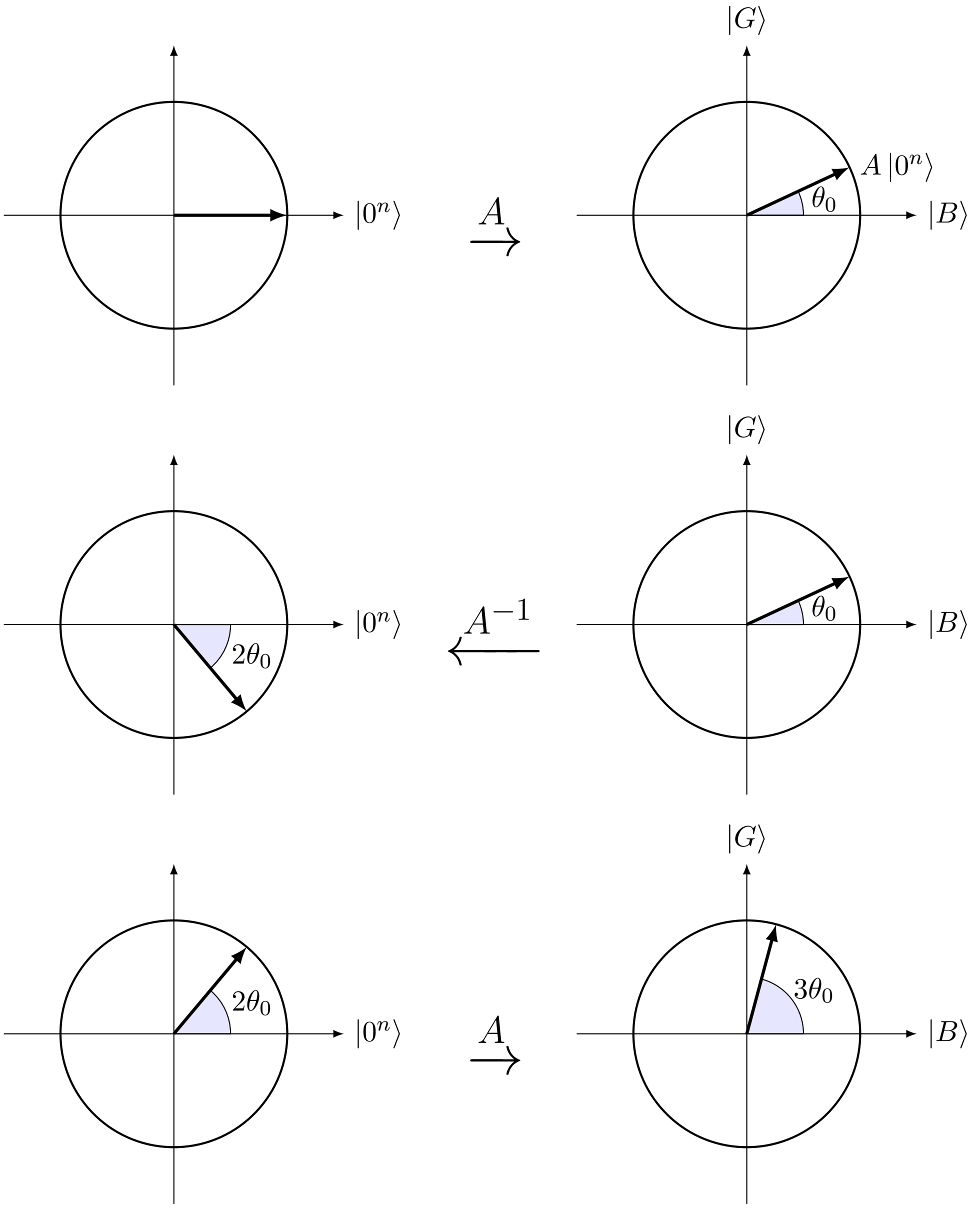

Figure 1: Amplitude amplification in the two-domain view

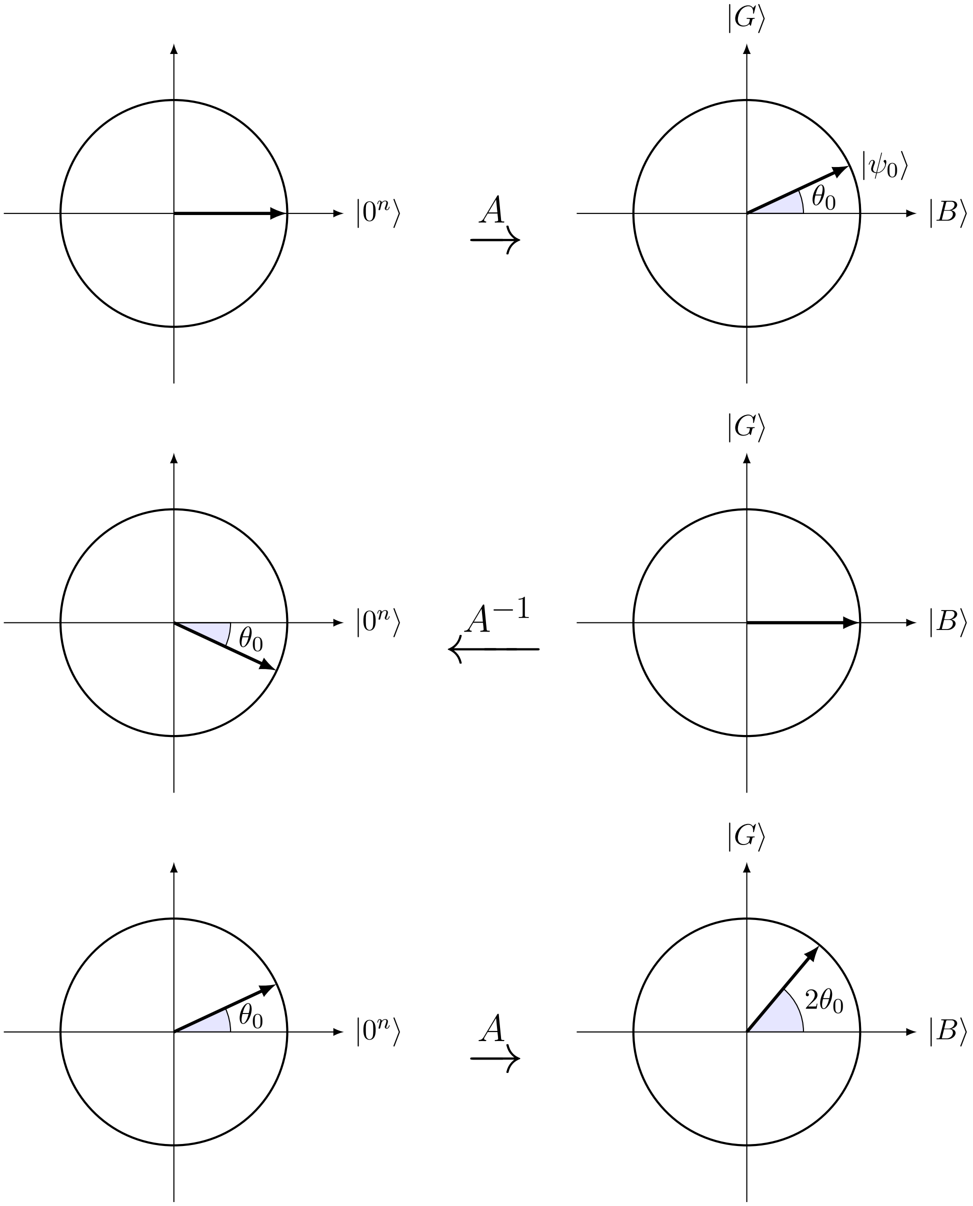

Figure 2: Exercise 1

Now, unlike the initial state, the probability of measuring something in $|B\rangle$ is $\cos^2(2\theta_0)$. Recalling that $\sin(\theta_0) = \sqrt{p}$ and applying the double angle formula $\cos(2\theta) = 1 - 2\sin^2(\theta)$, we can reduce this to $(1-2p)^2$.

Now we can answer the questions in the exercise.

1. The probability that we fail initially, that is before the first iteration, was $1-p$. Then after each subsequent iteration, the probability of failure is $(1-2p)^2$. Thus the probability of no successes after k iterations is $(1-p)(1-2p)^{2k}$.

2. The probability of initial success is p. In particular, this means that if $p = 1$ the expected number of iterations is 0. If $p < 1$, then we are essentially doing a Bernoulli experiment by repeating the process until we get a success. For any trial the probability of success is $q = 1 - (1-2p)^2 = 4p(1-p)$. In general, the expected number of trials until success for a Bernoulli experiment with probability of success q is $1/q$. So, assuming we don't have an initial success, we get an expected number of trials $p \cdot 0 + \frac{1-p}{q} = \frac{1}{4p}$. Note that in the case where $p = 1$, this formula would give us $\frac{1}{4}$ which is incorrect. This is because the first measurement is different and must be treated separately.

2 Oblivious Amplitude Amplification

2.1 Problem Statement

The oblivious amplitude amplification problem is similar to the amplitude amplification problem with one key difference. Again we assume we have blackbox access to $f : \{0,1\}^n \to \{0,1\}$ and a unitary circuit A. Now however, we do not assume that A is applied to $|0^n\rangle$, or even that we know how to generate the initial state at all. Instead we apply A to an arbitrary pure state $|\psi\rangle$ that is given to us and we assume that $A|\psi\rangle = \sum_x \alpha_x |x\rangle$ has nonzero amplitude α_x for some x with $f(x) = 1$. We again let $p(\psi) = \sum_{x:f(x)=1} |\alpha_x|^2$, which may or may not be known. Note that p may depend on the unknown ψ. We will use $|G(\psi)\rangle$ and $|B(\psi)\rangle$ as the natural analogs of $|G\rangle$ and $|B\rangle$ respectively. The goal of this problem is similar to amplitude amplification: we want to output $|G(\psi)\rangle$ with probability at least 1/2 and with a success indicator.

Unfortunately, there are a number of challenges that will make this problem more difficult. First, we only have a single copy of $|\psi\rangle$. In particular we do not have a quantum circuit that produces $|\psi\rangle$ and the no cloning theorem prevents us from making more copies. This causes a problem even for getting a $O(1/p)$ algorithm based on the classical one. The second challenge is that even if we did have many copies of $|\psi\rangle$ or $A|\psi\rangle$, we would need a way of efficiently computing the reflection $R_{initial}$ around $|\psi_0\rangle = A|\psi\rangle$ in order to create an $O(1/\sqrt{p})$ algorithm in the same way as we did before. Unfortunately, there is no such efficient calculation of this arbitrary reflection known so we need to make some adjustments to the setting and the algorithm.

2.2 Setting with Independent Initial Weight

The first adjustment we will make is to restrict the problem to the setting where the initial weight of $|G(\psi)\rangle$, $p(\psi) = p$, is independent of $|\psi\rangle$. However, as we will see, the only possibility then is for $p = 1$, which is trivial: just take a measurement, no amplification required. Because of this, we will additionally accept the promise that $|\psi\rangle$ is of the form $|0^\ell\rangle|\phi\rangle$ for some $\ell > 0$. While this restriction may at first seem contrived, it arises when purifying a quantum circuit (recall this is when we delay or defer partial measurement by use of additional ancilla qubits) where the $|0^\ell\rangle$ are the added ancillas. Independence of the initial weight in this purified setting arises naturally in certain zero-knowledge systems, which we will see later in the course. It also encompasses the paradigm of block encoding which we will explore at the end of the lecture.

Note that since p is independent of $|\psi\rangle$, we have for all $|\phi\rangle$ on $n-\ell$ qubits that $||P_1 A |0^\ell\rangle |\phi\rangle ||_2^2 = p$ where P_1 is the projection onto basis states $|x\rangle$ with $f(x) = 1$. This simply comes from the definition of p and the 2-norm.

2.3 Analysis

We begin our analysis from this last observation:

$$||P_1 A |0^\ell\rangle |\phi\rangle ||_2^2 = p \tag{*}$$

As the 2-norm can be calculated via an inner product with the complex transpose, we can rewrite this as

$$||P_1 A |0^\ell\rangle |\phi\rangle ||_2^2 = \langle 0^\ell| \langle\phi| A^* P_1 A |0^\ell\rangle |\phi\rangle = p.$$

Note we dropped the P_1^* as $P_1^* = P_1 = P_1^2$. Let M denote the matrix $M \doteq A^* P_1 A$. Note that $M^* = (A^* P_1 A)^* = A^* P_1^* (A^*)^* = A^* P_1 A = M$, so the matrix M is Hermitian. We can separate the rows an columns of M corresponding to basis vectors of the form $|0^\ell *\rangle$ in the top left part:

$$M = \begin{bmatrix} M_{LT} & M_{RT} \\ M_{LB} & M_{RB} \end{bmatrix}$$

Then when we work out $\langle 0^l| \langle\phi| M |0^l\rangle |\phi\rangle$, the zeros will cause everything except $\langle\phi| M_{TL} |\phi\rangle$ to cancel out. Thus (*) is equivalent to $\langle\phi| M_{TL} |\phi\rangle = p$.

Claim. $M_{LT} = p \cdot I$

Proof. Since (*) holds for all $|\phi\rangle$, all eigenvalues of M_{LT} are p. Additionally, since M is Hermitian and M_{LT} is the upper left square, it must also be Hermitian. This means it has a full basis of eigenvectors. The only matrix satisfying these two properties is $p \cdot I$. □

Another way to state this result is that

$$M = \begin{bmatrix} p \cdot I & M_{RT} \\ M_{LB} & M_{RB} \end{bmatrix}$$

This will be helpful for a few corollaries:

Corollary 1. *If $\ell = 0$ then $P_1 = I$ and $p = 1$*

Proof. Suppose, by way of contradiction, that there is some $|x\rangle$ such that $P_1 |x\rangle = \mathbf{0}$ (zero vector not $|0\rangle$). Then let $|\phi\rangle = A^{-1} |x\rangle$ so that $M |\phi\rangle = \mathbf{0}$. But $p \cdot I |x\rangle = p |x\rangle \neq \mathbf{0}$. Since $M = p \cdot I$, this is a contradiction. □

Corollary 2. *Let C denote the projection to clean ancillas (starting with $|0^\ell\rangle$) followed by renormalization. Then we have that applying M followed by C gives*

$$|0^\ell\rangle |\phi\rangle \xrightarrow{A} |\psi_0\rangle \xrightarrow{success} \frac{1}{\sqrt{p}} P_1 |\psi_0\rangle \xrightarrow{A^{-1}} |\psi_0'\rangle \xrightarrow{C} |0^\ell\rangle |\phi\rangle$$

Proof. All steps except for the last come from M. The last step happens because $M_{LT} = p \cdot I$ so multiplying $|0^\ell\rangle\, \phi$ by M and then projecting to clean ancillas leaves us with the p times the identity on the second part and all zeros on the first part. Renormalizing gets rid of the p so we end up back where we started. □

If we repeat all of this analysis starting from the observation similar to (*) that

$$||P_0 A\, |0^\ell\rangle\, |\phi\rangle\, ||_2^2 = 1 - p$$

and let $M' = A^* P_0 A$, we get to the analogous corollary

Corollary 3. *Let C denote the projection to clean ancillas (starting with $|0^\ell\rangle$) followed by renormalization. Then we have that applying M' followed by C gives*

$$|0^\ell\rangle\, |\phi\rangle \xrightarrow{A} |\psi_0\rangle \xrightarrow{failure} \frac{1}{\sqrt{1-p}} P_0\, |\psi_0\rangle \xrightarrow{A^{-1}} |\psi_0''\rangle \xrightarrow{C} |0^\ell\rangle\, |\phi\rangle$$

Now it becomes clear why we wrote the sequence out step-by-step. We can apply A, observe, apply A^{-1}, and apply C to return to where we started regardless the state in the middle. This is the key to performing oblivious amplitude amplification in $O(1/\sqrt{p})$ steps.

2.4 Algorithm

See figure 3 for the two-domain view diagram. We start in state $|0^\ell\rangle\, |\phi\rangle$ in the time domain and apply A to switch to the frequency domain. Then, we can apply R_{bad} as usual as this depends only on the blackbox function, and not at all on the initial position. We follow this by applying A^{-1} to return to the time domain. We can call the state we're at now $|\xi\rangle$. We would like to reflect over $|0^\ell\rangle\, |\phi\rangle$, but as was discussed earlier, we cannot. Instead we must use the fact that projecting any superposition of $A^{-1}\, |G(\phi)\rangle$ and $A^{-1}\, |B(\phi)\rangle$ gives (up to a scalar) $|0^l\rangle\, |\phi\rangle$. In particular, since $|\xi\rangle$ is such a superposition, its projection onto clean ancillas is (up to a scalar) $|0^\ell\rangle\, |\phi\rangle$. This tells us that the the superposition of all components of $|\xi\rangle$ that start with clean ancillas is $|0^\ell\rangle\, |\phi\rangle$. Because of this, reflecting over the subspace with basis $|0^\ell *\rangle$ will have exactly the same effect on $|\xi\rangle$ as reflecting over $|0^\ell\rangle\, |\phi\rangle$. Since the reflection $R_{|0^\ell *\rangle}$ is independent of $|\phi\rangle$ it can be computed in general efficiently. Finally we switch back to the frequency domain by applying A completing one iteration and rotating the state a total of $2\theta_0$ toward $|G(\phi)\rangle$ as before.

Since we are able to rotate toward $|G(\phi)\rangle$ by an angle of $2\theta_0$ each iteration, we can complete our goal of outputting $|G(\phi)\rangle$ with probability at least 1/2 and with a success indicator. This lecture doesn't do this part again as it is identical from this point to the last lecture. As a quick reminder, the success indicator was achieved by adding an additional ancilla and the fact that p is unknown was avoided by trying the algorithm with increasing values of k.

3 Block Encoding

In the final part of this lecture, we will make a connection between the previous analysis and the block encoding paradigm. Here we are assuming the function f is such that $f(x) = 1$ if and only if x is of the form $x = 0^\ell y$. This means that the reflection over the bad axis R_{bad} is exactly $R_{|0^\ell *\rangle}$. As before let P_1 be the projection onto the "good" states, which in this case means projecting onto

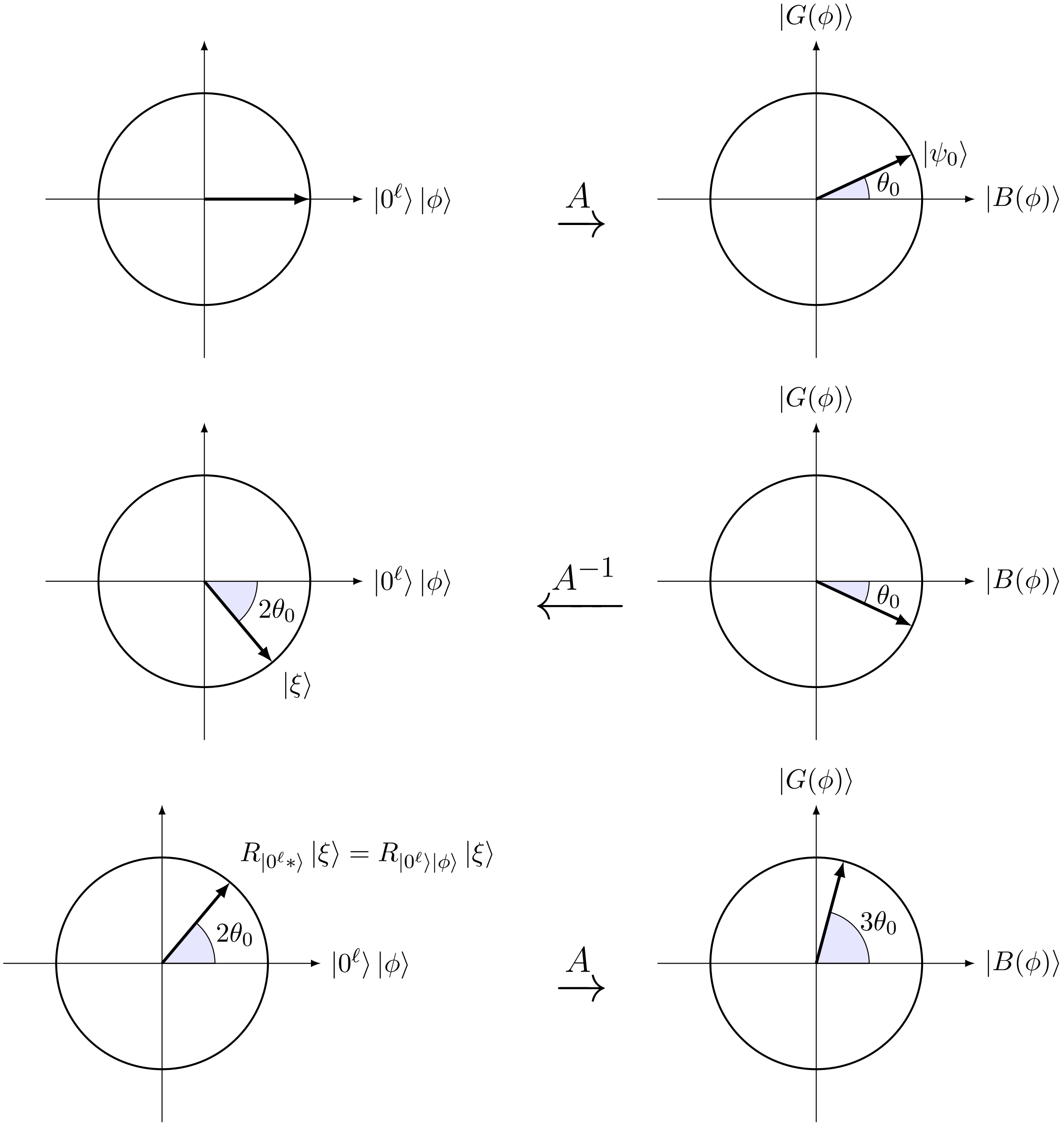

Figure 3: Oblivious Amplitude Amplification

clean ancillas. Now we will split the rows and columns of A corresponding to the basis vectors of $|0^\ell *\rangle$ in the top left corner as we did to M earlier. Then

$$A = \begin{bmatrix} A_{LT} & A_{RT} \\ A_{LB} & A_{RB} \end{bmatrix}$$

and

$$A^* = \begin{bmatrix} (A_{LT})^* & (A_{LB})^* \\ (A_{RT})^* & (A_{RB})^* \end{bmatrix}.$$

Then since P_1 projects to clean ancillas, observe that

$$A^* P_1 = \begin{bmatrix} (A_{LT})^* & 0 \\ (A_{RT})^* & 0 \end{bmatrix}$$

so

$$M = (A^* P_1)A = \begin{bmatrix} (A_{LT})^* & 0 \\ (A_{RT})^* & 0 \end{bmatrix} \begin{bmatrix} A_{LT} & A_{RT} \\ A_{LB} & A_{RB} \end{bmatrix} = \begin{bmatrix} (A_{LT})^* A_{LT} & \cdots \\ \cdots & \cdots \end{bmatrix}$$

Since we showed that

$$M = \begin{bmatrix} p \cdot I & \cdots \\ \cdots & \cdots \end{bmatrix}$$

in the previous section, we have that $(A_{LT})^* A_{LT} = p \cdot I$. From this we can conclude that $A_{LT} = \sqrt{p}U$ for some unitary U.

Definition 1. *A unitary A is a* ***block encoding*** *of a unitary U with success probability p if*

$$A = \begin{bmatrix} \sqrt{p}U & \cdots \\ \cdots & \cdots \end{bmatrix}$$

The block encodings are a fairly new paradigm for quantum algorithm design, and we will return to them later in the course. For now we can consider a few properties as they apply to things we have already seen. First, we can view A as a probabilistic implementation of U. We have a fairly simple success indicator: just measure the first ℓ qubits and if they are all 0, then U was successfully applied. Additionally, as we saw earlier in the lecture, we can do amplitude amplification on block encodings in $O(1/\sqrt{p})$ steps. This is important if p is small and we want to boost the probability that U is applied.