**Quantum Algorithms** 2/23/2023

# Lecture 10: Amplitude Amplification

Instructor: Dieter van Melkebeek

In this lecture, we generalize the principles of the previous lecture on quantum search and present amplitude amplification – a basic paradigm in quantum algorithms. We start by defining the problem and then present an algorithm with bounded error. We discuss how to eliminate error and how to relax the requirement that the total weight of "good" items is known ahead of time.

## 1 Amplitude Amplification

### 1.1 Problem Statement

The problem setup for amplitude amplification is similar to that of Grover's search algorithm discussed previously. In particular, we are given:

- A black-box Boolean function $f : \{0,1\}^n \to \{0,1\}$ which maps "good" inputs $x$ to 1 and "bad" inputs to 0,
- A unitary circuit $A$ operating on $n$ qubits such that $A|0^n\rangle = \sum_x \alpha_x |x\rangle$ produces nonzero amplitude $\alpha_x$ on at least one $x$ with $f(x) = 1$,
- The total weight $p \doteq \sum_{x:f(x)=1} |\alpha_x|^2$ of the good inputs.

The second assumption implies that $p > 0$. In Section 3 we relax the requirement that the total weight of "good" inputs $p$ is known ahead of time.

The state $A|0^n\rangle = \sum_x \alpha_x |x\rangle$ can be written as a superposition of the "good" states $|G\rangle = \frac{1}{\sqrt{p}} \sum_{x:f(x)=1} \alpha_x |x\rangle$ and "bad" states $|B\rangle = \frac{1}{\sqrt{1-p}} \sum_{x:f(x)=0} \alpha_x |x\rangle$:

$$A|0^n\rangle = \sqrt{1-p}\,|B\rangle + \sqrt{p}\,|G\rangle\,. \tag{1}$$

The goal of amplitude amplification is to output $|G\rangle$ in (1) with probability greater than or equal to 1/2 along with a success indicator (a bit which is 1 if $|G\rangle$ is the output and 0 otherwise). As was the case for quantum search, the high level idea is to amplify the weight of $|G\rangle$ while reducing the weight of $|B\rangle$.

**Remark 1.** The problem statement for amplitude amplification differs from that of quantum search in subtle yet distinct ways. In particular, for quantum search, the goal was to identify *a single* $x$ such that $f(x) = 1$ by amplifying the weights $\alpha_x$ for "good" inputs $x$ and then performing a measurement. Thus, the output of quantum search is *classical*. In this lecture, however, the output is instead the superposition of "good" states $|G\rangle$ itself. This means that amplitude amplification is actually a *quantum subroutine*, which can be used as a building block for larger quantum algorithms.

A second consequence of the difference in outputs is that for amplitude amplification we care about the individual coefficients in the state $|G\rangle$, not just that the weight of the "good" state is amplified so that a measurement yields a good one with high probability. In particular, amplitude amplification should proceed without affecting the relative weight between good inputs – the ratio of the amplitudes $\alpha_x$ and $\alpha_{x'}$ of good states should not change; only the ratio of the amplitudes of good versus bad states can change. While the algorithm we presented for quantum search satisfies this requirement, it is not important for quantum search in general, so long as a "good" $x$ ends up with a high probability after measurement.

Finally, note that quantum search as discussed in the previous lecture can simply be viewed as a special case of amplitude amplification with a measurement at the end. The specific case of $A = H^{\otimes n}$ is also noteworthy as it produces a uniform superposition over all $x$.

## 1.2 Algorithm

The algorithm for amplitude amplification proceeds in a similar fashion to the search algorithm discussed last lecture. The main difference comes when it is time to extract the output. As described in Algorithm 1, we start from the state $A|0^n\rangle$ then apply the following $k$ times: we first apply a phase flip on all "good" $x$ using $R_{bad}$ (which requires using the quantum version of the black-box function) and then reflect around the start state using $R_{initial}$.

---

**Algorithm 1** Amplitude Amplification.

---

Start with $A|0^n\rangle$
**for** $k$ times **do**
  Apply $R_{bad}$ (using $U_f$): Phase flip on all $|x\rangle$ with $f(x) = 1$
  Apply $R_{initial} \doteq AR_{|0^n\rangle}A^{-1}$: Reflection about the start state

---

**Two-dimensional state analysis.** The iterative procedure involved in amplitude amplification can be analyzed geometrically in two-dimensional space as done in the previous lecture. We summarize the idea here for completeness. First we express the state of the system after $k$ iterations as $|\psi_k\rangle = \beta_k |B\rangle + \gamma_k |G\rangle = \cos\theta_k |B\rangle + \sin\theta_k |G\rangle$. As such, the system corresponds to a point on the unit circle as shown in Figure 1. Initially we have that $\beta_0 = \sqrt{1-p} = \cos\theta_0$ and $\gamma_0 = \sqrt{p} = \sin\theta_0$. Then each iteration applies $R_{bad}$ followed by $R_{initial}$. Geometrically, $R_{bad}$ consists of a reflection about the $|B\rangle$ axis (as it flips the phase of all "good" states $x$). In the two-dimensional $|B\rangle$-$|G\rangle$ plane, $R_{initial}$ corresponds to a reflection about $A|0^n\rangle$. The combined result of the two operations is a counterclockwise rotation over $2\theta_0$ for each iteration as shown in Figure 1.

**Extracting $|G\rangle$.** Recall that the goal of amplitude amplification is to output $|G\rangle$. After $k$ iterations of the algorithm, we can do so as follows: We introduce a fresh ancilla qubit initially set to $|0\rangle$ creating the state $\sum_x \alpha_x^{(k)} |x\rangle |0\rangle$, with $\alpha_x^{(k)}$ the final amplitude for basis state $|x\rangle$ after the $k$-step amplification process. We then apply the black box function $U_f$ to the quantum state resulting in the state $\sum_x \alpha_x^{(k)} |x\rangle |f(x)\rangle$. The final step is to measure the ancilla qubit. If we observe 1 then the remaining quantum state collapses to those which are consistent with the measurement, i.e., states that satisfy $f(x) = 1$. This is exactly the state $|G\rangle$. Likewise if we observe 0 for the ancilla qubit then the remaining state is $|B\rangle$. Notice that the ancilla qubit is exactly the success indicator desired in our problem statement.

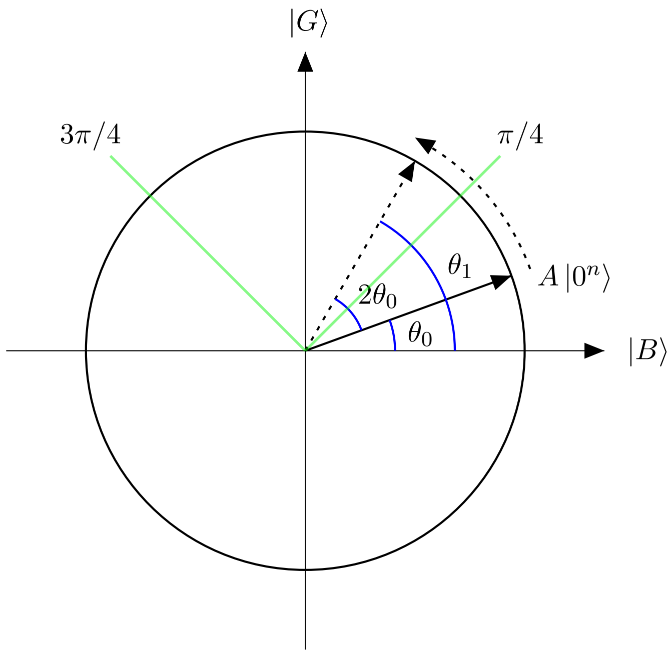


Figure 1: Depiction of amplitude amplification interation process in two-dimensional plane. The initial state starts with angle $\theta_0$. Each iteration rotates the state in the $|B\rangle$-$|G\rangle$ plane by $2\theta_0$. If the angle after $k$ iterations $\theta_k$ is between $[\pi/4, 3\pi/4]$ then the probability of extracting $|G\rangle$ is at least 1/2.

**Number of iterations and error bound.** In order to ensure that the probability of outputting $|G\rangle$ is greater than or equal to 1/2 we must ensure that after $k$ iterations $\gamma_k^2 \geq \beta_k^2$. In other words, in the two-dimensional state view, we wish for the final angle $\theta_k$ after $k$ iterations to end up as close to plus or minus $\pi/2$ as possible (Figure 1). This can be done by choosing $k$ as described in the previous lecture. Ideally $\theta_k = (2k+1)\theta_0 = \frac{\pi}{2}$. This implies that $k$ should be set to $k^* \doteq \frac{1}{2}(\frac{\pi}{2\theta_0} - 1)$. In general, $k^*$ is non-integral (in the case that it is an integer, then $|G\rangle$ can be extracted with probability 1). We take the number of iterations $k$ to be $\lfloor k^* \rceil$, the closest integer to $k^*$. This ensures that $\theta_k \in [\frac{\pi}{4}, \frac{3\pi}{4}]$ and thus that the probability of extracting $|G\rangle$ rather than $|B\rangle$ is greater than or equal to 1/2. Recall also that setting $k$ in this manner results in $O(\frac{1}{\sqrt{p}})$ iterations (and thus black-box queries) because $k^* = \Theta(1/\theta_0) = \Theta(1/\sqrt{p})$ since $\sin\theta_0 = \sqrt{p}$ and $\sin\theta \leq \theta$ for all nonnegative $\theta$.

## 2 Error Elimination

In this section we discuss how to eliminate the error in amplitude amplification, i.e., how to ensure we can produce $|G\rangle$ with certainty rather than just with probability greater than or equal to 1/2 (and a success indicator). This constitutes the solution to Exercise #8 posed in the previous lecture.

**Idea.** The problem with amplitude amplification as described above is that it is often impossible to rotate the initial angle $\theta_0$ to $\pi/2$ using an integral number $k$ of angle $2\theta_0$ rotations. Thus we are unable to produce a final state $|G\rangle$, and are instead only able to produce a superposition $|\psi_k\rangle = \beta_k |B\rangle + \gamma_k |G\rangle$ with $\gamma_k^2 \geq \beta_k^2$. To alleviate this issue, the key idea is to tweak $\theta_0$ a little bit to $\tilde{\theta}_0$ such that it is possible to end up exactly at $\pi/2$ after a integer sequence of iterations. We can reduce the angle by making use of an additional ancilla qubit.

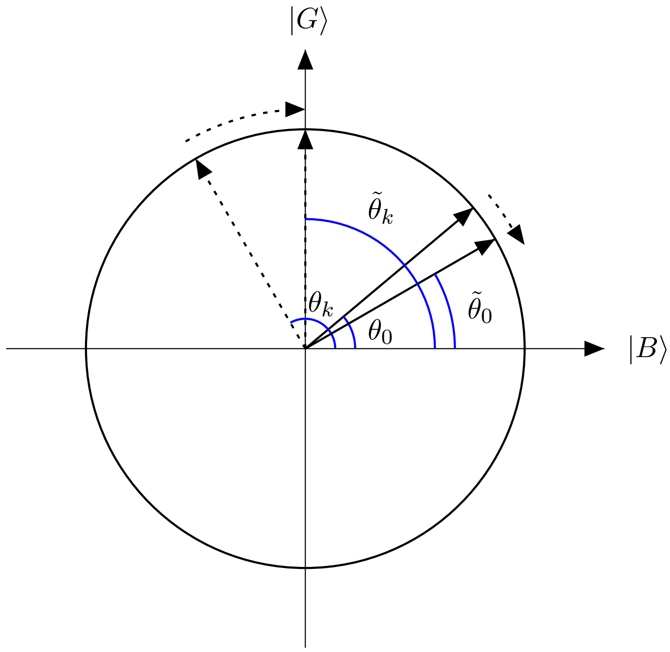


Figure 2: Depiction of the key idea in amplitude amplification error elimination: By reducing $\theta_0$ to $\tilde{\theta}_0$ we can move $\theta_k > \pi/2$ to $\tilde{\theta}_k = \pi/2$ and ensure that $|G\rangle$ is extracted with certainty.

**Solution.** Recall that we defined the quantity $k^*$ such that after $k^*$ rotations over $2\theta_0$, starting from an angle $\theta_0$, we arrive exactly at $\pi/2$: $k^* \doteq \frac{1}{2}\left(\frac{\pi}{2\theta_0} - 1\right)$. In general, $k^*$ is non-integral. Last lecture we rounded $k^*$ to the closest integer: $k = \lceil k^* \rfloor$. Now, consider rounding up $k^*$: $k = \lceil k^* \rceil$. Then we know that after $k$ iterations $\theta_k \geq \pi/2$. In other words, we may overshoot $|G\rangle$ in the two-dimensional $|B\rangle$-$|G\rangle$ plane. Following the key idea above, we aim to lower $\theta_0$ down to $\tilde{\theta}_0$ such that after $k$ iterations, instead of overshooting, we get exactly $\theta_k = \pi/2$. Lowering $\theta_0$ to $\tilde{\theta}_0$ means that $\tilde{\theta}_0 \leq \theta_0$ and implies that $\tilde{p} \doteq \sin^2(\tilde{\theta}_0) \leq \sin^2(\theta_0) = p$. Concretely, we wish to choose $\tilde{\theta}_0$ such that $(2k+1)\tilde{\theta}_0 = \frac{\pi}{2}$. The geometric interpretation is shown in Figure 2.

Lowering $\theta_0$ down to $\tilde{\theta}_0$ means lowering the weight of the good inputs from $p$ down to $\tilde{p}$. We do this by extending each input with an additional bit, giving the extension with 1 relative weight $\tilde{p'}/p$, and only considering the extensions with 1 as "good." We then apply amplitude amplification to the extended system. As the weight of the good inputs is now exactly $p'$, $k$ iterations yield success with certainty. Moreover, as the ratios of the amplitudes of good states $|x\rangle$ before and after extension is the same, the state of the extended system upon success is $|G\rangle$ extended with $|1\rangle$.

To implement the reduction from $p$ to $\tilde{p}$, we introduce an additional ancilla to capture the extension bit, start if from the basis state $|0\rangle$, and apply a unitary $U$ to it such that $U|0\rangle \doteq \alpha_0 |0\rangle +$

$\alpha_1 |1\rangle$ with $|\alpha_1|^2 = \tilde{p}/p$. We use the matrix

$$U = \begin{bmatrix} \alpha_0 & \alpha_1 \\ \alpha_1 & -\alpha_0 \end{bmatrix} \tag{2}$$

with $\alpha_1 = \sqrt{\tilde{p}/p}$ and $\alpha_0 = \sqrt{1 - \tilde{p}/p}$. Since $\tilde{p} \leq p$, $\alpha_1 \leq 1$ and we can chose $\alpha_0 \in \mathbb{R}$ such that the first column of $U$ has 2-norm one. The second column is chosen to be orthogonal to the first without consequence, as we are only interested in the application of $U$ to $|0\rangle$.

We apply amplitude amplification to a system with $n+1$ qubits and unitary operator $\tilde{A} = U \otimes A$. For convenience in drawing the resulting circuit, we put the extension ancilla in front. After expanding the size of the system from $n$ to $n+1$ qubits, we must also "expand" the function $f$ to operate on $n+1$ bits rather than $n$ bits. We define $\tilde{f}(bx) = b \cdot f(x)$, where $b$ is the additional ancilla bit. "Good" inputs are now those that have $\tilde{f}(bx) = 1$. These are the inputs for which $f(x) = 1$ and $b = 1$. Our new superposition of "good" inputs is now $|\tilde{G}\rangle = |1\rangle |G\rangle$, and the weight of $|\tilde{G}\rangle$ is $\tilde{p} = p \cdot |\alpha_1|^2$. Note that since $k = \lceil k^* \rceil \leq \lceil k^* \rfloor + 1$, the number of iterations and queries went up by at most 1. The final state is $|1\rangle |G\rangle$. We can return the ancilla to it original state of $|0\rangle$ and discard it if needed.

The only remaining issue is that the amplitude amplification procedure now needs the blackbox $U_{\tilde{f}}$ instead of $U_f$. We need to efficiently simulate $U_{\tilde{f}}$ using $U_f$ as a blackbox. Note that $U_{\tilde{f}}$ is equivalent to a controlled $U_f$ gate, where the extension qubit acts as the control: For $x \in \{0,1\}^n$ and $y \in \{0,1\}$:

$$\begin{aligned} |0\rangle |x\rangle |y\rangle &\mapsto |0\rangle |x\rangle |y\rangle = |0\rangle I(|x\rangle |y\rangle) \\ |1\rangle |x\rangle |y\rangle &\mapsto |1\rangle |x\rangle |y \oplus f(x)\rangle = |1\rangle U_f(|x\rangle |y\rangle). \end{aligned}$$

Since the only use of the blackbox in amplitude amplification is for the purposes of phase kickbacks, we only need to run $U_{\tilde{f}}$ when the last qubit is in state $|-\rangle$. Recall that $U_f$ maps $|x\rangle |-\rangle$ to $(-1)^{f(x)} |x\rangle |-\rangle$, and leaves $|x\rangle |+\rangle$ unaffected. Thus, we can simulate the controlled $U_f$ gate when the last qubit is in $|-\rangle$ as follows: If the control qubit is in $|0\rangle$, then apply a basis transformation to the last qubit that maps $|-\rangle$ to $|+\rangle$, apply $U_f$ to the controlled qubits, and revert the basis transformation. If the control qubit is in $|1\rangle$, we just apply $U_f$, without a basis transformation before and after. A unitary that realizes the desired basis transformation (and its inverse) is the $Z$ gate. Thus, we can simulate $U_{\tilde{f}}$ with the last qubit in state $|-\rangle$ by applying $U_f$ to all but the first qubit, preceded and followed by a controlled $Z$-gate where the first qubit acts at the control and the last qubit as the target. See Figure 3 for the resulting circuitry, where the controls with the unfilled circles indicate that the control is active when the qubit is 0.

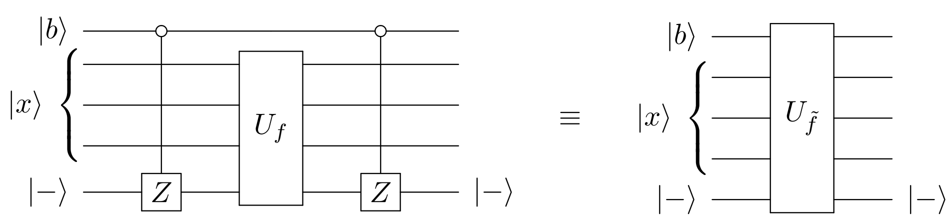


Figure 3: Simulating phase kickback for $\tilde{f}$ using $U_f$

The above solution uses two ancillas: one for the phase kickbacks in the applications of $U_f$ to effect $R_{bad}$, and one for reducing the weight from $p$ to $\tilde{p}$ in the applications of $U$. In fact, we can get by with a single ancilla that plays both roles. After the initialization, the ancilla is in state $U|0\rangle$. Each time we want to apply $R_{bad}$, we apply $H$ to the ancilla so as to switch from the standard basis $(|0\rangle, |1\rangle)$ to the Hadamard basis $(|+\rangle, |-\rangle)$, apply $U_f$ with the ancilla as the last qubit, and apply $H$ to switch back to the standard basis. The effect of the three steps is:

$$|x\rangle|0\rangle \mapsto |x\rangle|+\rangle \mapsto |x\rangle|+\rangle \mapsto |x\rangle|0\rangle$$
$$|x\rangle|1\rangle \mapsto |x\rangle|-\rangle \mapsto (-1)^{f(x)}|x\rangle|-\rangle \mapsto (-1)^{f(x)}|x\rangle|1\rangle.$$

This results in the circuit of Figure 4 for the overall process. Note that the ancilla, which starts in state $|0\rangle$, ends in state $|1\rangle$; if we want, we can return it to the initial state by applying an $X$ gate.

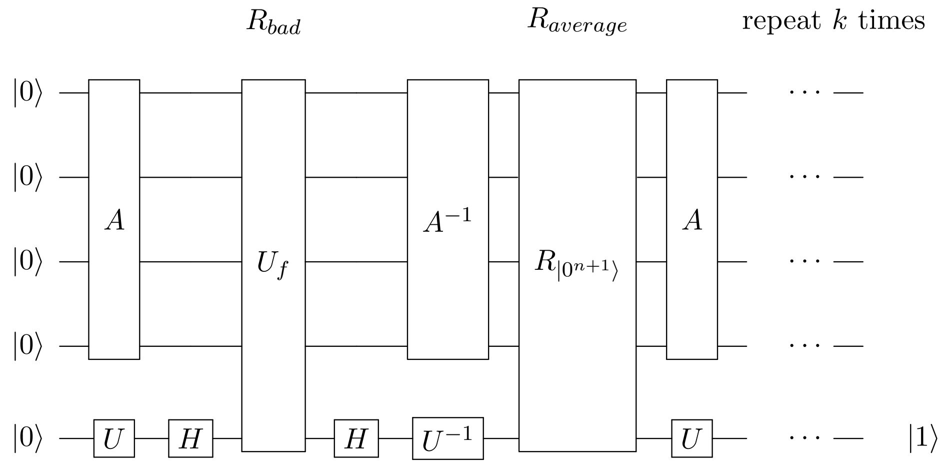


Figure 4: Final circuit for error-less amplitude amplification

**Remark 2.** Note that if all inputs $x$ are "good", then amplitude amplification requires zero iterations. Amplitude amplification only requires one iteration if $p \geq 1/4$. In this case $p$ will be reduced to $\tilde{p} = 1/4$ meaning $\tilde{\theta}_0 = \pi/6$. With one iteration, $\pi/6$ will then rotate to $3\pi/6 = \pi/2$.

## 3 Amplitude Amplification with Unknown Initial Weight

We now discuss how to perform amplitude amplification when the total weight of "good" inputs $p = \sum_{x:f(x)=1} |\alpha_x|^2$ is unknown ($p$ is still assumed to be larger than 0). The rest of the problem setup remains the same as before. Additionally, rather than focus on the number of black-box function calls required to output $|G\rangle$ with probability $\geq 1/2$, we now focus on the expected number of black-box function queries required to output $|G\rangle$. The two quantities differ by at most a factor of two: One direction follows because the expected number of trials required for success in a Bernoulli experiment with probability 1/2 is two. The other direction follows from Markov's inequality.

We develop the algorithm for amplitude amplification with unknown weight $p$ in a series of stages that tweak the algorithm presented above. Later in the course we will see how the final algorithm emerges more naturally. Note that we can still use $p$ in the analysis of the algorithm. We cannot, however, use it in the algorithm itself, e.g., to determine the number of iterations $k^*$ as done previously.

### 3.1 First Attempt

Since $p$ is unknown, we are unable to compute the number of iterations $k^*$ as described in Section 1.2 (unknown $p$ means unknown $\theta_0$). A natural first attempt then, consists of trying $k = 0, 1, 2, 3 \ldots$ iterations until there is a success (recall we have a success bit indicating when we extracted $|G\rangle$). For each value of $k$ we must first reset the system to $A|0^n\rangle$ and then perform the sequence of iterations (see the remark below).

To count the number of expected black-box queries required for this algorithm attempt, we can split the trial values of $k$ into those that are less than and those that are at least $k^*$ (even though $k^*$ is unknown, it can still be used for analysis). The number of queries for each trial is $k + 1$ ($k$ for the iterations, 1 for the ancilla qubit introduced to extract $|G\rangle$). Thus, the number of queries for the trials of all values of $k$ up to $k^*$ is:

$$\sum_{k=0}^{k^*}(k+1) = \Theta((k^*)^2) = \Theta((1/\sqrt{p})^2) = \Theta(1/p). \tag{3}$$

The first equality follows from evaluating the sum, while the second results from the fact that $k^* = \Theta(1/\theta_o) = \Theta(1/\sqrt{p})$ as mentioned in Section 1.2.

While the algorithm may succeed in outputting $|G\rangle$ before reaching $k^*$, it turns about that the above upper bound is not too far from reality and is correct within constant factors (see the exercise below). Notice that this attempt has lost the square root quantum speed up even before $k$ reaches $k^*$: the query complexity is $\Theta(1/p)$ rather than $\Theta(1/\sqrt{p})$. The former is the same as the classical setting. In this case we have gained no performance improvement by using amplitude amplification: If the initial state has a probability $p$ of being in $|G\rangle$ then we can simply try to extract $|G\rangle$ from the initial superposition, and will succeed with probability $p$. Viewing this as a Bernoulli experiment, we can expect to succeed after $1/p$ trials.

**Remark 3.** Retrying for different values of $k$, and more generally restarting in quantum algorithms, can be non-trivial. In general, algorithms can't go back to their initial quantum superpositions after a measurement has been made (in amplitude amplification the measurement happens to check success). This is one reason why quantum algorithms that have superpositions as inputs, often require that those superpositions can be generated by running a given unitary circuit $A$ on the basis state $|0^n\rangle$. In the case of amplitude amplification, there is another reason for this assumption: $A$ is needed to effect the reflection $R_{initial}$.

**Exercise (intended for theory students only).** Show that $\Pr[k^* \text{ is reached}] = \Omega(1)$. This means that the probability we must try values of $k$ up to $k^*$ is a constant value and thus that (3) gives a lower bound for the query complexity up to a constant factor.

### 3.2 Second Attempt

The problem with the above attempt that leads to $\Omega(1/p)$ query complexity is that the algorithm spends significant effort trying small values of $k < k^*$ which have small probabilities of success in extracting $|G\rangle$. A simple approach to mitigate this issue is to use a geometric sequence of $k$ values rather than an arithmetic sequence. In particular, we can try doubling $k$ each time until we successfully extract $|G\rangle$: $k = \lfloor 2^i \rfloor$ with $i = -1, 0, 1, 2, \ldots$ Additionally, by only doubling $k$ each time, we can actually ensure that we will not overshoot the region of the $|B\rangle$-$|G\rangle$ plane that measures $|G\rangle$ with probability at least 1/2. In other words, some value of $k$ will have $\theta_k \in [\frac{\pi}{4}, \frac{3\pi}{4}]$. To see why this is true, recall from Section 1.2 that $\theta_k = (2k+1)\theta_0$. Thus if we double $k$ to $2k$, $\theta_{2k} \leq 2\theta_k$. This implies that any $\theta_k < \pi/4$ will not result in a $\theta_{2k} > 3\pi/4$, and thus the high probability region $[\frac{\pi}{4}, \frac{3\pi}{4}]$ is not skipped over.

We can analyze the number of queries to the blackbox function for this attempt as done above. The number of queries for values of $k$ up to $k^*$ is:

$$\sum_{i=-1}^{\log(k^*)} (\lfloor 2^i \rfloor + 1) = O(k^* \cdot \sum_{i \geq 0} \frac{1}{\lfloor 2^i \rfloor}) = O(k^*) = O(1/\sqrt{p}). \tag{4}$$

While this attempt results in the desired query complexity for $k < k^*$, there still remains an issue for $k \geq k^*$. If $k^*$ is exceeded, then the number of queries in subsequent attempts grows very rapidly above our desired bound, so we need to control the probability of this happening. Note that the probability of success once $k^*$ is exceeded, can go down. For example, consider the case where $\theta_0 = \pi/3$. Then for $k = 0$ we have a success probability of 75%, so $k^* = 0$. With a probability of 25%, we fail and run the trial for $k = 1$. In this case $\theta_k = \theta_1 = \pi$, resulting in a zero probability of success!

### 3.3 Third Attempt

From the second attempt above, we know that a geometric sequence for trial values of $k$ leads to $O(1/\sqrt{p})$ black-box queries while $k < k^*$. For these successive values of $k$, the probability of success increases ($\theta_k$ approaches the region $[\frac{\pi}{4}, \frac{3\pi}{4}]$). However, we also know that if we fail to extract $|G\rangle$ before $k > k^*$ then the probability of success can drop considerably. How can we handle this problem? One solution is to view the geometric sequence as an *upper bound* on the number of amplitude amplification iterations for each successive trial, rather than the exact number of iterations itself. To that end, in this attempt, we pick the actual number of iterations $k$ for each trial to be uniformly from $\{0, 1, 2, \ldots \lfloor 2^i \rfloor \}$ with $i = -1, 0, 1, 2, \ldots$ and continue increasing $i$ until the first success. The intuition here is that, once $\lfloor 2^i \rfloor$ has reached $k^*$, then the probability of picking a number of iterations so that the final state has weight at least 1/2 on $|G\rangle$ is at least some constant. For this attempt, as $i$ goes up in increments of 1, we continue to double the maximum number of iterations (the "bound") between successive trials.

From the above analysis, we know that the total number of blackbox queries before $\lfloor 2^i \rfloor$ exceeds $k^*$ is $O(k^*) = O(1/\sqrt{p})$ (the uniform sampling can only lower the number of blackbox queries and does not change the overall scaling). Thus, we focus the analysis here on the case where $\lfloor 2^i \rfloor > k^*$ (i.e., we have not yet had success before $\lfloor 2^i \rfloor$ reaches $k^*$). Asymptotically, once $\lfloor 2^i \rfloor > k^*$, at least half of the range from which we uniformly sample the number of iterations lies in the "good region," i.e., in $[\frac{\pi}{4}, \frac{3\pi}{4}]$ or $[\frac{5\pi}{4}, \frac{7\pi}{4}]$ modulo $2\pi$. Since in this region the probability of measuring $|G\rangle$ is at least half, the probability of success for each trial with $\lfloor 2^i \rfloor > k^*$ is asymptotically: Pr[$k$ in good region] $\cdot$ Pr[success | in good region] $\geq \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$.

Thus, every trial with $\lfloor 2^i \rfloor > k^*$ is like a Bernoulli experiment with success probability at least 1/4. To analyze the expected number of blackbox queries after the bound $\lfloor 2^i \rfloor > k^*$, we introduce $i'$ to count the number of trials for which $\lfloor 2^i \rfloor > k^*$. In other words, $i = i^* + i'$ with $2^{i^*} = k^*$. The expected number of blackbox queries after the bound $\lfloor 2^i \rfloor > k^*$ is then the probability of requiring a certain number of trials $i'$ (the probability that all previous trials $i' = 0, 1, 2, \ldots$ failed) times the number of queries for that specific value of $i'$. This is order:

$$\sum_{i'\geq 0}(\frac{3}{4})^{i'}\cdot(2^i)=\sum_{i'\geq 0}(\frac{3}{4})^{i'}\cdot(k^*\cdot 2^{i'})=k^*\sum_{i'\geq 0}(\frac{3}{2})^{i'}. \tag{5}$$

We would like this complexity to be $O(k^*)=O(1/\sqrt{p})$, in which case the whole algorithm would be $O(1/\sqrt{p})$ (by combining the complexities before and after $k^*$). Unfortunately, the above geometric series diverges since $3/2>1$.

### 3.4 Final Attempt

Our final attempt for amplitude amplification with unknown initial weight $p$ (unknown $k^*$) consists of tweaking attempt three to ensure that the series for the number of queries after the bound $\lfloor 2^i \rfloor > k^*$ converges. The idea is to change from doubling the bound in each successive trial to some other multiplicative constant factor. Assume that the bound on the number of iterations increases by a constant factor $\lambda > 1$ (it must be larger than 1 for the bound to increase). I.e., we pick the number of iterations $k$ for each trial to be uniformly from $\{0,1,2,\ldots\lfloor\lambda^i\rfloor\ \}$ with $i=-1,0,1,2...$ and continue until the first success. Using a similar analysis as above in attempt three results in the expected number of blackbox queries after the bound $\lfloor\lambda^i\rfloor > k^*$ to be given by:

$$\sum_{i'\geq 0}(\frac{3}{4})^{i'}\cdot(k^*\cdot\lambda^{i'})=k^*\sum_{i'\geq 0}(\frac{3\lambda}{4})^{i'}. \tag{6}$$

This series converges for $3\lambda/4<1$. Thus we have that the number of applications of the blackbox is $O(k^*)=O(1/\sqrt{p})$ for $1<\lambda<3/4$. This final attempt results in an amplitude amplification algorithm with square root speed up even when the total weight $p$ of the "good" inputs $x$ is unknown.