Quantum Algorithms 2/28/2023

Lecture 11: Oblivious Amplitude Amplification

Instructor: Dieter van Melkebeek

In this lecture, we present the two-domain view of amplitude amplification, introduce oblivious amplitude amplification, and discuss the latter's relationship with block encoding. Oblivious amplitude amplification is a generalization of amplitude amplification in which the starting state $|0^n\rangle$ is replaced by some pure state $|\psi\rangle$. This problem is not solvable in general, but we give an algorithm to solve it in the *purified setting with independent initial weight*, in which the success probability is independent of $|\psi\rangle$ and we are promised that $|\psi\rangle = |0^\ell\rangle |\phi\rangle$ for some $\ell > 0$. The setting has close ties with the notion of block encoding, a quantum framework to effect linear transformations that we will use extensively in later parts of the course. We introduce the framework and show an equivalence between the purified setting with independent initial weights and block encodings of scaled unitary transformations.

1 Recap

Recall the setup for amplitude amplification. We are given blackbox access to a function $f : \{0,1\}^n \to \{0,1\}$, and an n-qubit unitary circuit A such that $A|0^n\rangle = \sum_x \alpha_x |x\rangle$ has $\alpha_x \neq 0$ for some *good* basis state $|x\rangle$, i.e., one satisfying $f(x) = 1$. The weight $p = \sum_{x:f(x)=1} |\alpha_x|^2$ of the good states may or may not be given; we saw how to handle both of those cases.

We write

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

where $|B\rangle = \frac{1}{\sqrt{1-p}} \sum_{x:f(x)=0} \alpha_x |x\rangle$ and $|G\rangle = \frac{1}{\sqrt{p}} \sum_{x:f(x)=1} \alpha_x |x\rangle$ are the renormalized initial superpositions of bad and good states, respectively. Last time we saw the following algorithm that outputs $|G\rangle$ with probability at least $\frac{1}{2}$ when p is know: Start from state $A|0^n\rangle$ and repeat the below two steps until arriving in a state with weight at least $\frac{1}{2}$ on $|G\rangle$.

1. R_{bad}: Use phase kickback to flip the sign of all good $|x\rangle$.
2. $R_{initial} = AR_{|0^n\rangle}A^{-1}$: reflect across $A|0^n\rangle$.

The number of repetitions needed is only $O(1\sqrt{p})$.

2 Two-Domain View

We next present the two-dimensional unit circle representation of amplitude amplification. We may view A as a sort of Fourier transform between the signal domain containing the initial input $|0^n\rangle$, and the frequency domain, with axes $|B\rangle$ and $|G\rangle$ for the bad and good components, respectively. Since A is unitary, whenever we rotate the state by an angle θ in the signal domain, the state in the frequency domain is rotated by the same angle θ and vice-versa. One iteration of the amplitude amplification algorithm is shown in Figure 1, with the signal domain on the left and the frequency domain on the right.

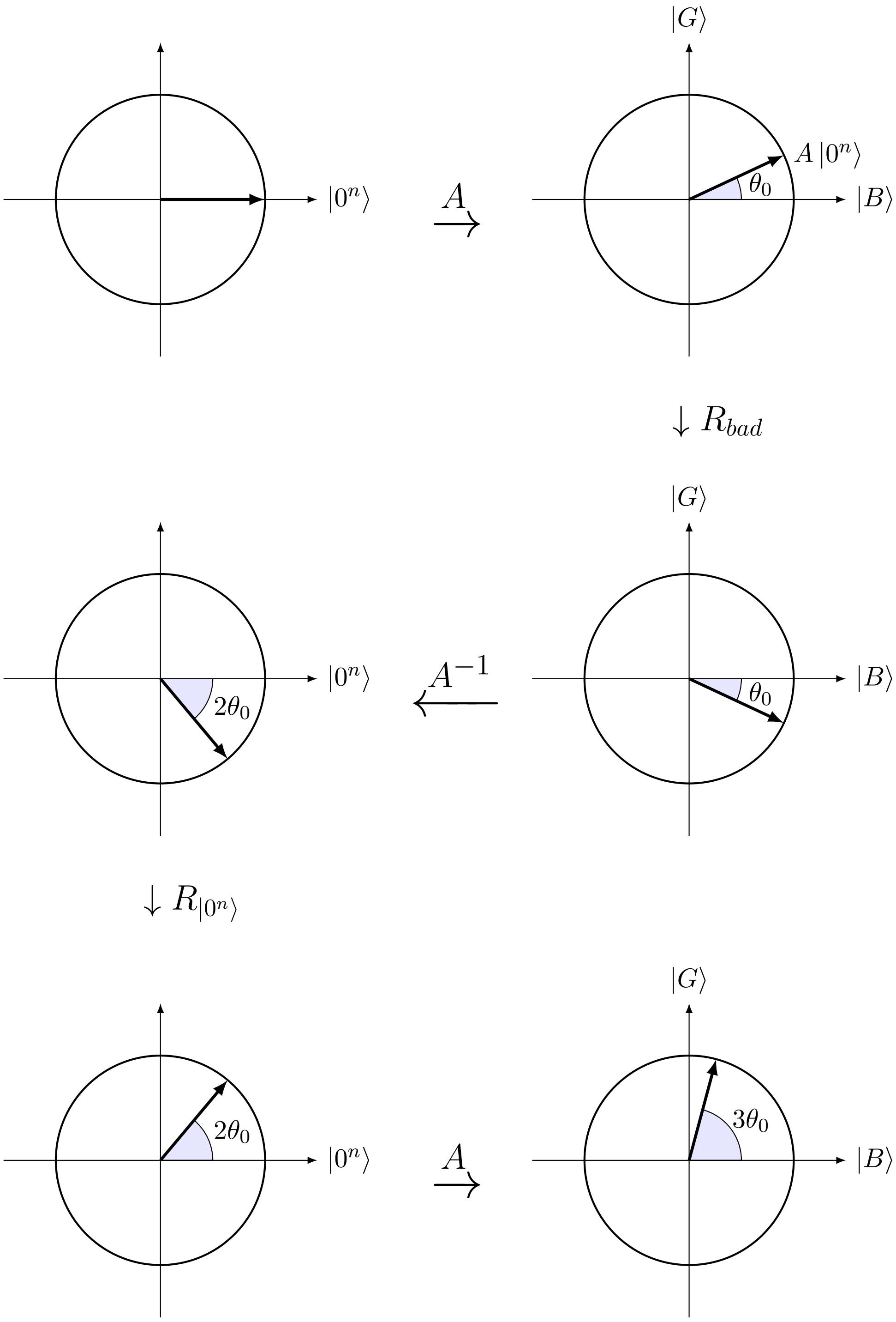

Figure 1: Amplitude amplification in the two-domain view

3 Oblivious Amplitude Amplification

Now consider the same amplitude amplification setup, but the starting state is an arbitrary pure state $|\psi\rangle$ instead of $|0^n\rangle$, so $A|\psi\rangle = \sum_x \alpha_x |x\rangle$ has $\alpha_x \neq 0$ for some x with $f(x) = 1$. The weight $p(\psi) = \sum_{x:f(x)=1} |\alpha_x|^2$ of the good states now depends on $|\psi\rangle$, as does

$$|G(\psi)\rangle = \frac{1}{\sqrt{p(\psi)}} \sum_{x:f(x)=1} \alpha_x |x\rangle .$$

We aim to output $|G(\psi)\rangle$ with probability at least $\frac{1}{2}$. As in normal amplitude amplification, we attach a success indicator qubit and apply U_f so that the success indicator stores $|f(x)\rangle$. If we measure 1 from the success indicator, we know the state has collapsed to $|G(\psi)\rangle$, as desired.

Replacing starting state $|0^n\rangle$ with $|\psi\rangle$ causes some issues not present in normal amplitude amplification. First, we're only given a single copy of $|\psi\rangle$ and can't clone or regenerate it, so you can no longer throw everything away and start the whole process over from $|0^n\rangle$ after a failed measurement. This is an issue for the naïve randomized $O\left(\frac{1}{p}\right)$ algorithm as well, which also relies on repeated trials. Furthermore, even if we somehow had access to many copies of $|\psi\rangle$ or $A|\psi\rangle$, we would still need an efficient replacement for the reflection $R_{initial}$ about $|\psi_0\rangle = A|\psi\rangle$. Recall in the original amplitude amplification setup we had $R_{initial} = AR_{|0^n\rangle}A^{-1}$, and reflecting across $|0^n\rangle$ was easy since $|0^n\rangle$ is a basis state - you merely need to flip the phase of all other basis states. It is not obvious, however, how to compute $R_{|\psi\rangle}$ efficiently given $|\psi\rangle$.

These difficulties make it too difficult to perform oblivious amplitude amplification in general. Even if we restrict $p(\psi) \equiv p$ to be independent of $|\psi\rangle$, we will see that oblivious amplitude amplification is still only possible in general when $p = 1$, in which case amplitude amplification is unneccessary because we're already guaranteed a good measurement.

3.1 Purified Setting with Independent Initial Weight

We now introduce the purified setting with independent initial weight, where oblivious amplitude amplification is possible. We have a promise that $|\psi\rangle$ is composed of ℓ ancilla qubits followed by a "real" state: that is, $|\psi\rangle = |0^\ell\rangle |\phi\rangle$. We also have that $p(\psi) = p$ is independent of ψ. States of the form $|0^\ell\rangle |\phi\rangle$ naturally arise when purifying quantum circuits. The type of independence naturally arises in certain settings, e.g., in some zero-knowledge systems that we will cover later in the course.

Formally, for state $|\phi\rangle$ on $m = n - \ell$ qubits, we have

$$p = \|P_1 A |0^\ell\rangle |\phi\rangle\|_2^2, \tag{1}$$

where P_1 is the projection onto the space spanned by the good basis states. In other words, the weight of the good states after applying A is independent of ϕ. Using the fact that P_1 is a projection operator (in particular, $P_1^* P_1 = P_1$), we may rewrite

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

for $M = A^* P_1 A$. M is a $2^n \times 2^n$ Hermitian (as P_1 is Hermitian) matrix indexed by basis states, represented as strings of n bits. Split up M into blocks based on whether its index bitstrings start with 0^ℓ:

$$M = \begin{bmatrix} M_{TL} & M_{TR} \\ M_{BL} & M_{BR} \end{bmatrix}, \tag{3}$$

where blocks M_{TL} and M_{TR} have row indices starting with 0^ℓ and M_{TL} and M_{BL} have column indices starting with 0^ℓ. For example, for $\ell = 1$, all four blocks have the same size, and for $\ell = n$, M_{TL} consists of only a single entry. $|0^\ell\rangle\,|\phi\rangle$ is a superposition of basis states starting with $|0^\ell\rangle$, so, in (2), $|0^\ell\rangle\,|\phi\rangle$ and $\langle 0^\ell|\,\langle\phi|$ select the columns and rows, respectively, of M whose indices begin with 0^ℓ. Thus we may rewrite (2) as

$$p = \langle\phi|\, M_{TL}\, |\phi\rangle \tag{4}$$

for all m-component pure states $|\phi\rangle$. $M_{TL} = pI$ satisfies (4). Furthermore, since M_{TL} is Hermitian, it has an orthonormal basis of eigenvectors. Letting $|\phi\rangle$ range over all eigenvectors of M_{TL}, we obtain from (4) that every eigenvalue of M_{TL} is p. Hence, writing any $v \in \mathbb{C}^{2^n}$ in the eigenvector basis, we find that $M_{TL}v = pv$. This implies that we in fact must have $M_{TL} = pI$, so we may rewrite

$$M = \begin{bmatrix} pI & M_{TR} \\ M_{BL} & M_{BR} \end{bmatrix}. \tag{5}$$

In particular, if $\ell = 0$, then $M = pI$. In this case, as $p > 0$, all the basis states need to be good for (1) to hold: If there were some $x^* \in \{0,1\}^n$ such that $f(x) = 0$, then $|\phi\rangle \doteq A^{-1}\,|x^*\rangle$ is well-defined (as A is unitary) and satisfies $P_1 A\,|\phi\rangle = 0$, violating (1). Thus, $P_1 = I$ and $M \doteq A^* P_1 A = I$, so $p = 1$. If $p = 1$ then we don't need to do amplification, so this procedure doesn't make sense with $\ell = 0$ ancilla qubits. In other words, the ancillas are necessary.

3.2 Analysis

Consider applying A to the start state $|0^\ell\rangle\,|\phi\rangle$, followed by a successful measurement of $f(x) = 1$ in the success indicator bit (effectively projecting the state onto the good states using P_1), followed by an application of $A^* = A^{-1}$:

$$|0^\ell\rangle\,|\phi\rangle \xrightarrow{A} |\psi_0\rangle \xrightarrow{\text{success}} |G\rangle = \frac{1}{\sqrt{p}} P_1\,|\psi_0\rangle \xrightarrow{A^{-1}} |\psi_0'\rangle\,.$$

Observe that we've applied A, then $\frac{1}{\sqrt{p}}P_1$, then A^*, so

$$|\psi_0'\rangle = \frac{1}{\sqrt{p}} A^* P_1 A\,|0^\ell\rangle\,|\phi\rangle = \frac{1}{\sqrt{p}} M\,|0^\ell\rangle\,|\phi\rangle\,. \tag{6}$$

$|0^\ell\rangle\,|\phi\rangle$ is a superposition of basis states starting with 0^ℓ, so only the left two blocks of M act on M's input in (6). Let C be the projection onto clean ancillas (components starting with 0^ℓ), followed by renormalization. When we apply C to $|\psi_0'\rangle$, we eliminate all of the output basis states from M's bottom two blocks. Hence we effectively only apply pI in (6), so $C\,|\phi_0'\rangle = |0^\ell\rangle\,|\phi\rangle$ (renormalization removes the scalar factor p). To recap, we have

$$|0^\ell\rangle\,|\phi\rangle \xrightarrow{A} |\psi_0\rangle \xrightarrow{\text{success}} |G\rangle = \frac{1}{\sqrt{p}} P_1\,|\psi_0\rangle \xrightarrow{A^{-1}} |\psi_0'\rangle \xrightarrow{C} |0^\ell\rangle\,|\phi\rangle\,. \tag{7}$$

The probability of failure is $\|P_0 A\,|0^\ell\rangle\,|\phi\rangle\,\|_2^2 = 1-p$ (where P_0 is the projection onto the bad states), which is also independent of $|\phi\rangle$. A similar analysis gives

$$|0^\ell\rangle\,|\phi\rangle \xrightarrow{A} |\psi_0\rangle \xrightarrow{\text{failure}} |B\rangle = \frac{1}{\sqrt{1-p}} P_0\,|\psi_0\rangle \xrightarrow{A^{-1}} |\psi_0''\rangle \xrightarrow{C} |0^\ell\rangle\,|\phi\rangle\,. \tag{8}$$

By linearity it follows that for any $\alpha_G, \alpha_B \in \mathbb{C}$, $CA^{-1}(\alpha_G |G\rangle + \alpha_B |B\rangle) = |0^\ell\rangle |\phi\rangle$. In words, if we run the unitary circuit A in reverse on $\alpha_G |G\rangle + \alpha_B |B\rangle$, then the projection of the state onto the clean ancillas equals the initial state $|0^\ell\rangle |\phi\rangle$ up to normalization. We next use this property to run the $O(1/\sqrt{p})$ procedure for quantum amplification in a way independent of $|\phi\rangle$.

3.3 Algorithm

The oblivious amplitude amplification algorithm is similar to our original amplitude amplification algorithm discussed in section 1 and illustrated in Figure 1.

After applying A to the inital state $|0^\ell\rangle |\phi\rangle$, we apply R_{bad} (which is implemented using phase kickback, hence we still have access to it), then A^{-1}, which rotates clockwise by $2\theta_0$ in the signal domain. At this point, our original amplitude amplification algorithm applied $R_{|0^n\rangle}$. Here we would like to reflect across $|0^\ell\rangle |\phi\rangle$, but we don't know $|\phi\rangle$. However, recall from (7) and (8) that the projection C of both $|G\rangle$ and $|B\rangle$, respectively, onto clean ancillas is exactly $|0^\ell\rangle |\phi\rangle$, up to normalization. The state is a superposition of $|B\rangle$ and $|G\rangle$, so its projection via C onto the basis states with clean ancillas is $|0^\ell\rangle |\phi\rangle$. Thus the reflection $R_{|0^\ell *\rangle}$ across the superposition of all basis vectors with clean ancillas has the same effect as reflection across $|0^\ell\rangle |\phi\rangle$, and is easy to implement – we simply flip the phase of all basis states that don't have clean ancillas.

Now we have rotated clockwise by $4\theta_0$ in the frequency domain, so upon applying A to bring us back to the signal domain, we find ourselves at angle $3\theta_0$ above the $|B(\phi)\rangle$-axis. Hence after one iteration we have rotated by $2\theta_0$ towards $|G(\phi)\rangle$. This is the same effect as one iteration of the original amplitude amplification algorithm, and the remaining analysis is identical. An illustration of this algorithm is given in Figure 2.

Exercise #9. Consider the algorithm for oblivious amplitude amplification from subsection 3.3 but evaluate the success indicator each time the frequency domain is reached, and only continue in case of no success.

(a) Determine the probability of no success within the first k iterations as a function of p.

(b) Determine the expected number of iterations until the first success as a function of p.

4 Block Encoding

The purified setting with independent initial weight is closely related to the notion of block encoding for scaled unitary transformations. A *block encoding* A of a matrix B that acts on m qubits with ℓ ancilla qubits is a unitary operator

$$A = \begin{bmatrix} B & * \\ * & * \end{bmatrix}$$

acting on $n = \ell + m$ qubits. The columns in A's left two blocks correspond to inputs beginning with $|0^\ell\rangle$ and the rows in A's top two blocks correspond to outputs beginning with $|0^\ell\rangle$. One way to obtain $B|v\rangle$ given $|v\rangle$, is to apply A to $|0^\ell\rangle |v\rangle$ and project onto the basis states beginning with $|0^\ell\rangle$. Up to normalization, this can be effected by the following quantum process, assuming a unitary circuit for A: Apply A to $|0^\ell\rangle |v\rangle$, and measure the first ℓ qubits. If they are all 0, then we were successful, and the remaining m qubits are now in state $B|v\rangle / \|B|v\rangle\|_2$. The probability of a success is $\|B|v\rangle\|_2^2$, the squared length of the projection, and the measurement outcome of

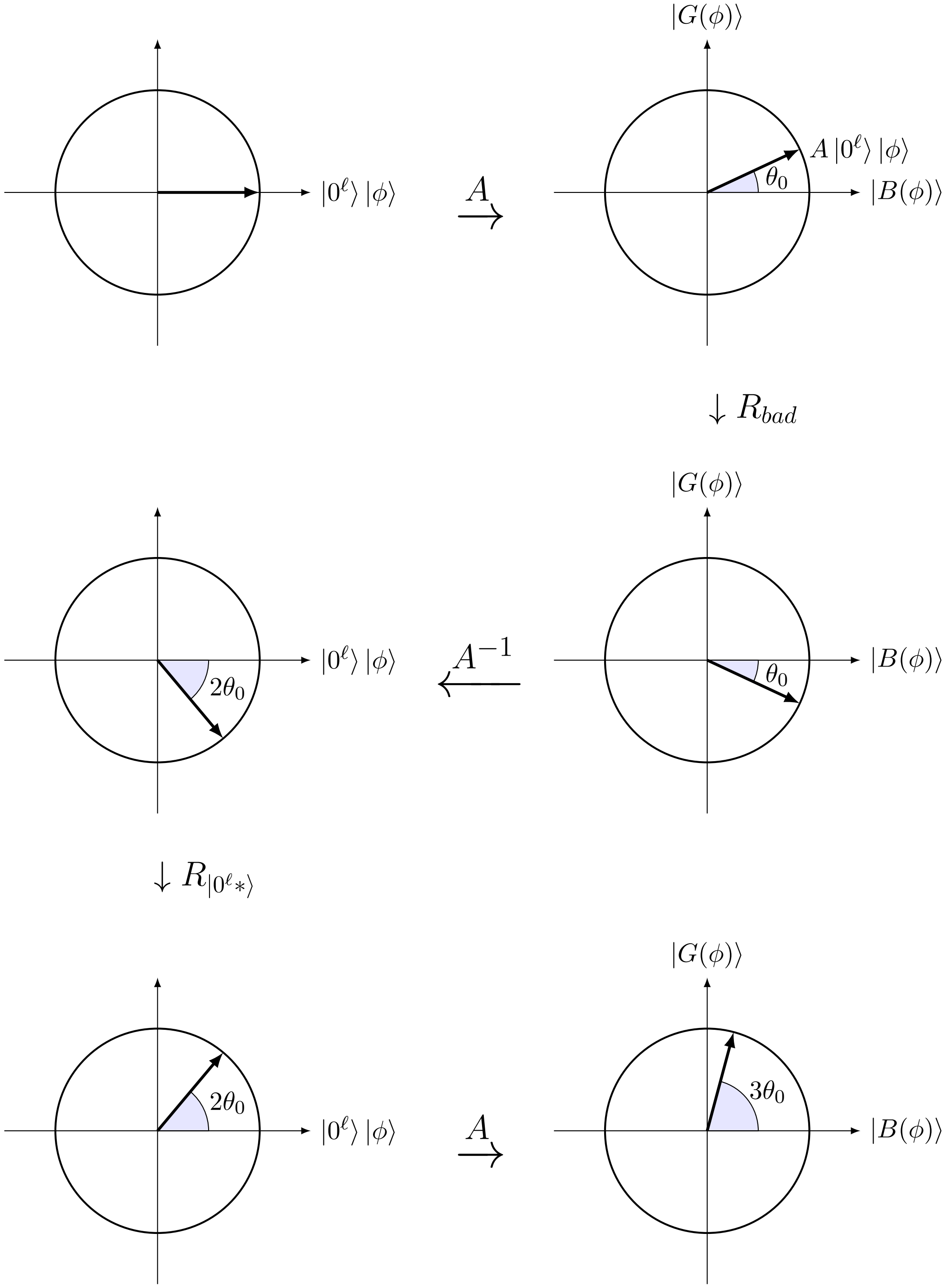

Figure 2: Oblivious amplitude amplification in the two-domain view

the first ℓ qubits acts as a success indicator. Thus, block encodings form a probabilistic way of realizing the linear transformation B (up to normalization) with a success indicator. Measuring the ancillas and only continuing in case they have a certain value (in this case 0^ℓ) if often referred to as "postselecting" that value.

The following two propositions show that block encodings of matrices B that are scaled unitaries are equivalent to the setting of amplitude amplification in the purified setting with independent initial weight where the good states are the ones where the ancillas are all 0.

Proposition 1. *A block encoding of a unitary matrix yields a purified setting with independent initial weight.*

Proof. Consider a block encoding A of $B = \sqrt{p}U$ for some unitary U. Consider amplitude amplification with A starting from a state of the form $|0^\ell\rangle\,|\phi\rangle$ with the following predicate:

$$f(x) = \begin{cases} 1 & x = 0^\ell y \text{ for some } y \\ 0 & \text{otherwise} \end{cases}. \tag{9}$$

$A\,|0^\ell\rangle\,|\phi\rangle$ has weight $\|B\,|\phi\rangle\,\|_2^2 = p$ on the "good" states, independent of $|\phi\rangle$. Thus the conditions for the purified setting with independent initial weight are satisfied. □

Note that, using oblivious amplitude amplification, we may boost p, the probability of successfully applying B.

Proposition 2. *A purified setting with independent initial weight yields a block encoding of a scaled unitary matrix.*

Proof. Again consider the purified setting with independent initial weight with f defined as in (9). Hence P_1 is the projection onto basis states of the form $|0^\ell *\rangle$. As in (3), we may write

$$A = \begin{bmatrix} A_{TL} & A_{TR} \\ A_{BL} & A_{BR} \end{bmatrix}.$$

Since P_1 projects onto the basis states indexing the left two blocks of A, defining M as in (2), we have

$$M = (A^* P_1)A = \begin{bmatrix} A_{TL}^* & 0 \\ A_{TR}^* & 0 \end{bmatrix} \begin{bmatrix} A_{TL} & A_{TR} \\ A_{BL} & A_{BR} \end{bmatrix} = \begin{bmatrix} A_{TL}^* A_{TL} & * \\ * & * \end{bmatrix}.$$

We saw in (5) that the upper left block of M is pI for any purified setting with independent initial weight, so $A_{TL}^* A_{TL} = pI$, hence A_{TL} is unitary up to factor $\sqrt{p}$. Thus A is a block encoding of a scaled unitary matrix. □

The equivalence given by Propositions 1 and 2 mean that, in the block encoding framework, the success probability p of the encoding can be amplified to 1/2 or more using only $O(1/\sqrt{p})$ applications of the original block encoding A and its inverse, provided the matrix B is unitary up to a scalar. If p is known, we can make the success probability 1.