Quantum Subroutine Composition

Stacey Jeffery[*]

CWI & QuSoft, the Netherlands

February 14, 2025

Abstract

An important tool in algorithm design is the ability to build algorithms from other algorithms that run as subroutines. In the case of quantum algorithms, a subroutine may be called on a superposition of different inputs, which complicates things. For example, a classical algorithm that calls a subroutine Q times, where the average probability of querying the subroutine on input i is $\bar{p}_i$, and the cost of the subroutine on input i is a random variable T_i, incurs expected cost $\mathsf{Q} \sum_i \bar{p}_i \mathbb{E}[T_i]$ from all subroutine queries. While this statement is obvious for classical algorithms, for quantum algorithms, it is much less so, since naively, if we run a quantum subroutine on a superposition of inputs, we need to wait for all branches of the superposition to terminate before we can apply the next operation. We nonetheless show an analogous quantum statement (*): If $\bar{q}_i$ is the average *query weight on i* over all queries, the cost from all quantum subroutine queries is $\mathsf{Q} \sum_i \bar{q}_i \mathbb{E}[T_i]$. Here the query weight on i for a particular query is the probability of measuring i in the input register if we were to measure right before the query.

We prove this result using the technique of multidimensional quantum walks, recently introduced in [Jeffery, Zur 2022]. We present a more general version of their quantum walk edge composition result, which yields *variable-time quantum walks*, generalizing variable-time quantum search, by, for example, replacing the update cost with $\sqrt{\sum_{u,v} \pi_u P_{u,v} \mathbb{E}[T_{u,v}^2]}$, where $T_{u,v}$ is the cost to move from vertex u to vertex v. The same technique that allows us to compose quantum subroutines in quantum walks can also be used to compose in any quantum algorithm, which is how we prove (*).

1 Introduction

Classical algorithms make extensive use of subroutines, which allow an algorithm designer to build up an algorithm in modular blocks. Analysis of such algorithms can also be done in a simple, modular way. Suppose an outer classical algorithm makes use of some subroutine at Q steps of the algorithm, and suppose the subroutine's stopping time on input i is some random variable T_i. Suppose $\bar{p}_i$ is the average probability, over all Q queries to the subroutine, that the subroutine is called on input i – so $\mathsf{Q}\bar{p}_i$ is the expected number of times the subroutine is called on input i, and $\sum_i \bar{p}_i = 1$. Let L denote the total number of *other* operations used by the outer algorithm. Then the expected running time of the classical algorithm is:

$$\mathsf{L} + \mathsf{Q} \sum_i \bar{p}_i \mathbb{E}[T_i]. \tag{1}$$

We can get a bounded-error algorithm with this asymptotic running time by simply stopping after 10 times the expected number of steps, if the algorithm has not already halted. At that point we can output an arbitrary (likely wrong) answer if the algorithm hasn't already halted, but this happens with probability at most some small constant.

While subroutines are also quite useful in quantum algorithms, it is not at all obvious how to do the above classical analysis in a quantum setting. Suppose an outer quantum algorithm makes use of

---

some subroutine. In general, the subroutine will be called on a superposition of inputs i, so if some branches of the superposition are finished early, it is not obvious how to exploit this in the overall complexity. Generally, if $\mathsf{T}_{\max}$ is the maximum time we must wait for the quantum subroutine to terminate, the total complexity of the quantum algorithm is:

$$O(\mathsf{L} + \mathsf{Q} \cdot \mathsf{T}_{\max}), \tag{2}$$

where again, Q is the number of queries to the subroutine, and L is the number of other operations made by the outer algorithm. Our main result is to give a quantum analogue of (1) for the special case where each subroutine outputs a bit. Given a quantum algorithm that decides f using L basic operations, as well as Q queries to a subroutine that runs in time T_i on input i to compute some bit[1] g_i, there is a bounded error quantum algorithm that computes $f(g)$ with complexity (neglecting log factors):

$$\mathsf{L} + \mathsf{Q} \sum_i \bar{q}_i \mathbb{E}[T_i],$$

where $\bar{q}_i$ is the *average query weight on i* – the squared norm on $|i\rangle$ averaged over all queries, so in particular, $\sum_i \bar{q}_i = 1$. We describe the result in more detail in Section 1.2. While such a result is obvious in classical algorithms, it is much less so in quantum algorithms. In fact, if the outer algorithm and subroutine are zero-error algorithms, and we want to compose them to get a zero-error algorithm with this *expected* running time, while again obvious in the classical case, this is *not* possible in general for quantum algorithms [BdW03].

If $\mathbb{E}[T_i] = \mu$ is a known constant (in i, meaning μ does not depend on i, though T_i could), then this result is much less interesting: we can always stop the subroutine after 10μ steps, and by Markov's inequality, this introduces at most $1/10$ additional probability of error (which we can reduce through logarithmic repetitions). However, if the values $\mathbb{E}[T_i]$ vary significantly in i, in contrast to the classical case, it is not obvious (to the best of our knowledge) that this result should hold for quantum algorithms, and special cases of this result have been the subject of significant effort.

For example, consider evaluating an unbalanced formula that is the OR of n ANDs, of arity $k_1, \ldots, k_n$:

$$f(x^{(1)}, \ldots, x^{(n)}) = \mathrm{OR}_n(\mathrm{AND}_{k_1}(x^{(1)}), \ldots, \mathrm{AND}_{k_n}(x^{(n)})),$$

where $x^{(j)} \in \{0,1\}^{k_j}$. Naively, using nested quantum search, one could search for a j such that $\mathrm{AND}_{k_j}(x^{(j)}) = 1$, where $\mathrm{AND}_{k_j}(x^{(j)})$ can be evaluated in $\sqrt{k_j}$ steps, for total cost $\sqrt{n} \cdot \max_{j \in [n]} \sqrt{k_j}$ (we neglect log factors in this paragraph), but this may be far from optimal. However, using highly non-trivial techniques, it has been shown how to evaluate any such formula in $\sqrt{n}\sqrt{\frac{1}{n}\sum_{j=1}^{n} k_j}$ quantum time or *any* AND-OR formula on N bits in $\sqrt{N}$ time [ACR+07, Rei11a].

This non-trivial upper bound for OR of ANDs can be seen as a special case of a technique called *variable-time quantum search*. Consider a setting in which a quantum subroutine runs in time T_i on input i, and one wants to decide if there exists an $i \in [n]$ on which the subroutine outputs 1. Classically, it is easy to see that this can be decided in cost $\sum_{i \in [n]} \mathsf{T}_i$. Naive use of quantum search gives an upper bound of $\sqrt{n}\mathsf{T}_{\max}$, where $\mathsf{T}_{\max} = \max_{i \in [n]} \mathsf{T}_i$. Ambainis showed, in a highly non-trivial way, an upper bound of $\widetilde{O}\left(\sqrt{\sum_{i=1}^{n} \mathsf{T}_i^2}\right)$ [Amb10]. This result assumes that if the i-th subroutine consists of unitaries $U_1^i, \ldots, U_{\mathsf{T}_{\max}}^i$, the operator $\sum_{t=1}^{\mathsf{T}_{\max}} |t\rangle\langle t| \otimes U_t^i$ can be implemented in "unit cost". We will make this assumption as well, throughout this paper, and use "unit cost" to describe a cost we are willing to accept as a multiplicative factor on all complexities (for example, $O(1)$ or polylogarithmic in some natural variable). This assumption is only reasonable if the unitaries each have unit cost. It does not hold for strict gate complexity when $U_1^i, \ldots, U_{\mathsf{T}_{\max}}^i$ are arbitrary gates (local unitaries), but it holds, for example, in the *fully quantum* QRAM model if $U_1^i, \ldots, U_{\mathsf{T}_{\max}}^i$ are stored as a list of gates in QCRAM[2], or if they satisfy certain uniformity conditions (see [JZ22, Section 2.2]).

---

[1] We can think of the bit g_i as being $g_i(x)$ for some function g_i, and x the full input to the composed algorithm.

[2] Classical memory to which a quantum computer has read-only superposition access, sometimes referred to as "QRAM" in previous literature.

In [Amb12], Ambainis considers a slightly different setting, where a quantum subroutine runs in time T, where T is a random variable on $\{1, \ldots, \mathsf{T}_{\max}\}$ – that is, the algorithm may stop at different times with different probabilities, outputting either a 0 or 1 in some answer register whenever it stops. Ambainis shows how to amplify the 1 part of the computation in complexity (neglecting log factors):

$$\mathsf{T}_{\max} + \sqrt{\frac{\mathbb{E}[T^2]}{\varepsilon}}$$

where ε is a lower bound on the probability that the algorithm outputs 1 (assuming this is non-zero).

We consider a combination of Ambainis' two settings, where a subroutine computes some bit g_i for any input i, and the time it takes is a random variable T_i on $\{1, \ldots, \mathsf{T}_{\max}\}$. When the expected stopping times $\mathbb{E}[T_i]$ are unknown, and ε is a lower bound on the probability in uniform input i that the subroutine outputs 1, we obtain (neglecting log factors)

$$\sqrt{\frac{1}{n\varepsilon} \sum_{i \in [n]} \mathbb{E}[T_i^2]}. \tag{3}$$

When the expected stopping times $\mathbb{E}[T_i]$ are known, we can replace $\mathbb{E}[T_i^2]$ with $\mathbb{E}[T_i]^2$, which is better when T_i has large variance, and improve, but not remove, the polylogarithmic dependence on $\mathsf{T}_{\max}$.

We also show two alternative variable-time search results that may be better in certain settings. Our variable-time search results are a special case of a new technique: variable-time quantum walk algorithms. Loosely speaking, if the cost of moving from a vertex u to a vertex v in a random walk on a graph G is some random variable $T_{u,v}$ on $\{1, \ldots, \mathsf{T}_{\max}\}$, previous quantum walk search algorithms would have incurred a multiplicative overhead on the complexity of $\mathsf{T}_{\max}$. We show how to improve this overhead to $\sqrt{\sum_{\{u,v\} \in E(G)} \pi(u) P_{u,v} \mathbb{E}[T_{u,v}^2]}$, where P is the transition matrix of the random walk, and π is its stationary distribution. We show a similar variable-time dependence on the *checking cost*. We describe our variable-time quantum search and variable-time quantum walk results in Section 1.1.

Techniques: Our results build on multidimensional quantum walks, recently described in [JZ22]. Given a collection of states, each of which can be efficiently prepared, and therefore reflected around, we can define an *overlap graph*, with one node for every state, and a pair of nodes adjacent if and only if the states overlap. If this graph is bipartite, with bipartition $V_{\mathcal{A}}, V_{\mathcal{B}}$, then reflections around the states in $V_{\mathcal{A}}$ (resp. $V_{\mathcal{B}}$) commute, and so we can also efficiently reflect around the span of all states in $V_{\mathcal{A}}$, and around the span of all states in $V_{\mathcal{B}}$. Then sufficiently precise phase estimation of the product of these two reflections can be used to reflect around the span of *all* states. When the states are *quantum walk states* for some graph G: $|\psi_\star^G(u)\rangle = \sum_{v : \{u,v\} \in E(G)} \sqrt{P_{u,v}}|\{u,v\}\rangle$, then $|\psi_\star^G(u)\rangle$ and $|\psi_\star^G(v)\rangle$ overlap if and only if $\{u, v\} \in G$, so the overlap graph of these states is exactly G. Quantum walk search algorithms, like those we describe shortly in Section 1.1, use analysis of the properties of the graph G to analyze what precision of phase estimation is needed to reflect around the span of all states.

Ref. [JZ22] extends this idea to a set of *spaces* that can be efficiently reflected around. If $\{\Psi_u\}_{u \in V}$ is a set of subspaces of some inner product space H, we can define an overlap graph on V, with an edge between u and v whenever Ψ_u and Ψ_v are non-orthogonal. Just as in the case when each space is one-dimensional, if the overlap graph is bipartite, a bipartition gives rise to two efficiently implementable reflections, whose product can be used to reflect around the span of all the spaces using sufficiently precise phase estimation. The properties of the overlap graph can be used to bound the required precision.

Ref. [JZ22] use this idea in two ways. The first is the technique of *alternative neighbourhoods*, which is orthogonal to the ideas of this paper: we do not deal with this technique here for simplicity, but there is no reason it could not be used simultaneously with our techniques. This work builds on the second technique, called *edge composition*[3]. Suppose there is a subroutine that implements a step of the quantum walk: for now imagine some reversible version of moving from u, to some neighbour v, in

---

[3]which is similar to the concurrent, independent technique of *graph composition* of span programs [Cor23].
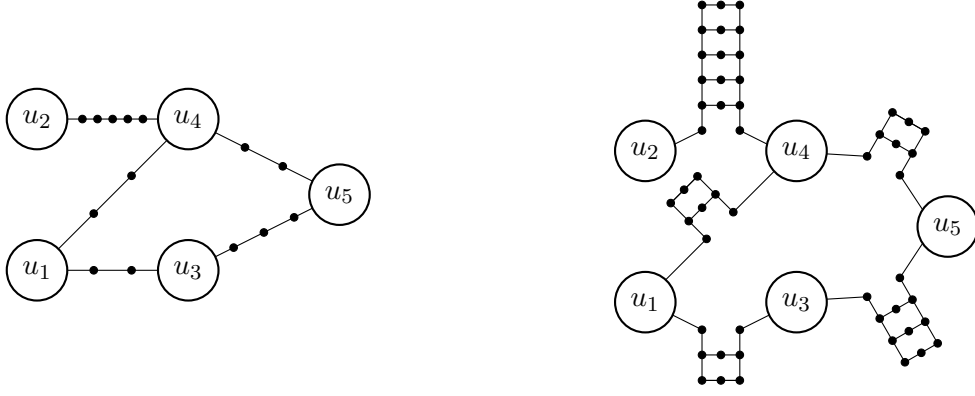
Figure 1: Instead of using a $T_{u,v}$ step quantum algorithm for the transition from u to v to build a "bridge" from u to v that functions like a path of length $T_{u,v}$ (left), we use a variable-time quantum algorithm to put a ladder-like gadget between u and v (right). The rungs of the ladder correspond to the steps of the quantum algorithm, and intuitively, we should think of the weight of each rung as corresponding to the probability that the algorithm terminates at that step.

cost $T_{u,v}$. Then implementing the reflection around the states $\{|\psi_\star^G(u)\rangle\}_{u \in V_\mathcal{A}}$ costs $\max_{u,v} T_{u,v}$. The edge composition technique uses the random walk and the subroutine to define, for each $\{u,v\} \in E$, a sequence of $T_{u,v}$ spaces whose overlap graph is similar to G, but each edge is replaced by a path of length $T_{u,v}$ (see Figure 1). Our quantum walk composition described in Section 1.1 is similar, except that we allow the subroutine to have a variable, unknown stopping time, and then we describe spaces whose overlap graph looks like G, but in place of an edge $\{u,v\}$, there is a ladder-like gadget (see Figure 1), where we can imagine getting from u to v by going up one side (representing computation) and that at various "rungs" with various probabilities, crossing to the other side of the ladder and uncomputing. By increasing or decreasing the "weights" ($\{\alpha_t\}_t$ in Theorem 4.1) as the ladder goes up, we can get various variable-time complexities without knowing the total lengths of the ladders.

Our result for general algorithmic composition (Section 1.2) uses the same ideas, but within an algorithm rather than a graph. This can be done by defining a sequence of spaces from the outer algorithm whose overlap graph is a line of length L, and then plugging in ladder-like gadgets like those in Figure 1 wherever there is a subroutine query (see Figure 4).

Comparison with Previous and Concurrent Work: Aside from the special cases of formula evaluation and variable-time search already mentioned, previous work in composing quantum algorithms in non-trivial ways has mainly been restricted to the model of query complexity. *Span programs* can be composed in various nice ways [RŠ12, Rei11b], for example, to evaluate formulas in various gate sets with a quantum query complexity that amortizes nicely – a fact exploited in [CKKD$^+$22] (concurrent with this work) to analyze the quantum query complexity of divide-and-conquer algorithms. It is also rather simple to get a version of Ambainis' variable-time search result that applies only to query complexity using span program composition. Ref. [CJOP20] showed how to extend span program composition to time complexity, also reproducing Ambainis' variable-time search in a time-efficient way using span program composition. The main idea of [CJOP20] allows span programs, also known as *dual adversary solutions*, to maintain time complexity structure where it is given, whereas previously span programs were only used to study query complexity. The techniques of [CJOP20] are the basis for characterizing time complexty in [JZ22], and now in this work as well.

In a concurrent, independent work, Belovs and Yolcu defined the notion of quantum Las Vegas query complexity, and in particular, showed a version of our main theorem (Theorem 1.1) that applies only to query complexity [BY23]. While their results do not apply to time complexity, they work with a significantly more general class of subroutines. While we assume subroutines compute a single bit (which easily extends to any classical function), their results work for subroutines that enact arbitrary state conversion, or even non-unitary maps. Their techniques are quite novel, and distinct from ours.

Subsequent Work: In subsequent work, our algorithmic composition results have been extended to quantum subroutines whose behaviour is described by an arbitrary unitary, rather than subroutines that are assumed to compute a single bit [BJY24]. This improved composition is obtained by a new tool *transducers*, that generalizes multidimensional quantum walks, using the ideas from this work to extend the constructions in [BY23] to capture time complexity. For subroutines that compute deterministic functions with bounded error – i.e. the setting where majority voting can be used to amplify the success probability – [BJY24] also shows how to get composition results like ours, but without the log factor overhead that would come from such majority voting.

1.1 Quantum Walk Composition

Quantum walk search algorithms were first introduced by Szegedy [Sze04] who showed the following. Let P be the transition matrix of a random walk on a graph G with stationary distribution π, and $M \subset V(G)$ a *marked set*, both of which may implicitly depend on some input x – we write, for example M_x when we want to make this dependence explicit. Let $\mathcal{H}$ be a known bound such that whenever $M \neq \emptyset$, we are promised that the hitting time to M – the expected number of steps needed by a classical random walk starting in the distribution π to reach some $u \in M$ – is at most $\mathcal{H}$. Then there is a quantum algorithm that decides, with bounded error, if $M = \emptyset$, which starts by generating an initial state $\sum_u \sqrt{\pi(u)}|u\rangle$, and then makes $\sqrt{\mathcal{H}}$ calls to a subroutine whose main components are: (1) for any $u \in V(G)$, generating the quantum walk state $\sum_v \sqrt{P_{u,v}}|u,v\rangle$, which can be seen as the quantum analogue of taking a random walk step (and are essentially the states $|\psi^G_\star(u)\rangle$ mentioned above); (2) for any $u \in V(G)$, checking if $u \in M$. In early work, the most important parameter in the complexity was $\sqrt{\mathcal{H}}$ – a generic quadratic speedup over the analogous classical algorithm – and it was assumed that other operations, such as generating the quantum walk states had unit cost. In later applications, these operations often had non-trivial costs, impacting the complexity in different ways. Generally S is used to denote the cost of generating $\sum_u \sqrt{\pi(u)}|u\rangle$, U the cost of generating the quantum walk states, and C the cost of checking if $u \in M$. In our work, we will subdivide the task of generating quantum walk states. We suppose (as is common in classical random walk algorithms) that for each $u \in V(G)$, there is a set of labels $L(u)$, for example, $L(u) = [d_u]$, and a mapping $f_u : L(u) \to V$ such that $f_u(i)$ is the i-th neighbour of u. Then the following two subroutines can be used in place of generating quantum walk states (this is not entirely obvious):

1. Sampling: For each $u \in V$, generate $\sum_{i \in L(u)} \sqrt{P_{u,f_u(i)}}|i\rangle$. We assume this can be done in unit cost.

2. Transitions: For $\{u,v\} \in E(G)$ with $f_u(i) = v$ and $f_v(j) = u$, map $|u,i\rangle \mapsto |v,j\rangle$, in cost $T_{u,v}$, which is a random variable on $[\mathsf{T}_{\max}]$.

In all applications that we know of, sampling has at most polylogarithmic cost, and it is the transition that may be expensive.[4] We can also suppose the checking cost is variable, with random variable C_u on $[\mathsf{C}_{\max}]$ representing the cost of checking if u is marked. With this notation in place, the complexity of a Szegedy walk is

$$\mathsf{S} + \sqrt{\mathcal{H}}(\mathsf{T}_{\max} + \mathsf{C}_{\max}). \tag{4}$$

Belovs generalized Szegedy's framework to the *electric network framework* [Bel13] (published in [BCJ$^+$13]), in which the walk may start in any distribution σ. If S_σ is the cost of generating $|\sigma\rangle = \sum_u \sqrt{\sigma(u)}|u\rangle$ (so $\mathsf{S} = \mathsf{S}_\pi$ above), and $\mathcal{C}_{\sigma,\mathsf{M}}$ is an upper bound on a quantity $\mathcal{C}_{\sigma,M}(G)$, to be described shortly, then Belovs exhibited a quantum algorithm for deciding if $M = \emptyset$ with complexity

$$\mathsf{S}_\sigma + \sqrt{\mathcal{C}_{\sigma,\mathsf{M}}}(\mathsf{T}_{\max} + \mathsf{C}_{\max}). \tag{5}$$

Let us say more about this quantity $\mathcal{C}_{\sigma,M}(G)$. If $\mathcal{W}(G)$ is the *total weight* of G (the number of edges when G is unweighted) and $\mathcal{R}_{\sigma,M}(G)$ is the *effective resistance* between σ and M (see Definition 2.4),

---

[4]For example, in algorithms for triangle finding there is often a quantum walk step where the sampling step consists of selecting a new vertex of the input graph (not the graph being walked on) to add to a stored set of vertices, but to complete the transition, it is necessary to find all neighbours of the new vertex in the already stored set.

then $\mathcal{C}_{\sigma,M}(G) = 2\mathcal{W}(G)\mathcal{R}_{\sigma,M}(G)$. This is probably not a very helpful definition, so we mention some special cases that give some intuition. If $\sigma = \pi$ is the stationary distribution, then $\mathcal{C}_{\pi,M}(G)$ is just the hitting time from π to M, and we recover the Szegedy framework. If σ is supported on a single vertex s, then $\mathcal{C}_{s,M}(G)$ is the *commute time* or expected number of steps starting from s, to get to any vertex in M and then return to s [CRR+96] (see [AGJ20] for a proof of the case when $|M| > 1$, which was previously folklore).

Another framework, incomparable to Szegedy's, is the MNRS framework [MNRS11], in which the walk must start in the stationary distribution π, but the cost is:

$$\mathsf{S}_\pi + \frac{1}{\sqrt{\varepsilon}} \left(\frac{1}{\sqrt{\delta}} \mathsf{T}_{\max} + \mathsf{C}_{\max} \right), \tag{6}$$

where δ is the *spectral gap* of P, and ε is a lower bound on $\pi(M) = \sum_{u \in M} \pi(u)$. This framework has different dependence on the cost of transitions and the cost of checking, since the checking subroutine is run less frequently. Since the hitting time is between $\frac{1}{\varepsilon}$ and $\frac{1}{\varepsilon\delta}$, this may be better or worse than a Szegedy walk, depending on the relative values of $\mathsf{T}_{\max}$ and $\mathsf{C}_{\max}$. In [AGJ20], the electric network framework was extended to be able to not only detect if $M \neq \emptyset$, but to *find* an element of M as well, and this extension was shown to also include the MNRS framework as a special case – simulating an MNRS walk with a walk whose hitting time is $1/\varepsilon$, and update cost is $\mathsf{T}_{\max}/\sqrt{\delta}$. In this paper, we do not consider finding a marked vertex, but only detecting if there exists a marked vertex.

Finally, a result of Dohotaru and Høyer [DH17] shows how to achieve an optimal number of both checking and update steps in the special case where $|M| \leq 1$, obtaining complexity:

$$\mathsf{S}_\pi + \sqrt{\mathcal{H}}\mathsf{T}_{\max} + \frac{1}{\sqrt{\varepsilon}}\mathsf{C}_{\max}. \tag{7}$$

Previously, quantum walk applications have used various tricks to handle the case when the transitions can have different complexities. The MNRS framework is already one means of handling the case when the checking cost is much larger than the update cost. Nested quantum walks [JKM12, CJKM13], where the checking or update is implemented by another quantum walk subroutine, can be seen as one large quantum walk with multiple types of update operations with different costs. Variable-time search has also been used *within* the checking subroutine of a quantum walk [LG14]. The most common method for handling variability in the transition or checking cost has been through the use of tail bounds, but this only works when the cost is highly concentrated around the average. In the special case of MNRS quantum walks on *Johnson graphs*, where subroutines are restricted to being in the framework of *extended learning graphs* and the measure of complexity is query complexity rather than time complexity, [CLM20] showed how to replace $\mathsf{T}_{\max}$ and $\mathsf{C}_{\max}$ with ℓ_2-averages, also getting an average in the setup cost S_π, in a way that is specific to Johnson graphs. While Johnson graphs are quite a specific case, almost all known MNRS quantum walk algorithms work on Johnson graphs.

The recent work of [JZ22], on whose techniques our results are heavily based, achieves an improvement on the electric network framework (and by extension, its special cases) as follows. Suppose the transition algorithm, rather than having random stopping times, stops after a known time $\mathsf{T}_{u,v} \leq \mathsf{T}_{\max}$ on input (u, i), where $f_u(i) = v$. The value $\mathsf{T}_{u,v}$ must be computable from (u, i), but also from (v, j), where $f_v(j) = u$. Suppose checking can be done in unit cost. Then there is a quantum algorithm that decides if $M = \emptyset$ in complexity

$$\mathsf{S}_\sigma + \sqrt{\mathcal{C}^{\mathsf{T}}_{\sigma,\mathsf{M}}}, \tag{8}$$

where $\mathcal{C}^{\mathsf{T}}_{\sigma,\mathsf{M}}$ is an upper bound on $\mathcal{C}_{\sigma,M}(G^{\mathsf{T}})$, where G^{T} is the graph G where every edge $\{u, v\}$ has been replaced by a path of length $\mathsf{T}_{u,v}$. This is already a huge improvement over having worst case $\mathsf{T}_{\max}$ complexity. The results presented in this section are based on a generalization of the results in [JZ22], and a careful analysis of their implications.

Variable-time Quantum Walks: We now state our results. For simplicity, we begin by stating our results in the case that checking has unit cost. This is without loss of generality, since we can always add extra edges to the graph from a vertex u to a new vertex that includes a bit indicating if

$u \in M$, although this change to the graph changes parameters such as hitting time. In that case, we can simply replace $\mathsf{T}_{\max}$ with a weighted ℓ_2 average, weighted by $\pi(u)P_{u,v}$, which is the proportion of time spent traversing edge (u, v) in an infinitely long random walk starting from any distribution, to get complexity (see Corollary 4.2):

$$\mathsf{S}_\sigma + \sqrt{\mathcal{C}_{\sigma,\mathsf{M}}} \sqrt{\sum_{(u,v) \in \overrightarrow{E}(G)} \pi(u)P_{u,v}\mathbb{E}[T_{u,v}^2]} \log^{1.5} \mathsf{T}_{\max}. \tag{9}$$

Above, $\overrightarrow{E}(G)$ is the set of all edges, assigned some arbitrary orientation (see Definition 2.1). If we further assume that the values $\mathbb{E}[T_{u,v}]$ are computable in some strong sense (see Corollary 4.2), then there is a quantum algorithm that decides if $M = \emptyset$ with bounded error in complexity:

$$\mathsf{S}_\sigma + \sqrt{\mathcal{C}_{\sigma,\mathsf{M}}} \sqrt{\sum_{(u,v) \in \overrightarrow{E}(G)} \pi(u)P_{u,v}\mathbb{E}[T_{u,v}]^2} \log \mathsf{T}_{\max}. \tag{10}$$

Aside from the log factor improvement, (10) is better when the variances of $T_{u,v}$ are large, as $\mathrm{Var}(T_{u,v}) = \mathbb{E}[T_{u,v}^2] - \mathbb{E}[T_{u,v}]^2 \geq 0$. The expressions in (9) and (10) should be compared with the electric network framework, (5) (or (4) in the special case when $\sigma = \pi$, and so $\mathcal{C}_{\sigma,\mathsf{M}} = \mathcal{H}$), and the graph composition in (8), although comparison with (8) is more difficult. These results follow from a more general statement (Theorem 4.1) of which a version of (8) where the transition subroutine may have variable stopping times, and the values $\mathbb{E}[T_{u,v}]$ need not be known in advance, also follows as a special case. As we discuss further below, in some cases, the expressions in (9) and (10) are better, whereas in some cases the expression in (8) is better, but analysis of (9) and (10) may be easier, since we can work with the original graph G, and need not analyze a modified graph G^T.

As stated above, the checking cost can be made part of the update cost by adding a new edge to each vertex u, whose other endpoint encodes u and a bit indicating if $u \in M$, and whose transition cost is C_u. For comparison with previous work, we work out what impact this might have on the complexity. For comparison with the MNRS framework, we restrict our attention to initial distribution $\sigma = \pi$. Then we can decide if $M = \emptyset$ in complexity (see Corollary 4.3):

$$\mathsf{S}_\pi + \frac{1}{\sqrt{\varepsilon}} \left(\frac{1}{\sqrt{\delta}} \sqrt{\sum_{(u,v) \in \overrightarrow{E}} \pi(u)P_{u,v}\mathbb{E}[T_{u,v}^2]} + \sqrt{\sum_{u \in V(G)} \pi(u)\mathbb{E}[C_u^2]} \right) \log^{.5} \frac{1}{\pi_{\min}} \log^{1.5}(\mathsf{T}_{\max}\mathsf{C}_{\max}), \tag{11}$$

where $\pi_{\min} = \min_u \pi(u)$. Compared to the complexity achieved by the MNRS framework in (6), ignoring log factors, we have replaced $\mathsf{T}_{\max}$ and $\mathsf{C}_{\max}$ with weighted ℓ_2-averages. This also generalizes the results of [CLM20] from Johnson graphs to general graphs, learning graphs to arbitrary subroutines, and query complexity to time complexity. As before, if the values $\mathbb{E}[C_u]$ and $\mathbb{E}[T_{u,v}]$ are computable in some strong sense, then we can replace $\mathbb{E}[T_{u,v}^2]$ with $\mathbb{E}[T_{u,v}]^2$ and $\mathbb{E}[C_u^2]$ with $\mathbb{E}[C_u]^2$, and shave off a $\sqrt{\log}$ factor. For more general initial distributions, see Corollary 4.2.

In the simpler case where $|M| \leq 1$, we can once again consider a more general starting distribution σ. In that case, we get a variable-time analogue of the results of Dohotaru and Høyer in (7), but even more general. Specifically, let τ be *any* distribution on $V(G)$ such that we are promised that if $M = \{m\} \neq \emptyset$, then $\tau(m) \geq \varepsilon$. Then we can decide if $M = \emptyset$ in complexity (see Corollary 4.2):

$$\mathsf{S}_\sigma + \left(\sqrt{\mathcal{C}_{\sigma,\mathsf{M}}} \sqrt{\sum_{(u,v) \in \overrightarrow{E}(G)} \pi(u)P_{u,v}\mathbb{E}[T_{u,v}^2]} + \frac{1}{\sqrt{\varepsilon}} \sum_{u \in V(G)} \tau(u)\mathbb{E}[C_u^2] \right) \log^{1.5}(\mathsf{T}_{\max}\mathsf{C}_{\max}).$$

When $\sigma = \tau = \pi$, so $\mathcal{C}_{\sigma,\mathsf{M}} = \mathcal{H}$, we get a variable-time version of the Dohotaru-Høyer result in (7). As before, if the values $\mathbb{E}[C_u]$ and $\mathbb{E}[T_{u,v}]$ are computable in some strong sense, then we can replace $\mathbb{E}[T_{u,v}^2]$ with $\mathbb{E}[T_{u,v}]^2$ and $\mathbb{E}[C_u^2]$ with $\mathbb{E}[C_u]^2$, and shave off a $\sqrt{\log}$ factor.

Variable-time Search: As a special case of our variable-time quantum walk results, we recover something similar to Ambainis variable-time search result, but allowing the stopping time on input i to be a random variable T_i, similar to the setting Ambainis considers separately in variable-time amplitude amplification. We restate our result (mentioned already in (3)) with slightly more detail. Let $M \subset [n]$ be some marked set, and let π be any distribution on $[n]$ such that if $M \neq \emptyset$, $\pi(M) \geq \varepsilon$ for some known bound ε. Leting T_i be the stopping time on input $i \in [n]$ of a subroutine that checks if $i \in M$, we can decide if $M = \emptyset$ with bounded error in the following complexities, assuming the values $\mathbb{E}[T_i]$ are unknown (first expression), or known in a certain strong sense (see Corollary 4.4):

$$(1) \text{ unknown: } \sqrt{\frac{1}{\varepsilon} \sum_{i \in [n]} \pi(i)\mathbb{E}[T_i^2]} \log^{1.5} \mathsf{T}_{\max} \qquad \text{known: } \sqrt{\frac{1}{\varepsilon} \sum_{i \in [n]} \pi(i)\mathbb{E}[T_i]^2} \log \mathsf{T}_{\max}.$$

The case when the values $\mathbb{E}[T_i]$ are known (computable from i) follows from Ambainis' variable-time search result, since in that case, on input i we can simply stop after $\mathsf{T}_i = 10\mathbb{E}[T_i]$ steps (repeating $O(\log n)$ times and taking a majority to reduce the error). In the case when these values are unknown, this would probably also follow from a combination of variable-time search and variable-time amplitude amplification. What is perhaps more interesting is that we obtain alternative complexities for variable-time search, by using different parameters. In the case of unknown expected stopping times, we can also choose either of the following two complexities (neglecting log factors).

$$(2) \quad \sqrt{\frac{\sum_{i \in [n]} \pi(i)\mathbb{E}[T_i]}{\min_{x:M_x \neq \emptyset} \sum_{i \in M} \frac{\pi(i)}{\mathbb{E}[T_i]}}} \quad \text{or} \quad (3) \quad \frac{1}{\sqrt{\min_{x:M_x \neq \emptyset} \sum_{i \in M} \frac{\pi(i)}{\mathbb{E}[T_i^2]}}}.$$

Above, we minimize over all allowed (non-empty) marked sets, assuming the marked set depends on some implicit input. Table 1 shows the three complexities we can obtain, and argues that there are settings where each of the three is the smallest complexity. We note that we can achieve analogous complexities to each of (1), (2) and (3) for more general quantum walk algorithms. The generalization of (1) is the results we presented above. The generalization of (2) is essentially what was shown in [JZ22] ((8)). We have not explored the generalization of (3). Other settings of the parameters in our main quantum walk theorem, Theorem 4.1, could lead to further alternatives, although it is unclear if these would be useful.

1.2 Quantum Algorithm Composition

We use a similar technique to compose an arbitrary quantum query algorithm with a variable-time subroutine. That is, fix functions $f : \{0,1\}^n \to \{0,1\}$ and $\{g_i : \{0,1\}^m \to \{0,1\}\}_{i \in [n]}$, and define $f \circ g : \{0,1\}^m \to \{0,1\}^n$ by $f \circ g(x) = f(g(x)) = f(g_1(x), \ldots, g_n(x))$. Suppose we have a quantum algorithm that decides f with bounded error in L time steps and Q queries to the input to f, and a quantum algorithm that decides $g(i, x) = g_i(x)$ in $\mathsf{T}_{\max}$ steps with some sufficiently small error. Then by composing these algorithms in a naive way, we get a bounded error quantum algorithm for $f \circ g$ with complexity:

$$\widetilde{O}(\mathsf{L} + \mathsf{Q} \cdot \mathsf{T}_{\max}).$$

As in the case of quantum walk algorithms, we show how we can do significantly better if the inner algorithm's running time varies significantly in the input i, as well as its internal randomness. We show the following (see Theorem 5.1):

Theorem 1.1 (Informal). *Let $\bar{\mathsf{q}}_i$ be the average (over all queries made by the outer algorithm) squared amplitude on querying index $i \in [n]$ (see (52)). Let ϵ_i be the error of the inner subroutine on input i, and T_i the stopping time of the inner subroutine on input i, which is a random variable. Let $\mathsf{T}_{\mathrm{avg}}$ be an upper bound such that:*

$$\sum_{i \in [n]} \bar{\mathsf{q}}_i \mathbb{E}[T_i] \leq \mathsf{T}_{\mathrm{avg}}$$

and suppose the subroutine's errors satisfy the following condition:

$$\epsilon_{\mathrm{avg}} := \sum_{i \in [n]} \bar{\mathsf{q}}_i \epsilon_i \leq \frac{1}{\mathsf{Q}(\mathsf{L} + \mathsf{Q} \cdot \mathsf{T}_{\mathrm{avg}})}.$$

Then there is a quantum algorithm that computes $f \circ g$ with bounded error in complexity

$$\widetilde{O}\left(\mathsf{L} + \mathsf{Q} \cdot \mathsf{T}_{\mathrm{avg}}\right).$$

We compare this with the naive upper bound of $\mathsf{L} + \mathsf{Q} \cdot \mathsf{T}_{\mathrm{max}}$. Since $\mathsf{T}_{\mathrm{max}}$ is the *maximum* running time of any subroutine, our new running time can be significantly better when $\mathbb{E}[T_i]$ varies widely over different i. If the values $\mathbb{E}[T_i]$ are *known* (i.e. computable from i), we can always truncate the inner algorithm, on input i, at some time $\mathsf{T}_i = 10\mathbb{E}[T_i]$, and then by Markov's inequality, the probability that the algorithm has output correctly is not impacted by more than a constant. In that case, Theorem 1.1 does better than naive composition if and only if the values $\mathbb{E}[T_i]$ vary.

We remark that our error dependence is not optimal. If $\epsilon_i = \epsilon$ for all i, then a subroutine called Q times needs $\epsilon \leq 1/\mathsf{Q}^2$ to achieve overall bounded error. In the usual way of calling a subroutine, there is no reason that the error bound on the subroutine queries should scale with anything other than the number of times the subroutine is called. In particular, it need not scale with the number of *other* steps of the algorithm (L) or the number of steps taken by a given call of the subroutine, since ϵ is already totalled over all these steps. There is also no reason to think that the error should scale with these things when we combine a subroutine and outer algorithm in the way we do in the proof of Theorem 5.1, so we suspect there is something here that can be improved. In practice it does not make much difference, since the error of a bounded error subroutine can be made as small as any ϵ at a multiplication cost of $\log \frac{1}{\epsilon}$, so the overhead is at most logarithmic. Still, it would be nice to understand, conceptually, how different errors add up, and at the moment it seems there is something missing from the picture. We leave this for future work.

1.3 Outline

The rest of this paper is organized as follows. In Section 2, we give preliminaries on graph theory and random walks (Section 2.1) as well as phase estimation algorithms, on which all of our algorithms are based (Section 2.2). In Section 3, we introduce our model of variable-time quantum algorithms, which is a slight generalization of the models of [Amb10] and [Amb12]. We define certain states and subspaces from such algorithms, and prove a number of properties of these that will be used in our composition results. In Section 4, we state and prove our results for composing variable-time subroutines in quantum walks, and in Section 5, we prove Theorem 1.1, by showing how to compose a variable-time subroutine into an arbitrary quantum algorithm.

2 Preliminaries

2.1 Graph Theory

In this section, we define graph theoretic concepts and notation, used in our results on quantum walks.

Definition 2.1 (Network). *A* network *is a weighted graph G with an (undirected) edge set $E(G)$, vertex set $V(G)$, and some weight function $\mathsf{w} : E(G) \to \mathbb{R}_{>0}$. Since edges are undirected, we can equivalently describe the edges by some set $\overrightarrow{E}(G)$ such that for all $\{u, v\} \in E(G)$, exactly one of (u, v) or (v, u) is in $\overrightarrow{E}(G)$. The choice of edge directions is arbitrary. Then we can view the weights as a function $\mathsf{w} : \overrightarrow{E}(G) \to \mathbb{R}_{>0}$, and for all $(u, v) \in \overrightarrow{E}(G)$, define $\mathsf{w}_{v,u} = \mathsf{w}_{u,v}$. For convenience, we will define $\mathsf{w}_{u,v} = 0$ for every pair of vertices such that $\{u, v\} \notin E(G)$. The* total weight *of G is*

$$\mathcal{W}(G) := \sum_{e \in \overrightarrow{E}(G)} \mathsf{w}_e.$$

For an implicit network G, and $u \in V(G)$, we will let $\Gamma(u)$ denote the *neighbourhood* of u:

$$\Gamma(u) := \{v \in V(G) : \{u, v\} \in E(G)\}.$$

We use the following notation for *the out- and in-neighbourhoods* of $u \in V(G)$:

$$\begin{aligned}
\Gamma^+(u) &:= \{v \in \Gamma(u) : (u, v) \in \overrightarrow{E}(G)\} \\
\Gamma^-(u) &:= \{v \in \Gamma(u) : (v, u) \in \overrightarrow{E}(G)\}.
\end{aligned} \tag{12}$$

Random Walks: A *Markov process* on a finite set V is specified by a stochastic *transition matrix* $P \in \mathbb{R}^{V \times V}$, where we interpret $P_{u,v}$ as the probability of moving to the state v when currently in the state u. We can define a Markov process P from a weighted graph G by letting:

$$P_{u,v} = \frac{\mathsf{w}_{u,v}}{\mathsf{w}_u}, \text{ where } \mathsf{w}_u := \sum_{v \in \Gamma(u)} \mathsf{w}_{u,v}.$$

Next, define

$$\forall u \in V(G), \ \pi(u) := \frac{\mathsf{w}_u}{2\mathcal{W}(G)}.$$

Then it is easy to check that $\sum_{u \in V(G)} \pi(u) = 1$. Furthermore, if G is connected, π is the unique left-1-eigenvector of P, called the *stationary distribution*. We note that for any $u, v \in V(G)$,

$$\pi(u) P_{u,v} = \frac{\mathsf{w}_u}{2\mathcal{W}(G)} \frac{\mathsf{w}_{u,v}}{\mathsf{w}_u} = \frac{\mathsf{w}_{u,v}}{2\mathcal{W}(G)} = \frac{\mathsf{w}_{v,u}}{2\mathcal{W}(G)} = \pi(v) P_{v,u}. \tag{13}$$

When P satisfies this condition, called *detailed balance*, we say that P is *reversible*. We have just seen that any random walk on an undirected weighted graph is reversible, and it is also true that any reversible Markov process can be modelled as a random walk on an undirected weighted graph, using edge weights $\mathsf{w}_{u,v} = \pi(u) P_{u,v}$.

Finally, we let δ denote the *spectral gap* of P, which is the smallest non-zero eigenvalue of $I - P$. For intuition, it is useful to know that $\frac{1}{\delta}$ is within a $\log |V(G)|$ factor of the mixing time of P (see, for example, [LWP09]).

Accessing G: In computations involving a (classical) random walk on a graph G, it is usually assumed that for any $u \in V(G)$, it is possible to sample a neighbour $v \in \Gamma(u)$ according to the distribution given by the u-th row of P. It is standard to assume this is broken into two steps: (1) sampling some $i \in [d_u]$, where $d_u := |\Gamma(u)|$ is the degree of u, and (2) computing the i-th neighbour of u. That is, we assume that for each $u \in V(G)$, there is an efficiently computable function $f_u : [d_u] \to V(G)$ such that $\mathrm{im}(f_u) = \Gamma(u)$, and we call $f_u(i)$ the *i-th neighbour of u*. In the quantum case (see Definition 2.2 below), we assume that sampling (1) can be done coherently, and we use a reversible version of the map $(u, i) \mapsto f_u(i)$. We will also find it convenient to suppose the indices i of the neighbours of u come from some more general set $L(u)$, which may equal $[d_u]$, or some other convenient set, which we call the *edge labels of u*. It is possible to have $|L(u)| > |\Gamma(u)| = d_u$, meaning that some elements of $L(u)$ do not label an edge adjacent to u (these labels should be sampled with probability 0). We assume we have a partition of $L(u)$ into disjoint $L^+(u)$ and $L^-(u)$ such that:

$$\begin{aligned}
L^+(u) &\supseteq \{i \in L(u) : (u, f_u(i)) \in \overrightarrow{E}(G)\} = \{i \in L(u) : f_u(i) \in \Gamma^+(u)\} \\
L^-(u) &\supseteq \{i \in L(u) : (f_u(i), u) \in \overrightarrow{E}(G)\} = \{i \in L(u) : f_u(i) \in \Gamma^-(u)\}.
\end{aligned}$$

Note that for any $(u, v) \in \overrightarrow{E}(G)$, with $i = f_u^{-1}(v)$ and $j = f_v^{-1}(u)$, any of (u, v), (v, u), (u, i), or (v, j) fully specify the edge. Thus, it will be convenient to denote the weight of the edge using any of the alternatives:

$$\mathsf{w}_{u,v} = \mathsf{w}_{v,u} = \mathsf{w}_{u,i} = \mathsf{w}_{v,j}.$$

For any $i \in L(u)$, we set $\mathsf{w}_{u,i} = 0$ if and only if $\{u, f_u(i)\} \notin E(G)$.

Definition 2.2 (Quantum Walk access to G). *For each $u \in V(G)$, let $L(u) = L^+(u) \cup L^-(u)$ be some finite set of edge labels, and $f_u : L(u) \to V(G)$ a function such that $\Gamma(u) \subseteq \mathrm{im}(f_u)$. A quantum algorithm has* quantum walk access *to G if it has access to the following subroutines:*

- *A subroutine that generates quantum samples from $L(u)$ by implementing a unitary $U_\star$ in cost $\mathsf{A}_\star$ that acts as:*

$$U_\star |u, 0\rangle \propto \sum_{i \in L^+(u)} \sqrt{\mathsf{w}_{u,i}} |u, i\rangle - \sum_{i \in L^-(u)} \sqrt{\mathsf{w}_{u,i}} |u, i\rangle =: |\psi_\star^G(u)\rangle.$$

- *A subroutine that implements the* transition map

$$|u, i\rangle \mapsto |v, j\rangle$$

(possibly with some error) where $i = f_u^{-1}(v)$ and $j = f_v^{-1}(u)$, with costs $\{\mathsf{T}_{u,i} = \mathsf{T}_{u,v}\}_{(u,v) \in \vec{E}(G)}$.

- *Query access to the total vertex weights $\mathsf{w}_u = \sum_{v \in \Gamma(u)} \mathsf{w}_{u,v}$.*

We call $\{\mathsf{T}_e\}_{e \in \vec{E}(G)}$ the set of transition costs *and $\mathsf{A}_\star$ the cost of generating the star states.*

Flows and Resistances: Electric networks of resistors can be modelled as weighted graphs, where the weights represent *conductances* (so their inverses represent *resistances*). This beautiful connection was expounded by Doyle and Snell [DS84], or see [LWP09] for a modern exposition. This connection inspires the following definitions.

Definition 2.3 (Flow, Circulation). *A flow on a network G is a real-valued function $\theta : \vec{E}(G) \to \mathbb{R}$, extended to edges in both directions by $\theta(u, v) = -\theta(v, u)$ for all $(u, v) \in \vec{E}(G)$. For any flow θ on G, and vertex $u \in V(G)$, we define $\theta(u) = \sum_{v \in \Gamma(u)} \theta(u, v)$ as the flow coming out of u. If $\theta(u) = 0$, we say flow is conserved at u. If flow is conserved at every vertex, we call θ a* circulation. *If $\theta(u) > 0$, we call u a* source, *and if $\theta(u) < 0$ we call u a* sink. *The set of sources and sinks is called the* boundary *of θ. A flow with unique source s and unique sink t is called an st-flow. If additionally $\theta(s) = 1$, we call θ a* unit st-flow. *The* energy *of θ is*

$$\mathcal{E}(\theta) := \sum_{(u,v) \in \vec{E}(G)} \frac{\theta(u, v)^2}{\mathsf{w}_{u,v}}.$$

Definition 2.4 (Effective Resistance). *Let σ and τ be distributions on $V(G)$. We define the* effective resistance from σ to τ:

$$\mathcal{R}_{\sigma,\tau}(G) := \min\{\mathcal{E}(\theta) : \forall u \in V(G), \theta(u) = \sigma(u) - \tau(u)\}.$$

When σ and τ have disjoint support, this is the minimum energy of a flow whose sources are exactly $\mathrm{supp}(\sigma)$, with $\theta(u) = \sigma(u)$, and whose sinks are exactly $\mathrm{supp}(\tau)$, with $\theta(u) = -\tau(u)$. In the special case when $\mathrm{supp}(\sigma) = \{s\}$ and $\mathrm{supp}(\tau) = \{t\}$, then θ ranges over all unit st-flows, and we denote $\mathcal{R}_{\sigma,\tau}(G)$ by $\mathcal{R}_{s,t}(G)$. Finally, for any $M \subseteq V(G) \setminus \mathrm{supp}(\sigma)$, we define:

$$\mathcal{R}_{\sigma,M}(G) := \min\{\mathcal{R}_{\sigma,\tau}(G) : \mathrm{supp}(\tau) \subseteq M\}.$$

That is, $\mathcal{R}_{\sigma,M}(G)$ is the minimum energy of a flow from σ to M (where any distribution on M is allowed).

There is always a unique θ from σ to τ (or σ to M) that achieves energy $\mathcal{R}_{\sigma,\tau}(G)$ (or $\mathcal{R}_{\sigma,M}(G)$): If two σ-τ flows achieve energy $\mathcal{E}$, then there is some affine combination of them that achieves energy strictly smaller than $\mathcal{E}$.

These concepts are related to random walks as follows.

Theorem 2.5 ([CRR⁺96]). *For any $s, t \in V(G)$, let $\mathcal{C}_{s,t}(G)$ be the* commute time *from s to t, which is the expected number of steps taken in a random walk starting at s, before the walker reaches t, and then returns to s. Then $\mathcal{C}_{s,t}(G) = 2\mathcal{W}(G)\mathcal{R}_{s,t}(G)$.*

Theorem 2.6 (Folklore, or see [AGJ20]). *For any set $M \subset V(G)$ and $s \in V(G)$, let $\mathcal{C}_{s,M}(G)$ be the* commute time *from s to M, which is the expected number of steps taken in a random walk starting at s, before the walker reaches M, and then returns to s. Then $\mathcal{C}_{s,M}(G) = 2\mathcal{W}(G)\mathcal{R}_{s,M}(G)$.*

Theorem 2.7 ([Bel13]). *For any set $M \subset V(G)$, let $\mathcal{H}_{\pi,M}(G)$ be the* hitting time *from the stationary distribution π to M, which is the expected number of steps taken in a random walk starting in a vertex sampled according to π, before the walker reaches M. Then $\mathcal{H}_{\pi,M}(G) = 2\mathcal{W}(G)\mathcal{R}_{\pi,M}(G)$.*

Motivated by the three theorems above, we define

$$\mathcal{C}_{\sigma,\tau}(G) := 2\mathcal{W}(G)\mathcal{R}_{\sigma,\tau}(G) \text{ and } \mathcal{C}_{\sigma,M}(G) := 2\mathcal{W}(G)\mathcal{R}_{\sigma,M}(G). \tag{14}$$

2.2 Phase Estimation Algorithms

We will use the precise notion of a *phase estimation algorithm* from [JZ22], which is based on the phase estimation technique of [Kit96]. We refer the reader to [JZ22, Section 3.1] for intuition about why such an algorithm works as stated here.

Definition 2.8 (Parameters of a Phase Estimation Algorithm). *For an implicit input $x \in \{0,1\}^*$, fix a finite-dimensional complex inner product space H, a unit vector $|\psi_0\rangle \in H$, and sets of vectors $\Psi^{\mathcal{A}}, \Psi^{\mathcal{B}} \subset H$. We further assume that $|\psi_0\rangle$ is orthogonal to every vector in $\Psi^{\mathcal{B}}$. Let $\Pi_{\mathcal{A}}$ be the orthogonal projector onto $\mathcal{A} = \mathrm{span}\{\Psi^{\mathcal{A}}\}$, and similarly for $\Pi_{\mathcal{B}}$.*

Let $U_{\mathcal{AB}} = (2\Pi_{\mathcal{A}} - I)(2\Pi_{\mathcal{B}} - I)$. The algorithm defined by $(H, |\psi_0\rangle, \Psi^{\mathcal{A}}, \Psi^{\mathcal{B}})$ performs phase estimation of $U_{\mathcal{AB}}$ on initial state $|\psi_0\rangle$, to sufficient precision that by measuring the phase register and checking if the output is 0, we can distinguish between a *negative case* and a *positive case*.

Definition 2.9 (Negative Witness). *A δ-negative witness for $(H, |\psi_0\rangle, \Psi^{\mathcal{A}}, \Psi^{\mathcal{B}})$ is a pair of vectors $|w_{\mathcal{A}}\rangle, |w_{\mathcal{B}}\rangle \in H$ such that $\|(I - \Pi_{\mathcal{A}})|w_{\mathcal{A}}\rangle\|^2 \leq \delta$, $\|(I - \Pi_{\mathcal{B}})|w_{\mathcal{B}}\rangle\|^2 \leq \delta$, and $|\psi_0\rangle = |w_{\mathcal{A}}\rangle + |w_{\mathcal{B}}\rangle$.*

Definition 2.10 (Positive Witness). *A δ-positive witness for $(H, |\psi_0\rangle, \Psi^{\mathcal{A}}, \Psi^{\mathcal{B}})$ is a vector $|w\rangle \in H$ such that $\langle\psi_0|w\rangle \neq 0$ and $|w\rangle$ is almost orthogonal to all $|\psi\rangle \in \Psi^{\mathcal{A}} \cup \Psi^{\mathcal{B}}$, in the sense that $\|\Pi_{\mathcal{A}}|w\rangle\|^2 \leq \delta \||w\rangle\|^2$ and $\|\Pi_{\mathcal{B}}|w\rangle\|^2 \leq \delta \||w\rangle\|^2$.[5]*

Theorem 2.11 ([JZ22]). *Fix $(H, |\psi_0\rangle, \Psi^{\mathcal{A}}, \Psi^{\mathcal{B}})$ as in Definition 2.8. Suppose we can generate the state $|\psi_0\rangle$ in cost S, and implement $U_{\mathcal{AB}} = (2\Pi_{\mathcal{A}} - I)(2\Pi_{\mathcal{B}} - I)$ in cost A.*
Let $c_+ \in [1, 50]$ be some constant, and let $\mathcal{C}_- \geq 1$ be a positive real number that may scale with $|x|$. Let δ and δ' be positive real parameters such that

$$\delta \leq \frac{1}{(8c_+)^3\pi^8\mathcal{C}_-} \quad and \quad \delta' \leq \frac{3}{4}\frac{1}{\pi^4 c_+}.$$

Suppose we are guaranteed that exactly one of the following holds:

Positive Condition: *There is a δ-positive witness $|w\rangle$ s.t. $\frac{|\langle w|\psi_0\rangle|^2}{\||w\rangle\|^2} \geq \frac{1}{c_+}$.*

Negative Condition: *There is a δ'-negative witness $|w_{\mathcal{A}}\rangle, |w_{\mathcal{B}}\rangle$ s.t. $\||w_{\mathcal{A}}\rangle\|^2 \leq \mathcal{C}_-$.*

Then there is a quantum algorithm that distinguishes these two cases with bounded error in cost

$$O\left(\mathsf{S} + \sqrt{\mathcal{C}_-}\mathsf{A}\right).$$

---

[5] We note that for technical reasons, positive witness error is defined multiplicatively (relative error), whereas negative witness error is defined additively.

3 Variable-Time Subroutines

We formally define a *variable-time subroutine*, which is an algorithm whose runtime is a random variable, which may also depend on some input i, assumed to be called in the context of some outer algorithm. Formally, a variable-time subroutine is a sequence of unitaries $U_1, \ldots, U_\mathsf{T}$, for some $\mathsf{T} = \mathsf{T}_{\max}$, acting on the space

$$H_\mathcal{I} \otimes H_\mathcal{A} \otimes H_\mathcal{Z} = \operatorname{span}\{|i\rangle|a\rangle|z\rangle : i \in \mathcal{I}, a \in \mathcal{A}, z \in \mathcal{Z}\}$$

for finite sets $\mathcal{I}$, representing inputs to the subroutine; $\mathcal{A}$, representing answers the subroutine may output; and $\mathcal{Z}$, representing states of the algorithm's workspace. Both $\mathcal{A}$ and $\mathcal{Z}$ can be assumed to be initialized to $|0\rangle$. The unitaries are controlled on $H_\mathcal{I}$, so we can express $U_t = \sum_{i \in \mathcal{I}} |i\rangle\langle i| \otimes U_t^i$ for some unitaries U_t^i. In other words, the input register is read only. Following [Amb12], we make the following assumptions.

1. There are subspaces

$$\{0\} = H_0 \subseteq H_1 \subseteq \cdots \subseteq H_\mathsf{T} = H_\mathcal{Z}$$

   such that U_t leaves H_{t-1} invariant, so any part of the algorithm in H_{t-1} after U_{t-1} has been applied, does not get changed any more. Letting $\Pi_{\leq t-1}$ denote the orthogonal projector onto H_{t-1}, we can express this mathematically as:

$$U_t \Pi_{\leq t-1} = \Pi_{\leq t-1}. \tag{15}$$

   For any t, we assume we can measure a bit indicating if we are in H_t in unit cost. [6]

2. We can implement $\sum_{t=1}^\mathsf{T} |t\rangle\langle t| \otimes U_t$ in unit cost.

We say that t is a potential stopping time if $\overline{H}_t := H_t \cap H_{t-1}^\perp \neq \{0\}$, so $H_\mathcal{Z} := \overline{H}_1 \oplus \cdots \oplus \overline{H}_\mathsf{T}$. We can imagine an algorithm that, for every potential stopping time t, after applying U_t, measures to see if the algorithm is in the space H_t, and if so, the algorithm is done and the answer register can also be measured, and if not, the algorithm continues. Since the later unitaries leave H_t invariant, these measurements do not affect the computation.

Let $\{|z\rangle : z \in \mathcal{Z}_t\}$ be an orthonormal basis for $\overline{H}_t$, for some disjoint sets $\mathcal{Z} = \bigcup_t \mathcal{Z}_t$. Note that $\mathcal{Z}_t \neq \emptyset$ precisely when t is a potential stopping time. Since we assume we can measure if the algorithm is in each space, we assume we can generate each of these bases.

For $t \in [\mathsf{T}]$, we let $\bar{p}_i(t)$ be the probability of stopping at time t (i.e. measuring "1" right after U_t is applied) on input i. Then

$$p_i(t) = \bar{p}_i(0) + \cdots + \bar{p}_i(t-1) \tag{16}$$

is the probability that the algorithm halts some time before U_t is applied. We let T_i be the random variable such that $\Pr[T_i = t] = \bar{p}_i(t)$, so T_i is the stopping time on input i.

Suppose the algorithm is meant to compute some function $g : \mathcal{I} \to \mathcal{A}$. Then if we let Π_t be the orthogonal projector onto $\overline{H}_t$, and Ξ_a the orthogonal projector onto a in the answer register, the probability that the algorithm outputs an incorrect answer on input i, given that it stops at time t, is:

$$\epsilon_i^t := \frac{1}{\bar{p}_i(t)} \left\| ((I - \Xi_{g(i)}) \otimes \Pi_t) U_t \ldots U_1 |i, 0, 0\rangle \right\|^2. \tag{17}$$

Reversible Variable-Time Subroutines: So far this is just a generalization of Ambainis' notion of a variable stopping time algorithm in [Amb12] to allowing the output to be more than just a bit. We will extend it to what we call *Reversible Variable-Time Subroutines*. Building on the above notation, suppose what we are *actually* interested in is computing $|i\rangle \mapsto A|i\rangle$, for some isometry A. A simple example we will consider (Section 5) is $A|i\rangle = (-1)^{g(i)}|i\rangle$ for some $g : [n] \to \{0, 1\}$. In the other example we will see, in Section 4, $A|u, i\rangle = |v, j\rangle$ for vertices $(u, v) \in \overrightarrow{E}(G)$, with $f_u(i) = v$ and

---

[6] For example, imagine having T single-qubit registers, $F_1, \ldots, F_\mathsf{T}$, where we set a 1 in F_t after U_t has been applied to indicate that the algorithm is ready to halt, and then all subsequent unitaries are controlled on no 1s yet being set.

$f_v(j) = u$ (see Definition 2.2). Whatever the case, we will suppose that computing A can be reduced to computing some auxiliary information $g(i) \in \mathcal{A}$, for which we have a variable-time subroutine. To this end, we say that a variable-time subroutine *reversibly computes* A if it computes $g(i)$, and satisfies the following additional assumptions:

1. For $t \in \{1, \ldots, \mathsf{T}\}$, letting
$$\tilde{U}_t = \sum_{i \in \mathcal{I}} A|i\rangle\langle i|A^\dagger \otimes U_t^i,$$
the unitary $\sum_{t=1}^{\mathsf{T}} |t\rangle\langle t| \otimes \tilde{U}_t$ can be implemented in unit cost.

2. There exists a unitary A' of the following form that can be implemented in unit cost:
$$A' = \sum_{a \in \mathcal{A}} |a\rangle\langle a| \otimes A_a,$$
and such that $A_{g(i)} = A$.

3. For all $i, i' \in \mathcal{I}$ such that $i \neq i'$, and all $a \in \mathcal{A}$, $\langle i'|A^\dagger A_a|i\rangle = 0$.

If A were easy to implement, then we would not need a subroutine. Instead, we are assuming (condition 2) that there is an easy to implement A' that computes A given the auxiliary information $g(i)$ computed by the variable-time subroutine. Given a reversible variable-time subroutine for computing A, we can implement the map $|i\rangle \mapsto A|i\rangle$ by running the variable-time subroutine $U_1, \ldots, U_\mathsf{T}$ until the algorithm halts at some step T_i, applying A', and then uncomputing by running $\tilde{U}_{T_i}, \ldots, \tilde{U}_1$ (condition 1). The last condition is not strictly necessary, but the total error of algorithms that use the subroutine (see Section 4 and Section 5) is different if we omit it. The condition basically says that while the subroutine might output the wrong answer, it will not output an answer that could interfere with the correct part of another branch of the superposition running the algorithm on a different input.

For example, suppose $A|i\rangle|0\rangle = |i\rangle|g(i)\rangle$, so A just adds information to the state – not in an information theoretic sense, but in a computational sense. In that case, it is still possible to control on i to reverse the computation after we have applied A (conditions 1 and 2), and condition 3 is satisfied, since $\langle i', g(i')|i, a\rangle = 0$ whenever $i \neq i'$, regardless of a. Similarly, if $g(i)$ is a single bit, and $A|i\rangle = (-1)^{g(i)}|i\rangle$, then a variable-time subroutine for g is trivially reversible, by setting $\tilde{U}_t = U_t$ for all t, and $A_a|i\rangle = (-1)^a|i\rangle$.

3.1 Transition States and Algorithm States

We define several states associated with a variable stopping time algorithm, and prove some properties that will be useful later. We first define some states called *transition states* from the subroutine, and sort them into two sets Ψ_0 and Ψ_1, such that each set is pairwise orthogonal. These sets could be used to define a phase estimation algorithm as described in Section 2.2. There would be little point in turning an algorithm into a phase estimation algorithm – if we already have an algorithm, we do not need another one – but we will later (in Section 4 and Section 5) combine these states with some other ones to get a more complicated phase estimation algorithm that uses the variable stopping time subroutine as a building block.

Definition 3.1. *Fix some set of positive weights* $\{\alpha_t\}_{t=1}^{\mathsf{T}}$ *such that* $\alpha_0 = 1$. *The* transition states *of a reversible variable stopping time subroutine are defined as follows. Let* $\mathcal{Z}_{>t} = \bigcup_{t'=1}^{t-1} \mathcal{Z}_{t'}$. *The* forward transitions states *are defined:*

$\forall i \in \mathcal{I}, t \in \{0, \ldots, \mathsf{T} - 1\}$
$$\Psi_t^{i,\rightarrow} := \left\{ |\psi_{a,z,t}^{i,\rightarrow}\rangle := |\rightarrow\rangle|i\rangle \left(\sqrt{\alpha_t}|a,z\rangle|t\rangle - \sqrt{\alpha_{t+1}}U_{t+1}^i|a,z\rangle|t+1\rangle\right) : a \in \mathcal{A}, z \in \mathcal{Z}_{>t} \right\}.$$

The backward transition states *are defined:*

$\forall i \in \mathcal{I}, t \in \{0, \ldots, \mathsf{T} - 1\}$
$$\Psi_t^{i,\leftarrow} := \left\{ |\psi_{a,z,t}^{i,\leftarrow}\rangle := |\leftarrow\rangle A|i\rangle \left(\sqrt{\alpha_t}|a,z\rangle|t\rangle - \sqrt{\alpha_{t+1}}U_{t+1}^i|a,z\rangle|t+1\rangle\right) : a \in \mathcal{A}, z \in \mathcal{Z}_{>t} \right\}.$$
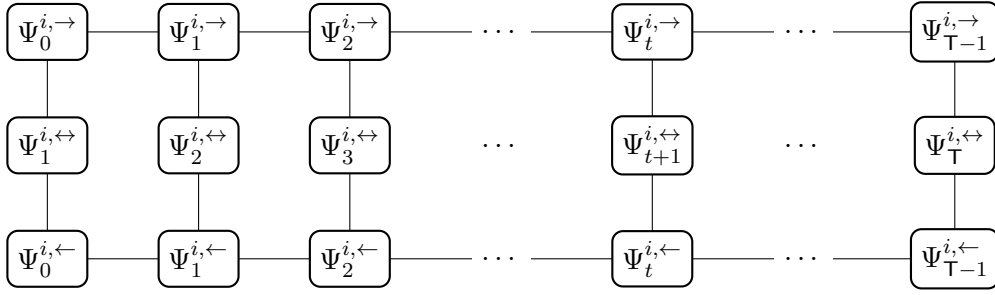
14

Figure 2: The overlap graph of the sets of states defined in Definition 3.4 for some fixed i. Each node represents a set of states that are pairwise orthogonal. Two nodes share an edge if and only if their sets contain overlapping states. For different values of i, all states are orthogonal. One can imagine an algorithm starting in the state $|\rightarrow\rangle|i\rangle|0,0\rangle|0\rangle$, which only overlaps $\Psi_0^{i,\rightarrow}$. The state of the algorithm moves through the graph until some part of it is of the form $|\leftarrow\rangle A_a|i\rangle|a,z\rangle|t\rangle$, hopefully for $a = g(i)$, then on that part, it uncomputes to move back down the other side of the ladder to a state that only overlaps $\Psi_0^{i,\leftarrow}$. The length of the algorithm's path depends on which "rung" t of the ladder it uses to move from the $\rightarrow$ to the $\leftarrow$ part, which will depend on the stopping probabilities at various steps.

The reversal states *are defined:*

$$\forall i \in \mathcal{I}, t \in \{1, \ldots, \mathsf{T}\} \qquad \Psi_t^{i,\leftrightarrow} := \left\{|\psi_{a,z,t}^{i,\leftrightarrow}\rangle := \sqrt{\alpha_t}\left(|\rightarrow\rangle|i\rangle - |\leftarrow\rangle A_a|i\rangle\right)|a,z\rangle|t\rangle : a \in \mathcal{A}, z \in \mathcal{Z}_t\right\}.$$

We finally define $\Psi_0^{i,\leftrightarrow} = \Psi_\mathsf{T}^{i,\rightarrow} = \Psi_\mathsf{T}^{i,\leftarrow} = \emptyset$*, and:*

$$\Psi_0 = \bigcup_{i \in \mathcal{I}} \bigcup_{t=0:t \text{ even}}^{\mathsf{T}} \left(\Psi_t^{i,\rightarrow} \cup \Psi_t^{i,\leftarrow} \cup \Psi_t^{i,\leftrightarrow}\right) \quad and \quad \Psi_1 = \bigcup_{i \in \mathcal{I}} \bigcup_{t=0:t \text{ odd}}^{\mathsf{T}} \left(\Psi_t^{i,\rightarrow} \cup \Psi_t^{i,\leftarrow} \cup \Psi_t^{i,\leftrightarrow}\right).$$

Claim 3.2. *For* $b \in \{0,1\}$*, the states of* Ψ_b *are pairwise orthogonal.*

To convince oneself of Claim 3.2, one should notice that the states in each set $\Psi_t^{i,d}$ are pairwise orthogonal, and the only overlaps between sets are those shown in the graph in Figure 2. Then since Ψ_0 and Ψ_1 are formed from a bipartition of this graph, they are each pairwise orthogonal.

Lemma 3.3. *For* $b \in \{0,1\}$*, let* Π_{Ψ_b} *denote the orthogonal projector onto the span of* Ψ_b*. Then* $(2\Pi_{\Psi_b} - I)$ *can be implemented in* $O(\log \mathsf{T})$ *complexity.*

Proof. The basic idea is that for $b \in \{0,1\}$, we can implement a unitary U_b that acts, for all $i \in \mathcal{I}$, $a \in \{0,1\}$, and t of parity b, as:

1. $|\rightarrow\rangle|i\rangle|a,z,t\rangle \overset{U_b}{\mapsto} \propto |\psi_{a,z,t}^{i,\rightarrow}\rangle$ for all $z \in \mathcal{Z}_{>t}$

2. $|\leftarrow\rangle A|i\rangle|a,z,t\rangle \overset{U_b}{\mapsto} \propto |\psi_{a,z,t}^{i,\leftarrow}\rangle$ for all $z \in \mathcal{Z}_{>t}$

3. $|\leftrightarrow\rangle|i\rangle|a,z,t\rangle \overset{U_b}{\mapsto} \propto |\psi_{a,z,t}^{i,\leftrightarrow}\rangle$ for all $z \in \mathcal{Z}_t$

using, respectively:

1. The ability to implement $\sum_{t=1}^{\mathsf{T}} |t\rangle\langle t| \otimes U_t$ in unit cost.

2. The ability to implement $\sum_{t=1}^{\mathsf{T}} |t\rangle\langle t| \otimes \tilde{U}_t$ in unit cost.

3. The ability to implement $\sum_{a \in \mathcal{A}} |a\rangle\langle a| \otimes A_a$ in unit cost.

We implement U_b in two parts: the first is conditioned on $z \in \mathcal{Z}_{>t}$, which will ensure (1) and (2); the second is conditioned on $z \in \mathcal{Z}_t$, which will ensure condition (3). First, conditioned on $z \in \mathcal{Z}_{>t}$, do the following. Append an auxiliary qubit initialized to $|0\rangle$. Controlled on t, rotate this qubit to

15

$\propto \sqrt{\alpha_t}|0\rangle - \sqrt{\alpha_{t+1}}|1\rangle$. When $z \in \mathcal{Z}_{>t}$, we only care about the cases where the first qubit is $|\rightarrow\rangle$ or $|\leftarrow\rangle$, and in these cases we now have:

$$(\rightarrow)\text{-case: } \sqrt{\alpha_t}|0\rangle|\rightarrow\rangle|i\rangle|a, z, t\rangle - \sqrt{\alpha_{t+1}}|1\rangle|\rightarrow\rangle|i\rangle|a, z, t\rangle,$$
$$(\leftarrow)\text{-case: } \sqrt{\alpha_t}|0\rangle|\leftarrow\rangle A|i\rangle|a, z, t\rangle - \sqrt{\alpha_{t+1}}|1\rangle|\leftarrow\rangle A|i\rangle|a, z, t\rangle.$$

Controlled on the auxiliary qubit being $|1\rangle$, increment the last register, and then (still controlled on $|1\rangle$) controlled on $|\rightarrow\rangle$, apply $\sum_{t=1}^{\mathsf{T}} |t\rangle\langle t| \otimes U_t$, and controlled on $|\leftarrow\rangle$, apply $\sum_{t=1}^{\mathsf{T}} |t\rangle\langle t| \otimes \tilde{U}_t$ to get:

$$(\rightarrow)\text{-case: } \sqrt{\alpha_t}|0\rangle|\rightarrow\rangle|i\rangle|a, z, t\rangle - \sqrt{\alpha_{t+1}}|1\rangle|\rightarrow\rangle|i\rangle U_{t+1}|a, z\rangle|t+1\rangle,$$
$$(\leftarrow)\text{-case: } \sqrt{\alpha_t}|0\rangle|\leftarrow\rangle A|i\rangle|a, z, t\rangle - \sqrt{\alpha_{t+1}}|1\rangle|\leftarrow\rangle A|i\rangle U_{t+1}|a, z\rangle|t+1\rangle.$$

Since we only care about the action of U_b when t has parity b, we can uncompute the auxiliary using the parity of the last register.

Next, conditioned on $z \in \mathcal{Z}_t$, do the following. Map $|\leftrightarrow\rangle$ to $(|\rightarrow\rangle - |\leftarrow\rangle)/\sqrt{2}$, and then conditioned on $|\leftarrow\rangle$, apply $\sum_a |a\rangle\langle a| \otimes A_a$ to get:

$$\frac{1}{\sqrt{2}}(|\rightarrow\rangle|i\rangle|a, z, t\rangle - |\leftarrow\rangle A_a|i\rangle|a, z, t\rangle) \propto |\psi_{a,z,t}^{i,\leftrightarrow}\rangle.$$

The map U_b can be combined with a reflection R_b that checks, for any state $|d\rangle|i\rangle|a, z\rangle|t\rangle$: t has parity b; if $d \in \{\leftarrow, \rightarrow\}$, then $z \in \mathcal{Z}_{>t}$ and $t < \mathsf{T}$; and if $d = \leftrightarrow$, then $z \in \mathcal{Z}_t$ and $t > 0$. Then $(2\Pi_{\Psi_b} - I) = U_b R_b U_b^\dagger$. $\qquad\square$

We now define some states which are essentially positive and negative witnesses for a phase estimation algorithm defined by Ψ_0 and Ψ_1. When we combine Ψ_0 and Ψ_1 with other states to get more complex phase estimation algorithms in Section 4 and Section 5, these will be building blocks for the positive and negative witnesses for those phase estimation algorithms, so we prove several claims analyzing these states here.

Definition 3.4 (Algorithm States). *Fix a reversible variable-time subroutine $U_1, \ldots, U_\mathsf{T}, A$ and some set of positive weights $\{\alpha_t\}_t$. Let $\Pi_{\geq t} := \Pi_t + \cdots + \Pi_\mathsf{T}$ be the orthogonal projector onto $H_{t-1}^\perp = \overline{H}_t \oplus \cdots \oplus \overline{H}_\mathsf{T}$. Define the following* algorithm states *in $H_A \otimes H_Z$, for all $i \in \mathcal{I}$:*

$$|w^0(i)\rangle = |0, 0\rangle$$
$$\forall t \in [\mathsf{T}], \ |w^t(i)\rangle = U_t^i \Pi_{\geq t}|w^{t-1}(i)\rangle.$$

The positive history state *of the algorithm on input $i \in \mathcal{I}$ is defined:*

$$|w_+(i)\rangle = (|\rightarrow\rangle|i\rangle + |\leftarrow\rangle A|i\rangle) \sum_{t=0}^{\mathsf{T}} \frac{1}{\sqrt{\alpha_t}}|w^t(i)\rangle|t\rangle.$$

The negative history state *of the algorithm on input $i \in \mathcal{I}$ is defined:*

$$|w_-(i)\rangle = (|\rightarrow\rangle|i\rangle - |\leftarrow\rangle A|i\rangle) \sum_{t=0}^{\mathsf{T}} \sqrt{\alpha_t}(-1)^t|w^t(i)\rangle|t\rangle.$$

We have defined the algorithm states in this way because $\Pi_{\geq t}|w^{t-1}(i)\rangle$ is the part of the algorithm that has not yet output before we apply U_t^i. The following claim makes clear that $|w^t(i)\rangle$ consists of two orthogonal parts: the part in $\overline{H}_t$, which is output at this step, and the part in $H_t^\perp = \overline{H}_{t+1} \oplus \cdots \oplus \overline{H}_\mathsf{T}$, which has yet to halt.

Claim 3.5. *For all $i \in \mathcal{I}$ and $t \in \{0, \ldots, \mathsf{T}\}$,*

$$|w^t(i)\rangle = \Pi_{\geq t} U_t^i \Pi_{\geq t-1} U_{t-1}^i \ldots \Pi_{\geq 1} U_1^i |0, 0\rangle = \Pi_{\geq t} U_t^i \ldots U_1^i |0, 0\rangle.$$

Proof. First, note that since $U_t^i \Pi_{\leq t-1} = \Pi_{\leq t-1}$ (by (15)), we have:

$$U_t^i \Pi_{\geq t} = U_t^i (I - \Pi_{\leq t-1}) = U_t^i - \Pi_{\leq t-1} = (I - \Pi_{\leq t-1})U_t^i = \Pi_{\geq t} U_t^i, \qquad (18)$$

so the first equality follows easily from the definition of $|w^t(i)\rangle$.

We prove the second equality by induction. For the base step, we have: $|w^0(i)\rangle = |0,0\rangle = \Pi_{\geq 0}|0,0\rangle$. For the induction step, assuming $|w^{t-1}(i)\rangle = \Pi_{\geq t-1} U_{t-1}^i \dots U_1^i |0,0\rangle$, we have:

$$|w^t(i)\rangle = U_t \Pi_{\geq t}|w^{t-1}(i)\rangle = U_t \Pi_{\geq t}\Pi_{\geq t-1} U_{t-1}^i \dots U_1^i |0,0\rangle = U_t \Pi_{\geq t} U_{t-1}^i \dots U_1^i |0,0\rangle$$
$$= \Pi_{\geq t} U_t U_{t-1}^i \dots U_1^i |0,0\rangle. \qquad \square$$

Corollary 3.6. *Let $i \in \mathcal{I}$. Recall the definitions of $\bar{p}_i(t)$ and $p_i(t)$ ((16)). Then we have:*

$$\left\| \Pi_t |w^t(i)\rangle \right\|^2 = \bar{p}_i(t) \ and \ \left\| |w^t(i)\rangle \right\|^2 = 1 - p_i(t).$$

Proof. By Claim 3.5, we have $\Pi_t |w^t(i)\rangle = \Pi_t U_t^i \dots U_1^i |0,0\rangle$, which is the part of the state after applying U_t that did not already output, but does output in the measurement after U_t, so $\bar{p}_i(t) = \left\| \Pi_t |w^t(i)\rangle \right\|^2$. For the second part, we have

$$\left\| |w^t(i)\rangle \right\|^2 = \left\| U_t^i \dots U_1^i |0,0\rangle \right\|^2 - \left\| \Pi_{<t} U_t^i \dots U_1^i |0,0\rangle \right\|^2 = 1 - p_i(t),$$

since $\Pi_{<t} U_t^i \dots U_1^i |0,0\rangle$ is the part of the state that has already halted before U_t was applied. $\square$

Corollary 3.7. *For all $i \in \mathcal{I}$,*

$$\left\| |w_-(i)\rangle \right\|^2 = 2 \sum_{t=0}^{\mathsf{T}} \alpha_t (1 - p_i(t)) = 2\mathbb{E}\left[\sum_{t=0}^{T_i} \alpha_t \right]$$
$$and \ \left\| |w_+(i)\rangle \right\|^2 = 2 \sum_{t=0}^{\mathsf{T}} \frac{1}{\alpha_t}(1 - p_i(t)) = 2\mathbb{E}\left[\sum_{t=0}^{T_i} \frac{1}{\alpha_t} \right].$$

Proof. By (16) and the fact that $\sum_{t=1}^{\mathsf{T}} \bar{p}_i(t) = 1$, and letting $\bar{p}_i(0) = 0$, we have:

$$\left\| |w_-(i)\rangle \right\|^2 = \sum_{t=0}^{\mathsf{T}} \alpha_t(1 - p_i(t)) = \sum_{t=0}^{\mathsf{T}} \alpha_t \sum_{t'=t}^{\mathsf{T}} \bar{p}_i(t') = \sum_{t'=0}^{\mathsf{T}} \bar{p}_i(t') \sum_{t=0}^{t'} \alpha_t = \mathbb{E}\left[\sum_{t=0}^{T_i} \alpha_t \right].$$

A similar computation applies in the positive case. $\square$

Claim 3.8 (Positive Witness). *Assume T is odd. For all i, $|w_+(i)\rangle$ is orthogonal to all forward and backward transition states, and all reversal states $|\psi_{a,z,t}^{i',\leftrightarrow}\rangle$ such that $i' \neq i$ or $a' \neq g(i)$ (see Definition 3.1). Moreover, if Π_{Ψ_b} is the orthogonal projector onto $\mathrm{span}\{\Psi_b\}$, then*

$$\left\| \Pi_{\Psi_0}|w_+(i)\rangle \right\|^2 \leq 2 \sum_{\substack{t=0 \\ even}}^{\mathsf{T}-1} \frac{1}{\alpha_t} \bar{p}_i(t)\epsilon_i^t \leq 2\mathbb{E}\left[\frac{\epsilon_i^{T_i}}{\alpha_{T_i}} \right] \ and \ \left\| \Pi_{\Psi_1}|w_+(i)\rangle \right\|^2 \leq 2 \sum_{\substack{t=1 \\ odd}}^{\mathsf{T}} \frac{1}{\alpha_t} \bar{p}_i(t)\epsilon_i^t \leq 2\mathbb{E}\left[\frac{\epsilon_i^{T_i}}{\alpha_{T_i}} \right].$$

Proof. We start with the **forward transition states**. We always have $\langle w_+(i)|\psi_{a,z,t}^{i',\rightarrow}\rangle = 0$ when $i \neq i'$. For all $z \in \mathcal{Z}_{>t}$, we have:

$$\langle \psi_{a,z,t}^{i,\rightarrow}|w_+(i)\rangle = \langle a,z|w^t(i)\rangle - \langle a,z|(U_{t+1}^i)^\dagger|w^{t+1}(i)\rangle = \langle a,z|w^t(i)\rangle - \langle a,z|(U_{t+1}^i)^\dagger U_{t+1}^i \Pi_{\geq t+1}|w^t(i)\rangle = 0,$$

since $\langle a,z|\Pi_{\geq t} = \langle a,z|$ for all $z \in \mathcal{Z}_{>t}$. A nearly identical argument holds for the **backward transition states**.

Next consider the **reversal states**. We have

$$\langle \psi_{a,z,t}^{i,\leftrightarrow}|w_+(i)\rangle = \left(1 - \langle i|A_a^\dagger A|i\rangle\right)\langle a,z|w^t(i)\rangle. \qquad (19)$$

When $a = g(i)$, so $A_a = A_{g(i)} = A$, this is 0, and otherwise we have $|\langle \psi_{a,z,t}^{i,\leftrightarrow}|w_+(i)\rangle|^2 \le 4|\langle a, z|w^t(i)\rangle|^2$. Thus:

$$\Pi_{\Psi_0}|w_+(i)\rangle = \sum_{\substack{t=0 \\ \text{even}}}^{\mathsf{T}-1} \sum_{\substack{a \in \mathcal{A}\setminus\{g(i)\} \\ z \in \mathcal{Z}_t}} \frac{|\psi_{a,z,t}^{i,\leftrightarrow}\rangle\langle\psi_{a,z,t}^{i,\leftrightarrow}|}{\left\||\psi_{a,z,t}^{i,\leftrightarrow}\rangle\right\|^2}|w_+(i)\rangle$$

$$\|\Pi_{\Psi_0}|w_+(i)\rangle\|^2 = \sum_{\substack{t=0 \\ \text{even}}}^{\mathsf{T}-1} \sum_{\substack{a \in \mathcal{A}\setminus\{g(i)\} \\ z \in \mathcal{Z}_t}} \frac{|\langle\psi_{a,z,t}^{i,\leftrightarrow}|w_+(i)\rangle|^2}{\left\||\psi_{a,z,t}^{i,\leftrightarrow}\rangle\right\|^2} \le \sum_{\substack{t=0 \\ \text{even}}}^{\mathsf{T}-1} \sum_{\substack{a \in \mathcal{A}\setminus\{g(i)\} \\ z \in \mathcal{Z}_t}} \frac{4|\langle a, z|w^t(i)\rangle|^2}{2\alpha_t}.$$

We have, for any t,

$$\sum_{\substack{a \in \mathcal{A}\setminus\{g(i)\} \\ z \in \mathcal{Z}_t}} |\langle a, z|w^t(i)\rangle|^2 = \bar{p}_i(t)\epsilon_i^t$$

is the probability that the algorithm outputs at time t, and the answer is incorrect (not equal to $g(i)$). Thus:

$$\|\Pi_{\Psi_0}|w_+(i)\rangle\|^2 \le 2\sum_{\substack{t=0 \\ \text{even}}}^{\mathsf{T}-1} \bar{p}_i(t)\frac{1}{\alpha_t}\epsilon_i^t \le 2\sum_{t=0}^{\mathsf{T}} \bar{p}_i(t)\frac{1}{\alpha_t}\epsilon_i^t = 2\mathbb{E}\left[\frac{\epsilon_i^{T_i}}{\alpha_{T_i}}\right],$$

and similarly for Ψ_1. $\qquad\square$

Claim 3.9 (Negative witness). *Assume T is odd. For $i \in \mathcal{I}$, let $|w_-(i)\rangle$ be the negative history state defined in Definition 3.4. Then, letting Π_{Ψ_b} be the orthogonal projector onto $\mathrm{span}\{\Psi_b\}$, we have:*

$$\|(I - \Pi_{\Psi_0})|w_-(i)\rangle\|^2 \le 2\sum_{\substack{t=0 \\ \text{even}}}^{\mathsf{T}-1} \alpha_t\bar{p}_i(t)\epsilon_i^t \le 2\mathbb{E}\left[\alpha_{T_i}\epsilon_i^{T_i}\right]$$

and $\|(I - \Pi_{\Psi_1})\left(|w_-(i)\rangle - (|\!\rightarrow\rangle|i\rangle - |\!\leftarrow\rangle A|i\rangle)|0,0\rangle|0\rangle\right)\|^2 \le 2\sum_{\substack{t=1 \\ \text{odd}}}^{\mathsf{T}} \alpha_t\bar{p}_i(t)\epsilon_i^t \le 2\mathbb{E}\left[\alpha_{T_i}\epsilon_i^{T_i}\right].$

Proof. Recall from Definition 3.4 that

$$|w_-(i)\rangle = (|\!\rightarrow\rangle|i\rangle - |\!\leftarrow\rangle A|i\rangle)\sum_{t=0}^{\mathsf{T}} \sqrt{\alpha_t}(-1)^t|w^t(i)\rangle|t\rangle,$$

where $|w^t(i)\rangle = \Pi_{\ge t}|w^t(i)\rangle$ by Claim 3.5. We have:

$$\sum_{t=0}^{\mathsf{T}} \sqrt{\alpha_t}(-1)^t|w^t(i)\rangle|t\rangle = \sum_{\substack{t=0 \\ \text{even}}}^{\mathsf{T}-1} \sqrt{\alpha_t}|w^t(i)\rangle|t\rangle - \sum_{\substack{t=1 \\ \text{odd}}}^{\mathsf{T}} \sqrt{\alpha_t}|w^t(i)\rangle|t\rangle$$

$$= \sum_{\substack{t=0 \\ \text{even}}}^{\mathsf{T}-1} \sqrt{\alpha_t}\Pi_t|w^t(i)\rangle|t\rangle + \sum_{\substack{t=0 \\ \text{even}}}^{\mathsf{T}-1} \sqrt{\alpha_t}\Pi_{\ge t+1}|w^t(i)\rangle|t\rangle - \sum_{\substack{t=1 \\ \text{odd}}}^{\mathsf{T}} \sqrt{\alpha_t}\Pi_{\ge t}U_t^i|w^{t-1}(i)\rangle|t\rangle$$

$$= \sum_{\substack{t=0 \\ \text{even}}}^{\mathsf{T}-1} \sqrt{\alpha_t}\Pi_t|w^t(i)\rangle|t\rangle + \sum_{\substack{t=0 \\ \text{even}}}^{\mathsf{T}-1} \left(\sqrt{\alpha_t}\Pi_{\ge t+1}|w^t(i)\rangle|t\rangle - \sqrt{\alpha_{t+1}}U_{t+1}^i\Pi_{\ge t+1}|w^t(i)\rangle|t+1\rangle\right) \quad \text{by (18).}$$

Then since $\Pi_{\ge t+1}|w^t(i)\rangle$ is supported on states $|a, z\rangle$ such that $z \in \mathcal{Z}_{>t}$, we have:

$$(|\!\rightarrow\rangle|i\rangle - |\!\leftarrow\rangle A|i\rangle)\sum_{\substack{t=0 \\ \text{even}}}^{\mathsf{T}-1} \left(\sqrt{\alpha_t}\Pi_{\ge t+1}|w^t(i)\rangle|t\rangle - \sqrt{\alpha_{t+1}}U_{t+1}^i\Pi_{\ge t+1}|w^t(i)\rangle|t+1\rangle\right) \in \mathrm{span}\{\Psi_0\}.$$

18

Let Ξ_a be the orthogonal projector onto states with a in the answer register. We can thus see that:

$$(I - \Pi_{\Psi_0})|w_-(i)\rangle = (I - \Pi_{\Psi_0}) \left(|\rightarrow\rangle|i\rangle - |\leftarrow\rangle A|i\rangle \right) \sum_{\substack{t=0 \\ \text{even}}}^{\mathsf{T}-1} \sqrt{\alpha_t} \Pi_t |w^t(i)\rangle|t\rangle$$

$$- (I - \Pi_{\Psi_0}) \left(|\rightarrow\rangle|i\rangle - |\leftarrow\rangle A_{g(i)}|i\rangle \right) \underbrace{\sum_{\substack{t=0 \\ \text{even}}}^{\mathsf{T}-1} \sqrt{\alpha_t} \Xi_{g(i)} \Pi_t |w^t(i)\rangle|t\rangle}_{\in \Psi_0} \quad (20)$$

$$= (I - \Pi_{\Psi_0}) \left(|\rightarrow\rangle|i\rangle - |\leftarrow\rangle A_{g(i)}|i\rangle \right) \sum_{\substack{t=0 \\ \text{even}}}^{\mathsf{T}-1} \sqrt{\alpha_t} \left(I - \Xi_{g(i)} \right) \Pi_t |w^t(i)\rangle|t\rangle,$$

since $A|i\rangle = A_{g(i)}|i\rangle$. Thus:

$$\left\| (I - \Pi_{\Psi_0})|w_-(i)\rangle \right\|^2 \leq \left\| \left(|\rightarrow\rangle|i\rangle - |\leftarrow\rangle A_{g(i)}|i\rangle \right) \sum_{\substack{t=0 \\ \text{even}}}^{\mathsf{T}-1} \sqrt{\alpha_t} \left(I - \Xi_{g(i)} \right) \Pi_t |w^t(i)\rangle|t\rangle \right\|^2$$

$$= 2 \sum_{\substack{t=0 \\ \text{even}}}^{\mathsf{T}-1} \alpha_t \left\| \left(I - \Xi_{g(i)} \right) \Pi_t |w^t(i)\rangle \right\|^2 = 2 \sum_{\substack{t=0 \\ \text{even}}}^{\mathsf{T}-1} \bar{p}_i(t) \alpha_t \epsilon_i^t \leq 2\mathbb{E}\left[\alpha_{T_i} \epsilon_i^{T_i} \right].$$

For the second part of the claim, we note that:

$$\sum_{t=0}^{\mathsf{T}} \sqrt{\alpha_t}(-1)^t |w^t(i)\rangle|t\rangle$$

$$= \sqrt{\alpha_0}|w^0(i)\rangle|0\rangle - \sqrt{\alpha_{\mathsf{T}}}|w^{\mathsf{T}}(i)\rangle|\mathsf{T}\rangle$$

$$- \sum_{\substack{t=1 \\ \text{odd}}}^{\mathsf{T}-2} \sqrt{\alpha_t} \Pi_t |w^t(i)\rangle|t\rangle - \sum_{\substack{t=1 \\ \text{odd}}}^{\mathsf{T}-2} \left(\sqrt{\alpha_t} \Pi_{\geq t+1} |w^t(i)\rangle|t\rangle - \sqrt{\alpha_{t+1}} U_{t+1}^i \Pi_{\geq t+1} |w^t(i)\rangle|t+1\rangle \right)$$

$$= |0,0\rangle|0\rangle - \sum_{\substack{t=1 \\ \text{odd}}}^{\mathsf{T}} \sqrt{\alpha_t} \Pi_t |w^t(i)\rangle|t\rangle - \sum_{\substack{t=1 \\ \text{odd}}}^{\mathsf{T}-2} \left(\sqrt{\alpha_t} \Pi_{\geq t+1} |w^t(i)\rangle|t\rangle - \sqrt{\alpha_{t+1}} U_{t+1}^i \Pi_{\geq t+1} |w^t(i)\rangle|t+1\rangle \right),$$

since $\alpha_0 = 1$, $|w^0(i)\rangle = |0,0\rangle$, and $|w^{\mathsf{T}}(i)\rangle$ is in the support of $\Pi_{\geq \mathsf{T}} = \Pi_{\mathsf{T}}$. Similar to above, we have

$$(|\rightarrow\rangle|i\rangle - |\leftarrow\rangle A|i\rangle) \sum_{\substack{t=1 \\ \text{odd}}}^{\mathsf{T}-2} \left(\sqrt{\alpha_t} \Pi_{\geq t+1} |w^t(i)\rangle|t\rangle - \sqrt{\alpha_{t+1}} U_{t+1}^i \Pi_{\geq t+1} |w^t(i)\rangle|t+1\rangle \right) \in \Psi_1,$$

and just as in (20),

$$(I - \Pi_{\Psi_1})(|\rightarrow\rangle|i\rangle - |\leftarrow\rangle A|i\rangle) \sum_{\substack{t=1 \\ \text{odd}}}^{\mathsf{T}} \sqrt{\alpha_t} \Pi_t |w^t(i)\rangle|t\rangle$$

$$= (I - \Pi_{\Psi_1})(|\rightarrow\rangle|i\rangle - |\leftarrow\rangle A|i\rangle) \sum_{\substack{t=1 \\ \text{odd}}}^{\mathsf{T}} \sqrt{\alpha_t} \Pi_t (I - \Xi_{g(i)}) |w^t(i)\rangle|t\rangle.$$

Thus, we can see that:

$$(I - \Pi_{\Psi_1}) \left(|w_-(i)\rangle - (|\rightarrow\rangle|i\rangle - |\leftarrow\rangle A|i\rangle)|0,0\rangle|0\rangle \right)$$

$$= (I - \Pi_{\Psi_1})(|\rightarrow\rangle|i\rangle - |\leftarrow\rangle A|i\rangle) \left(|0,0\rangle|0\rangle - \sum_{\substack{t=1 \\ \text{odd}}}^{\mathsf{T}} \sqrt{\alpha_t} \Pi_t (I - \Xi_{g(i)}) |w^t(i)\rangle|t\rangle - |0,0\rangle|0\rangle \right)$$

so $\|(I - \Pi_{\Psi_1})(|w_-(i)\rangle - (|\rightarrow\rangle|i\rangle - |\leftarrow\rangle A|i\rangle)|0,0\rangle|0\rangle)\|^2$

$$\leq 2 \sum_{\substack{t=1 \\ \text{odd}}}^{\mathsf{T}} \alpha_t \left\| \Pi_t (I - \Xi_{g(i)})|w^t(i)\rangle \right\|^2 = 2 \sum_{\substack{t=1 \\ \text{odd}}}^{\mathsf{T}} \bar{p}_i(t) \alpha_t \epsilon_t^i \leq 2\mathbb{E}\left[\alpha_{T_i} \epsilon_i^{T_i} \right]. \quad \square$$

4 Composition in Quantum Walks

In this section, we state and prove Theorem 4.1, which extends the quantum walk electric network framework by allowing a variable-time subroutine to implement the edge transition subroutine. This is similar to [JZ22, Theorem 3.10], with the following differences:

1. We require that our transition subroutine be *reversible*, meaning that the mapping $|u,i\rangle \mapsto |v,j\rangle$ must be done by first computing some auxiliary information $g(u,i)$, from which the transition becomes easy (see Section 3). This is not required in [JZ22], but it is not a particularly stringent constraint, and we know of no quantum walk application where the transition is not done this way.

2. [JZ22] does not allow the running time $T_{u,v}$ for some fixed (u,v) to be a random variable, but instead depends on a parameter $\mathsf{T}_{u,v}$, which is an upper bound on the running time on input (u,v), and requires that this value be efficiently computable from (u,i) and (v,j) (where $f_u(i) = v$ and $f_v(j) = u$). We do not require $\mathbb{E}[T_{u,v}]$ to be efficiently computable.

3. In [JZ22], there are no parameters $\{\alpha_t\}_t$ (equivalently, $\alpha_t = 1$ for all t).

4. [JZ22] also extends the electric network framework by allowing the use of *alternative neighbourhoods*. For simplicity, we do not incude this extension here, but there is no reason it could not be combined with our results.

We first state Theorem 4.1, which is too general to be easily understood. Immediately afterwards, we discuss some more easily understood consequences of this theorem, for quantum walks (Corollary 4.2 and Corollary 4.3) and the special case of black-box search (Corollary 4.4 and Table 1), before proving the theorem in Section 4.1.

Theorem 4.1 (Quantum Walks with Edge Composition). *Fix the following, which may implicitly depend on some input x:*

- *a network G with disjoint sets $V_0, V_{\mathsf{M}} \subset V(G)$ such that for any vertex, checking if $v \in V_0$ (resp. if $v \in V_{\mathsf{M}}$) can be done in unit complexity;*

- *a reversible variable-time subroutine (see Section 3) that implements the transition map (see Definition 2.2) with stopping times $\{T_{u,v}\}_{(u,v)\in\vec{E}(G)}$, which are random variables on $[\mathsf{T}]$, and errors $\{\epsilon_{u,v}^t\}_{(u,v)\in\vec{E}(G),t\in[\mathsf{T}]}$;*

- *a subset $M \subseteq V_{\mathsf{M}}$, and distribution σ on V_0.*

Fix some positive real numbers $\{\alpha_t\}_{t=0}^{\mathsf{T}}$ such that $\alpha_0 = 1$; and positive real-valued $\mathcal{W}$ and $\mathcal{R}$, that may scale with $|x|$. Suppose the following conditions hold.

Setup Subroutine: *The state $|\sigma\rangle = \sum_{u\in V_0} \sqrt{\sigma(u)}|u\rangle$ can be generated in cost S, and furthermore, for any $u \in V_0$, $\sigma(u)$ can be computed in unit cost.*

Star State Generation Subroutine: *There is a subroutine that generates $\{|\psi_\star^G(u)\rangle\}_{u\in V(G)}$ in unit cost (see Definition 2.2).*

Checking Subroutine: *There is an algorithm that checks, for any $u \in V_{\mathsf{M}}$, if $u \in M$, in unit cost.*

Positive Condition: *If $M \neq \emptyset$, then there exists a flow θ on G (see Definition 2.3) such that:*

P1 *The boundary of θ is in $V_0 \cup M$, and $\sum_{u \in V_0} \theta(u) = 1$.*[7]

P2 *$\frac{1}{3} \le \sum_{u \in V_0} \frac{\theta(u)^2}{\sigma(u)} \le 3$.*

P3 $\displaystyle\sum_{(u,v) \in \overrightarrow{E}(G)} \frac{\theta(u,v)^2}{\mathsf{w}_{u,v}} \mathbb{E}\left[\sum_{t=0}^{T_{u,v}} \frac{1}{\alpha_t}\right] \le \mathcal{R}.$

P4 $\displaystyle\sum_{(u,v) \in \overrightarrow{E}(G)} \frac{\theta(u,v)^2}{\mathsf{w}_{u,v}} \mathbb{E}\left[\frac{\epsilon_{u,v}^{T_{u,v}}}{\alpha_{T_{u,v}}}\right] = o(1/\mathcal{W});$

Negative Condition: *If $M = \emptyset$, then*

N1 $\displaystyle\sum_{(u,v) \in \overrightarrow{E}(G)} \mathsf{w}_{u,v} \mathbb{E}\left[\sum_{t=0}^{T_{u,v}-1} \alpha_t\right] \le \mathcal{W}.$

N2 $\displaystyle\sum_{(u,v) \in \overrightarrow{E}(G)} \mathsf{w}_{u,v} \mathbb{E}\left[\alpha_{T_{u,v}} \epsilon_{u,v}^{T_{u,v}}\right] = o(1/\mathcal{R});$

Then there is a quantum algorithm that decides if $M = \emptyset$ or not with bounded error in complexity:

$$O\left(\mathsf{S} + \sqrt{\mathcal{R}\mathcal{W}} \cdot \log \mathsf{T}\right).$$

We prove Theorem 4.1 in Section 4.1 by defining a phase estimation algorithm and analysing it using Theorem 2.11. First, we give some less general, easier to digest, corollaries of Theorem 4.1. One thing that probably makes Theorem 4.1 particularly hard to parse is the parameters $\{\alpha_t\}_{t=0}^{\mathsf{T}}$, so we mention three cases of special interest. First, if we set $\alpha_t = 1$ for all t, then we recover a version of [JZ22, Theorem 3.10]: the resulting complexities are as if we had replaced the graph G with a graph G^{T}, in which each edge (u,v) is replaced by a path of length $\mathbb{E}[T_{u,v}]+1$ (or an upper bound on $T_{u,v}+1$, in the case of [JZ22]), but now each transition has unit cost.

We mention two other interesting choices of $\{\alpha_t\}_t$ that lead to two distinct "variable-time" quantum walk results. First, if we set $\alpha_t = t + 1$, we satisfy the requirement $\alpha_0 = 1$, and we get:

$$\mathbb{E}\left[\sum_{t=0}^{T_{u,v}} \frac{1}{\alpha_t}\right] = \mathbb{E}\left[\sum_{t=1}^{T_{u,v}+1} \frac{1}{t}\right] = \Theta(\mathbb{E}[\log T_{u,v}]) \text{ and } \mathbb{E}\left[\sum_{t=0}^{T_{u,v}} \alpha_t\right] = \mathbb{E}\left[\sum_{t=1}^{T_{u,v}+1} t\right] = \Theta(\mathbb{E}[T_{u,v}^2]).$$

Thus **P3** has only a logarithmic dependence on the transition times, whereas we get a weighted ℓ_2-average in **N1**. On the other hand, by setting $\alpha_t = \frac{1}{1+t}$, $\mathbb{E}[\log T_{u,v}]$ and $\mathbb{E}[T_{u,v}^2]$ are swapped: now **P3** has a (different!) weighted ℓ_2-average, whereas **N1** just has a log factor.

If we additionally know the values $\mathbb{E}[T_{u,v}]$, then we can use $\alpha_t = 1$ for all t, and scale $\mathsf{w}_{u,v}$ by a factor of $\mathbb{E}[T_{u,v}]$ to get:

$$\mathbf{P3}: \sum_{(u,v) \in \overrightarrow{E}(G)} \frac{\theta(u,v)^2}{\mathsf{w}_{u,v}} \le \mathcal{R} \quad \mathbf{N1}: \sum_{(u,v) \in \overrightarrow{E}(G)} \mathsf{w}_{u,v} \mathbb{E}[T_{u,i}]^2 \le \mathcal{W},$$

or alternatively, scaling $\mathsf{w}_{u,v}$ by $1/\mathbb{E}[T_{u,v}]$ gives:

$$\mathbf{P3}: \sum_{(u,v) \in \overrightarrow{E}(G)} \frac{\theta(u,v)^2}{\mathsf{w}_{u,v}} \mathbb{E}[T_{u,i}]^2 \le \mathcal{R} \quad \mathbf{N1}: \sum_{(u,v) \in \overrightarrow{E}(G)} \mathsf{w}_{u,v} \le \mathcal{W}.$$

Thus, when the expected stopping times are known, we can save a log factor, and also replace $\mathbb{E}[T_{u,v}^2]$ with $\mathbb{E}[T_{u,v}]^2$, which is better the higher the variance. We formally state these results for $\alpha_t = t + 1$

---

[7]A σ-M-flow always satisfies conditions **P1** and **P2**. While we do not make it a strict requirement that all sources are in V_0 and all sinks in M, **P1** implies that we do not simply have flow coming in at V_0 and then leaving again at V_0. **P2** implies that the flow in V_0 is roughly distributed as σ.

in the following corollary (the $\alpha_t = 1/(t+1)$ case can be worked out from Theorem 4.1, but is less elegantly stated). We simplify things by only considering flows that are exactly from σ to M, and assuming our subroutine has no error, but by Theorem 4.1, it is clear that some deviation would also be tolerable.[8]

Corollary 4.2 (Variable-time quantum walks ($\alpha_t = t+1$)). *Fix a network G; marked set $M \subset V(G)$, and initial distribution σ on $V(G)$, all of which may implicitly depend on an input x, although for simplicity, we assume $\mathcal{W}(G)$ only depends on $|x|$. Let P be the transition matrix for the random walk on G, and π its stationary distribution. Let S be the cost of generating $|\sigma\rangle$, and suppose $\{|\psi_\star^G(u)\rangle\}_{u \in V(G)}$ can be generated in unit cost. Let $\mathcal{C}_{\sigma,\mathsf{M}}$ be an upper bound on $\mathcal{R}_{\sigma,M}(G)\mathcal{W}(G)$ (see Definition 2.3) whenever $M \neq \emptyset$.*

Suppose there is a variable-time subroutine[9] that checks, for any $u \in V(G)$, if $u \in M$, with stopping times $\{C_u\}_{u \in V(G)}$, which are random variables on $[\mathsf{C}]$, and all errors zero. Fix some (known) terminal distribution τ on $V(G)$ such that $\tau(M) \geq \varepsilon$ whenever $M \neq \emptyset$, and let $\mathsf{C}_{\mathrm{avg}}$ and $\mathsf{C}'_{\mathrm{avg}}$ be upper bounds such that whenever $M = \emptyset$,

$$\sqrt{\sum_{u \in V(G)} \tau(u)\mathbb{E}[C_u^2]} \leq \mathsf{C}_{\mathrm{avg}}, \quad and \quad \sqrt{\sum_{u \in V(G)} \tau(u)\mathbb{E}[C_u]^2} \leq \mathsf{C}'_{\mathrm{avg}}.$$

Suppose there is a reversible variable-time subroutine that implements the transition map with stopping times $\{T_{u,v}\}_{(u,v) \in \vec{E}(G)}$, which are random variables on $[\mathsf{T}]$, and all errors zero. Let $\mathsf{T}_{\mathrm{avg}}$ and $\mathsf{T}'_{\mathrm{avg}}$ be upper bounds such that whenever $M = \emptyset$,

$$\sqrt{\sum_{(u,v) \in \vec{E}(G)} \pi(u)P_{u,v}\mathbb{E}[T_{u,v}^2]} \leq \mathsf{T}_{\mathrm{avg}}, \quad and \quad \sqrt{\sum_{(u,v) \in \vec{E}(G)} \pi(u)P_{u,v}\mathbb{E}[T_{u,v}]^2} \leq \mathsf{T}'_{\mathrm{avg}}.$$

Claim 0: *Suppose the checking cost is trivial, meaning for all $u \in V(G)$, $\mathbb{E}[C_u] = O(1)$. Then there is a quantum algorithm that detects if $M = \emptyset$ with bounded error in complexity:*

$$O\left(\mathsf{S} + \sqrt{\mathcal{C}_{\sigma,\mathsf{M}}}\mathsf{T}_{\mathrm{avg}}\log^{1.5}\mathsf{T}\right).$$

Suppose in addition that the values $\mathbb{E}[T_{u,v}] = \mathbb{E}[T_{u,i}]$ are computable in the strong sense that for any u, we can generate a superposition proportional to $\sum_{i \in L(u)} \sqrt{\mathsf{w}_{u,i}\mathbb{E}[T_{u,i}]}|i\rangle$, and we have query access to $\mathsf{w}'_u = \sum_{i \in L(u)} \mathsf{w}_{u,i}\mathbb{E}[T_{u,i}]$. Then there is a quantum algorithm that detects if $M = \emptyset$ with bounded error in complexity:

$$O\left(\mathsf{S} + \sqrt{\mathcal{C}_{\sigma,\mathsf{M}}}\mathsf{T}'_{\mathrm{avg}}\log\mathsf{T}\right).$$

Claim I: *Let $\mathcal{D}_{\mathsf{M}}$ be an upper bound such that $\sum_{u \in M} \theta(u)^2/\tau_M(u) \leq \mathcal{D}_{\mathsf{M}}$, where θ is the unit flow from σ to M with minimal energy and $\tau_M(u) = \tau(u)/\tau(M)$ is the normalized restriction of τ to M, whenever $M \neq \emptyset$. Then there is a quantum algorithm that detects if $M = \emptyset$ with bounded error in complexity:*

$$O\left(\mathsf{S} + \left(\sqrt{\mathcal{C}_{\sigma,\mathsf{M}}}\mathsf{T}_{\mathrm{avg}} + \sqrt{\frac{\mathcal{D}_{\mathsf{M}}}{\varepsilon}}\mathsf{C}_{\mathrm{avg}}\right)\log^{1.5}(\mathsf{TC})\right).$$

Suppose in addition that the values $\mathbb{E}[T_{u,v}]$ are computable in the same strong sense as Claim 0, and $\mathbb{E}[C_u]$ are computable in the the sense that for any $u \in V(G)$, we can query $\mathbb{E}[C_u]$ in unit cost. Then there is a quantum algorithm that detects in $M = \emptyset$ with bounded error in complexity:

$$O\left(\mathsf{S} + \left(\sqrt{\mathcal{C}_{\sigma,\mathsf{M}}}\mathsf{T}'_{\mathrm{avg}} + \sqrt{\frac{\mathcal{D}_{\mathsf{M}}}{\varepsilon}}\mathsf{C}'_{\mathrm{avg}}\right)\log(\mathsf{TC})\right).$$

---

[8]In all the applications discussed in this section, we assume the subroutine has no error, for simplicity. We believe it is also interesting to further investigate the different ways the error can propagate depending on the choice of $\{\alpha_t\}_t$, but we leave this for future work. It is also possible, as in the case of the results of Section 5, that our error terms can be improved.

[9]A variable-time subroutine that computes a bit is always reversible.

Claim II: *Suppose we are promised that either $M = \emptyset$, or $|M| = 1$. Then there is a quantum algorithm that detects if $M = \emptyset$ with bounded error in complexity:*

$$O\left(\mathsf{S} + \left(\sqrt{\mathcal{C}_{\sigma,\mathsf{M}}}\mathsf{T}_{\mathrm{avg}} + \frac{1}{\sqrt{\varepsilon}}\mathsf{C}_{\mathrm{avg}}\right)\log^{1.5}(\mathsf{TC})\right).$$

Suppose in addition that the values $\mathbb{E}[T_{u,v}]$ and $\mathbb{E}[C_u]$ are computable as in Claim I. Then there is a quantum algorithm that detects in $M = \emptyset$ with bounded error in complexity:

$$O\left(\mathsf{S} + \left(\sqrt{\mathcal{C}_{\sigma,\mathsf{M}}}\mathsf{T}'_{\mathrm{avg}} + \frac{1}{\sqrt{\varepsilon}}\mathsf{C}'_{\mathrm{avg}}\right)\log(\mathsf{TC})\right).$$

Claim III: *Let $\mathcal{C}_{\sigma,\tau}$ be an upper bound on $\mathcal{W}(G)\mathcal{R}_{\sigma,\tau_M}(G)$ (see Definition 2.3). Then there is a quantum algorithm that detects if $M = \emptyset$ with bounded error in complexity:*

$$O\left(\mathsf{S} + \left(\sqrt{\mathcal{C}_{\sigma,\tau}}\mathsf{T}_{\mathrm{avg}} + \frac{1}{\sqrt{\varepsilon}}\mathsf{C}_{\mathrm{avg}}\right)\log^{1.5}\mathsf{TC}\right).$$

Suppose in addition that the values $\mathbb{E}[T_{u,v}]$ and $\mathbb{E}[C_u]$ are computable as in Claim I. Then there is a quantum algorithm that detects in $M = \emptyset$ with bounded error in complexity:

$$O\left(\mathsf{S} + \left(\sqrt{\mathcal{C}_{\sigma,\tau}}\mathsf{T}'_{\mathrm{avg}} + \frac{1}{\sqrt{\varepsilon}}\mathsf{C}'_{\mathrm{avg}}\right)\log\mathsf{TC}\right).$$

Proof. Let $g_u = 1$ if $u \in M$, and $g_u = 0$ else. We will make a new graph G', which is obtained from G by adding, for each vertex $u \in V(G)$, a new vertex (u, g_u) connected only to u by an edge of weight $\mathsf{w}_{u,\mathsf{M}}$, to be defined. That is,

$$V(G') := \underbrace{V(G)}_{=:V_0} \cup \underbrace{\{(u, g_u) : u \in V(G)\}}_{=:V_\mathsf{M}} \text{ and } \overrightarrow{E}(G') := \overrightarrow{E}(G) \cup \{(u, (u, g_u)) : u \in V(G)\},$$

with $\mathsf{w}_{u,(u,g_u)} = \mathsf{w}_{u,\mathsf{M}}$. Then let $M' = M \times \{1\} \subseteq V_\mathsf{M}$ be the marked set of G'. The new edges $(u, (u, g_u))$ will have label 0, in both directions, so: $f_u(0) = (u, g_u)$ and $f_{(u,g_u)}(0) = u$. Since we can generate $\{|\psi_\star^G(u)\rangle\}_{u \in V(G)}$ in unit cost, we can also generate

$$\{|\psi_\star^{G'}(u)\rangle = |\psi_\star^G(u)\rangle + \sqrt{\mathsf{w}_{u,\mathsf{M}}}|0\rangle\}_{u \in V(G)}$$

in unit cost, assuming $\mathsf{w}_{u,\mathsf{M}}$ is chosen so that we can prepare a state proportional to $\sqrt{\mathsf{w}_u}|1\rangle + \sqrt{\mathsf{w}_{u,\mathsf{M}}}|0\rangle$. For any $u \in V_\mathsf{M}$, we can check if $u \in M'$ in unit cost by examining g_u, and generate $|\sigma\rangle$ in cost S, by assumption. Thus, to apply Theorem 4.1, we just need to check the positive and negative conditions. **P4** and **N2** are satisfied because we assume no error, and for the positive condition, we will always choose θ to be a σ-M' flow, so **P1** and **P2** are automatically satisifed. We thus only need to check **P3** and **N1**. We will use $\alpha_t = t + 1$ for the first half of the proof of each claim, and $\alpha_t = 1$ for the second half. We note that we immediately have a variable-time subroutine for the transition map with $T_{u,(u,g_u)} = C_u$.

Proof of Claim 0: To prove Claim 0, we let $\mathsf{w}_{u,\mathsf{M}} := \mathsf{w}_u$. Note that this makes preparing a state proportional to $\sqrt{\mathsf{w}_u}|0\rangle + \sqrt{\mathsf{w}_{u,\mathsf{M}}}|1\rangle$ trivial.

P3: Let θ be the flow from σ to M with minimal energy $\mathcal{R}_{\sigma,M}(G)$ (see Definition 2.3), and extend it to a flow on G' from σ to M' by setting $\theta(u, (u, g_u)) = \theta(u)$ for all $u \in M$. Then:

$$\sum_{(u,v) \in \overrightarrow{E}(G')} \frac{\theta(u,v)^2}{\mathsf{w}_{u,v}} \mathbb{E}\left[\sum_{t=0}^{T_{u,v}} \frac{1}{t+1}\right] = \sum_{(u,v) \in \overrightarrow{E}(G)} \frac{\theta(u,v)^2}{\mathsf{w}_{u,v}} \mathbb{E}\left[\sum_{t=0}^{T_{u,v}} \frac{1}{t+1}\right] + \sum_{u \in M} \frac{\theta(u)^2}{\mathsf{w}_{u,\mathsf{M}}} \mathbb{E}\left[\sum_{t=0}^{C_u} \frac{1}{t+1}\right]$$

$$\leq \mathcal{R}_{\sigma,M}(G)(\ln(\mathsf{T}+1) + 1) + \frac{3}{2} \sum_{u \in M} \frac{\theta(u)^2}{\mathsf{w}_u}$$

(21)

since $C_u = 1$ for all u, by assumption. By Jensen's inequality, we have:

$$\left(\frac{\sum_{v\in\Gamma(u)} \mathsf{w}_{u,v} \frac{\theta(u,v)}{\mathsf{w}_{u,v}}}{\sum_{v\in\Gamma(u)} \mathsf{w}_{u,v}}\right)^2 \leq \frac{\sum_{v\in\Gamma(u)} \mathsf{w}_{u,v} \left(\frac{\theta(u,v)}{\mathsf{w}_{u,v}}\right)^2}{\sum_{v\in\Gamma(u)} \mathsf{w}_{u,v}}$$

$$\frac{\left(\sum_{v\in\Gamma(u)} \theta(u,v)\right)^2}{\sum_{v\in\Gamma(u)} \mathsf{w}_{u,v}} \leq \sum_{v\in\Gamma(u)} \frac{\theta(u,v)^2}{\mathsf{w}_{u,v}}$$

$$\sum_{u\in M} \frac{\theta(u)^2}{\mathsf{w}_u} \leq \sum_{u\in M}\sum_{v\in\Gamma(u)} \frac{\theta(u,v)^2}{\mathsf{w}_{u,v}} \leq \mathcal{R}_{\sigma,M}(G),$$

so we can continue from (21), using $\mathcal{W}(G)\mathcal{R}_{\sigma,M}(G) \leq \mathcal{C}_{\sigma,\mathsf{M}}$:

$$\sum_{(u,v)\in\overrightarrow{E}(G')} \frac{\theta(u,v)^2}{\mathsf{w}_{u,v}} \mathbb{E}\left[\sum_{t=0}^{T_{u,v}} \frac{1}{t+1}\right] \leq 4\frac{\mathcal{C}_{\sigma,\mathsf{M}}}{\mathcal{W}(G)} \ln \mathsf{T} =: \mathcal{R}. \tag{22}$$

N1: We use $\pi(u)P_{u,v} = \frac{\mathsf{w}_{u,v}}{2\mathcal{W}(G)}$ from (13). If $M = \emptyset$, we have:

$$
\begin{aligned}
\sum_{(u,v)\in\overrightarrow{E}(G')} \mathsf{w}_{u,v}\mathbb{E}\left[\sum_{t=0}^{T_{u,v}}(t+1)\right] &= \sum_{(u,v)\in\overrightarrow{E}(G)} \mathsf{w}_{u,v}\mathbb{E}\left[\sum_{t=0}^{T_{u,v}}(t+1)\right] + \sum_{u\in V(G)} \mathsf{w}_{u,\mathsf{M}}\mathbb{E}\left[\sum_{t=0}^{C_u}(t+1)\right] \\
&\leq 2\mathcal{W}(G)\sum_{(u,v)\in\overrightarrow{E}(G)} \pi(u)P_{u,v}\mathbb{E}[T_{u,v}^2] + 2\sum_{u\in V(G)} \mathsf{w}_u \\
&\leq 2\mathcal{W}(G)\mathsf{T}_{\mathrm{avg}}^2 + 2\mathcal{W}(G) =: \mathcal{W}.
\end{aligned}
\tag{23}
$$

To complete the first half of the proof of Claim 0, we apply Theorem 4.1 using the fact that:

$$\sqrt{\mathcal{R}\mathcal{W}} = O\left(\sqrt{\mathcal{C}_{\sigma,\mathsf{M}}\mathsf{T}_{\mathrm{avg}}^2 \log \mathsf{T}}\right).$$

Known $\mathbb{E}[T_{u,v}]$: For the second half of Claim 0, we scale the weights of G' as follows:

$$\forall(u,v)\in\overrightarrow{E}(G), \ \mathsf{w}'_{u,v} := \mathsf{w}_{u,v}\mathbb{E}[T_{u,v}] \quad \text{and} \quad \forall u\in V(G), \ \mathsf{w}'_{u,\mathsf{M}} := \mathsf{w}_{u,\mathsf{M}} = \mathsf{w}_u.$$

Note that by changing the weights, we have changed the star states, but we can generate them by first generating a state proportional to $\sum_{i\in L(u)} \sqrt{\mathsf{w}_{u,i}\mathbb{E}[T_{u,i}]}|i\rangle$, which we assume can be done in unit cost, and then putting a phase in front of each $|i\rangle$ depending on if $i \in L^+(u)$ or $i \in L^-(u)$. We have also changed the total weight of each vertex, $\mathsf{w}'_u = \sum_i \mathsf{w}_{u,i}\mathbb{E}[T_{u,i}]$, but we assume we still have query access to these values.

Then for **P3**, now using $\alpha_t = 1$, in place of (21) and (22), we have:

$$
\begin{aligned}
\sum_{(u,v)\in\overrightarrow{E}(G')} \frac{\theta(u,v)^2}{\mathsf{w}'_{u,v}}\mathbb{E}\left[\sum_{t=0}^{T_{u,v}}1\right] &= \sum_{(u,v)\in\overrightarrow{E}(G)} \frac{\theta(u,v)^2}{\mathsf{w}_{u,v}\mathbb{E}[T_{u,v}]}\mathbb{E}\left[T_{u,v}+1\right] + \sum_{u\in M} \frac{\theta(u)^2}{\mathsf{w}_u}\mathbb{E}\left[C_u+1\right] \\
&\leq 2\mathcal{R}_{\sigma,M}(G) + 2\mathcal{R}_{\sigma,M}(G) \leq 4\frac{\mathcal{C}_{\sigma,\mathsf{M}}}{\mathcal{W}(G)} =: \mathcal{R}',
\end{aligned}
\tag{24}
$$

and for **N1**, in place of (23), we have:

$$
\begin{aligned}
\sum_{(u,v)\in\overrightarrow{E}(G')} \mathsf{w}'_{u,v}\mathbb{E}\left[\sum_{t=0}^{T_{u,v}}1\right] &= \sum_{(u,v)\in\overrightarrow{E}(G)} \mathsf{w}_{u,v}\mathbb{E}[T_{u,v}]\mathbb{E}\left[T_{u,v}+1\right] + \sum_{u\in V(G)} \mathsf{w}_u\mathbb{E}\left[C_u+1\right] \\
&\leq 4\mathcal{W}(G)\sum_{(u,v)\in\overrightarrow{E}(G)} \pi(u)P_{u,v}\mathbb{E}[T_{u,v}]^2 + 2\mathcal{W}(G) \\
&\leq 4\mathcal{W}(G)(\mathsf{T}'_{\mathrm{avg}})^2 + 2\mathcal{W}(G) =: \mathcal{W}'.
\end{aligned}
\tag{25}
$$

Then a simple calculation of $\sqrt{\mathcal{R}'\mathcal{W}'}$ shows that the second part of Claim 0 follows from Theorem 4.1.

Proof of Claim I: For the proof of Claim I, we now let

$$\mathsf{w}_{u,\mathsf{M}} := \frac{\mathcal{W}(G)\mathcal{D}_\mathsf{M}\tau(u)}{\varepsilon\mathcal{C}_{\sigma,\mathsf{M}}} = \frac{\mathcal{W}(G)\mathcal{D}_\mathsf{M}\tau_M(u)\tau(M)}{\varepsilon\mathcal{C}_{\sigma,\mathsf{M}}} \text{ if } u \in M.$$

With this choice of $\mathsf{w}_{u,\mathsf{M}}$, preparing a state proportional to $\sqrt{\mathsf{w}_u}|1\rangle + \sqrt{\mathsf{w}_{u,\mathsf{M}}}|0\rangle$ can be done using a single qubit rotation that depends on $\frac{\mathcal{W}(G)\mathcal{D}_\mathsf{M}}{\varepsilon\mathcal{C}_{\sigma,\mathsf{M}}}\frac{\tau(u)}{\mathsf{w}_u}$.

P3: Let θ be the flow from σ to M with minimal energy $\mathcal{R}_{\sigma,M}(G)$ (see Definition 2.3), and extend it to a flow on G' from σ to M' by setting $\theta(u,(u,g_u)) = \theta(u)$ for all $u \in M$. Then we have:

$$\sum_{(u,v)\in\overrightarrow{E}(G')} \frac{\theta(u,v)^2}{\mathsf{w}_{u,v}}\mathbb{E}\left[\sum_{t=0}^{T_{u,v}}\frac{1}{t+1}\right] = \sum_{(u,v)\in\overrightarrow{E}(G)} \frac{\theta(u,v)^2}{\mathsf{w}_{u,v}}\mathbb{E}\left[\sum_{t=0}^{T_{u,v}}\frac{1}{t+1}\right] + \sum_{u\in M}\frac{\theta(u)^2}{\mathsf{w}_{u,\mathsf{M}}}\mathbb{E}\left[\sum_{t=0}^{C_u}\frac{1}{t+1}\right]$$

$$\leq 2\mathcal{R}_{\sigma,M}(G)\ln\mathsf{T} + 2\frac{\mathcal{C}_{\sigma,\mathsf{M}}}{\mathcal{W}(G)\mathcal{D}_\mathsf{M}}\frac{\varepsilon}{\tau(M)}\sum_{u\in M}\frac{\theta(u)^2}{\tau_M(u)}\ln\mathsf{C} \qquad (26)$$

$$\leq 2\frac{\mathcal{C}_{\sigma,\mathsf{M}}}{\mathcal{W}(G)}\ln\mathsf{T} + 2\frac{\mathcal{C}_{\sigma,\mathsf{M}}}{\mathcal{W}(G)}\ln\mathsf{C} =: \mathcal{R},$$

where we used $\mathcal{W}(G)\mathcal{R}_{\sigma,M} \leq \mathcal{C}_{\sigma,\mathsf{M}}$, $\sum_{u\in M}\theta(u)^2/\tau_M(u) \leq \mathcal{D}_\mathsf{M}$ and $\varepsilon \leq \tau(M)$.

N1: If $M = \emptyset$, we have:

$$\sum_{(u,v)\in\overrightarrow{E}(G')} \mathsf{w}_{u,v}\mathbb{E}\left[\sum_{t=0}^{T_{u,v}}(t+1)\right] = \sum_{(u,v)\in\overrightarrow{E}(G)} \mathsf{w}_{u,v}\mathbb{E}\left[\sum_{t=0}^{T_{u,v}}(t+1)\right] + \sum_{u\in V(G)}\mathsf{w}_{u,\mathsf{M}}\mathbb{E}\left[\sum_{t=0}^{C_u}(t+1)\right]$$

$$\leq 2\mathcal{W}(G)\sum_{(u,v)\in\overrightarrow{E}(G)}\pi(u)P_{u,v}\mathbb{E}[T_{u,v}^2] + \sum_{u\in V(G)}\frac{\mathcal{W}(G)\mathcal{D}_\mathsf{M}\tau(u)}{\varepsilon\mathcal{C}_{\sigma,\mathsf{M}}}\mathbb{E}[C_u^2] \quad (27)$$

$$\leq 2\mathcal{W}(G)\mathsf{T}_{\text{avg}}^2 + \frac{\mathcal{W}(G)\mathcal{D}_\mathsf{M}}{\mathcal{C}_{\sigma,\mathsf{M}}}\frac{1}{\varepsilon}\mathsf{C}_{\text{avg}}^2 =: \mathcal{W}.$$

To complete the first half of the proof of Claim I, we apply Theorem 4.1 using the fact that:

$$\sqrt{\mathcal{R}\mathcal{W}} = O\left(\sqrt{\left(\mathcal{C}_{\sigma,\mathsf{M}}\mathsf{T}_{\text{avg}}^2 + \frac{\mathcal{D}_\mathsf{M}}{\varepsilon}\mathsf{C}_{\text{avg}}^2\right)\log\mathsf{TC}}\right).$$

Known $\mathbb{E}[T_{u,v}]$ and $\mathbb{E}[C_u]$: For the second half of Claim I, we scale the weights of G' as follows:

$$\forall(u,v)\in\overrightarrow{E}(G),\ \mathsf{w}_{u,v}' := \mathsf{w}_{u,v}\mathbb{E}[T_{u,v}] \text{ and } \forall u \in V(G),\ \mathsf{w}_{u,\mathsf{M}}' := \mathsf{w}_{u,\mathsf{M}}\mathbb{E}[C_u] = \frac{\mathcal{W}(G)\mathcal{D}_\mathsf{M}\tau(u)\mathbb{E}[C_u]}{\varepsilon\mathcal{C}_{\sigma,\mathsf{M}}}.$$

$$(28)$$

Note that by changing the weights, we have changed the star states, but we can generate them as follows. Let $\mathsf{w}_u' = \sum_{i\in L(u)}\mathsf{w}_{u,i}\mathbb{E}[T_{u,i}]$, which we can query, by assumption. Then we first prepare a state proportional to $\sqrt{\mathsf{w}_u'}|1\rangle + \sqrt{\mathsf{w}_{u,\mathsf{M}}\mathbb{E}[C_u]}|0\rangle$, using a single qubit rotation, and then map $|1\rangle$ to a state proportional to $\sum_{i\in L(u)}\sqrt{\mathsf{w}_{u,i}\mathbb{E}[T_{u,i}]}|i\rangle$, which we assume can be done in unit cost. Then all that remains to be done is to put a phase in front of each $|i\rangle$ depending on if $i \in L^+(u)$ or $i \in L^-(u)$.

Then for the positive analysis, now using $\alpha_t = 1$, in place of (26), we have:

$$\sum_{(u,v)\in\overrightarrow{E}(G')} \frac{\theta(u,v)^2}{\mathsf{w}_{u,v}'}\mathbb{E}\left[\sum_{t=0}^{T_{u,v}}1\right] = \sum_{(u,v)\in\overrightarrow{E}(G)}\frac{\theta(u,v)^2}{\mathsf{w}_{u,v}\mathbb{E}[T_{u,v}]}\mathbb{E}[T_{u,v}+1] + \sum_{u\in M}\frac{\theta(u)^2}{\mathsf{w}_{u,\mathsf{M}}\mathbb{E}[C_u]}\mathbb{E}[C_u+1]$$

$$\leq 2\mathcal{R}_{\sigma,M}(G) + 2\frac{\mathcal{C}_{\sigma,\mathsf{M}}}{\mathcal{W}(G)\mathcal{D}_\mathsf{M}}\frac{\varepsilon}{\tau(M)}\sum_{u\in M}\frac{\theta(u)^2}{\tau_M(u)} \qquad (29)$$

$$\leq 2\frac{\mathcal{C}_{\sigma,\mathsf{M}}}{\mathcal{W}(G)} + 2\frac{\mathcal{C}_{\sigma,\mathsf{M}}}{\mathcal{W}(G)} =: \mathcal{R}',$$

25

and for the negative analysis, in place of (27), we have:

$$\sum_{(u,v)\in\overrightarrow{E}(G')} \mathsf{w}'_{u,v}\mathbb{E}\left[\sum_{t=0}^{T_{u,v}}1\right] = \sum_{(u,v)\in\overrightarrow{E}(G)} \mathsf{w}_{u,v}\mathbb{E}[T_{u,v}]\mathbb{E}\left[T_{u,v}+1\right] + \sum_{u\in V(G)} \mathsf{w}_{u,\mathsf{M}}\mathbb{E}[C_u]\mathbb{E}\left[C_u+1\right]$$

$$\leq 4\mathcal{W}(G)\sum_{(u,v)\in\overrightarrow{E}(G)} \pi(u)P_{u,v}\mathbb{E}[T_{u,v}]^2 + 2\sum_{u\in V(G)} \frac{\mathcal{W}(G)\mathcal{D}_{\mathsf{M}}}{\varepsilon\mathcal{C}_{\sigma,\mathsf{M}}}\tau(u)\mathbb{E}[C_u]^2 \quad (30)$$

$$\leq 4\mathcal{W}(G)(\mathsf{T}'_{\mathrm{avg}})^2 + 2\frac{\mathcal{W}(G)\mathcal{D}_{\mathsf{M}}}{\mathcal{C}_{\sigma,\mathsf{M}}}\frac{1}{\varepsilon}(\mathsf{C}'_{\mathrm{avg}})^2 =: \mathcal{W}'.$$

Then a simple calculation of $\sqrt{\mathcal{R}'\mathcal{W}'}$ show that the second part of Claim I follows from Theorem 4.1.

Proof of Claim II: Next, we consider the special case where there is a unique marked element if $M \neq \emptyset$. In that case, when $M = \{m\}$ is non-empty, we have, for any σ-M flow, $\theta(m) = -1$ (all flow leaves at m), so:

$$\sum_{u\in M} \frac{\theta(u)^2}{\tau_M(u)} = \frac{1}{\tau_M(m)} = \frac{\tau(M)}{\tau(m)} = 1.$$

Thus Claim II follows from applying Claim I with $\mathcal{D}_{\mathsf{M}} = 1$.

Proof of Claim III: Finally, let τ be the known distribution described in Claim III. If such a τ is known, we can define

$$\mathsf{w}_{u,\mathsf{M}} := \frac{\tau(u)\mathcal{W}(G)}{\varepsilon\mathcal{C}_{\sigma,\tau}} = \frac{\tau_M(u)\tau(M)\mathcal{W}(G)}{\varepsilon\mathcal{C}_{\sigma,\tau}} \text{ when } u \in M.$$

With this choice of $\mathsf{w}_{u,\mathsf{M}}$, preparing a state proportional to $\sqrt{\mathsf{w}_u}|1\rangle + \sqrt{\mathsf{w}_{u,\mathsf{M}}}|0\rangle$ can be done using a single qubit rotation that depends on $\frac{\mathcal{W}(G)}{\varepsilon\mathcal{C}_{\sigma,\tau}}\frac{\tau(u)}{\mathsf{w}_u}$.

P3: Let θ be the flow from σ to τ_M with minimal energy $\mathcal{R}_{\sigma,\tau_M}(G)$ (see Definition 2.4), and extend it to a flow on G' from σ to M' by setting $\theta(u,(u,g_u)) = \tau_M(u)$ for all $u \in M$. Then, similar to (26), we have:

$$\sum_{(u,v)\in\overrightarrow{E}(G')} \frac{\theta(u,v)^2}{\mathsf{w}_{u,v}}\mathbb{E}\left[\sum_{t=0}^{T_{u,v}}\frac{1}{t+1}\right] \leq 2\sum_{(u,v)\in\overrightarrow{E}(G)} \frac{\theta(u,v)^2}{\mathsf{w}_{u,v}}\ln\mathsf{T} + 2\sum_{u\in M} \frac{\tau_M(u)^2}{\mathsf{w}_{u,\mathsf{M}}}\ln\mathsf{C}$$

$$= 2\mathcal{R}_{\sigma,\tau_M}(G)\ln\mathsf{T} + 2\frac{\varepsilon}{\tau(M)}\frac{\mathcal{C}_{\sigma,\tau}}{\mathcal{W}(G)}\sum_{u\in M} \frac{\tau_M(u)^2}{\tau_M(u)}\ln\mathsf{C}$$

$$\leq 2\frac{\mathcal{C}_{\sigma,\tau}}{\mathcal{W}(G)}\ln\mathsf{T} + 2\frac{\mathcal{C}_{\sigma,\tau}}{\mathcal{W}(G)}\ln\mathsf{C} =: \mathcal{R}.$$

N1: If $M = \emptyset$, we have, similar to (27):

$$\sum_{(u,v)\in\overrightarrow{E}(G')} \mathsf{w}_{u,v}\mathbb{E}\left[\sum_{t=0}^{T_{u,v}}t+1\right] \leq 2\sum_{(u,v)\in\overrightarrow{E}(G)} \mathsf{w}_{u,v}\mathbb{E}[T_{u,v}^2] + 2\sum_{u\in V(G)} \mathsf{w}_{u,\mathsf{M}}\mathbb{E}[C_u^2]$$

$$\leq 4\mathcal{W}(G)\mathsf{T}_{\mathrm{avg}}^2 + \frac{\mathcal{W}(G)}{\mathcal{C}_{\sigma,\tau}}\frac{1}{\varepsilon}\mathsf{C}_{\mathrm{avg}}^2 =: \mathcal{W}.$$

To complete the proof of the first half of Claim III, we apply Theorem 4.1 using the fact that

$$\sqrt{\mathcal{R}\mathcal{W}} = O\left(\left(\mathcal{C}_{\sigma,\tau}\mathsf{T}_{\mathrm{avg}}^2 + \frac{1}{\varepsilon}\mathsf{C}_{\mathrm{avg}}^2\right)\log\mathsf{T}\mathsf{C}\right).$$

The proof of the second half of Claim III proceeds just as in Claim I. $\qquad\square$

The expressions we achieve in Corollary 4.2 in the case when checking is non-trivial and M is not restricted to being a singleton are somewhat complicated. However, in the special case when $\sigma = \pi$, we can get a bound comparable to the MNRS framework. The following corollary also gives an alternative proof to the one in [AGJ20] that the MNRS framework is a special case of the electric network framework (up to $\log 1/\pi_{\min}$ factors, which we suspect can be removed).

Corollary 4.3 (Variable-time MNRS). *Fix a network G and marked set $M \subset V(G)$ that may implicitly depend on an input x. Let P be the transition matrix for the random walk on G, and π its stationary distribution. Let S be the cost of generating $|\pi\rangle$, and suppose $\{|\psi^G_\star(u)\rangle\}_{u \in V(G)}$ can be generated in unit cost. Let δ be a lower bound on the spectral gap of P, and ε a lower bound on $\pi(M)$ whenever $M \neq \emptyset$. Let $\pi_{\min} = \min_{u \in V(G)} \pi(u)$.*

Suppose there is a variable-time subroutine that checks, for any $u \in V(G)$, if $u \in M$, with stopping times $\{C_u\}_{u \in V(G)}$, which are random variables on $[\mathsf{C}]$, and all errors zero. Let $\mathsf{C}_{\mathrm{avg}}$ and $\mathsf{C}'_{\mathrm{avg}}$ be upper bounds such that whenever $M = \emptyset$,

$$\sqrt{\sum_{u \in V(G)} \pi(u)\mathbb{E}[C_u^2]} \leq \mathsf{C}_{\mathrm{avg}}, \quad and \quad \sqrt{\sum_{u \in V(G)} \pi(u)\mathbb{E}[C_u]^2} \leq \mathsf{C}'_{\mathrm{avg}}.$$

Suppose there is a reversible variable-time subroutine that implements the transition map with stopping times $\{T_{u,v}\}_{(u,v) \in \overrightarrow{E}(G)}$, which are random variables on $[\mathsf{T}]$, and all errors zero. Let $\mathsf{T}_{\mathrm{avg}}$ and $\mathsf{T}'_{\mathrm{avg}}$ be upper bounds such that whenever $M = \emptyset$,

$$\sqrt{\sum_{(u,v) \in \overrightarrow{E}(G)} \pi(u)P_{u,v}\mathbb{E}[T_{u,v}^2]} \leq \mathsf{T}_{\mathrm{avg}}, \quad and \quad \sqrt{\sum_{(u,v) \in \overrightarrow{E}(G)} \pi(u)P_{u,v}\mathbb{E}[T_{u,v}]^2} \leq \mathsf{T}'_{\mathrm{avg}}.$$

Then there is a quantum algorithm that detects if $M = \emptyset$ with bounded error in complexity:

$$O\left(\mathsf{S} + \frac{1}{\sqrt{\varepsilon}}\left(\frac{1}{\sqrt{\delta}}\mathsf{T}_{\mathrm{avg}} + \mathsf{C}_{\mathrm{avg}}\right)\sqrt{\log\frac{1}{\pi_{\min}}}\log^{1.5}(\mathsf{TC})\right).$$

Suppose in addition that the values $\mathbb{E}[T_{u,v}] = \mathbb{E}[T_{u,i}]$ are computable in the strong sense that for any u, we can generate a superposition proportional to $\sum_{i \in L(u)} \sqrt{\mathsf{w}_{u,i}\mathbb{E}[T_{u,i}]}|i\rangle$, and we have query access to $\mathsf{w}'_u = \sum_{i \in L(u)} \mathsf{w}_{u,i}\mathbb{E}[T_{u,i}]$ and $\mathbb{E}[C_u]$. Then there is a quantum algorithm that decides if $M = \emptyset$ with bounded error in complexity:

$$O\left(\mathsf{S} + \frac{1}{\sqrt{\varepsilon}}\left(\frac{1}{\sqrt{\delta}}\mathsf{T}'_{\mathrm{avg}} + \mathsf{C}'_{\mathrm{avg}}\right)\sqrt{\log\frac{1}{\pi_{\min}}}\log(\mathsf{TC})\right).$$

Proof. As in the proof of Corollary 4.2, we will apply Theorem 4.1 to a modified graph G' in which we connect each $u \in V(G)$ to a new vertex (u, g_u), where $g_u = 1$ if and only if $u \in M$, by an edge of weight $\mathsf{w}_{u,\mathsf{M}} = \delta\mathsf{w}_u$.

We will use the following process – which is similar to a random walk on G' starting from π except for the log factor in p – to design a flow on G'.

1. Sample a vertex u from π.

2. Repeat:

    (a) With probability $p = \frac{\delta}{\log(2/\pi_{\min}^2)}$, check if $u \in M$, and if so, halt.

    (b) Take a step to some v according to P (the random walk on G) by setting $u \leftarrow v$.

This process gives rise to a Markov chain $X_1, \ldots, X_\tau$ on $V(G)$ for a random variable τ with

$$\mathbb{E}[\tau] = O\left(\frac{1}{p\pi(M)}\right) = O\left(\frac{\log\frac{1}{\pi_{\min}}}{\varepsilon\delta}\right).$$

From this process, we can define a flow on G as follows. Let $[u \to v]$ be the number of times the above process moves from u to v and define for any $\{u, v\} \in E(G)$:

$$\theta(u, v) := \mathbb{E}[u \to v] - \mathbb{E}[v \to u].$$

This satisfies $\theta(u, v) = -\theta(v, u)$, and letting $N_u^{\to}$ be the number of times we leave u, and $N_u^{\leftarrow}$ be the number of times we enter u, we have:

$$\theta(u) = \sum_v \theta(u, v) = (\mathbb{E}[N_u^{\to}] - \mathbb{E}[N_u^{\leftarrow}]) = \Pr[X_1 = u] - \Pr[X_\tau = u].$$

It is clear that $\Pr[X_1 = u] = \pi(u)$. We also only ever output if $u \in M$, so $\Pr[X_\tau = u] = 0$ unless $u \in M$. We can extend this to a flow on G' by setting $\theta(u, (u, 1)) = \theta(u)$ for all $u \in M$. Let $\tilde{\pi}$ be the distribution on M defined $\tilde{\pi}(u) = \pi(u) - \theta(u)$ for all $u \in M$ – in other words, $\tilde{\pi}(u)$ is the probability that $X_\tau = u$. Since there are $1/p = \frac{1}{\delta} \log \frac{1}{\pi_{\min}^2}$ steps between checks, in expectation, each time we check if $u \in M$, we are distributed according to some distribution π' that is expected $\pi_{\min}/2$-close to π [LWP09, Theorem 12.3], meaning that for each $u \in M$

$$\mathbb{E}[\pi'_M(u)] \leq \frac{\pi(u) + \pi_{\min}/2}{\pi(M) - \pi_{\min}/2} \leq \frac{\frac{3}{2}\pi(u)}{\frac{1}{2}\pi(M)} = 3\pi_M(u)$$

and so $\tilde{\pi}(u) \leq 3\pi_M(u)$, and thus:

$$|\theta(u)| \leq \tilde{\pi}(u) + \pi(u) \leq 3\pi_M(u) + \pi(M)\pi_M(u) \leq 4\pi_M(u).$$

Then we have (using $\alpha_t = t + 1$):

$$
\begin{aligned}
2\mathcal{W}(G) \sum_{(u,v) \in \vec{E}(G')} \frac{\theta(u,v)^2}{\mathsf{w}_{u,v}} \mathbb{E}\left[\sum_{t=0}^{T_{u,v}} \frac{1}{t+1}\right] &\leq \sum_{(u,v) \in \vec{E}(G)} \frac{\theta(u,v)^2}{\pi(u)P_{u,v}} \mathbb{E}[\log T_{u,v}] + \sum_{u \in M} \frac{\theta(u)^2}{\delta\pi(u)} \mathbb{E}[\log C_u] \\
&\leq \sum_{(u,v) \in \vec{E}(G)} \frac{\theta(u,v)^2}{\pi(u)P_{u,v}} \log \mathsf{T} + \frac{16}{\delta\pi(M)^2} \sum_{u \in M} \pi(u) \log \mathsf{C} \\
&\leq \sum_{(u,v) \in \vec{E}(G)} \frac{\theta(u,v)^2}{\pi(u)P_{u,v}} \log \mathsf{T} + \frac{16}{\delta\varepsilon} \log \mathsf{C}.
\end{aligned}
$$

Above, we used $\mathsf{w}_{u,v} = \pi(u)P_{u,v}/(2\mathcal{W}(G))$, $\theta(u)^2 \leq 16\pi_M(u)^2 = 16\pi(u)^2/\pi(M)^2$, and $\varepsilon \leq \pi(M) = \sum_{u \in M} \pi(u)$. We will show that,

$$\sum_{(u,v) \in \vec{E}(G)} \frac{\theta(u,v)^2}{\pi(u)P_{u,v}} \leq \mathbb{E}[\tau] \leq \frac{c \log \frac{1}{\pi_{\min}}}{\varepsilon\delta} \tag{31}$$

for some constant c, from which it follows that

$$\sum_{(u,v) \in \vec{E}(G')} \frac{\theta(u,v)^2}{\mathsf{w}_{u,v}} \mathbb{E}\left[\sum_{t=0}^{T_{u,v}} \frac{1}{t+1}\right] \leq \frac{1}{2\mathcal{W}(G)} \frac{1}{\varepsilon\delta}\left(c \log \frac{1}{\pi_{\min}} \log \mathsf{T} + 16 \log \mathsf{C}\right) =: \mathcal{R}.$$

Then since:

$$\sum_{(u,v) \in \vec{E}(G')} \mathsf{w}_{u,v} \mathbb{E}\left[\sum_{t=0}^{T_{u,v}}(t+1)\right] \leq 2\mathcal{W}(G)\left(\sum_{(u,v) \in \vec{E}(G)} \pi(u)P_{u,v}\mathbb{E}[T_{u,v}^2] + \sum_{u \in V(G)} \delta\pi(u)\mathbb{E}[C_u^2]\right) =: \mathcal{W}$$

28

we have

$$\mathcal{RW} = O\left(\left(\frac{1}{\varepsilon\delta}\sum_{(u,v)\in\vec{E}(G)}\pi(u)P_{u,v}\mathbb{E}[T_{u,v}^2] + \frac{1}{\varepsilon}\sum_{u\in V(G)}\pi(u)\mathbb{E}[C_u^2]\right)\log\frac{1}{\pi_{\min}}\log\mathsf{TC}\right)$$

and so the claim follows from Theorem 4.1. Thus, it remains only to establish (31), which we now undertake.

Let $[u\to v]_t$ be the event that we move from u to v in the t-th step, and let $E_t(u)$ be the event that $X_t = u$, and $\tau > t$ (i.e. we don't stop there). Then:

$$\mathbb{E}[u\to v] = \sum_{t=1}^{\infty}\Pr[[u\to v]_t] = \sum_{t=1}^{\infty}\Pr[E_t(u)]P_{u,v}.$$

Then since $\theta(u,v) = \mathbb{E}[u\to v] - \mathbb{E}[v\to u]$, by definition, we have:

$$\begin{aligned}
\sum_{(u,v)\in\vec{E}(G)}\frac{\theta(u,v)^2}{\pi(u)P_{u,v}} &= \frac{1}{2}\sum_{u\in V(G)}\sum_{v\in\Gamma(u)}\pi(u)P_{u,v}\left(\frac{\mathbb{E}[u\to v]}{\pi(u)P_{u,v}} - \frac{\mathbb{E}[v\to u]}{\pi(v)P_{v,u}}\right)^2 \\
&= \frac{1}{2}\sum_{u\in V(G)}\sum_{v\in\Gamma(u)}\pi(u)P_{u,v}\left(\frac{1}{\pi(u)}\sum_{t=1}^{\infty}\Pr[E_t(u)] - \frac{1}{\pi(v)}\sum_{t=1}^{\infty}\Pr[E_t(v)]\right)^2 \quad (32) \\
&= \frac{1}{2}\sum_{u\in V(G)}\sum_{v\in\Gamma(u)}\pi(u)P_{u,v}\left(\frac{\mathbb{E}[N_u^{\to}]}{\pi(u)} - \frac{\mathbb{E}[N_v^{\to}]}{\pi(v)}\right)^2.
\end{aligned}$$

From (32) and the detailed balance condition $\pi(u)P_{u,v} = \pi(v)P_{v,u}$, it follows that:

$$\begin{aligned}
\sum_{(u,v)\in\vec{E}(G)}\frac{\theta(u,v)^2}{\pi(u)P_{u,v}} &= \frac{1}{2}\sum_{u\in V(G)}\sum_{v\in\Gamma(u)}\pi(u)P_{u,v}\left(\frac{\mathbb{E}[N_u^{\to}]^2}{\pi(u)^2} + \frac{\mathbb{E}[N_v^{\to}]^2}{\pi(v)^2} - 2\frac{\mathbb{E}[N_u^{\to}]}{\pi(u)}\frac{\mathbb{E}[N_v^{\to}]}{\pi(v)}\right) \\
&= \sum_{u\in V(G)}\sum_{v\in\Gamma(u)}\pi(u)P_{u,v}\frac{\mathbb{E}[N_u^{\to}]^2}{\pi(u)^2} - \sum_{u\in V(G)}\sum_{v\in\Gamma(u)}\pi(u)P_{u,v}\frac{\mathbb{E}[N_u^{\to}]}{\pi(u)}\frac{\mathbb{E}[N_v^{\to}]}{\pi(v)} \quad (33) \\
&= \sum_{u\in V(G)}\sum_{v\in\Gamma(u)}P_{u,v}\mathbb{E}[N_u^{\to}]\left(\frac{\mathbb{E}[N_u^{\to}]}{\pi(u)} - \frac{\mathbb{E}[N_v^{\to}]}{\pi(v)}\right).
\end{aligned}$$

The number of visits to u depends on if it is visited at step 1, which happens with probability $\pi(u)$, as well as the number of visits to its neighbours. Every time we leave a neighbour v of u, we visit u next with probability $P_{v,u}$. Thus, the expected number of visits to u is:

$$\pi(u) + \sum_{v\in\Gamma(u)}P_{v,u}\mathbb{E}[N_v^{\to}].$$

Then since we leave u if we visit and don't halt, the expected number of times we leave $u\in\overline{M}$ is

$$\mathbb{E}[N_u^{\to}] = \pi(u) + \sum_{v\in\Gamma(u)}P_{v,u}\mathbb{E}[N_v^{\to}], \quad (34)$$

since we never halt on $u\in\overline{M}$; and for $u\in M$, since every visits ends in halting with probability p,

$$\mathbb{E}[N_u^{\to}] = (1-p)\pi(u) + (1-p)\sum_{v\in\Gamma(u)}P_{v,u}\mathbb{E}[N_v^{\to}]$$

so $\mathbb{E}[N_u^{\to}] - \sum_{v\in\Gamma(u)}P_{v,u}\mathbb{E}[N_v^{\to}] = \pi(u) - \frac{p}{1-p}\mathbb{E}[N_u^{\to}]. \quad (35)$

Thus:

$$\sum_{v \in \Gamma(u)} P_{u,v} \left(\frac{\mathbb{E}[N_u^{\rightarrow}]}{\pi(u)} - \frac{\mathbb{E}[N_v^{\rightarrow}]}{\pi(v)} \right) = \frac{\mathbb{E}[N_u^{\rightarrow}]}{\pi(u)} - \sum_{v \in \Gamma(u)} P_{u,v} \frac{\mathbb{E}[N_v^{\rightarrow}]}{\pi(v)}$$

$$= \frac{1}{\pi(u)} \left(\mathbb{E}[N_u^{\rightarrow}] - \sum_{v \in \Gamma(u)} P_{v,u} \mathbb{E}[N_v^{\rightarrow}] \right) \qquad \text{since } \frac{P_{u,v}}{\pi(v)} = \frac{P_{v,u}}{\pi(u)}$$

$$= \begin{cases} 1 & \text{if } u \in \overline{M}, \text{ by (34)} \\ 1 - \frac{p}{1-p} \frac{\mathbb{E}[N_u^{\rightarrow}]}{\pi(u)} & \text{if } u \in M, \text{ by (35).} \end{cases}$$

Thus, continuing from (33), we have

$$\sum_{(u,v) \in \overrightarrow{E}(G)} \frac{\theta(u,v)^2}{\pi(u) P_{u,v}} = \sum_{u \in \overline{M}} \mathbb{E}[N_u^{\rightarrow}] + \sum_{u \in M} \mathbb{E}[N_u^{\rightarrow}] \left(1 - \frac{p}{1-p} \frac{\mathbb{E}[N_u^{\rightarrow}]}{\pi(u)} \right)$$

$$\leq \sum_{u \in V(G)} \mathbb{E}[N_u^{\rightarrow}] = \mathbb{E}[\tau],$$

establishing (31). $\square$

We now turn to the special case of search, where we have the following corollary, again assuming, for simplicity, that there are no errors:

Corollary 4.4 (Variable-time Search $(\alpha_t = t+1)$)**.** *Fix a search space $[n]$ and a distribution π on $[n]$ such that we can generate $|\pi\rangle$ in unit cost. Suppose we have a variable-time subroutine that computes, for any $i \in [n]$, a bit $g(i)$, in time T_i, which is a random variable on $[\mathsf{T}]$. Assume all errors are 0. Suppose $\varepsilon > 0$ is a lower bound on $\sum_{i:g(i)=1} \pi(i)$ whenever $\bigvee_{i \in [n]} g(i) = 1$, and T_{avg} and T'_{avg} are upper bounds such that whenever $\bigvee_{i \in [n]} g(i) = 0$:*

$$\sqrt{\sum_{i \in [n]} \pi(i) \mathbb{E}[T_i^2]} \leq \mathsf{T}_{\text{avg}} \quad and \quad \sqrt{\sum_{i \in [n]} \pi(i) \mathbb{E}[T_i]^2} \leq \mathsf{T}'_{\text{avg}}.$$

Then there is a quantum algorithm that decides $\bigvee_{i \in [n]} g(i)$ with bounded error in complexity:

$$O \left(\frac{1}{\sqrt{\varepsilon}} \mathsf{T}_{\text{avg}} \log^{1.5} \mathsf{T} \right).$$

If, in addition, we assume the values $\mathbb{E}[T_i]$ are known, in the strong sense that we can generate a superposition proportional to $\sum_{i \in [n]} \sqrt{\pi(i) \mathbb{E}[T_i]} |i\rangle$ in unit cost, there is a quantum algorithm that decides $\bigvee_{i \in [n]} g(i)$ with bounded error in complexity:

$$O \left(\frac{1}{\sqrt{\varepsilon}} \mathsf{T}'_{\text{avg}} \log \mathsf{T} \right).$$

Proof. First note that, letting $A|i\rangle = (-1)^{g(i)}$, a variable-time algorithm for g is trivially a reversible variable-time algorithm (just use $\tilde{U}_t = U_t$). We will define a graph G as follows. Let $V(G) = \{u_0, \ldots, u_n\}$, where $u_0 = (0,0)$ and for all $i \in [n]$, $u_i = (i, g(i))$. Let $V_0 = \{u_0\}$, so σ is just a point function, and $M = \{u_i : g(i) = 1\}$. The graph will be a star with u_0 the centre, connected to each u_i by an edge of weight $\pi(i)$. Define label sets $L(u_0) = L^+(u_0) = [n]$, and for all $i \in [n]$, $L(u_i) = L^-(u_i) = \{\leftarrow\}$, and transition function $f_{u_0}(i) = (i, g(i))$, $f_{u_i}(\leftarrow) = u_0$ for $i \in [n]$. Thus, the variable-time subroutine can be used to implement the edge transition subroutine in this graph.

We note that it is trivial to generate $|\sigma\rangle = |u_0\rangle$, as well as to check if (i, b) is marked, by checking if $b = 1$. Generating star states is trivial for all $\{u_i : i \in [n]\}$, each of which has a single neighbour, and generating $|\psi_\star(u_0)\rangle = |\pi\rangle$ is assumed to have unit cost. Thus, to apply Theorem 4.1, we just need to check the positive and negative conditions. We will use $\alpha_t = t+1$ for the first part of the proof.

Positive Condition: Define $\pi(M) := \sum_{u_i \in M} \pi(i) \geq \varepsilon$. Let θ be defined as $\theta(u_0, u_i) = \frac{\pi(i)}{\pi(M)}$ for all $u_i \in M$, and 0 otherwise. Then it is easy to see that θ satisfies **P1** and **P2**. **P4** is trivially satisfied, because all errors are 0. For **P3**, we have, using $\alpha_t = t + 1$:

$$\sum_{(u,v) \in \overrightarrow{E}(G)} \frac{\theta(u,v)^2}{\mathsf{w}_{u,v}} \mathbb{E}\left[\sum_{t=0}^{T_{u,v}} \frac{1}{t+1}\right] \leq \frac{2}{\pi(M)^2} \sum_{u_i \in M} \frac{\pi(i)^2}{\pi(i)} \mathbb{E}[\ln T_{u,v}] \leq \frac{2}{\varepsilon} \ln \mathsf{T} =: \mathcal{R}.$$

Negative Condition: N2 is trivially satisfied because all errors are 0. For **N1**, we have:

$$\sum_{(u,v) \in \overrightarrow{E}(G)} \mathsf{w}_{u,v} \mathbb{E}\left[\sum_{t=0}^{T_{u,v}} (t+1)\right] \leq \sum_{i \in [n]} \pi(i) \mathbb{E}\left[T_{u,v}^2\right] \leq \mathsf{T}_{\mathrm{avg}}^2 =: \mathcal{W}.$$

The stated complexity follows from Theorem 4.1

For the case when the expected stopping times are known, we can set $\mathsf{w}_{u_0,u_i} = \pi(i)\mathbb{E}[T_i]$, but this means that we need to be able to generate a state proportional to $|\psi_\star(u_0)\rangle = \sum_{i \in [n]} \sqrt{\pi(i)\mathbb{E}[T_i]}|i\rangle$ in unit time, which is precisely what we assume. Then to apply Theorem 4.1, we just need to check the positive and negative conditions. We will now use $\alpha_t = 1$.

Positive Condition: Let θ be defined as $\theta(u_0, u_i) = \frac{\pi(i)}{\pi(M)}$ for all $u_i \in M$, and 0 otherwise. Then it is easy to see that θ satisfies **P1** and **P2**. **P4** is trivially satisfied, because all errors are 0. For **P3**, we have:

$$\sum_{(u,v) \in \overrightarrow{E}(G)} \frac{\theta(u,v)^2}{\mathsf{w}_{u,v}} \mathbb{E}\left[\sum_{t=0}^{T_{u,v}} 1\right] = \frac{1}{\pi(M)^2} \sum_{u_i \in M} \frac{\pi(i)^2}{\pi(i)\mathbb{E}[T_i]} \mathbb{E}\left[T_i + 1\right] \leq \frac{2}{\varepsilon} =: \mathcal{R}'.$$

Negative Condition: N2 is trivially satisfied because all errors are 0. For **N1**, we have:

$$\sum_{(u,v) \in \overrightarrow{E}(G)} \mathsf{w}_{u,v} \mathbb{E}\left[\sum_{t=0}^{T_{u,v}} 1\right] = \sum_{i \in [n]} \pi(i) \mathbb{E}\left[T_{u,v}\right] \mathbb{E}\left[T_{u,v} + 1\right] \leq 2(\mathsf{T}_{\mathrm{avg}}')^2 =: \mathcal{W}'.$$

The stated complexity follows from Theorem 4.1. $\qquad\qquad\qquad\qquad\qquad\qquad\square$

We proved the first part of Corollary 4.4 (i.e. $\mathbb{E}[T_i]$ unknown) by applying Theorem 4.1 with the setting $\alpha_t = t + 1$. By setting $\alpha_t = 1$, or $\alpha_t = 1/(t+1)$, we can get similar results with different complexities, as shown in Table 1. The proofs proceed as in the $\alpha_t = t + 1$ case, except that we use different flow settings, θ, to upper bound **P3**. For the $\alpha_t = 1$ case, we obtain the expression in Table 1 row (2) by setting $\theta(u_0, u_i) = \frac{\pi(i)/\mathbb{E}[T_i]}{\sum_{j \in M} \pi(j)/\mathbb{E}[T_j]}$. For the $\alpha_t = 1/(t+1)$ case, we obtain the expression in Table 1 row (3) by setting $\theta(u_0, u_i) = \frac{\pi(i)/\mathbb{E}[T_i^2]}{\sum_{j \in M} \pi(j)/\mathbb{E}[T_j^2]}$. One can verify that these settings of θ satsify the conditions of Theorem 4.1. Analogous to the second part of Corollary 4.4, if the values $\mathbb{E}[T_i]$ are known to the extent that we can efficiently generate a state proportional to $\sum_{i \in [n]} \sqrt{\pi(i)/\mathbb{E}[T_i]}|i\rangle$, which allows the edge weights to be set to $\mathsf{w}_{u_0,u_i} = \pi(i)/\mathbb{E}[T_i]$, then by using $\alpha_t = 1$ and $\theta(u_0, u_i) = \frac{\pi(i)/\mathbb{E}[T_i]^2}{\sum_{j \in M} \pi(j)/\mathbb{E}[T_j]^2}$, we can get a complexity like the one in (3), but better by a $\sqrt{\log \mathsf{T}}$ factor, and with $\mathbb{E}[T_i]^2$ in place of $\mathbb{E}[T_i^2]$.

We stress that these results could also be applied to other quantum walk search algorithms, for example, in a version of element distinctness in which the query costs are variable. We leave specific applications for future work.

	α_t	Complexity	Special case $M=\{m\}$, $\pi(m)\geq\varepsilon$
(1)	$\alpha_t = t+1$	$\sqrt{\dfrac{\sum_{i\in[n]}\pi(i)\mathbb{E}[T_i^2]}{\min_{M:M\neq\emptyset}\sum_{i\in M}\pi(i)}}$	$\sqrt{\dfrac{1}{\varepsilon}\sum_{i\in[n]}\pi(i)\mathbb{E}[T_i^2]}$
(2)	$\alpha_t = 1$	$\sqrt{\dfrac{\sum_{i\in[n]}\pi(i)\mathbb{E}[T_i]}{\min_{M:M\neq\emptyset}\sum_{i\in M}\frac{\pi(i)}{\mathbb{E}[T_i]}}}$	$\sqrt{\dfrac{1}{\varepsilon}\sum_{i\in[n]}\pi(i)\mathbb{E}[T_i]\mathbb{E}[T_m]}$
(3)	$\alpha_t = 1/(t+1)$	$\dfrac{1}{\sqrt{\min_{M:M\neq\emptyset}\sum_{i\in M}\frac{\pi(i)}{\mathbb{E}[T_i^2]}}}$	$\sqrt{\dfrac{1}{\varepsilon}\mathbb{E}[T_m^2]}$

Table 1: Here we show three different versions of Corollary 4.4 that can be obtained by different settings of α_t. (1) corresponds to Corollary 4.4. We neglect log factors. For (3), as with (1) (Corollary 4.4) if the weights are known, we can get a similar expression with $\mathbb{E}[T_i]^2$ instead of $\mathbb{E}[T_i^2]$, which is better when the variance of T_i is large, and a slightly better dependence on log T. To allow a comparison of these three complexities, we show the special case where there is a single marked element $M = \{m\}$, promised to have weight at least ε. In that case, we can see that which of the three expressions is optimal depends on how $\mathbb{E}[T_m]$ compares to the average: Suppose for simplicity that the stopping times have 0 variance, so the T_i are just some natural numbers. If $T_m > \sum_{i\in[n]}\pi(i)T_i^2 / \sum_{i\in[n]}\pi(i)T_i$, then (1) is the smallest. If $\sum_{i\in[n]}\pi(i)T_i < T_m < \sum_{i\in[n]}\pi(i)T_i^2 / \sum_{i\in[n]}\pi(i)T_i$, then (2) is the smallest. And finally, if $T_m < \sum_{i\in[n]}\pi(i)T_i$, then (3) is the smallest. This does not guarantee that each of the three variable-time search results is useful. This is somehow just saying that if we know T_m is relatively small for the marked m, then we should not spend too much time in any subroutine, whereas if we know T_m is relatively large, then we should. Perhaps there would be a simpler way to exploit such information.

4.1 Proof of Theorem 4.1

We start by defining the parameters of the phase estimation algorithm (see Section 2.2) that will prove Theorem 4.1. Our algorithm will work on the space:

$$H = \mathrm{span}\left\{|\!\rightarrow\rangle|u,i\rangle|0,0\rangle|0\rangle, |\!\leftarrow\rangle|u,j\rangle|0,0\rangle|0\rangle : u \in V(G), i \in L^+(u), j \in L^-(u)\right\}$$

$$\oplus \bigoplus_{t=1}^{\mathsf{T}} \mathrm{span}\left\{|\!\rightarrow\rangle|u,i\rangle|a,z\rangle|t\rangle, |\!\leftarrow\rangle|u,j\rangle|a,z\rangle|t\rangle : u \in V(G), i \in L^+(u), j \in L^-(u), a \in \mathcal{A}, z \in \mathcal{Z}_{\geq t}\right\}$$

$$\subset \mathrm{span}\left\{|d\rangle|u,i\rangle|a,z\rangle|t\rangle : d \in \{\leftarrow, \rightarrow\}, u \in V(G), i \in L(u), a \in \mathcal{A}, z \in \mathcal{Z}, t \in \{0,\dots,\mathsf{T}\}\right\}. \quad (36)$$

We first define the *star states* in H. To this end, suppose we obtain a graph G' from G by adding a single new vertex, v_0, connected to each vertex $u \in V_0$ by an edge of weight $\mathsf{w}_0\sigma(u)$ and each vertex in M by an edge of weight w_M (for w_0 and w_M to be determined later). We can suppose each of these new edges have the label 0, and so we set:

$$L^+_{G'}(u) = \begin{cases} L^+_G(u) \cup \{0\} & \text{if } u \in V_0 \cup M \\ L^+_G(u) & \text{else.} \end{cases}$$

Then we define the star states:

$$\forall u \in V(G),\ |\psi_\star(u)\rangle = \sum_{i\in L^+_{G'}(u)} \sqrt{\mathsf{w}_{u,i}}\,|\!\rightarrow\rangle|u,i\rangle|0,0\rangle|0\rangle - \sum_{j\in L^-_G(u)} \sqrt{\mathsf{w}_{u,j}}\,|\!\leftarrow\rangle|u,j\rangle|0,0\rangle|0\rangle. \quad (37)$$

Throughout this section, we will let $L^+(u) = L^+_{G'}(u)$, so that we are always talking about the neighbours with respect to G' unless explicitly stated otherwise.

Our algorithm will be based on the star states, as well as the *transition states* $\Psi_0 \cup \Psi_1$ (see Definition 3.1) of the variable stopping time transition subroutine we are assuming. We recall the form of those states in the setting of the graph transition algorithm, where

$$\mathcal{I} = \{(u,i) : u \in V(G), i \in L^+(i)\} \quad \text{and} \quad A|u,i\rangle = |v,j\rangle$$
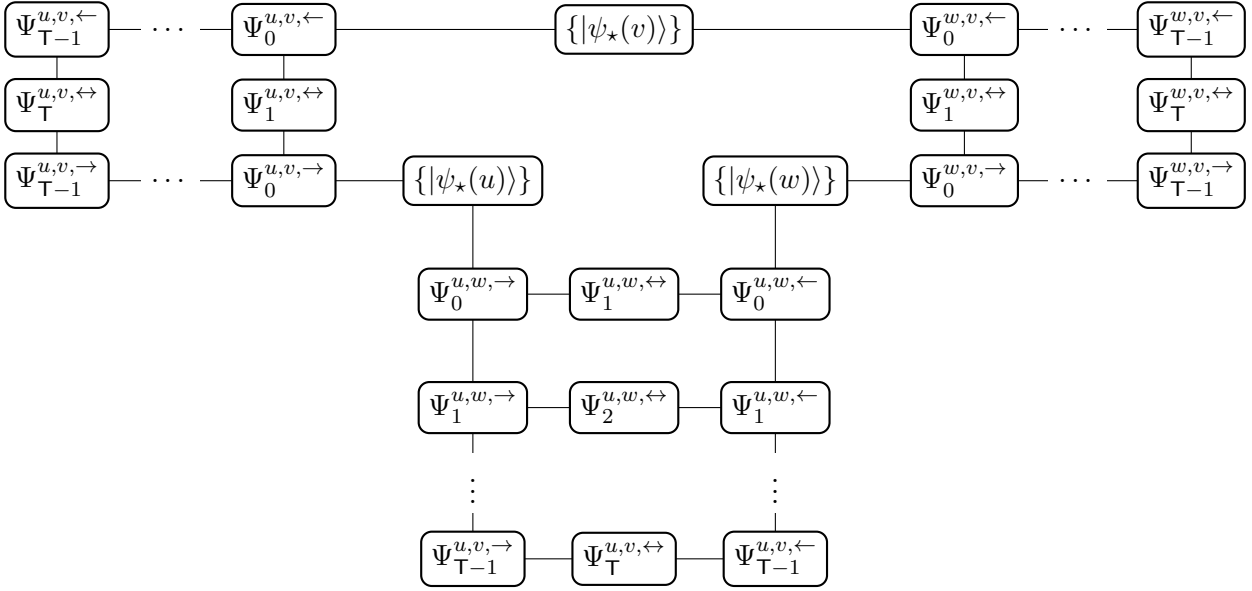
Figure 3: If G is a triangle with vertices u, v and w, and $\overrightarrow{E}(G) = \{(u,v),(w,v),(u,w)\}$, this figure shows the overlap graph for the spaces that make up $\Psi^{\mathcal{A}} \cup \Psi^{\mathcal{B}}$. If G is any graph, we can get a similar overlap graph by replacing each edge of G with a "ladder" gadget, like those shown here. While the height of each ladder is $\mathsf{T} = \mathsf{T}_{\max}$, if there is some probability of stopping at an earlier time, we can get flow through earlier rungs, thus saving time. The weights of the rungs should correspond to the probability that the algorithm halts at that time.

where $v = f_u(i)$, and $j = f_v^{-1}(u)$:

$$\Psi_t^{u,v,\rightarrow} := \Psi_t^{u,i,\rightarrow} = \left\{ |\psi_{a,z,t}^{u,i,\rightarrow}\rangle = |\rightarrow\rangle|u,i\rangle(\sqrt{\alpha_t}|a,z\rangle|t\rangle - \sqrt{\alpha_{t+1}}U_{t+1}^{u,i}|a,z\rangle|t+1\rangle) : a \in \mathcal{A}, z \in \mathcal{Z}_{>t} \right\}$$

$$\Psi_t^{u,v,\leftarrow} := \Psi_t^{u,i,\leftarrow} = \left\{ |\psi_{a,z,t}^{u,i,\leftarrow}\rangle = |\leftarrow\rangle|v,j\rangle(\sqrt{\alpha_t}|a,z\rangle|t\rangle - \sqrt{\alpha_{t+1}}U_{t+1}^{u,i}|a,z\rangle|t+1\rangle) : a \in \mathcal{A}, z \in \mathcal{Z}_{>t} \right\}$$

$$\Psi_t^{u,v,\leftrightarrow} := \Psi_t^{u,i,\leftrightarrow} = \left\{ |\psi_{a,z,t}^{u,i,\leftrightarrow}\rangle = \sqrt{\alpha_t}\left(|\rightarrow\rangle|u,i\rangle - |\leftarrow\rangle A_a|u,i\rangle\right)|a,z\rangle|t\rangle : a \in \mathcal{A}, z \in \mathcal{Z}_t \right\}.$$

(38)

Also recall from Definition 3.1:

$$\Psi_b = \bigcup_{\substack{u \in V(G) \\ i \in L^+(u)}} \bigcup_{\substack{t=0 \\ t=b \bmod 2}}^{\mathsf{T}-1} \left(\Psi_t^{u,i,\rightarrow} \cup \Psi_t^{u,i,\leftarrow} \right) \cup \bigcup_{\substack{t=1 \\ t=b \bmod 2}}^{\mathsf{T}} \Psi_t^{u,i,\leftrightarrow}.$$

Then we define:

$$\Psi^{\mathcal{A}} := \Psi_0 \qquad \text{and} \qquad \Psi^{\mathcal{B}} := \Psi_1 \cup \underbrace{\{|\psi_\star(u)\rangle : u \in V(G)\}}_{=:\Psi_\star}.$$

(39)

Each of these sets is pairwise orthogonal. This follows from the fact that each of Ψ_0 and Ψ_1 is a pairwise orthogonal set (Claim 3.2), the star states are pairwise orthogonal (easily verified), and each star state is orthogonal to Ψ_1 (easily verified). Another way of seeing this orthogonality is that starting from the graph G, we can form an orthogonality graph, where we replace each edge of G with a *ladder gadget* consisting of the orthogonality graph in Figure 2. An example of such a graph when G is a triangle is shown in Figure 3. Each node of the graph represents a set of pairwise orthogonal vectors, and there is an edge between two nodes if and only if states in the sets overlap. Since this graph is bipartite, and $\Psi^{\mathcal{A}}$ and $\Psi^{\mathcal{B}}$ are a bipartition (i.e. they are each independent sets) they are each pairwise orthogonal.

The algorithm in Theorem 4.1 will be a phase estimation algorithm, a la Theorem 2.11, on a unitary $U_{\mathcal{AB}} = (2\Pi_{\mathcal{A}} - I)(2\Pi_{\mathcal{B}} - I)$, where $\mathcal{A} = \text{span}\{\Psi^{\mathcal{A}}\}$ and $\mathcal{B} = \text{span}\{\Psi^{\mathcal{B}}\}$, on initial state:

$$|\psi_0\rangle = |\rightarrow\rangle|\sigma\rangle|0,0\rangle|0\rangle.$$

(40)

33

We have the following corollary of Lemma 3.3, and the fact that $\Psi_\star$ can be generated in unit cost:

Corollary 4.5. *The unitary $U_{\mathcal{AB}}$ can be implemented in $O(\log \mathsf{T})$ complexity.*

4.1.1 Positive Analysis

Suppose there is a flow θ satisfying conditions **P1-P4** of Theorem 4.1. We can extend θ to a circulation (see Definition 2.3) on G' by assigning:

$$\forall u \in V_0 \cup M, \ \theta(v_0, u) = \theta(u).$$

Then letting $\theta(u, i) := \theta(u, f_u(i))$, we define:

$$|w\rangle := \sum_{u \in V(G), i \in L^+_{G'}(u)} \frac{\theta(u, i)}{\sqrt{\mathsf{w}_{u,i}}} |w_+(u, i)\rangle, \tag{41}$$

where

$$|w_+(u, i)\rangle = (|\rightarrow\rangle|u, i\rangle + |\leftarrow\rangle|v, j\rangle) \sum_{t=0}^{\mathsf{T}} \frac{1}{\sqrt{\alpha_t}} |w^t(u, i)\rangle|t\rangle$$

is the subroutine's positive history state on input (u, i), as defined in Definition 3.4. We will prove that this is a positive witness.

Lemma 4.6. *The state $|w\rangle$ defined in (41) is a δ-positive witness for*

$$\delta = 3\mathsf{w}_0 \sum_{u \in V(G), i \in L^+(u)} \frac{\theta(u, i)^2}{\mathsf{w}_{u,i}} \mathbb{E}\left[\frac{\epsilon^{T_{u,i}}_{u,i}}{\alpha_{T_{u,i}}}\right].$$

Letting $\mathsf{w}_0 = \mathcal{R}^{-1}$, for sufficiently large w_M, $|w\rangle$ has complexity

$$\frac{\||w\rangle\|^2}{|\langle\psi_0|w\rangle|^2} \le 2\mathsf{w}_0 \sum_{(u,v) \in \vec{E}(G)} \frac{\theta(u, v)^2}{\mathsf{w}_{u,v}} \mathbb{E}\left[\sum_{t=0}^{T_{u,i}-1} \frac{1}{\alpha_t}\right] + 4 \le 6 =: c_+.$$

Proof. We begin by showing that $|w\rangle$ is almost orthogonal to $\mathcal{A} + \mathcal{B}$. We first note that by Claim 3.8 and the definition of $|w\rangle$, $|w\rangle$ is orthogonal to all $|\psi^{\leftarrow, u, i}_{a,z,t}\rangle$ and all $|\psi^{\rightarrow, u, i}_{a,z,t}\rangle$. We now show that $|w\rangle$ is also orthogonal to all star states. First note that star states have $|0\rangle$ in the last (time) register. Referring to Definition 3.4, and using $A|u, i\rangle = |v, j\rangle$, we have:

$$(I \otimes |0\rangle\langle 0|)|w_+(u, i)\rangle = \frac{1}{\sqrt{\alpha_0}}(|\rightarrow\rangle|u, i\rangle + |\leftarrow\rangle|v, j\rangle)|w^0(u, i)\rangle|0\rangle.$$

Since $|w^0(u, i)\rangle = |0, 0\rangle$ and $\alpha_0 = 1$, we have, using the notation $\overrightarrow{(v, j)} = (u, i)$ for $i \in L^+(u)$, $v = f_u(i)$, and $j = f_v^{-1}(u)$:

$$\langle\psi_\star(u)|w\rangle = \sum_{i \in L^+(u)} \sqrt{\mathsf{w}_{u,i}} \frac{\theta(u, i)}{\sqrt{\mathsf{w}_{u,i}}} \frac{1}{\sqrt{\alpha_0}} \langle 0, 0|w^0(u, i)\rangle + \sum_{j \in L^-(u)} \sqrt{\mathsf{w}_{u,j}} \frac{\theta(u, j)}{\sqrt{\mathsf{w}_{u,j}}} \frac{1}{\sqrt{\alpha_0}} \langle 0, 0|w^0\overrightarrow{(u, j)}\rangle$$

$$= \sum_{i \in L(u)} \theta(u, i) = 0,$$

since θ is a circulation on G'. Thus, the only states in $\Psi^{\mathcal{A}} \cup \Psi^{\mathcal{B}}$ not orthogonal to $|w\rangle$ are the states $|\psi^{u,i,\leftrightarrow}_{a,z,t}\rangle$ in (38). We claim that such a state only overlaps $|w\rangle$ in the $|w^+(u, i)\rangle$ part. That is because $|w^+(u, i)\rangle$ has $(|\rightarrow\rangle|u, i\rangle + |\leftarrow\rangle|v, j\rangle)$ in the first register, whereas $|\psi^{\leftrightarrow, u, i}_{a,z,t}\rangle$ has $(|\rightarrow\rangle|u, i\rangle - |\leftarrow\rangle A_a|u, i\rangle)$, and we are assuming that $\langle v', j'|A_a|u, i\rangle = 0$ for all a whenever $(v', j') \ne (v, j)$, where $v = f_u(i)$ and $j = f_v^{-1}(u)$. Thus:

$$\langle\psi^{u,i,\leftrightarrow}_{a,z,t}|w\rangle = \frac{\theta(u, i)}{\sqrt{\mathsf{w}_{u,i}}} \langle\psi^{u,i,\leftrightarrow}_{a,z,t}|w_+(u, i)\rangle = \frac{\theta(u, i)}{\sqrt{\mathsf{w}_{u,i}}} \left(1 - \langle u, i|A_a^\dagger|v, j\rangle\right) \langle a, z|w^t(u, i)\rangle$$

by (19). From this we get

$$|\langle \psi^{u,i,\leftrightarrow}_{a,z,t}|w\rangle|^2 \leq \begin{cases} 0 & \text{if } a = g(u,i) \\ 4\frac{\theta(u,i)^2}{\mathsf{w}_{u,i}}|\langle a,z|w^t(u,i)\rangle|^2 & \text{else.} \end{cases}$$

Thus, using the fact that reversal states are pairwise orthogonal, we have:

$$\|\Pi_\mathcal{A}|w\rangle + \Pi_\mathcal{B}|w\rangle\|^2 = \sum_{\substack{u\in V(G),\\ i\in L^+(u)}} \sum_{\substack{a\in\mathcal{A},\\ t\in\{0,\dots,\mathsf{T}\},\\ z\in\mathcal{Z}_t}} \frac{|\langle\psi^{u,i,\leftrightarrow}_{a,z,t}|w\rangle|^2}{\left\|\,|\psi^{u,i,\leftrightarrow}_{a,z,t}\rangle\right\|^2} = 4\sum_{\substack{u\in V(G),\\ i\in L^+(u)}} \sum_{\substack{a\in\mathcal{A}\setminus\{g(u,i)\},\\ t\in\{0,\dots,\mathsf{T}\},\\ z\in\mathcal{Z}_t}} \frac{\theta(u,i)^2}{\mathsf{w}_{u,i}}\frac{|\langle a,z|w^t(u,i)\rangle|^2}{2\alpha_t}.$$

Letting Λ^{bad}_t be the projector onto the part of the state that outputs at time t and is incorrect, we have

$$\sum_{\substack{a\in\mathcal{A}\setminus\{g(u,i)\},\\ z\in\mathcal{Z}_t}} |\langle a,z|w^t(u,i)\rangle|^2 = \left\|\Lambda^{\mathrm{bad}}_t|w^t(u,i)\rangle\right\|^2 = \bar{p}_{u,i}(t)\epsilon^t_{u,i},$$

where $\bar{p}_{u,i}(t)$ is the probability that the algorithm outputs at time t, and $\epsilon^t_{u,i}$ is the probability that the algorithm errs on input u,i given that it outputs at time t. Thus:

$$\|\Pi_\mathcal{A}|w\rangle + \Pi_\mathcal{B}|w\rangle\|^2 = 2\sum_{\substack{u\in V(G),\\ i\in L^+(u)}} \frac{\theta(u,i)^2}{\mathsf{w}_{u,i}} \sum_{t=0}^{\mathsf{T}} \frac{\bar{p}_{u,i}(t)\epsilon^t_{u,i}}{\alpha_t} = 2\sum_{\substack{u\in V(G),\\ i\in L^+(u)}} \frac{\theta(u,i)^2}{\mathsf{w}_{u,i}}\mathbb{E}\left[\frac{\epsilon^{T_{u,i}}_{u,i}}{\alpha_{T_{u,i}}}\right],$$

where $T_{u,i} = T_{u,v}$ is the stopping time of the subroutine on input (u,i). We can conclude that $|w\rangle$ is a δ-positive witness for any δ such that

$$\delta \geq \frac{2}{\||w\rangle\|^2}\sum_{\substack{u\in V(G),\\ i\in L^+(u)}} \frac{\theta(u,i)^2}{\mathsf{w}_{u,i}}\mathbb{E}\left[\frac{\epsilon^{T_{u,i}}_{u,i}}{\alpha_{T_{u,i}}}\right]. \tag{42}$$

Next we compute $\||w\rangle\|^2$. First we have, by Corollary 3.7:

$$\||w_+(u,i)\rangle\|^2 = 2\mathbb{E}\left[\sum_{t=0}^{T_{u,i}-1}\frac{1}{\alpha_t}\right],$$

from which we can compute:

$$\||w\rangle\|^2 = 2\sum_{u\in V(G),i\in L^+_G(u)} \frac{\theta(u,i)^2}{\mathsf{w}_{u,i}}\mathbb{E}\left[\sum_{t=0}^{T_{u,i}-1}\frac{1}{\alpha_t}\right] + \sum_{u\in V_0}\frac{\theta(u)^2}{\mathsf{w}_0\sigma(u)} + \sum_{u\in M}\frac{\theta(u)^2}{\mathsf{w}_\mathsf{M}}. \tag{43}$$

This already gives a lower bound of $\||w\rangle\|^2 \geq \frac{1}{\mathsf{w}_0}\sum_{u\in V_0}\frac{\theta(u)^2}{\sigma(u)} \geq \frac{1}{3\mathsf{w}_0}$ by **P2**, which, combined with (42), yields the desired bound on δ. Continuing from (43), and using $\sum_{u\in V_0}\frac{\theta(u)^2}{\sigma(u)} \leq 3$ (**P2**), we have:

$$\||w\rangle\|^2 \leq 2\sum_{u\in V(G),i\in L^+_G(u)} \frac{\theta(u,i)^2}{\mathsf{w}_{u,i}}\mathbb{E}\left[\sum_{t=0}^{T_{u,i}-1}\frac{1}{\alpha_t}\right] + \frac{3}{\mathsf{w}_0} + \frac{1}{\mathsf{w}_0} \tag{44}$$

for sufficiently large w_M. Finally, referring to (40), we compute

$$\langle\psi_0|w\rangle = \sum_{u\in V_0}\sqrt{\sigma(u)}\frac{\theta(u)}{\sqrt{\mathsf{w}_0\sigma(u)}} = \frac{1}{\sqrt{\mathsf{w}_0}}$$

by **P1**, which, combined with (44) and **P3**, yields the desired upper bound on the complexity. $\qquad\square$

4.1.2 Negative Analysis

Define

$$|w_{\mathcal{A}}\rangle := -\frac{1}{\sqrt{\mathsf{w}_0}} \sum_{\substack{u \in V(G) \\ i \in L_G^+(u)}} \sqrt{\mathsf{w}_{u,i}} |w_-(u,i)\rangle$$

$$|w_{\mathcal{B}}\rangle := -\frac{1}{\sqrt{\mathsf{w}_0}} \sum_{\substack{u \in V(G) \\ i \in L_G^+(u)}} \sqrt{\mathsf{w}_{u,i}} \left((|\rightarrow\rangle|u,i\rangle - |\leftarrow\rangle A|u,i\rangle)|0,0\rangle|0\rangle - |w_-(u,i)\rangle \right) + \frac{1}{\sqrt{\mathsf{w}_0}} \sum_{u \in V(G)} |\psi_\star(u)\rangle \tag{45}$$

where

$$|w_-(u,i)\rangle = (|\rightarrow\rangle|u,i\rangle - |\leftarrow\rangle|v,j\rangle) \sum_{t=0}^{\mathsf{T}} \alpha_t (-1)^t |w^t(u,i)\rangle|t\rangle$$

is the subroutine's negative history state on input (u,i), as defined in Definition 3.4. We will prove that $|w_{\mathcal{A}}\rangle, |w_{\mathcal{B}}\rangle$ are a negative witness. We start by showing the following.

Claim 4.7.

$$\|(I - \Pi_{\mathcal{A}})|w_{\mathcal{A}}\rangle\|^2 \leq 2 \sum_{u \in V(G), i \in L_G^+(u)} \frac{\mathsf{w}_{u,i}}{\mathsf{w}_0} \mathbb{E}\left[\alpha_{T_{u,i}} \epsilon_{u,i}^{T_{u,i}}\right]$$

$$\|(I - \Pi_{\mathcal{B}})|w_{\mathcal{B}}\rangle\|^2 \leq 2 \sum_{u \in V(G), i \in L_G^+(u)} \frac{\mathsf{w}_{u,i}}{\mathsf{w}_0} \mathbb{E}\left[\alpha_{T_{u,i}} \epsilon_{u,i}^{T_{u,i}}\right].$$

Proof. Refer to (39) for the definitions of $\Psi^{\mathcal{A}}$ and $\Psi^{\mathcal{B}}$, from which $\mathcal{A}$ and $\mathcal{B}$ are defined. By Claim 3.9,

$$\|(I - \Pi_{\Psi_0})|w_-(u,i)\rangle\|^2 \leq 2\mathbb{E}\left[\alpha_{T_{u,i}} \epsilon_{u,i}^{T_{u,i}}\right],$$

where $\epsilon_{u,i}^t$ is the probability of erring given that the algorithm stops at time t on input (u,i). Thus, since $\Psi^{\mathcal{A}} = \Psi_0$:

$$\|(I - \Pi_{\mathcal{A}})|w_{\mathcal{A}}\rangle\|^2 = \sum_{u \in V(G), i \in L_G^+(u)} \frac{\mathsf{w}_{u,i}}{\mathsf{w}_0} \|(I - \Pi_{\Psi_0})|w_-(u,i)\rangle\|^2 \leq 2 \sum_{u \in V(G), i \in L_G^+(u)} \frac{\mathsf{w}_{u,i}}{\mathsf{w}_0} \mathbb{E}\left[\alpha_{T_{u,i}} \epsilon_{u,i}^{T_{u,i}}\right],$$

where we used the fact that the $|w_-(u,i)\rangle$ are pairwise orthogonal, because of the state in their first register (see Definition 3.4).

We have $|\psi_\star(u)\rangle \in \mathcal{B}$ for all $u \in V(G)$. Combining this with $\Psi^{\mathcal{B}} = \Psi_1 \cup \Psi_\star$, we have:

$$\|(I - \Pi_{\mathcal{B}})|w_{\mathcal{B}}\rangle\|^2 = \sum_{u \in V(G), i \in L_G^+(u)} \frac{\mathsf{w}_{u,i}}{\mathsf{w}_0} \|(I - \Pi_{\Psi_1})\left((|\rightarrow\rangle|u,i\rangle - |\leftarrow\rangle A|u,i\rangle)|0,0\rangle|0\rangle - |w_-(u,i)\rangle\right)\|^2$$

$$\leq 2 \sum_{u \in V(G), i \in L_G^+(u)} \frac{\mathsf{w}_{u,i}}{\mathsf{w}_0} \mathbb{E}\left[\alpha_{T_{u,i}} \epsilon_{u,i}^{T_{u,i}}\right], \qquad \text{by Claim 3.9.} \quad \square$$

Lemma 4.8. *The states* $|w_{\mathcal{A}}\rangle, |w_{\mathcal{B}}\rangle$ *defined in (45) are a* δ' *negative witness for*

$$\delta' = 2 \sum_{u \in V(G), i \in L^+(u)} \frac{\mathsf{w}_{u,i}}{\mathsf{w}_0} \mathbb{E}\left[\alpha_{T_{u,i}} \epsilon_{u,i}^{T_{u,i}}\right].$$

Letting $\mathsf{w}_0 = \mathcal{R}^{-1}$, *as in Lemma 4.6,* $|w_{\mathcal{A}}\rangle$ *has complexity:* $\||w_{\mathcal{A}}\rangle\|^2 = 2\mathcal{R}\mathcal{W} =: \mathcal{C}_-$.

Proof. We first note that:

$$\sqrt{\mathsf{w}_0}(|w_{\mathcal{A}}\rangle + |w_{\mathcal{B}}\rangle) = - \sum_{u \in V(G), i \in L_G^+(u)} \sqrt{\mathsf{w}_{u,i}} \left(|\rightarrow\rangle|u,i\rangle - |\leftarrow\rangle A|u,i\rangle\right)|0,0\rangle|0\rangle + \sum_{u \in V(G)} |\psi_\star(u)\rangle. \tag{46}$$

Using the fact that $A|u,i\rangle = |v,j\rangle$ whenever $i \in L^+(u)$, $f_u(i) = v$ and $j = f_v^{-1}(u)$, we have:

$$\sum_{u \in V(G), i \in L_G^+(u)} \sqrt{\mathsf{w}_{u,i}} \left(|\rightarrow\rangle|u,i\rangle - |\leftarrow\rangle A|u,i\rangle\right)$$

$$= \sum_{u \in V(G), i \in L_G^+(u)} \sqrt{\mathsf{w}_{u,i}}|\rightarrow\rangle|u,i\rangle - \sum_{u \in V(G), j \in L_G^-(u)} \sqrt{\mathsf{w}_{u,j}}|\leftarrow\rangle|v,j\rangle. \tag{47}$$

On the other hand, referring to (37):

$$\sum_{u \in V(G)} |\psi_\star(u)\rangle = \sum_{u \in V(G)} \left(\sum_{i \in L_{G'}^+(u)} \sqrt{\mathsf{w}_{u,i}}|\rightarrow\rangle|u,i\rangle - \sum_{j \in L_G^-(u)} \sqrt{\mathsf{w}_{u,j}}|\leftarrow\rangle|u,j\rangle\right)|0,0\rangle|0\rangle. \tag{48}$$

Recall that G' only differs from G in that vertices in $M \cup V_0 = V_0$ (since $M = \emptyset$ by assumption) are connected to an additonal vertex v_0 by an edge labelled $0 \in L^+$. Thus, putting (47) and (48) into (46), we have: we have:

$$\sqrt{\mathsf{w}_0}(|w_\mathcal{A}\rangle + |w_\mathcal{B}\rangle) = \sum_{u \in V(G)} \sum_{i \in L_{G'}^+(u) \setminus L_G^+(u)} \sqrt{\mathsf{w}_{u,i}}|\rightarrow\rangle|u,i\rangle|0,0\rangle|0\rangle$$

$$= \sum_{u \in V_0} \sqrt{\mathsf{w}_0 \sigma(u)}|\rightarrow\rangle|u,0\rangle|0,0\rangle|0\rangle = \sqrt{\mathsf{w}_0}|\psi_0\rangle.$$

Combining this with Claim 4.7, we see that $|w_\mathcal{A}\rangle, |w_\mathcal{B}\rangle$ form a δ'-negative witness in the sense of Definition 2.9. We upper bound the complexity:

$$\||w_\mathcal{A}\rangle\|^2 = \frac{1}{\mathsf{w}_0} \sum_{u \in V(G), i \in L^+(u)} \mathsf{w}_{u,i}\||w_-(u,i)\rangle\|^2 = \frac{2}{\mathsf{w}_0} \sum_{u \in V(G), i \in L^+(u)} \mathsf{w}_{u,i}\mathbb{E}\left[\sum_{t=0}^{T_{u,i}-1} \alpha_t\right] \leq \frac{2\mathcal{W}}{\mathsf{w}_0},$$

by Corollary 3.7, and condition **N1** of Theorem 4.1. The result follows. $\square$

4.1.3 Conclusion of Proof of Theorem 4.1

Fix H as in (36), $|\psi_0\rangle$ as in (40), and $\Psi^\mathcal{A}$ and $\Psi^\mathcal{B}$ as in (39). Since $|\psi_0\rangle = |\rightarrow\rangle|\sigma\rangle|0,0\rangle|0\rangle$ and we assume we can generate $|\sigma\rangle$ in cost S, we can generate $|\psi_0\rangle$ in cost $\mathsf{S}' = \mathsf{S} + O(\log\mathsf{T})$, since we need to initialize $O(\log\mathsf{T})$ qubits in addition to $|\sigma\rangle$. By Corollary 4.5, we can implement $U_{\mathcal{AB}}$ in cost $\mathsf{A} = \log\mathsf{T}$. Let

$$c_+ = 6 \text{ and } \mathcal{C}_- = 2\mathcal{RW}$$

and let

$$\delta = 3\mathcal{R}^{-1} \sum_{(u,v) \in \vec{E}(G)} \frac{\theta(u,v)^2}{\mathsf{w}_{u,v}} \mathbb{E}\left[\frac{\epsilon_{u,v}^{T_{u,v}}}{\alpha_{T_{u,v}}}\right] \text{ and } \delta' = 2\mathcal{R} \sum_{(u,v) \in \mathbb{E}(G)} \mathsf{w}_{u,v}\mathbb{E}\left[\alpha_{T_{u,v}}\epsilon_{u,v}^{T_{u,v}}\right]$$

as in Lemma 4.6 and Lemma 4.8. Then we can verify that

$$\delta = \mathcal{R}^{-1}o(1/\mathcal{W}) \leq \frac{1}{(8c_+)^3\pi^8\mathcal{C}_-}$$

by condition **P4** of Theorem 4.1, and

$$\delta' = \mathcal{R}o(1/\mathcal{R}) \leq \frac{3}{4}\frac{1}{\pi^4 c_+},$$

by **N2**. Furthermore:

Positive Condition: by Lemma 4.6, if $M \neq \emptyset$, there is a δ-positive witness with $\frac{\||w\rangle\|^2}{|\langle w|\psi_0\rangle|^2} \leq c_+$; and

Negative Condition: by Lemma 4.8, if $M = \emptyset$, there is a δ'-negative witness with $\||w_\mathcal{A}\rangle\|^2 \leq \mathcal{C}_-$.

Thus, by Theorem 2.11, there is a quantum algorithm that can distinguish these two cases with bounded error in complexity:

$$O\left(\mathsf{S}' + \sqrt{\mathcal{C}_-}\mathsf{A}\right) = O\left(\mathsf{S} + \sqrt{\mathcal{RW}}\log\mathsf{T}\right).$$

5 Composition in Quantum Algorithms

Fix $f : \{0,1\}^n \to \{0,1\}$, and for each $i \in [n]$, fix $g_i : \{0,1\}^m \to \{0,1\}$. Then we can define $g : [n] \times \{0,1\}^m \to \{0,1\}$ by $g(i,x) = g_i(x)$, and $f \circ g : \{0,1\}^m \to \{0,1\}$ by $f \circ g(x) = f(g_1(x), \ldots, g_n(x))$. Given a bounded error quantum algorithm for f that takes time L and Q queries, and a quantum algorithm for g with sufficiently small error, that takes time at most $\mathsf{T} = \mathsf{T}_{\max}$, we can compose these algorithms to get a quantum algorithm for the composed function $f \circ g$ that takes time $\mathsf{L} + \mathsf{Q} \cdot \mathsf{T}$. In this section, we present a better way to compose these algorithms when the subroutine's running time varies in i, and its own randomness. Before we state our main theorem, we fix some notation.

Outer Algorithm: We suppose we have a quantum algorithm that computes $f : \{0,1\}^n \to \{0,1\}$ using quantum queries to an input $g \in \{0,1\}^n$. Let $V_1, \ldots, V_{\mathsf{L}}$ be some unitaries acting on a space:

$$H_{\mathcal{I}} \otimes H_{\mathcal{A}} \otimes H_{\mathcal{Y}} = \mathrm{span}\{|i\rangle|b,y\rangle : i \in [n] \cup \{0\}, b \in \{0,1\}, y \in \mathcal{Y}\},$$

where $H_{\mathcal{I}}$ is the query register, $H_{\mathcal{A}}$ is the answer register, and $H_{\mathcal{Y}}$ is the workspace. Let $\mathcal{Q} \subset [\mathsf{L}]$ be such that:

- If $\ell \in \mathcal{Q}$, $V_{\ell+1} = \mathcal{O}_g$, where $\mathcal{O}_g|i\rangle|b,y\rangle = (-1)^{g_i}|i\rangle|b,y\rangle$.

- If $\ell \notin \mathcal{Q}$, $V_{\ell+1}$ is an input-independent unitary that can be implemented in unit cost.

Let $\mathsf{Q} = |\mathcal{Q}|$. We suppose this algorithm computes f with bounded error ε_O. We assume, without loss of generality, that L is even, and that $\ell \in \mathcal{Q}$ only if ℓ is even.

We fix some notation for discussing this outer algorithm. First, define:

$$\forall \ell \in \{0, \ldots, \mathsf{L}\}, \ |w_O^\ell\rangle := V_\ell \ldots V_1 |0,0,0\rangle, \tag{49}$$

the algorithm's state at time ℓ. For $\ell \in \{0, \ldots, \mathsf{L}\}$, $y \in \mathcal{Y}$, $b \in \{0,1\}$, define:

$$\beta_{i,b,y}^\ell := \langle i,b,y|w_O^\ell\rangle. \tag{50}$$

Of special interest will be the *query weight* on i at time ℓ, for $\ell \in \mathcal{Q}$:

$$\mathsf{q}_{i,\ell} := \left\| (|i\rangle\langle i| \otimes I)|w_O^\ell\rangle \right\|^2 = \sum_{b \in \{0,1\}, y \in \mathcal{Y}} |\beta_{i,b,y}^\ell|^2. \tag{51}$$

Then the *average query weight* of $i \in [n]$ is:

$$\bar{\mathsf{q}}_i := \frac{1}{\mathsf{Q}} \sum_{\ell \in \mathcal{Q}} \mathsf{q}_{i,\ell}. \tag{52}$$

Note that $\sum_{i \in [n]} \bar{\mathsf{q}}_i = 1$.

Similar to variable-time algorithms, we assume we can implement, in unit cost, the unitary:

$$\sum_{\ell \in \{0, \ldots, \mathsf{L}-1\} \setminus \mathcal{Q}} |\ell\rangle\langle\ell| \otimes V_{\ell+1} + \sum_{\ell \in \mathcal{Q} \cup \{\mathsf{L}\}} |\ell\rangle\langle\ell| \otimes I.$$

Inner Subroutine: In addition to the outer algorithm, we suppose we have an inner variable-time subroutine, as defined in [Section 3](#), that computes the function $g(i) = g_i = g_i(x)$ with maximum time $\mathsf{T} = \mathsf{T}_{\max}$. For $i \in [n]$, let ϵ_i denote the error on input i, so in the notation of [Section 3](#), we have:

$$\epsilon_i = \sum_{t=0}^{\mathsf{T}} \bar{p}_t(i) \epsilon_i^t = \mathbb{E}[\epsilon_i^{T_i}].$$

As in [Section 3](#), we let T_i denote the time at which the inner subroutine stops on input i, which is a random variable on $\{1, \ldots, \mathsf{T}\}$. Then our main theorem of this section is the following.

38

Theorem 5.1. *Fix functions $f : \{0,1\}^n \to \{0,1\}$ and $\{g_i : \{0,1\}^m \to \{0,1\}\}_{i\in[n]}$, and an outer algorithm for f with time complexity L, query complexity Q, and bounded error ε_O; and inner subroutine for g with maximum time $\mathsf{T} = \mathsf{T}_{\max}$, as described above. Let $\mathsf{T}_{\mathrm{avg}}$ be an upper bound such that:*

$$\sum_{i\in[n]} \bar{\mathsf{q}}_i \mathbb{E}[T_i] \leq \mathsf{T}_{\mathrm{avg}}$$

and suppose the subroutine's errors satisfy the following condition:

$$\epsilon_{\mathrm{avg}} := \sum_{i\in[n]} \bar{\mathsf{q}}_i \epsilon_i \leq \frac{\eta}{\mathsf{Q}(\mathsf{L} + \mathsf{Q}\mathsf{T}_{\mathrm{avg}})}.$$

Then if ε_O and η are sufficiently small constants, there is a quantum algorithm that computes $f \circ g$ with bounded error in complexity

$$O\left((\mathsf{L} + \mathsf{Q}\mathsf{T}_{\mathrm{avg}}) \log(\mathsf{L}\mathsf{T})\right).$$

The remainder of this section is devoted to proving Theorem 5.1.

5.1 Parameters of the Phase Estimation Algorithm

We let

$$H = \mathrm{span}\{|d\rangle|i\rangle|a,z\rangle|t\rangle|b,y\rangle|\ell\rangle : d \in \{\leftarrow, \rightarrow\}, i \in [n], a, b \in \{0,1\}, z \in \mathcal{Z}, y \in \mathcal{Y},$$
$$t \in \{0, \ldots, \mathsf{T}\}, \ell \in \{0, \ldots, \mathsf{L}\}\}$$
$$\oplus \mathrm{span}\{|\perp\rangle|i\rangle|0,0\rangle|0\rangle|b,y\rangle|\ell\rangle : i \in [n], b \in \{0,1\}, z \in \mathcal{Z}, y \in \mathcal{Y}, \ell \in \{0, \ldots, \mathsf{L}\}\}$$
$$\oplus \mathrm{span}\{|\circ\rangle|i\rangle|0,0\rangle|0\rangle|b,y\rangle|\mathsf{L}\rangle : i \in [n], b \in \{0,1\}, y \in \mathcal{Y}\} \oplus \mathrm{span}\{|\circ\rangle|0\rangle|0,0\rangle|0\rangle|0,0\rangle|0\rangle\}. \tag{53}$$

We first define outer transition states, extending the unitaries V_ℓ to H by tensoring in identities:

$$\forall i \in [n], \ell \in \{0, \ldots, \mathsf{L}-1\} \setminus \mathcal{Q}$$
$$\Psi_{i,\ell}^O := \left\{|\psi_{i,b,y,\ell}^O\rangle := |\perp\rangle|i\rangle|0,0\rangle|0\rangle|b,y\rangle|\ell\rangle - |\perp\rangle V_{\ell+1}(|i\rangle|0,0\rangle|0\rangle|b,y\rangle)|\ell+1\rangle : b \in \{0,1\}, y \in \mathcal{Y}\right\}. \tag{54}$$

In addition to the outer algorithm, we are assuming we have an inner subroutine, which is a variable-time subroutine $\{U_t^i\}_{i\in[n],t\in[\mathsf{T}]}$ as in Section 3, computing $g(i) = g_i$ – that is, we have input space $\mathcal{I} = [n]$, and answer space $\mathcal{A} = \{0,1\}$. Note that letting $A|i\rangle = (-1)^{g_i}|i\rangle$, and $A_a|i\rangle = (-1)^a|i\rangle$ for $a \in \{0,1\}$, it is easy to verify that this variable-time subroutine *reversibly computes* A, simply letting $\tilde{U}_t = U_t$. We will use *inner transition states*, which are basically the transition states of the subroutine (see Definition 3.1), but we carry around the extra information of the current state of the outer algorithm, $|b,y\rangle|\ell\rangle$ as well. Throughout this section, we will use $\alpha_t = 1$ for all t, so we have:

$$\forall i \in [n], t \in \{0, \ldots, \mathsf{T}-1\}, \ell \in \mathcal{Q}$$
$$\Psi_{t,\ell}^{i,\rightarrow} := \left\{|\psi_{a,z,t}^{i,\rightarrow}\rangle|b,y\rangle|\ell\rangle = |\rightarrow\rangle|i\rangle\left(|a,z\rangle|t\rangle - U_{t+1}^i|a,z\rangle|t+1\rangle\right)|b,y\rangle|\ell\rangle\right\}_{\substack{a\in\{0,1\},z\in\mathcal{Z}_{>t}, \\ b\in\{0,1\},y\in\mathcal{Y}}}$$
$$\Psi_{t,\ell}^{i,\leftarrow} := \left\{|\psi_{a,z,t}^{i,\leftarrow}\rangle|b,y\rangle|\ell\rangle = |\leftarrow\rangle|i\rangle\left(|a,z\rangle|t\rangle - U_{t+1}^i|a,z\rangle|t+1\rangle\right)|b,y\rangle|\ell\rangle\right\}_{\substack{a\in\{0,1\},z\in\mathcal{Z}_{>t}, \\ b\in\{0,1\},y\in\mathcal{Y}}} \tag{55}$$
$$\forall i \in [n], t \in [\mathsf{T}], \ell \in \mathcal{Q}$$
$$\Psi_{t,\ell}^{i,\leftrightarrow} := \left\{|\psi_{a,z,t}^{i,\leftrightarrow}\rangle|b,y\rangle|\ell\rangle = (|\rightarrow\rangle - (-1)^a|\leftarrow\rangle)|i\rangle|a,z\rangle|t\rangle|b,y\rangle|\ell\rangle\right\}_{\substack{a\in\{0,1\},z\in\mathcal{Z}_t, \\ b\in\{0,1\},y\in\mathcal{Y}}}.$$

We note that for $d \in \{\leftarrow, \rightarrow, \leftrightarrow\}$, we have

$$\Psi_{t,\ell}^{i,d} = \bigcup_{b\in\{0,1\},y\in\mathcal{Y}} \Psi_t^{i,d} \otimes |b,y\rangle|\ell\rangle,$$
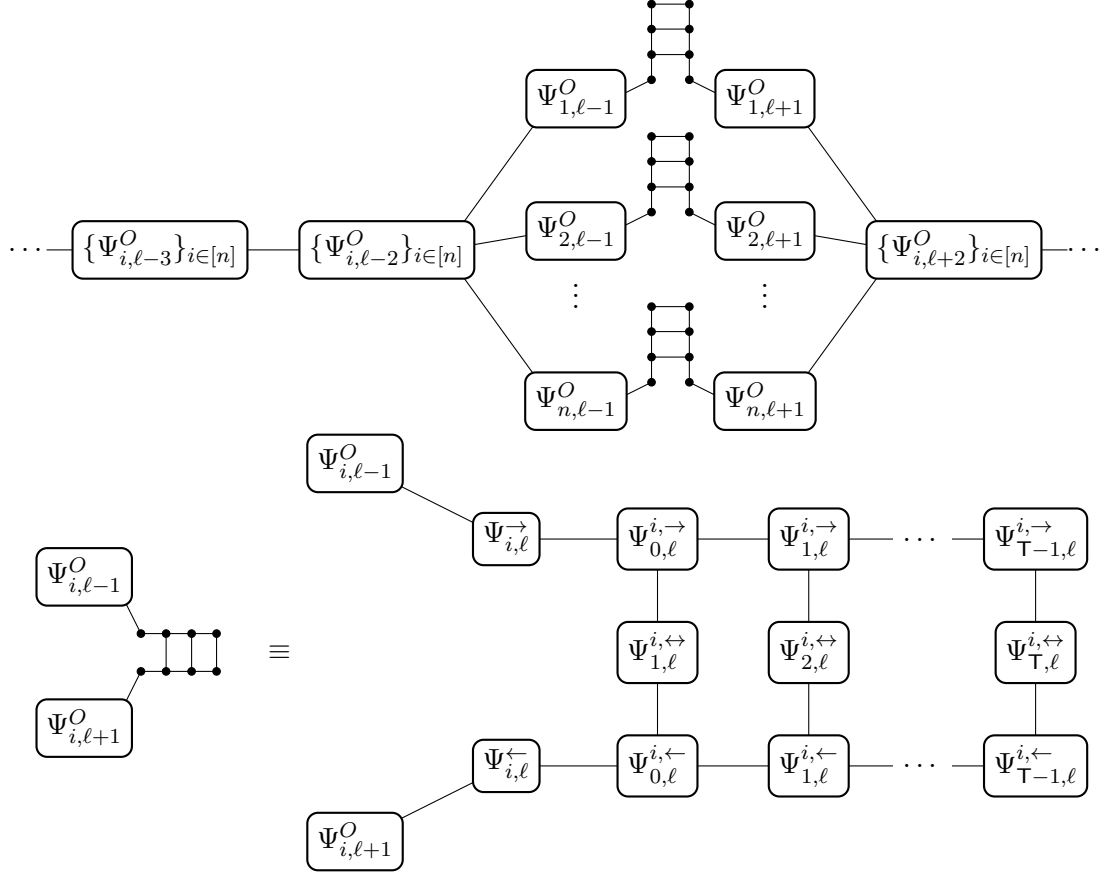
where $\Psi_t^{i,d}$ are as in Definition 3.1.

Figure 4: A graph showing the overlap between the states in $\Psi^{\mathcal{A}} \cup \Psi^{\mathcal{B}}$. Each node of the graph represents a set of pairwise orthogonal states. An edge between two nodes indicates that the sets are not orthogonal. The states in $\Psi^{\mathcal{A}}$ (or $\Psi^{\mathcal{B}}$) are an independent set of this graph. Here we show just a part of the graph, for some $\ell \in \mathcal{Q}$.

Next, we define *connecting states*, which represent transitioning into and out of the inner subroutine:

$\forall i \in [n], \ell \in \mathcal{Q}$
$$\Psi_{i,\ell}^{\rightarrow} := \left\{ |\psi_{i,b,y,\ell}^{\rightarrow}\rangle := |\bot\rangle|i\rangle|0,0\rangle|0\rangle|b,y\rangle|\ell\rangle - |\rightarrow\rangle|i\rangle|0,0\rangle|0\rangle|b,y\rangle|\ell\rangle \right\}_{b\in\{0,1\}, y\in\mathcal{Y}} \qquad (56)$$
$$\Psi_{i,\ell}^{\leftarrow} := \left\{ |\psi_{i,b,y,\ell}^{\leftarrow}\rangle := |\leftarrow\rangle|i\rangle|0,0\rangle|0\rangle|b,y\rangle|\ell\rangle - |\bot\rangle|i\rangle|0,0\rangle|0\rangle|b,y\rangle|\ell+1\rangle \right\}_{b\in\{0,1\}, y\in\mathcal{Y}}$$

Analogous to how we let star states for $u \in V_0$ overlap with the initial state, we define an *initial transition state*:
$$|\psi_{0,v_0}\rangle := \sqrt{\mathsf{w}_0}|\circ\rangle|0\rangle|0,0\rangle|0\rangle|0,0\rangle|0\rangle - |\bot\rangle|0\rangle|0,0\rangle|0\rangle|0,0\rangle|0\rangle \qquad (57)$$

Finally, analogous to how we let star states of marked vertices in Section 4 have an extra edge coming out of them (to v_0), we will consider a state at time L to be marked, but we would like it to be "more marked" if it has $b = 1$ in the answer register than not, which we reflect through the use of different weights on the extra edge. With this intuition, we include the following *terminal states*, for some weights $\mathsf{w}_{1,\text{out}} > \mathsf{w}_{0,\text{out}}$:

$$\Psi_{\text{out}} := \left\{ |\psi_{i,b,y}^{\text{out}}\rangle := |\bot\rangle|i\rangle|0,0\rangle|0\rangle|b,y\rangle|\mathsf{L}\rangle - \sqrt{\mathsf{w}_{b,\text{out}}}|\circ\rangle|i\rangle|0,0\rangle|0\rangle|b,y\rangle|\mathsf{L}\rangle : i \in [n], b \in \{0,1\}, y \in \mathcal{Y} \right\}. \qquad (58)$$

Recall that the transition states of the subroutine, $|\psi_{a,z,t}^{i,\rightarrow}\rangle$, $|\psi_{a,z,t}^{i,\leftarrow}\rangle$, and $|\psi_{a,z,t}^{i,\leftrightarrow}\rangle$, can be divided into two pairwise orthogonal sets, Ψ_0 and Ψ_1, based on the parity of t (see Definition 3.1). We define the following two sets of states.

$$\Psi^{\mathcal{A}} = \bigcup_{i\in[n]} \bigcup_{\ell\in\{0,\ldots,\mathsf{L}-1\}\backslash\mathcal{Q} \text{ even}} \Psi_{i,\ell}^O \cup \bigcup_{i\in[n]} \bigcup_{\ell\in\mathcal{Q}} \left(\Psi_{i,\ell}^{\rightarrow} \cup \Psi_{i,\ell}^{\leftarrow} \right) \cup \Psi_{\text{out}} \cup \bigcup_{\substack{b\in\{0,1\}\\y\in\mathcal{Y}}} \bigcup_{\ell\in\mathcal{Q}} \left(\Psi_1 \otimes |b,y\rangle|\ell\rangle \right)$$
$$\Psi^{\mathcal{B}} = \bigcup_{i\in[n]} \bigcup_{\ell\in\{0,\ldots,\mathsf{L}-1\}\backslash\mathcal{Q} \text{ odd}} \Psi_{i,\ell}^O \cup \{|\psi_{0,v_0}\rangle\} \cup \bigcup_{\substack{b\in\{0,1\}\\y\in\mathcal{Y}}} \bigcup_{\ell\in\mathcal{Q}} \left(\Psi_0 \otimes |b,y\rangle|\ell\rangle \right). \qquad (59)$$

We claim that the states in $\Psi^{\mathcal{A}}$ (or $\Psi^{\mathcal{B}}$) are pairwise orthogonal, for which the assumptions that L is even, and every $\ell \in \mathcal{Q}$ is even are necessary. To illustrate this, Figure 4 shows a diagram of which of the states in $\Psi^{\mathcal{A}} \cup \Psi^{\mathcal{B}}$ overlap.

The algorithm in Theorem 5.1 will be a phase estimation algorithm, a la Theorem 2.11, on a unitary $U_{\mathcal{AB}} = (2\Pi_{\mathcal{A}} - I)(2\Pi_{\mathcal{B}} - I)$, where $\mathcal{A} = \text{span}\{\Psi^{\mathcal{A}}\}$ and $\mathcal{B} = \text{span}\{\Psi^{\mathcal{B}}\}$, on initial state:

$$|\psi_0\rangle = |\circ\rangle|0\rangle|0,0\rangle|0\rangle|0,0\rangle|0\rangle. \qquad (60)$$

Then we have the following:

Lemma 5.2. $U_{\mathcal{AB}}$ *can be implemented in complexity* $O(\log \mathsf{T} + \log \mathsf{L})$.

Proof. We describe how to reflect around the states in $\Psi^{\mathcal{A}}$ one part at a time. We use the fact that reflecting around a set of pairwise orthogonal states can be reduced to generating the states and then performing a much simpler reflection, as in Lemma 3.3.

As in Lemma 3.3, we can reflect around Ψ_1 in complexity $O(\log \mathsf{T})$, from which it follows that we can reflect around $\bigcup_{b,y,\ell\in\mathcal{Q}}(\Psi_1 \otimes |b,y\rangle|\ell\rangle)$ in this cost. By a similar argument to Lemma 3.3, we can reflect around the outer algorithm states with ℓ even in complexity $O(\log \mathsf{L})$. We can reflect around the two types of connecting states in $O(\log \mathsf{L})$, using the maps $|\bot\rangle \mapsto \propto |\bot\rangle - |\rightarrow\rangle$ and $|\bot\rangle|\ell\rangle \mapsto \propto |\rightarrow\rangle|\ell\rangle - |\bot\rangle|\ell+1\rangle$. Finally, we can reflect around the states in Ψ_{out} using the $O(1)$ complexity map $|\bot\rangle|b\rangle \mapsto \propto |\bot\rangle - \sqrt{\mathsf{w}_{b,\text{out}}}|\circ\rangle$.

The case for $\Psi^{\mathcal{B}}$ is similar, but we note that we can reflect around $|\psi_{0,v_0}\rangle$ using the $O(1)$ complexity map $|\circ\rangle \mapsto \propto \sqrt{\mathsf{w}_0}|\circ\rangle - |\bot\rangle$. $\qquad \square$

5.2 Positive Analysis

Recall from Definition 3.4, the positive history states of the inner subroutine (using $\alpha_t = 1$):

$$|w_+(i)\rangle = (|\rightarrow\rangle|i\rangle + (-1)^{g_i}|\leftarrow\rangle|i\rangle) \sum_{t=0}^{\mathsf{T}} |w^t(i)\rangle|t\rangle,$$

where $|w^t(i)\rangle$ is the state of the inner subroutine at time t on input i. Our positive witness will be the following, where $\beta_{i,b,y}^{\ell} = \langle i, b, y|w_O^{\ell}\rangle$ are the amplitudes of the outer algorithm states, defined in (50):

$$|w\rangle = \frac{1}{\sqrt{\mathsf{w}_0}}|\circ\rangle|0\rangle|0,0\rangle|0\rangle|0,0\rangle|0\rangle + |\perp\rangle \sum_{\ell=0}^{\mathsf{L}} \sum_{\substack{i\in[n],b\in\{0,1\},y\in\mathcal{Y}}} \beta_{i,b,y}^{\ell}|i\rangle|0,0\rangle|0\rangle|b,y\rangle|\ell\rangle$$

$$+ \sum_{\ell\in\mathcal{Q}} \sum_{\substack{i\in[n],b\in\{0,1\},y\in\mathcal{Y}}} \beta_{i,b,y}^{\ell}|w_+(i)\rangle|b,y\rangle|\ell\rangle + |\circ\rangle \sum_{\substack{i\in[n],b\in\{0,1\},y\in\mathcal{Y}}} \frac{1}{\sqrt{\mathsf{w}_{b,\mathrm{out}}}} \beta_{i,b,y}^{\mathsf{L}}|i\rangle|0,0\rangle|0\rangle|b,y\rangle|\mathsf{L}\rangle. \quad (61)$$

Ignoring the first and last term, this can be seen (up to algorithm errors) as a history state of the algorithm obtained by composing the inner and outer algorithms.

Claim 5.3. $|w\rangle$ *is orthogonal to all outer transition states (see (54)), connecting states (see (56)), terminal states (see (58)), and the initial transition state (see (57)).*

Proof. **Outer Transition States:** For outer algorithms states, we fix some $\ell \notin \mathcal{Q}$. The first register of $|\psi_{i,b,y,\ell}^{O}\rangle = |\perp\rangle(|i\rangle|0,0\rangle|b,y\rangle|\ell\rangle - V_{\ell+1}|i\rangle|0,0\rangle|b,y\rangle|\ell+1\rangle)$ is only supported on $|\perp\rangle$, so we have:

$$\langle\psi_{i,b,y,\ell}^{O}|w\rangle = \langle\psi_{i,b,y,\ell}^{O}|\sum_{\ell'=0}^{\mathsf{L}} |\perp\rangle \sum_{\substack{i'\in[n]\\b'\in\{0,1\},y'\in\mathcal{Y}}} \beta_{i',b',y'}^{\ell'}|i'\rangle|0,0\rangle|0\rangle|b',y'\rangle|\ell\rangle$$

$$= \langle i,b,y| \sum_{\substack{i'\in[n]\\b'\in\{0,1\},y'\in\mathcal{Y}}} \beta_{i',b',y'}^{\ell}|i'\rangle|b',y'\rangle - \langle i,b,y|V_{\ell+1}^{\dagger} \sum_{\substack{i'\in[n]\\b'\in\{0,1\},y'\in\mathcal{Y}}} \beta_{i',b',y'}^{\ell+1}|i'\rangle|b',y'\rangle$$

$$= \langle i,b,y|w_O^{\ell}\rangle - \langle i,b,y|V_{\ell+1}^{\dagger}|w_O^{\ell+1}\rangle = \langle i,b,y|w_O^{\ell}\rangle - \langle i,b,y|V_{\ell+1}^{\dagger}V_{\ell+1}|w_O^{\ell}\rangle = 0.$$

Connecting States: For any $i\in[n]$, $b\in\{0,1\}$, $y\in\mathcal{Y}$, $\ell\in\mathcal{Q}$, we can see by referring to (56) that $|\psi_{i,b,y,\ell}^{\rightarrow}\rangle = (|\perp\rangle - |\rightarrow\rangle)|i\rangle|0,0\rangle|b,y\rangle|\ell\rangle$, so:

$$\langle\psi_{i,b,y,\ell}^{\rightarrow}|w\rangle = \beta_{i,b,y}^{\ell} - \beta_{i,b,y}^{\ell}\langle\rightarrow,i,0,0,0|w_+(i)\rangle = \beta_{i,b,y}^{\ell} - \beta_{i,b,y}^{\ell}\langle 0,0|w^0(i)\rangle = 0,$$

where we have used $|w^0(i)\rangle = |0,0\rangle$ (see Definition 3.4). Similarly, $|\psi_{i,b,y,\ell}^{\leftarrow}\rangle = |\leftarrow\rangle|i\rangle|0,0\rangle|b,y\rangle|\ell\rangle - |\perp\rangle|i\rangle|0,0\rangle|b,y\rangle|\ell+1\rangle$, so:

$$\langle\psi_{i,b,y,\ell}^{\leftarrow}|w\rangle = \beta_{i,b,y}^{\ell}\langle\leftarrow,i,0,0,0|w_+(i)\rangle - \beta_{i,b,y}^{\ell+1} = (-1)^{g_i}\beta_{i,b,y}^{\ell}\langle 0,0|w^0(i)\rangle - \beta_{i,b,y}^{\ell+1} = (-1)^{g_i}\beta_{i,b,y}^{\ell} - \beta_{i,b,y}^{\ell+1}.$$

To see that this is 0, we use the fact that since $\ell\in\mathcal{Q}$, $V_{\ell+1} = \mathcal{O}_g$ is a query, so:

$$\beta_{i,b,y}^{\ell+1} = \langle i,b,y|w_O^{\ell+1}\rangle = \langle i,b,y|V_{\ell+1}|w_O^{\ell}\rangle = (-1)^{g_i}\langle i,b,y|w_O^{\ell}\rangle = (-1)^{g_i}\beta_{i,b,y}^{\ell}. \quad (62)$$

Initial Transition State $|\psi_{0,v_0}\rangle = \sqrt{\mathsf{w}_0}|\circ\rangle|0\rangle|0,0\rangle|0\rangle|0,0\rangle|0\rangle - |\perp\rangle|0\rangle|0,0\rangle|0\rangle|0,0\rangle|0\rangle$:

$$\langle\psi_{0,v_0}|w\rangle = \sqrt{\mathsf{w}_0}\frac{1}{\sqrt{\mathsf{w}_0}}\langle\circ,0,0,0,0,0,0,0|\circ,0,0,0,0,0,0,0\rangle - \beta_{0,0,0}^0 = 0,$$

since $|w_O^0\rangle = |0\rangle|0,0\rangle$, so $\beta_{0,0,0}^0 = \langle 0,0,0|w_O^0\rangle = 1$.

Terminal States: For $i\in[n]$, $b\in\{0,1\}$, and $y\in\mathcal{Y}$, $|\psi_{i,b,y}^{\mathrm{out}}\rangle = (|\perp\rangle - \sqrt{\mathsf{w}_{b,\mathrm{out}}}|\circ\rangle)|i\rangle|0,0\rangle|0\rangle|b,y\rangle|\mathsf{L}\rangle$, so:

$$\langle\psi_{i,y}^{M}|w\rangle = \langle\perp|\perp\rangle\beta_{i,b,y}^{\mathsf{L}} - \langle\circ|\circ\rangle\sqrt{\mathsf{w}_{b,\mathrm{out}}}\frac{1}{\sqrt{\mathsf{w}_{b,\mathrm{out}}}}\beta_{i,b,y}^{\mathsf{L}} = 0. \qquad \square$$

Lemma 5.4 (Positive Witness). *Suppose $f \circ g(x) = 1$, and let $|w\rangle$ be as in (61), and $|\psi_0\rangle$ as in (60). Then letting $\delta = 2\mathsf{w}_0 \mathsf{Q} \epsilon_{\mathrm{avg}}$, $|w\rangle$ is a δ-positive witness with*

$$\frac{\||w\rangle\|^2}{|\langle w|\psi_0\rangle|^2} \leq 1 + \mathsf{w}_0 \left(\mathsf{L} + 1 + 2\mathsf{Q}\left(\mathsf{T}_{\mathrm{avg}} + 1\right) + \frac{\varepsilon_O}{\mathsf{w}_{0,\mathrm{out}}} + \frac{1 - \varepsilon_O}{\mathsf{w}_{1,\mathrm{out}}} \right).$$

Proof. Let Ξ_b be the orthogonal projector onto states with b in the outer algorithm's answer register, and p_0 the probability that the outer algorithm outputs 0 on input $g(x)$ (assuming $\mathcal{O}_g$ is implemented with no error), so in particular, $p_0 \leq \varepsilon_O$. Then we have:

$$\||w\rangle\|^2 = \frac{1}{\mathsf{w}_0} + \sum_{\ell=0}^{\mathsf{L}} \left\||w_O^\ell\rangle\right\|^2 + \sum_{\ell \in \mathcal{Q}} \sum_{\substack{i \in [n] \\ b \in \{0,1\}, y \in \mathcal{Y}}} |\beta_{i,b,y}^\ell|^2 \||w_+(i)\rangle\|^2 + \frac{\left\|\Xi_0|w_O^\mathsf{L}\rangle\right\|^2}{\mathsf{w}_{0,\mathrm{out}}} + \frac{\left\|\Xi_1|w_O^\mathsf{L}\rangle\right\|^2}{\mathsf{w}_{1,\mathrm{out}}} \tag{63}$$

$$= \frac{1}{\mathsf{w}_0} + \mathsf{L} + 1 + \sum_{\ell \in \mathcal{Q}} \sum_{\substack{i \in [n] \\ b \in \{0,1\}, y \in \mathcal{Y}}} |\beta_{i,b,y}^\ell|^2 2\mathbb{E}\left[T_i + 1\right] + \frac{p_0}{\mathsf{w}_{0,\mathrm{out}}} + \frac{1 - p_0}{\mathsf{w}_{1,\mathrm{out}}} \qquad \text{Corollary 3.7}$$

$$\leq \frac{1}{\mathsf{w}_0} + \mathsf{L} + 1 + 2 \sum_{i \in [n]} \sum_{\ell \in \mathcal{Q}} \mathsf{q}_{i,\ell} \mathbb{E}[T_i + 1] + \frac{\varepsilon_O}{\mathsf{w}_{0,\mathrm{out}}} + \frac{1 - \varepsilon_O}{\mathsf{w}_{1,\mathrm{out}}} \qquad \mathsf{w}_{1,\mathrm{out}} > \mathsf{w}_{0,\mathrm{out}} \text{ and } (51)$$

$$= \frac{1}{\mathsf{w}_0} + \mathsf{L} + 1 + 2\mathsf{Q} \sum_{i \in [n]} \bar{\mathsf{q}}_i(\mathbb{E}[T_i] + 1) + \frac{\varepsilon_O}{\mathsf{w}_{0,\mathrm{out}}} + \frac{1 - \varepsilon_O}{\mathsf{w}_{1,\mathrm{out}}} \qquad \text{by (52)}$$

$$\leq \frac{1}{\mathsf{w}_0} + \mathsf{L} + 1 + 2\mathsf{Q}\left(\mathsf{T}_{\mathrm{avg}} + 1\right) + \frac{\varepsilon_O}{\mathsf{w}_{0,\mathrm{out}}} + \frac{1 - \varepsilon_O}{\mathsf{w}_{1,\mathrm{out}}},$$

where in the last line we have used the assumption from the theorem statement that $\mathsf{T}_{\mathrm{avg}}$ upper bounds the average stopping time. Combined with the observation that $\langle \psi_0 | w \rangle = \frac{1}{\sqrt{\mathsf{w}_0}}$, we get the desired complexity bound.

To complete the proof, we upper bound $\|\Pi_{\mathcal{A}}|w\rangle\|^2$ and $\|\Pi_{\mathcal{B}}|w\rangle\|^2$, to show that $|w\rangle$ is a δ-positive witness, as defined in Definition 2.10. By Claim 5.3, the only states in $\Psi^{\mathcal{A}} \cup \Psi^{\mathcal{B}}$ (see (59)) that overlap $|w\rangle$ are the inner algorithm states (see (55)). Since the inner algorithm states do not have $\perp$ or $\circ$ in the first register, we have:

$$\Pi_{\mathcal{A}}|w\rangle = \sum_{\substack{b \in \{0,1\} \\ y \in \mathcal{Y}}} \left(\sum_{\ell \in \mathcal{Q}} \Pi_{\Psi_1} \otimes |b, y, \ell\rangle\langle b, y, \ell| \right) \sum_{\ell \in \mathcal{Q}} \sum_{i \in [n]} \beta_{i,b,y}^\ell |w_+(i)\rangle|b, y\rangle|\ell\rangle$$

$$= \sum_{i \in [n]} \sum_{\ell \in \mathcal{Q}} \Pi_{\Psi_1}|w_+(i)\rangle \otimes \sum_{b \in \{0,1\}, y \in \mathcal{Y}} \beta_{i,b,y}^\ell |b, y\rangle|\ell\rangle.$$

Then applying Claim 3.8, which says that (using $\alpha_t = 1$)

$$\|\Pi_{\Psi_1}|w_+(i)\rangle\|^2 \leq 2\mathbb{E}\left[\epsilon_i^{T_i}\right] = 2\epsilon_i,$$

and since $|w_+(i)\rangle$ are orthogonal for different i, we have:

$$\|\Pi_{\mathcal{A}}|w\rangle\|^2 = \sum_{i \in [n]} \sum_{\ell \in \mathcal{Q}} \|\Pi_{\Psi_1}|w_+(i)\rangle\|^2 \sum_{b \in \{0,1\}, y \in \mathcal{Y}} |\beta_{i,b,y}^\ell|^2 \leq 2\mathsf{Q} \sum_{i \in [n]} \bar{\mathsf{q}}_i \epsilon_i \qquad \text{by (51) and (52)}$$

$$\leq 2\mathsf{Q}\epsilon_{\mathrm{avg}} \leq 2\mathsf{w}_0 \mathsf{Q} \epsilon_{\mathrm{avg}} \||w\rangle\|^2 = \delta \||w\rangle\|^2,$$

since, by (63), $\||w\rangle\|^2 \geq 1/\mathsf{w}_0$. By a nearly identical proof where we simply swap Π_{Ψ_0} with Π_{Ψ_1}, we have a similar upper bound on $\|\Pi_{\mathcal{B}}|w\rangle\|^2$, and thus, $|w\rangle$ is a δ-positive witness. $\qquad \square$

5.3 Negative Analysis

We will use the negative history states $|w_-(i)\rangle$ defined in Definition 3.4, and the amplitudes of the states of the outer algorithm at time ℓ, $\beta_{i,b,y} = \langle i, b, y | w_O^\ell \rangle$, defined in (50), to define a negative witness (see Definition 2.9). We first define:

$$|\tilde{w}_{\mathcal{A}}\rangle := \sum_{\substack{\ell \in \{0,\dots,\mathsf{L}-1\}\setminus\mathcal{Q} \\ \text{even}}} \sum_{i,b,y} \beta_{i,b,y}^\ell |\psi_{i,b,y,\ell}^O\rangle + \sum_{\ell\in\mathcal{Q}} \sum_{i,b,y} \beta_{i,b,y}^\ell \left(|\psi_{i,b,y,\ell}^{\rightarrow}\rangle + (-1)^{g_i}|\psi_{i,b,y,\ell}^{\leftarrow}\rangle \right.$$

$$\left. - |w_-(i)\rangle|b,y\rangle|\ell\rangle + (|\rightarrow\rangle|i\rangle - (-1)^{g_i}|\leftarrow\rangle|i\rangle)|0,0\rangle|0\rangle|b,y\rangle|\ell\rangle \right) \tag{64}$$

$$\text{and } |\tilde{w}_{\mathcal{B}}\rangle := \sum_{\substack{\ell\in\{0,\dots,\mathsf{L}-1\}\setminus\mathcal{Q} \\ \text{odd}}} \sum_{i,b,y} \beta_{i,b,y}^\ell |\psi_{i,b,y,\ell}^O\rangle + \sum_{\ell\in\mathcal{Q}} \sum_{i,b,y} \beta_{i,b,y}^\ell |w_-(i)\rangle|b,y\rangle|\ell\rangle,$$

where the outer transition states $|\psi_{i,b,y,\ell}^O\rangle$ are defined in (54), and the connecting states $|\psi_{i,b,y,\ell}^{\rightarrow}\rangle, |\psi_{i,b,y,\ell}^{\leftarrow}\rangle$ are defined in (56). Recall that we assume that all $\ell \in \mathcal{Q}$ are even.

Claim 5.5. $\|(I - \Pi_{\mathcal{A}})|\tilde{w}_{\mathcal{A}}\rangle\|^2 \le 2\mathsf{Q}\epsilon_{\text{avg}}$ and $\|(I - \Pi_{\mathcal{B}})|\tilde{w}_{\mathcal{B}}\rangle\|^2 \le 2\mathsf{Q}\epsilon_{\text{avg}}$.

Proof. Referring to (59), we can see that $|\psi_{i,b,y,\ell}^O\rangle \in \Psi^{\mathcal{A}}$ whenever ℓ is even, and in $\Psi^{\mathcal{B}}$ whenever ℓ is odd. Furthermore, whenever $\ell \in \mathcal{Q}$ (and so by assumption ℓ is even), $|\psi_{i,b,y,\ell}^{\rightarrow}\rangle, |\psi_{i,b,y,\ell}^{\leftarrow}\rangle \in \Psi^{\mathcal{A}}$, $\Psi_1 \otimes |b,y\rangle|\ell\rangle \subset \Psi^{\mathcal{A}}$, and $\Psi_0 \otimes |b,y\rangle|\ell\rangle \subset \Psi^{\mathcal{B}}$, so:

$$\|(I - \Pi_{\mathcal{A}})|\tilde{w}_{\mathcal{A}}\rangle\|^2 \le \sum_{\ell\in\mathcal{Q}}\sum_{i,b,y} |\beta_{i,b,y}^\ell|^2 \, \|(I - \Pi_{\Psi_1})(|w_-(i)\rangle - (|\rightarrow\rangle|i\rangle - (-1)^{g_i}|\leftarrow\rangle|i\rangle)|0,0\rangle|0\rangle)\|^2$$

$$\le \sum_{\ell\in\mathcal{Q}}\sum_{i,b,y} |\beta_{i,b,y}^\ell|^2 2\mathbb{E}[\epsilon_i^{T_i}] \qquad\qquad \text{by Claim 3.9,}$$

$$= 2\sum_{\ell\in\mathcal{Q}}\sum_{i\in[n]} \mathsf{q}_{i,\ell}\epsilon_i = 2\mathsf{Q}\sum_{i\in[n]} \bar{\mathsf{q}}_i\epsilon_i = 2\mathsf{Q}\epsilon_{\text{avg}} \qquad \text{by (51) and (52).}$$

Similarly, again applying Claim 3.9:

$$\|(I - \Pi_{\mathcal{B}})|\tilde{w}_{\mathcal{B}}\rangle\|^2 \le \sum_{\ell\in\mathcal{Q}}\sum_{i,b,y} |\beta_{i,b,y}^\ell|^2 \, \|(I - \Pi_{\Psi_0})|w_-(i)\rangle\|^2 \le \sum_{\ell\in\mathcal{Q}}\sum_{i,b,y} |\beta_{i,b,y}^\ell|^2 2\mathbb{E}[\epsilon_i^{T_i}] = 2\mathsf{Q}\epsilon_{\text{avg}}. \qquad \square$$

Claim 5.6. Let $|\tilde{w}_{\mathcal{A}}\rangle$ and $|\tilde{w}_{\mathcal{B}}\rangle$ be as defined in (64). For $\ell \in \{0,\dots,\mathsf{L}\}$, define

$$|\bar{w}_O^\ell\rangle := \sum_{i,b,y} \beta_{i,b,y}^\ell |i\rangle|0,0\rangle|0\rangle|b,y\rangle,$$

which is just the state of the outer algorithm at time ℓ, $|w_O^\ell\rangle$, with extra registers set to 0. Then

$$|\tilde{w}_{\mathcal{A}}\rangle + |\tilde{w}_{\mathcal{B}}\rangle = |\bot\rangle|0\rangle|0,0\rangle|0\rangle|0,0\rangle|0\rangle - |\bot\rangle|\bar{w}_O^{\mathsf{L}}\rangle|\mathsf{L}\rangle.$$

Proof. Referring to (64), we have:

$$|\tilde{w}_{\mathcal{A}}\rangle + |\tilde{w}_{\mathcal{B}}\rangle = \sum_{\ell=0}^{\mathsf{L}-1}\sum_{i,b,y} \beta_{i,b,y}^\ell |\psi_{i,b,y,\ell}^O\rangle$$

$$+ \sum_{\ell\in\mathcal{Q}}\sum_{i,b,y} \beta_{i,b,y}^\ell \left(|\psi_{i,b,y,\ell}^{\rightarrow}\rangle + (-1)^{g_i}|\psi_{i,b,y,\ell}^{\leftarrow}\rangle + (|\rightarrow\rangle - (-1)^{g_i}|\leftarrow\rangle)|i\rangle|0,0\rangle|0\rangle|b,y\rangle|\ell\rangle \right). \tag{65}$$

Referring to (56), we can compute, for any $\ell \in \mathcal{Q}$:

$$\sum_{i,b,y} \beta_{i,b,y}^\ell |\psi_{i,b,y,\ell}^{\rightarrow}\rangle = \sum_{i,b,y} \beta_{i,b,y}^\ell (|\bot\rangle|i\rangle|0,0\rangle|0\rangle|b,y\rangle|\ell\rangle - |\rightarrow\rangle|i\rangle|0,0\rangle|0\rangle|b,y\rangle|\ell\rangle) = (|\bot\rangle - |\rightarrow\rangle)|\bar{w}_O^\ell\rangle|\ell\rangle.$$

$$\tag{66}$$

Similarly, since $\ell \in \mathcal{Q}$, $V_{\ell+1}$ is a query, so we have $\beta_{i,b,y}^{\ell+1} = (-1)^{g_i}\beta_{i,b,y}^{\ell}$ (see (62)). Thus we have, again referring to (56):

$$\sum_{i,b,y}\beta_{i,b,y}^{\ell}(-1)^{g_i}|\psi_{i,b,y,\ell}^{\leftarrow}\rangle = \sum_{i,b,y}\beta_{i,b,y}^{\ell+1}(|\leftarrow\rangle|i\rangle|0,0\rangle|0\rangle|b,y\rangle|\ell\rangle - |\perp\rangle|i\rangle|0,0\rangle|0\rangle|b,y\rangle|\ell+1\rangle)$$

$$= |\leftarrow\rangle|\bar{w}_O^{\ell+1}\rangle|\ell\rangle - |\perp\rangle|\bar{w}_O^{\ell+1}\rangle|\ell+1\rangle. \tag{67}$$

Still assuming $\ell \in \mathcal{Q}$, we have:

$$\sum_{i,b,y}\beta_{i,b,y}^{\ell}\left(|\rightarrow\rangle|i\rangle - (-1)^{g_i}|\leftarrow\rangle|i\rangle\right)|0,0\rangle|0\rangle|b,y\rangle|\ell\rangle$$

$$= |\rightarrow\rangle\sum_{i,b,y}\beta_{i,b,y}^{\ell}|i\rangle|0,0\rangle|0\rangle|b,y\rangle|\ell\rangle - |\leftarrow\rangle\sum_{i,b,y}\beta_{i,b,y}^{\ell+1}|i\rangle|0,0\rangle|0\rangle|b,y\rangle|\ell\rangle = |\rightarrow\rangle|\bar{w}_O^{\ell}\rangle|\ell\rangle - |\leftarrow\rangle|\bar{w}_O^{\ell+1}\rangle|\ell\rangle. \tag{68}$$

Thus, combining (66), (67) and (68), we have that for all $\ell \in \mathcal{Q}$,

$$\sum_{i,b,y}\beta_{i,b,y}^{\ell}\left(|\psi_{i,b,y,\ell}^{\rightarrow}\rangle + (-1)^{g_i}|\psi_{i,b,y,\ell}^{\leftarrow}\rangle + (|\rightarrow\rangle|i\rangle - (-1)^{g_i}|\leftarrow\rangle|i\rangle)|0,0\rangle|0\rangle|b,y\rangle|\ell\rangle\right)$$

$$= |\perp\rangle\left(|\bar{w}_O^{\ell}\rangle|\ell\rangle - |\bar{w}_O^{\ell+1}\rangle|\ell+1\rangle\right). \tag{69}$$

Next, referring to (54), we can compute:

$$\sum_{\ell\in\{0,\ldots,L-1\}\setminus\mathcal{Q}}\sum_{i,b,y}\beta_{i,b,y}^{\ell}|\psi_{i,b,y,\ell}^{O}\rangle$$

$$= \sum_{\ell\in\{0,\ldots,L-1\}\setminus\mathcal{Q}}\sum_{i,b,y}\beta_{i,b,y}^{\ell}|\perp\rangle\left(|i\rangle|0,0\rangle|0\rangle|b,y\rangle|\ell\rangle - V_{\ell+1}(|i\rangle|0,0\rangle|0\rangle|b,y\rangle)|\ell+1\rangle\right)$$

$$= |\perp\rangle\sum_{\ell\in\{0,\ldots,L-1\}\setminus\mathcal{Q}}|\bar{w}_O^{\ell}\rangle|\ell\rangle - |\perp\rangle\sum_{\ell\in\{0,\ldots,L-1\}\setminus\mathcal{Q}}V_{\ell+1}|\bar{w}_O^{\ell}\rangle|\ell+1\rangle \tag{70}$$

$$= |\perp\rangle\sum_{\ell\in\{0,\ldots,L-1\}\setminus\mathcal{Q}}|\bar{w}_O^{\ell}\rangle|\ell\rangle - |\perp\rangle\sum_{\ell\in\{0,\ldots,L-1\}\setminus\mathcal{Q}}|\bar{w}_O^{\ell+1}\rangle|\ell+1\rangle.$$

Combining (69) and (70) in (65), we can compute:

$$|\tilde{w}_{\mathcal{A}}\rangle + |\tilde{w}_{\mathcal{B}}\rangle = |\perp\rangle\left(\sum_{\ell\in\{0,\ldots,L-1\}\setminus\mathcal{Q}}\left(|\bar{w}_O^{\ell}\rangle|\ell\rangle - |\bar{w}_O^{\ell+1}\rangle|\ell+1\rangle\right) + \sum_{\ell\in\mathcal{Q}}\left(|\bar{w}_O^{\ell}\rangle|\ell\rangle - |\bar{w}_O^{\ell+1}\rangle|\ell+1\rangle\right)\right)$$

$$= \sum_{\ell=0}^{L-1}|\bar{w}_O^{\ell}\rangle|\ell\rangle - \sum_{\ell=0}^{L-1}|\bar{w}_O^{\ell+1}\rangle|\ell+1\rangle = |\perp\rangle|\bar{w}_O^0\rangle|0\rangle - |\perp\rangle|\bar{w}_O^L\rangle|L\rangle.$$

The proof is completed by observing that $|\bar{w}_O^0\rangle = |0\rangle|0,0\rangle|0\rangle|0,0\rangle$ (see (49)). $\qquad\square$

We are now ready to define the negative witness. We let:

$$|w_{\mathcal{A}}\rangle = \frac{1}{\sqrt{w_0}}\left(|\tilde{w}_{\mathcal{A}}\rangle + |\perp\rangle|\bar{w}_O^L\rangle|L\rangle\right) \text{ and } |w_{\mathcal{B}}\rangle = \frac{1}{\sqrt{w_0}}\left(|\psi_{0,v_0}\rangle + |\tilde{w}_{\mathcal{B}}\rangle\right), \tag{71}$$

where $|\psi_{0,v_0}\rangle$ is defined in (57), $|\bar{w}_O^L\rangle$ is defined in Claim 5.6, and $|\tilde{w}_{\mathcal{A}}\rangle$ and $|\tilde{w}_{\mathcal{B}}\rangle$ are defined in (64).

Lemma 5.7. *Suppose $f(g(x)) = 0$. Let*

$$\delta' = \frac{4}{w_0}\left(2Q\epsilon_{\text{avg}} + w_{1,\text{out}}\varepsilon_O + w_{0,\text{out}}(1 - \varepsilon_O)\right).$$

Then $|w_{\mathcal{A}}\rangle$, $|w_{\mathcal{B}}\rangle$ is a δ'-negative witness (see Definition 2.9) with complexity:

$$\||w_{\mathcal{A}}\rangle\|^2 \leq \frac{4}{w_0}\left(L + 2Q(T_{\text{avg}} + 1) + 1\right).$$

Proof. We first verify that $|w_{\mathcal{A}}\rangle + |w_{\mathcal{B}}\rangle = |\psi_0\rangle$ (see (60)), letting $|0^6\rangle = |0\rangle|0,0\rangle|0\rangle|0,0\rangle$:

$$
\begin{aligned}
\sqrt{\mathsf{w}_0}(|w_{\mathcal{A}}\rangle + |w_{\mathcal{B}}\rangle) &= |\psi_{0,v_0}\rangle + |\perp\rangle|0^6\rangle|0\rangle - |\perp\rangle|\bar{w}_O^{\mathsf{L}}\rangle|\mathsf{L}\rangle + |\perp\rangle|\bar{w}_O^{\mathsf{L}}\rangle|\mathsf{L}\rangle && \text{by Claim 5.6} \\
&= \sqrt{\mathsf{w}_0}|\circ\rangle|0^6\rangle|0\rangle - |\perp\rangle|0^6\rangle|0\rangle + |\perp\rangle|0^6\rangle|0\rangle && \text{by (57)} \\
&= \sqrt{\mathsf{w}_0}|\psi_0\rangle && \text{by (60)}.
\end{aligned}
$$

We thus analyze the error of this negative witness. We have:

$$
\|(I - \Pi_{\mathcal{A}})|w_{\mathcal{A}}\rangle\|^2 \leq \frac{4}{\mathsf{w}_0}\left(\|(I - \Pi_{\mathcal{A}})|\tilde{w}_{\mathcal{A}}\rangle\|^2 + \left\||(I - \Pi_{\mathcal{A}})|\perp\rangle|\bar{w}_O^{\mathsf{L}}\rangle|\mathsf{L}\rangle\right\|^2\right). \tag{72}
$$

For the second term, we note that since $|\psi_{i,b,y}^{\text{out}}\rangle \in \Psi^{\mathcal{A}}$, we have:

$$
(I - \Pi_{\mathcal{A}})|\perp\rangle|\bar{w}_O^{\mathsf{L}}\rangle|\mathsf{L}\rangle = (I - \Pi_{\mathcal{A}})\left(|\perp\rangle|\bar{w}_O^{\mathsf{L}}\rangle|\mathsf{L}\rangle - \sum_{i,b,y}\beta_{i,b,y}^{\mathsf{L}}|\psi_{i,b,y}^{\text{out}}\rangle\right)
$$

and, letting Ξ_b denote the orthogonal projector onto states with b in the outer algorithm's answer register:

$$
\begin{aligned}
&|\perp\rangle|\bar{w}_O^{\mathsf{L}}\rangle|\mathsf{L}\rangle - \sum_{i,b,y}\beta_{i,b,y}^{\mathsf{L}}|\psi_{i,b,y}^{\text{out}}\rangle \\
&= |\perp\rangle\sum_{i,b,y}\beta_{i,b,y}^{\mathsf{L}}|i\rangle|0,0\rangle|0\rangle|b,y\rangle|\mathsf{L}\rangle - \sum_{i,b,y}\beta_{i,b,y}^{\mathsf{L}}\left(|\perp\rangle - \sqrt{\mathsf{w}_{b,\text{out}}}|\circ\rangle\right)|i\rangle|0,0\rangle|0\rangle|b,y\rangle|\mathsf{L}\rangle \\
&= \sqrt{\mathsf{w}_{1,\text{out}}}|\circ\rangle\Xi_1|\bar{w}_O^{\mathsf{L}}\rangle|\mathsf{L}\rangle + \sqrt{\mathsf{w}_{0,\text{out}}}|\circ\rangle\Xi_0|\bar{w}_O^{\mathsf{L}}\rangle|\mathsf{L}\rangle.
\end{aligned}
$$

Since $f(g) = 0$, we have $\left\|\Xi_1|w_O^{\mathsf{L}}\rangle\right\|^2 \leq \varepsilon_O$, since this is the probability that the outer algorithm outputs 1 (when $\mathcal{O}_g$ is implemented perfectly). Thus, using $\mathsf{w}_{1,\text{out}} > \mathsf{w}_{0,\text{out}}$:

$$
\left\||(I - \Pi_{\mathcal{A}})|\perp\rangle|\bar{w}_O^{\mathsf{L}}\rangle|\mathsf{L}\rangle\right\|^2 \leq \mathsf{w}_{1,\text{out}}\left\|\Xi_1|\bar{w}_O^{\mathsf{L}}\rangle\right\|^2 + \mathsf{w}_{0,\text{out}}\left(1 - \left\|\Xi_1|\bar{w}_O^{\mathsf{L}}\rangle\right\|^2\right) \leq \mathsf{w}_{1,\text{out}}\varepsilon_O + \mathsf{w}_{0,\text{out}}(1 - \varepsilon_O).
$$

Combining this with Claim 5.5, which gives the upper bound $\|(I - \Pi_{\mathcal{A}})|\tilde{w}_{\mathcal{A}}\rangle\|^2 \leq 2\mathsf{Q}\epsilon_{\text{avg}}$, in (72) we have:

$$
\|(I - \Pi_{\mathcal{A}})|w_{\mathcal{A}}\rangle\|^2 \leq \frac{4}{\mathsf{w}_0}\left(2\mathsf{Q}\epsilon_{\text{avg}} + \mathsf{w}_{1,\text{out}}\varepsilon_O + \mathsf{w}_{0,\text{out}}(1 - \varepsilon_O)\right) = \delta'.
$$

Similarly, since $|\psi_{0,v_0}\rangle \in \mathcal{B}$, applying Claim 5.5, we have:

$$
\|(I - \Pi_{\mathcal{B}})|w_{\mathcal{B}}\rangle\|^2 \leq \frac{1}{\mathsf{w}_0}\|(I - \Pi_{\mathcal{B}})|\tilde{w}_{\mathcal{B}}\rangle\|^2 \leq \frac{2\mathsf{Q}}{\mathsf{w}_0}\epsilon_{\text{avg}} < \delta'.
$$

To complete the proof, we give an upper bound on $\||w_{\mathcal{A}}\rangle\|^2$. By (71), we have:

$$
\||w_{\mathcal{A}}\rangle\|^2 \leq \frac{4}{\mathsf{w}_0}\left(\||\tilde{w}_{\mathcal{A}}\rangle\|^2 + \left\||\perp\rangle|\bar{w}_O^{\mathsf{L}}\rangle|\mathsf{L}\rangle\right\|^2\right) = \frac{4}{\mathsf{w}_0}\left(\||\tilde{w}_{\mathcal{A}}\rangle\|^2 + 1\right). \tag{73}
$$

Using the fact (see (56)):

$$
\begin{aligned}
&|\psi_{i,b,y,\ell}^{\rightarrow}\rangle + (-1)^{g_i}|\psi_{i,b,y,\ell}^{\leftarrow}\rangle + (|\rightarrow\rangle|i\rangle - (-1)^{g_i}|\leftarrow\rangle|i\rangle)|0,0\rangle|0\rangle|b,y\rangle|\ell\rangle \\
&= |\perp\rangle|i\rangle|0,0\rangle|0\rangle|b,y\rangle|\ell\rangle - (-1)^{g_i}|\perp\rangle|i\rangle|0,0\rangle|0\rangle|b,y\rangle|\ell+1\rangle
\end{aligned}
$$

and referring to (64), we have:

$$\||\tilde{w}_{\mathcal{A}}\rangle\|^2 \leq \sum_{\substack{\ell \in \{0,\ldots,L-1\}\backslash\mathcal{Q} \\ \text{even}}} \sum_{i,b,y} |\beta_{i,b,y}^\ell|^2 \||\psi_{i,b,y,\ell}^O\rangle\|^2$$

$$+ \sum_{\ell \in \mathcal{Q}} \sum_{i,b,y} |\beta_{i,b,y}^\ell|^2 \||\bot\rangle|i\rangle|0,0\rangle|0\rangle|b,y\rangle(|\ell\rangle - (-1)^{g_i}|\ell+1\rangle) - |w_-(i)\rangle|b,y\rangle|\ell\rangle\|^2$$

$$= 2 \sum_{\substack{\ell \in \{0,\ldots,L-1\}\backslash\mathcal{Q} \\ \text{even}}} \left\||w_O^\ell\rangle\right\|^2 + \sum_{\ell \in \mathcal{Q}} \sum_{i \in [n]} \mathsf{q}_{i,\ell} \left(2 + \||w_-(i)\rangle\|^2\right) \qquad \text{by (51)}$$

$$= 2(L/2 - Q) + 2Q + \sum_{i \in [n]} Q\bar{\mathsf{q}}_i \||w_-(i)\rangle\|^2 \leq L + \sum_{i \in [n]} Q\bar{\mathsf{q}}_i 2\mathbb{E}\left[T_i + 1\right],$$

by (52) and Corollary 3.7, using $\alpha_t = 1$. Plugging this into (73), we have:

$$\||w_{\mathcal{A}}\rangle\|^2 \leq \frac{4}{\mathsf{w}_0}\left(L + 2Q\left(\sum_{i \in [n]} \bar{\mathsf{q}}_i \mathbb{E}[T_i] + 1\right) + 1\right) \leq \frac{4}{\mathsf{w}_0}\left(L + 2Q(T_{\text{avg}} + 1) + 1\right),$$

by the assumption of Theorem 5.1. $\qquad\square$

5.4 Conclusion of Proof of Theorem 5.1

We set the parameters w_0, $\mathsf{w}_{1,\text{out}}$ and $\mathsf{w}_{0,\text{out}}$ as follows:

$$\mathsf{w}_0 := \frac{1}{L + 1 + 2Q(T_{\text{avg}} + 1)}, \quad \mathsf{w}_{1,\text{out}} := \frac{1 - \varepsilon_O}{8}\mathsf{w}_0, \quad \text{and} \quad \mathsf{w}_{0,\text{out}} := \frac{\varepsilon_O}{8}\mathsf{w}_0.$$

Note that $\mathsf{w}_{1,\text{out}} > \mathsf{w}_{0,\text{out}}$ is satisfied.

We apply Theorem 2.11 with H as in (53), $|\psi_0\rangle = |\circ\rangle|0\rangle|0,0\rangle|0\rangle|0,0\rangle|0\rangle$ (as in (60)), and $\Psi^{\mathcal{A}}$ and $\Psi^{\mathcal{B}}$ as in (59). First note that we can easily generate $|\psi_0\rangle$ in cost $\mathsf{S} = O(1)$. By Lemma 5.2 we can implement $U_{\mathcal{AB}}$ in cost $\mathsf{A} = \log T + \log L$. Let

$$c_+ = 18 \quad \text{and} \quad \mathcal{C}_- = 4(L + 1 + 2Q(T_{\text{avg}} + 1))^2$$

and let

$$\delta = 2\mathsf{w}_0 Q\epsilon_{\text{avg}} = \frac{2Q\epsilon_{\text{avg}}}{L + 1 + 2Q(T_{\text{avg}} + 1)}$$

and $\delta' = \frac{4}{\mathsf{w}_0}\left(2Q\epsilon_{\text{avg}} + \mathsf{w}_{1,\text{out}}\varepsilon_O + \mathsf{w}_{0,\text{out}}(1 - \varepsilon_O)\right) = 8Q\epsilon_{\text{avg}}(L + 1 + 2Q(T_{\text{avg}} + 1)) + \varepsilon_O(1 - \varepsilon_O)$

as in Lemma 5.4 and Lemma 5.7. Then, using the assumption in the theorem statement that $\epsilon_{\text{avg}} \leq \eta/(Q(L + QT_{\text{avg}}))$, we can verify that:

$$\delta = 2Q\epsilon_{\text{avg}}\frac{4(L + 1 + 2Q(T_{\text{avg}} + 1))^2}{L + 1 + 2Q(T_{\text{avg}} + 1)}\frac{1}{\mathcal{C}_-} \leq \frac{2Q\eta}{Q(L + QT_{\text{avg}})}4(L + 1 + 2Q(T_{\text{avg}} + 1))\frac{1}{\mathcal{C}_-} \leq \frac{1}{(8c_+)^3\pi^8}\frac{1}{\mathcal{C}_-}$$

when η is a sufficiently small constant, since $T_{\text{avg}} \geq 1$; and

$$\delta' \leq 8Q\eta\frac{L + 1 + 2Q(T_{\text{avg}} + 1)}{Q(L + QT_{\text{avg}})} + \varepsilon_O(1 - \varepsilon_O) \leq \frac{3}{4}\frac{1}{\pi^4 c_+}$$

for sufficiently small constants ε_O and η. Furthermore:

Positive Condition: By Lemma 5.4, if $f(g) = 1$, there is a δ-positive witness with

$$\frac{\||w\rangle\|^2}{|\langle w|\psi_0\rangle|^2} \leq 1 + \mathsf{w}_0\left(L + 1 + 2Q(T_{\text{avg}} + 1)\right) + \frac{\mathsf{w}_0}{\mathsf{w}_{0,\text{out}}}\varepsilon_O + \frac{\mathsf{w}_0}{\mathsf{w}_{1,\text{out}}}(1 - \varepsilon_O)$$

$$= 1 + 1 + 8 + 8 = 18 = c_+.$$

Negative Condition: By Lemma 5.7, if $f(g) = 0$, there is a δ'-negative witness with $\||w_{\mathcal{A}}\rangle\|^2 \leq \mathcal{C}_-$.

Thus, by Theorem 2.11, there is a quantum algorithm that decides $f \circ g$ with bounded error in complexity:

$$O(\mathsf{S} + \sqrt{\mathcal{C}_-}\mathsf{A}) = O((\mathsf{L} + \mathsf{QT}_{\text{avg}})(\log \mathsf{L} + \log \mathsf{T})).$$

6 Acknowledgements

References

[ACR+07] Andris Ambainis, Andrew M. Childs, Ben W. Reichardt, Robert Špalek, and Shengyu Zhang. Every AND-OR formula of size n can be evaluated in time $n^{\frac{1}{2}+o(1)}$ on a quantum computer. In *Proceedings of the 48th IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 363–372, 2007. 2

[AGJ20] Simon Apers, András Gilyén, and Stacey Jeffery. A unified framework of quantum walk search. In *Proceedings of the 38th Symposium on Theoretical Aspects of Computer Science (STACS)*, pages 6:1–6:13, 2020. arXiv: 1912.04233 6, 12, 27

[Amb10] Andris Ambainis. Quantum search with variable times. *Theory of Computing Systems*, 47:786–807, 2010. arXiv: quant-ph/0609168 2, 9

[Amb12] Andris Ambainis. Variable time amplitude amplification and quantum algorithms for linear algebra problems. In *Proceedings of the 29th Symposium on Theoretical Aspects of Computer Science (STACS)*, pages 636–647, 2012. arXiv: 1010.4458 3, 9, 13

[BCJ+13] Aleksandrs Belovs, Andrew M. Childs, Stacey Jeffery, Robin Kothari, and Frédéric Magniez. Time-efficient quantum walks for 3-distinctness. In *Proceedings of the 40th International Colloquium on Automata, Languages, and Programming (ICALP)*, pages 105–122, 2013. 5

[Bel13] Aleksandrs Belovs. Quantum walks and electric networks. arXiv: 1302.3143, 2013. 5, 12

[BJY24] Aleksandrs Belovs, Stacey Jeffery, and Duyal Yolcu. Taming quantum time complexity. *Quantum*, page 1444, 2024. arXiv: 2311.15873 5

[BdW03] Harry Buhrman and Ronald de Wolf. Quantum zero-error algorithms cannot be composed. *Information Processing Letters*, 87(2):79–84, 2003. arXiv: quant-ph/0211029 2

[BY23] Aleksandrs Belovs and Duyal Yolcu. One-way ticket to as vegas and the quantum adversary. arXiv: 2301.02003, 2023. 4, 5

[CJKM13] Andrew M. Childs, Stacey Jeffery, Robin Kothari, and Frédéric Magniez. A time-efficient quantum walk for 3-distinctness using nested updates. arXiv: 1302.7316, 2013. 6

[CJOP20] Arjan Cornelissen, Stacey Jeffery, Maris Ozols, and Alvaro Piedrafita. Span programs and quantum time complexity. In *Proceedings of the 45th International Symposium on Mathematical Foundations of Computer Science (MFCS)*, pages 21:1–26:14, 2020. arXiv: 2005.01323 4

[CKKD+22] Andrew M. Childs, Robin Kothari, Matt Kovacs-Deak, Aarthi Sundaram, and Daochen Wang. Quantum divide and conquer. arXiv: 2210.06419, 2022. 4

[CLM20]   Titouan Carette, Mathieu Laurière, and Frédéric Magniez. Extended learning graphs for triangle finding. *Algorithmica*, 82:980–1005, 2020. 6, 7

[Cor23]   Arjan Cornelissen. *Quantum multivariate estimation and span program algorithms*. PhD thesis, University of Amsterdam, 2023. 3

[CRR+96]   Ashok K. Chandra, Prabhakar Raghavan, Walter L. Ruzzo, Roman Smolensky, and Prasoon Tiwari. The electrical resistance of a graph captures its commute and cover times. *Computational Complexity*, 6(4):312–340, 1996. 6, 12

[DH17]   Cătălin Dohotaru and Peter Høyer. Controlled quantum amplification. In *Proceedings of the 44th International Colloquium on Automata, Languages, and Programming (ICALP)*, pages 18:1–18:13, 2017. 6

[DS84]   Peter G. Doyle and J. Laurie Snell. *Random walks and electric networks*. Mathematical Association of America, 1984. arXiv: math/0001057 11

[JKM12]   Stacey Jeffery, Robin Kothari, and Frédéric Magniez. Nested quantum walks with quantum data structures. In *Proceedings of the 24th ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1474–1485, 2012. arXiv: 1210.1199 6

[JZ22]   Stacey Jeffery and Sebastian Zur. Multidimensional quantum walks and application to k-distinctness. arXiv: 2208.13492, 2022. 2, 3, 4, 6, 8, 12, 20, 21

[Kit96]   Alexei Y. Kitaev. Quantum measurements and the Abelian stabilizer problem. *ECCC*, TR96-003, 1996. arXiv: quant-ph/9511026 12

[LG14]   François Le Gall. Improved quantum algorithm for triangle finding via combinatorial arguments. In *Proceedings of the 55th IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 216–225, 2014. arXiv: 1407.0085 6

[LWP09]   David A. Levin, Elizabeth Wilmer, and Yuval Peres. *Markov chains and mixing times*. AMS, Providence, RI, USA, 1st edition, 2009. 10, 11, 28

[MNRS11]   Frédéric Magniez, Ashwin Nayak, Jérémie Roland, and Miklos Santha. Search via quantum walk. *SIAM Journal on Computing*, 40(1):142–164, 2011. Earlier version in STOC'07. arXiv: quant-ph/0608026 6

[Rei11a]   Ben W. Reichardt. Faster quantum algorithm for evaluating game trees. In *Proceedings of the 22nd ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 546–559, 2011. 2

[Rei11b]   Ben W. Reichardt. Span-program-based quantum algorithm for evaluating unbalanced formulas. In *Proceedings of the 9th Conference on the Theory of Quantum Computation, Communication, and Cryptography (TQC)*, pages 73–103, 2011. 4

[RŠ12]   Ben W. Reichardt and Robert Špalek. Span-program-based quantum algorithm for evaluating formulas. *Theory of Computing*, 8(13):291–319, 2012. 4

[Sze04]   Mario Szegedy. Quantum speed-up of Markov chain based algorithms. In *Proceedings of the 45th IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 32–41, 2004. arXiv: quant-ph/0401053 5