On proving the robustness of algorithms for early fault-tolerant quantum computers

Rutuja Kshirsagar, Amara Katabarwa, and Peter D. Johnson

Zapata Computing Inc., 100 Federal Street, Boston, MA 02110, USA

arXiv:2209.11322v3 [quant-ph] 13 Feb 2025

The hope of the quantum computing field is that quantum architectures are able to scale up and realize fault-tolerant quantum computing. Due to engineering challenges, such "cheap" error correction may be decades away. In the meantime, we anticipate an era of "costly" error correction, or *early fault-tolerant quantum computing*. Costly error correction might warrant settling for error-prone quantum computations. This motivates the development of quantum algorithms which are robust to some degree of error as well as methods to analyze their performance in the presence of error. Several such algorithms have recently been developed; what is missing is a methodology to analyze their robustness. To this end, we introduce a randomized algorithm for the task of phase estimation and give an analysis of its performance under two simple noise models. In both cases the analysis leads to a noise threshold, below which arbitrarily high accuracy can be achieved by increasing the number of samples used in the algorithm. As an application of this general analysis, we compute the maximum ratio of the largest circuit depth and the dephasing scale such that performance guarantees hold. We calculate that the randomized algorithm can succeed with arbitrarily high probability as long as the required circuit depth is less than 0.916 times the dephasing scale.

Note: Equation (4) and output of Algorithm1 have been updated to accommodate for a missed negative sign. This does not affect any analysis presented in this paper.

Rutuja Kshirsagar: rkshirsagar@fujitsu.com, The author has moved to Fujitsu Research of America, Inc.

Amara Katabarwa: amara@zapata.ai

Peter D. Johnson: peter@zapatacomputing.com

1 Introduction

For several decades researchers have been in pursuit of quantum algorithms that would solve important problems in areas such as chemical and materials simulation [6], optimization [7], and cryptography [10]. It has been assumed that the machine needed to solve such problems is a large-scale fault-tolerant quantum computer [11]. Recent work has sought to develop methods for solving these problems using smaller, error-prone quantum devices [12, 13, 14]. Unfortunately, for many application areas, there is mounting evidence that some degree of quantum error correction will be needed to achieve quantum advantage in those areas [2, 15].

Quantum error correction is used to reduce error rates of quantum operations at the cost of overhead in number of physical qubits and operating time [16]. In the long-term the quantum computing field hopes to realize large-scale fault-tolerant quantum architectures that are able to implement (nearly) error-free, large quantum computations with relatively little quantum error correction overhead [17, 18]. It is possible, however, that for the early quantum computers that might be used to achieve quantum advantage, error reduction with quantum error correction will require large overhead [19] due to the challenges of scaling up system sizes. For instance, although the engineering achievements are impressive, recent advances in experimental demonstrations of quantum error correction [20] realized a only 4% reduction in logical state initialization error rates as the overhead is increased from a distance-3 to a distance-5 surface code. This is in contrast to predictions [21] of realizing a 90% to 96% reduction in error rate for the same increase in surface code distance in the future.

In the era of early fault-tolerant quantum computing (EFTQC), when error reduction is costly, it will be important to use error correction resources prudently. If a quantum algorithm is able to tolerate a degree of error, then this tolerance should be exploited to reduce the error correction resources consumed in

the computation. Quantum advantage will be realized when the quantum resources needed to solve useful computational problems can be supplied by quantum hardware. Over time there will be an increase in the quantum resources that can be supplied. But, considering the desire to realize quantum advantage as soon as possible, it is pertinent to determine the minimal quantum resources needed to run a given quantum algorithm. Considering that these quantum resources are used to reduce error, *what is the minimal error reduction required to ensure that a given algorithm will succeed?* To answer this question, we must investigate how error in quantum computations impacts the performance of quantum algorithms.

Examples of previous work on analyzing the robustness of quantum algorithms have been empirical [22], have focused on asymptotic scaling statements [23], or have considered error that is disconnected from a physical model [24]. More recently, there has been increasing focus [25, 26, 27] on the development of quantum algorithms that have a built-in robustness. Through such work, empirical performance models have been developed [15] and tested [28]. However, for the purposes of large-scale resource estimates, which help to predict the onset of quantum advantage with early fault-tolerant quantum computers, there is a need for provable robustness results which are more closely connected with realistic error models.

In this work, we develop an algorithm for the archetypal task of *phase estimation* (see Section 2). The motivating feature of our algorithm is the fact that it facilitates a simple analysis of its robustness under various algorithmic noise models (see Section 3). While a robustness analysis can be carried out for other quantum phase estimation algorithms, a goal of this work is to develop a foundation for the robustness analysis of algorithms beyond quantum phase estimation. The key insight allowing for this is that important quantum algorithms can be recast in terms of classical signal processing. Tools from classical signal processing have facilitated the development of recent quantum algorithms for early fault-tolerant quantum computing [29, 30, 31]. In these algorithms and ours, the quantum computer is used to generate a stochastic estimate of a time signal. The data from this estimated time signal is then processed with classical computing to return the output of the computation[1]. Casting this procedure as signal processing, we can answer questions about the performance of this signal processing when the input signal is subject to noise. This signal processing perspective serves as a foundation for analyzing the robustness of quantum algorithms more generally and a gateway to exploring application to algorithms such as amplitude estimation [25], ground state energy estimation [34], ground state property estimation [30], linear systems solvers [35], and factoring [10].

As with several recently developed quantum algorithms [29, 36], the one developed in this work is randomized. Such algorithms generate random variables, which, in expectation are the desired quantity. In the case of our algorithm, the desired quantity is the discrete Fourier transform of the function $g(k) = e^{ik\theta}$ and the randomization is over the "time variable" k. As such, we refer to the algorithm as *randomized Fourier estimation (RFE)*. The algorithm locates the peak of the estimated Fourier spectrum to yield an estimate of θ. The fact that the algorithm is randomized is what facilitates a simple analysis of its performance in the presence of device error.

In order to analyze robustness we must model noise. It is very challenging to accurately model the impact of device error on quantum algorithms. Accordingly, we will take an intermediate approach. The effect of device error on a quantum algorithm's performance is mediated through the way that device error affects quantum measurement outcome probabilities. Rather than propose a model for, say, gate-level quantum circuit errors, we will develop simple, general models for the deviation of outcome probabilities. We will refer to these as *algorithm error models*. Similar considerations affecting the outcome probabilities have been made in [57]. The bounds on the error tolerated by our algorithm are weaker as compared to the bounds on the additive error tolerated by RPE [Theorem 1, [57]]. However, our algorithm can be extended to developing other robust algorithms for amplitude estimation, ground state energy estimation, etc. That is, by rewriting the input ansatz as a linear combination of eigenstates, the Fourier spectrum generated by our algorithm corresponds to a time signal with multiple peaks which can be detected by RFE by choosing the maximum phase in specified intervals. We leave this analysis for future work. We anticipate that the error will be larger when the initial state is not an exact eigenvec-

[1] As a technical note, such methods vary in the *kernel function* they use to process the time signal estimate. Lin and Tong [29] use a Heaviside function, Wang et al. [31] use a Gaussian derivative function, and this work uses a Dirichlet kernel function (see Eq. 7).

	Expected Runtime
RFE (Noiseless) Thm. 2.1	$\frac{81\pi^3}{2\epsilon}\ln\left(\frac{8\pi}{\delta\epsilon}\right)$
RFE (Adversarial Noise) Thm. 3.1	$\frac{81\pi^3}{2\epsilon}\ln\left(\frac{8\pi}{\delta\epsilon}\right)\left(1-\frac{9\pi}{2\sqrt{2}}\bar{\eta}\right)^{-2}$
RFE (Random Noise) Cor. 3.2	$\frac{81\pi^3}{2\epsilon}\ln\left(\frac{16\pi}{\delta\epsilon}\right)\left(1-\frac{9\sigma}{8}\sqrt{\frac{\epsilon\pi}{\ln\left(\frac{16\pi}{\delta\epsilon}\right)}}\right)^{-2}$
RPE [57]	$\frac{2\pi}{\epsilon}\log\left(\frac{1}{\sqrt{\pi}\delta}\right)\frac{\log\left(\frac{1}{2}\left(1-\sqrt{2}\bar{\eta}\right)^2\right)}{\log\left(1-\frac{1}{2}\left(1-\sqrt{2}\bar{\eta}\right)^2\right)}$

Table 1: Comparison of the performance of the randomized Fourier estimation (RFE) algorithm in the noiseless setting and under the two noise settings analyzed in this work. In each case, the expected circuit depth (measured in number of c-U operations) is $\frac{\pi}{\epsilon}$. The primary feature of the expected runtime in the noisy settings is an additional factor that is greater than 1 and goes to infinity as the noise parameters $\bar{\eta}$ and σ reach a certain threshold. For fixed target accuracy ϵ and success probability δ, as long as the noise parameters are less than a certain value, the algorithm can be guaranteed to succeed with $1-\delta$ probability by taking this factor increase in samples. These algorithms and settings are also compared to the runtime of robust phase estimation (RPE) algorithm of [57]. For the sake of comparison, we have derived a runtime upper bound for this algorithm in terms of its accuracy and failure rate, and using the fact that $\bar{\eta}$, as defined in Sec. 3.1, is twice δ', as defined in Eq. V. 20 of [57]. While RPE is more performant than RFE for the task of phase estimation, the methodology we use to analyze RFE is generalizable to a broader class of algorithms for early fault-tolerant quantum computers.

tor. The error tolerance in such a case is currently being analyzed.

The remaining paper is structured as follows. In Section 2 we give a gentle introduction to the phase estimation algorithm and its performance in the noiseless setting. In Section 3 we develop two different algorithm error models and analyze the performance of randomized Fourier estimation under these conditions. In Section 4 we discuss the implications of our analysis and provide an outlook to future work. For ease of reading, we have reserved the lemmas and proofs of the main theorems to the appendix.

2 Randomized Fourier estimation

The central task we consider is the estimation of the phase angle θ, defined from

$$U|\psi\rangle = e^{i\theta}|\psi\rangle,$$

where $|\psi\rangle$ is an eigenstate of U. We assume that at our disposal is a preparation of the state $|\psi\rangle$ and the ability to apply the controlled unitary $c-U := |0\rangle\langle 0|\otimes\mathbb{I} + |1\rangle\langle 1|\otimes U$. Phase estimation is a fundamental subroutine of several quantum algorithms[2] including Shor's factoring algorithm [10], Harrow, Hassidim, and Lloyd's linear systems algorithm [37], and the quantum counting algorithm [38]. Furthermore, the robust estimation of quantum phases is closely related to the task of parameter estimation in the context of quantum metrology [39] and Hamiltonian learning [40, 41]. Many techniques have been developed in these contexts [9, 48, 42] and some have been implemented in experiments [43, 44]. Continued research at the interface of quantum computing and quantum metrology will be valuable in pushing the capabilities of quantum computers.

Our goal, slightly different than these above works, is to develop a framework for analyzing the robustness of quantum algorithms for early fault-tolerant quantum computing more broadly. Towards this, we develop a simple algorithm for the task of phase estimation that shares structural similarity with several other algorithms designed for early fault-tolerant quantum computing including those for ground state energy estimation [29, 31] and ground state property estimation (GSPE) [30]. This framework applies to algorithms which use collective data obtained from single-qubit measurement outcomes. Our hope then is that this framework and methods can be helpful in analyzing the robustness of such quantum algorithms.

Before introducing the algorithm, we note that, after this manuscript originally appeared, several related EFTQC algorithms have been developed and their robustness analyzed with alternative approaches. Ni et al. [45] developed a robust quantum algorithm for phase estimation that can succeed even when the input state is imperfect. Li et al. [46] developed a robust quantum algorithm for multiparameter phase estimation. A variant of the algorithm introduced below was analyzed in [47], proving it to be robust to "exponential decay error" and show-

[2] An important note is that, in their basic forms, these algorithms require a coherent implementation of quantum phase estimation, whereas this work and others consider incoherent versions of the algorithm.

ing that the algorithm can reduce quantum resources compared to the traditional quantum phase estimation algorithm. Most recently, an algorithm for ground state energy estimation [8] was developed based on the QCELS method [34] and also shown to be robust to exponential decay error.

With this context established, we introduce the robust Fourier estimation algorithm. The algorithm generates and processes single-qubit measurement data obtained from Hadamard tests as shown in Figure 1.

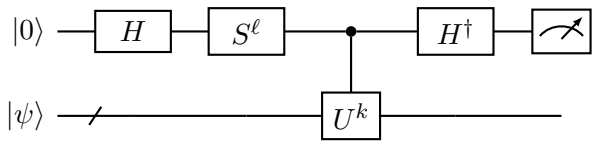

Figure 1: Circuit diagram of the Hadamard, where the application of phase gate $S = \begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix}$ is toggled by $\ell = 0, 1$.

In the case that $\ell = 0$, for Hadamard test of U^k the quantum computer gives us $c = \pm 1$ outcomes with probability

$$\Pr(c|k) = \frac{1 + c\cos(k\theta)}{2}, \tag{1}$$

where $\mathrm{Re}(\langle\psi|U|\psi\rangle) = \cos(\theta) = \Pi$. We call this procedure a *real Hadamard test*. If a phase gate S is placed before the ancilla measurement (that is $\ell = 1$), the probability of $s = \pm 1$ is

$$\Pr(s|k) = \frac{1 + s\sin(k\theta)}{2}. \tag{2}$$

We call this procedure an *imaginary Hadamard test*.

A simple strategy for estimating $\cos\theta$ (and therefore θ) is to set $k = 1$, make a number of Hadamard test measurements with $\ell = 0$, and compute the sample mean of c. This approach is often called the *prepare-and-measure* strategy. Using Hadamard tests with $k \geq 1$, we can gain more information about θ per measurement compared to the $k = 1$ case. The reason why data from these Hadamard tests gives more information about θ is that, for large k, the above likelihoods are highly sensitive to the value of θ.

We introduce the randomized Fourier estimation algorithm by drawing an analogy between the current setting and a signal processing task. Consider k to be a "time variable". The biases in the probabilities of the real and imaginary Hadamard tests then encode oscillating time signals,

$$\cos k\theta = \Pr(c = 1|k) - \Pr(c = -1|k) \tag{3}$$
$$\sin k\theta = \Pr(s = -1|k) - \Pr(s = 1|k). \tag{4}$$

These are the expected values of c and s. If we consider these values to be the real and imaginary parts of a complex time signal $g(k)$, we have

$$g(k) = \cos(k\theta) + i\sin(k\theta) = e^{ik\theta}. \tag{5}$$

For fixed k, if we take a real Hadamard test sample c and an imaginary Hadamard test sample s, we can form a random variable whose expected value is the time signal at times k,

$$\mathbb{E}(c + is) = e^{ik\theta} = g(k). \tag{6}$$

Choosing some large maximum time K, if we were to collect enough samples to get an accurate estimate of the time signal for times $k = 0, \ldots, K-1$, then we might expect to be able to recover the frequency, θ, from this time signal estimate.

One way to recover the frequency is to calculate the discrete Fourier transform (DFT) of the estimated time signal. In the infinite-sample limit, when the time signal estimate is $g(k) = e^{ik\theta}$, the DFT will peak around a value corresponding to θ (due to the finite time window, there will be "leakage" effects which create additional smaller false peaks (see blue curves in the Fourier spectrum plot of Figure 2). When using a finite, but sufficiently large, number of samples, we expect the DFT of the estimated time signal to still peak around the true peak. The time signal has K different times. So, getting a constant-accuracy estimate of the time signal requires $\Omega(K)$ samples. An average sample takes time $\Omega(K)$, leading to a total runtime of $\Omega(K^2)$. Fortunately, we do not need a very accurate estimate of the time signal in order to get an accurate estimate of the location of the true peak in the Fourier domain. With just $O(\log(K))$ samples, the peak can be estimated to within $O(\frac{1}{K})$ accuracy with high probability. The total (expected) runtime is then $O(K\log(K))$.

A rigorous description of this algorithm is given in Algorithm 1.

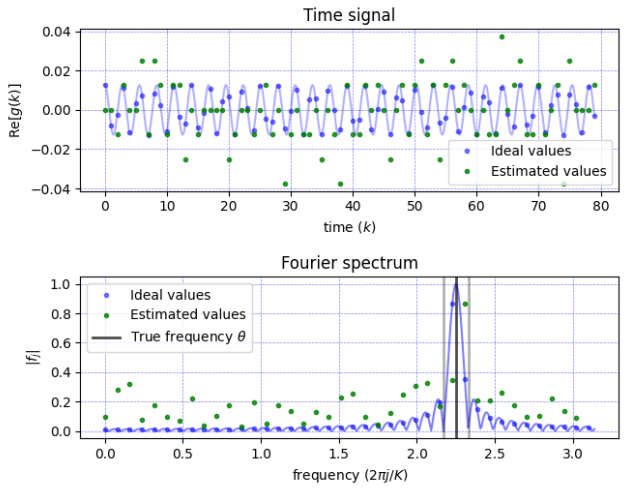

Figure 2: This figure depicts a successful instance of the randomized Fourier estimation algorithm. Despite the fact that the estimate of the time signal (green dots in upper panel) is rather poor, the estimate of the Fourier spectrum (green dots in lower panel) exhibits a spike within $\epsilon = 0.08$ (the grey vertical lines) of the true frequency of $\theta = 2.25$ (the black vertical line). Here we have chosen to take $M = 80$ samples, which is 40 times fewer than the number considered in our algorithm upper bound (c.f. Theorem 2.1). The blue dots and blue curves denote the ideal values and continuous functions that define them, respectively.

Algorithm1 Randomized Fourier Estimation

Input: M, $K \in \mathbb{N}$ ▷ Number of samples and grid size, respectively.

Output: $\hat{\theta} \in [0, \pi]$

$\hat{f} \leftarrow \vec{0} \in \mathbb{C}^K$ ▷ Initialize estimate of Fourier coefficients.

for $i \in \{1, \ldots, M\}$ **do**

Draw k_i uniformly from the set $\{0, \ldots, K-1\}$.

Call HADAMARD$[|\psi\rangle, U^{k_i}]$ to output (c_i, s_i).

for $j \in \{0, \ldots, K-1\}$ **do**

$\hat{f}_j \leftarrow \hat{f}_j + (c_i + is_i)e^{-2\pi i k_i j/K}/M$ ▷ Update estimate of Fourier coefficient.

end for

end for

$\hat{\theta} \leftarrow 2\pi - \frac{2\pi}{K} \arg\max_j |\hat{f}_j|$ ▷ Locate the largest-magnitude Fourier coefficient.

We give a simple explanation for how Algorithm 1 achieves this performance. The core concept of the algorithm is the fact that, with very few samples, the peak of the discrete Fourier transform can be detected. Each call to the subroutine HADAMARD$[|\psi\rangle, U^{k_i}]$ returns the outcomes c and s of the real and imaginary Hadamard tests, respectively, as shown in Figure 1. Surprisingly, from each real-imaginary Hadamard test sample, we can construct an unbiased estimate of the discrete Fourier transform. This is achieved with a randomization over the choice of $k \in \{0, \ldots, K-1\}$. Drawing k from the uniform distribution over these values, we form the jth discrete Fourier coefficient estimate $\hat{f}_j = (c + is)e^{\frac{-2\pi ijk}{K}}$, where $j \in \{0, \ldots, K-1\}$. As seen in the following calculation, the expected value of this random variable is the discrete Fourier transform of the time signal $g(k)$,

$$\begin{aligned} \mathbb{E}\hat{f}_j &= \mathbb{E}_k \mathbb{E}_{c|k} \mathbb{E}_{s|k} (c + is) e^{\frac{-2\pi ijk}{K}} \\ &= \frac{1}{K} \sum_{k=0}^{K-1} \exp(ik\theta)\, e^{\frac{-2\pi ijk}{K}} \\ &= \frac{1}{K} \frac{1 - \exp\left(iK(\theta - \frac{2\pi j}{K})\right)}{1 - \exp\left(i(\theta - \frac{2\pi j}{K})\right)}. \end{aligned} \tag{7}$$

We define $f_j := \mathbb{E}\hat{f}_j$. Viewed as a continuous function in j, the magnitude of f_j peaks at $j = \frac{K\theta}{2\pi}$. The algorithm works by averaging M independent Fourier coefficient estimates and then locating the index, or "frequency", j for which the magnitude of the estimated coefficient is largest. This frequency estimate j corresponds to the estimate of θ via $\hat{\theta} = \frac{2\pi j}{K}$. We will refer to the j as "frequencies", as they index the frequency dimension, while k indexes the time dimension. The granularity of the estimates is governed by K; the two frequencies closest to θ will be off by no more than $\frac{2\pi}{K}$. Choosing $K = \lceil \frac{2\pi}{\epsilon} \rceil$ then ensures that outputting either of these "adjacent" frequencies gives $|\hat{\theta} - \theta| \leq \epsilon$. The frequency "closest" to θ (or frequencies closest, in the case they are equidistant from θ) will be off by no more than $\frac{\pi}{K}$. Our analysis will rely on the fact that this close frequency (or these close frequencies) will have a large expected Fourier coefficient, while the non-adjacent frequencies will have small expected Fourier coefficients[3].

[3]The algorithm succeeds when either close frequency is the output estimate. Yet, we choose to disregard the properties of the frequency that is (or frequencies that are) adjacent, yet not close (i.e. j satisfying $\frac{1}{2} < |j - \frac{K\theta}{2\pi}| \leq 1$). Although the performance bound is looser, in our analysis we only count the algorithm as succeeding when the magnitude of the estimate of the close frequency Fourier component estimate is larger than a fixed value *and* the magnitude of the estimate of all non-adjacent frequency Fourier component estimates are smaller than that fixed value.

Why will this estimate be close to θ with high probability? The analysis uses two facts about the expected Fourier coefficients: 1) $|f_j|$ of a close frequency is relatively large (Corollary A.4 shows it is greater than $\frac{2}{\pi}$) and 2) for all frequencies j that would give estimates of θ with error greater than $\frac{2\pi}{K} = \epsilon$, "non-adjacent frequencies", $|f_j|$ is relatively small (Corollary A.6 shows they are less than $\frac{10}{9\pi}$). This gap of $\frac{8}{9\pi}$ between $|f_j|$ of close and non-adjacent frequencies gives a buffer for the statistical estimates of f_j; even if all of the estimates $\hat{f}_j$ are off by up to $\frac{4}{9\pi}$, then the magnitude of $\hat{f}_j$ for the close frequency will still be larger than those of the non-adjacent frequencies. In other words, inaccurate estimates (up to $\frac{4}{9\pi}$ off) still lead the algorithm to succeed.

The final step in the analysis is to assess the likelihood of the close frequency being chosen. Loosely, the Hoeffding bound ensures that a mean over M samples will deviate by $O(1)$ with no more than $O(\exp(-M))$ chance. Combining this with a union bound gives that the likelihood that any of K such sample means deviates by more than $O(1)$ is no more than $O(K\exp(-M))$. Therefore, by setting $K = O(\frac{1}{\epsilon})$, $M = O(\log\left(\frac{1}{\delta\epsilon}\right))$ samples will suffice to ensure that the algorithm succeeds (i.e. no non-adjacent frequency is chosen) with greater than $1-\delta$ chance. This is formalized in the following Theorem.

Theorem 2.1 (Algorithm success)**.** *Given that the inputs are $M \geq \lceil \frac{81\pi^2}{2} \ln\left(\frac{8\pi}{\delta\epsilon}\right) \rceil$ and $K \geq \lceil \frac{2\pi}{\epsilon} \rceil$, Algorithm 1 gives an ϵ-accurate estimate of θ with success probability greater than $1-\delta$.*

Proof. If $\epsilon \geq \frac{\pi}{2}$, then $\hat{\theta} = \frac{\pi}{2}$ will be a $\frac{\pi}{2}$-accurate estimate for any θ, so we assume $\epsilon < \frac{\pi}{2}$. Lemma A.1 ensures that the algorithm succeeds if $K = \lceil \frac{2\pi}{\epsilon} \rceil > 4$ and all estimated coefficients are in spec (i.e. within $\frac{4}{9\pi}$ of their expected value). Lemma A.7 ensures that the likelihood of all coefficients being in spec is greater than $1 - 4K\exp\left(-\frac{2M}{81\pi^2}\right)$. To ensure that this likelihood is greater than $1-\delta$, we set $M = \lceil \frac{81\pi^2}{2} \ln\left(\frac{8\pi}{\delta\epsilon}\right) \rceil$. □

Since Algorithm 1 chooses k uniformly from 0 to $K-1$, the expected runtime of a single circuit is $O(K)$. With this observation, the above theorem implies the following corollary.

Corollary 2.2. *Algorithm 1 can be used to give an ϵ-accurate estimate with success probability greater than $1-\delta$ with expected runtime $O(\frac{1}{\epsilon}\log\left(\frac{1}{\delta\epsilon}\right))$.*

In the following section we investigate the performance of Algorithm 1 in the presence of noise. We develop several noise models and give a rigorous analysis of the algorithm's success probability.

3 Provable Robustness of Randomized Fourier Estimation

Here we analyze our algorithm when subject to error in the quantum computation. We will show that Algorithm 1 is robust to a degree of noise under quite general noise models. The models will accommodate any type of noise source, including imperfect preparation of $|\psi\rangle$, imperfect quantum gate implementation, measurement error, and error introduced by approximating circuit operations (e.g. compilation error or Hamiltonian simulation approximation error). Whatever the source of these errors, their impact on the algorithm performance is mediated by the way they alter the Hadamard test probabilities:

$$\Pr(c|k) = \frac{1 + c(\cos(k\theta) + \eta_{1,k})}{2} \tag{8}$$

and

$$\Pr(s|k) = \frac{1 + s(\sin(k\theta) + \eta_{2,k})}{2}, \tag{9}$$

where $\eta_{1,k}, \eta_{2,k}$ represent the deviations from the ideal. We refer to a model of these alterations as an *algorithm error model*. In this setting, returning to the time signal analogy as described in Section 2, we can view our signal as being corrupted by some error source

$$g(k) \to \tilde{g}(k) = e^{ik\theta} + \eta_{1,k} + i\eta_{2,k}. \tag{10}$$

Much like signal recovery in signal processing, we are interested in understanding the degree to which Algorithm 1 can reconstruct the frequency θ from this noisy signal.

First, we clarify our approach to algorithm error modeling. We will assume that these deviations, labeled η, are drawn from some real distribution $\Pr(\eta)$, which includes the case that the η are pre-determined values chosen by an adversary. Under this model the expected Fourier coefficients will deviate from their noiseless values

$$\begin{aligned}\mathbb{E}\hat{f}_j &= \mathbb{E}_\eta \mathbb{E}_k \mathbb{E}_{c|k} \mathbb{E}_{s|k} \hat{f}_j \\ &= \mathbb{E}_\eta \mathbb{E}_k (e^{ik\theta} + \eta_{1,k} + i\eta_{2,k}) e^{-2\pi ijk/K} \\ &= f_j + \eta_j, \end{aligned} \tag{11}$$

where we define $\eta_j := \mathbb{E}_\eta \mathbb{E}_k (\eta_{1,k} + i\eta_{2,k}) e^{-2\pi i jk/K}$ and $\hat{\eta}_j := \mathbb{E}_k (\eta_{1,k} + i\eta_{2,k}) e^{-2\pi i jk/K}$ for the contribution to the jth Fourier coefficient from the drawn noise η. Moreover, f_j is the expectation of noiseless Fourier coefficient as computed in equation (7). We emphasize that the impact of noise as modeled above is quite general in that it accommodates any quantum algorithm that processes the outcomes of binary measurement outcomes. This includes algorithms for ground state energy estimation [29, 31] and ground state property estimation [30].

In the following two subsections we develop two different noise models (i.e. choices of $\Pr(\eta)$) and analyze the performance of Algorithm 1 when subject to them.

3.1 Bounded adversarial noise model

We first present a worst case analysis under the condition that the noise is bounded. We give two physical settings in which the BAN model can applied to:

1. In the first setting (Appendix B) we assume Pauli Errors in the state preparation and model the errors occuring in the unitary as the noisy unitary being $\tilde{U} = \mathcal{E} \circ U$ where $\mathcal{E}$ is the Pauli Channel and U is the noiseless unitary. The analogues of equations (8) and (9) are of course long and unwieldy but the general form can be subsumed with in the BAN model by providing what the η is.

2. The second is a simpler noise model and easier for analysis and is analyzed in this section. It assumes the more dominant source of noise comes for just dephasing errors. Possible realistic settings in which it might be relevant are discussed.

This setting allows the adversary to choose any deviation of the Hadamard test probabilities as long as they are within some limit. We can provide performance guarantees as long as this limit is under the bound. Any results we establish that assume bounded adversarial noise will be quite general because we have made limited assumptions about the nature of the noise. The bounded adversarial noise model assumes the following: once the parameters of the algorithm have been set (e.g. number of samples M, number of discrete points K, and the true value of the parameter θ), an adversary is able to set $\eta_{1,k}$ and $\eta_{2,k}$ to any value in the interval $[-\bar{\eta}, \bar{\eta}]$. The performance of any algorithm will be dependent on $\bar{\eta}$. For the purpose of our analysis, we assume $\bar{\eta} < \frac{2\sqrt{2}}{9\pi}$. While these assumptions allow for a more general setting, in a more realistic noise setting; in practice, noise tends not to act adversarially, but randomly. The following theorem summarizes our analysis of the algorithm in the bounded adversarial noise setting.

Theorem 3.1 (BAN Model Analysis)**.** *Let the Hadamard tests be subject to the bounded adversarial noise model with bound $\bar{\eta}$. Given that the inputs are*

$$M \geq \left\lceil \frac{81\pi^2}{2} \left(1 - \frac{9\pi}{2\sqrt{2}}\bar{\eta}\right)^{-2} \ln\left(\frac{8\pi}{\delta\epsilon}\right) \right\rceil \tag{12}$$

and $K \geq \lceil \frac{2\pi}{\epsilon} \rceil$, Algorithm 1 gives an ϵ-accurate estimate of any $\theta \in [0, \pi]$ with success probability greater than $1 - \delta$.

Proof. Lemma A.1 ensures that the algorithm succeeds if $K = \lceil 2\pi/\epsilon \rceil > 4$ and all estimated coefficients are in spec (i.e. within $\frac{4}{9\pi}$ of their expected value). Lemma A.8 ensures that, under the bounded adversarial noise model with bound $\bar{\eta}$, Algorithm 1 will have all coefficients in spec with likelihood at least $1 - 4K\exp\left(-\frac{2M\left(1-\frac{9\pi}{2\sqrt{2}}\bar{\eta}\right)^2}{81\pi^2}\right)$. To ensure that this likelihood is greater than $1 - \delta$, we set $M \geq \lceil \frac{81\pi^2}{2}\left(1 - \frac{9\pi}{2\sqrt{2}}\bar{\eta}\right)^{-2} \ln\left(\frac{8\pi}{\delta\epsilon}\right) \rceil$ and $K \geq \lceil \frac{2\pi}{\epsilon} \rceil$. □

Here we have chosen a very general noise model. We have shown that the algorithm can still give an accurate estimate despite a degree of error. To accommodate this error, the algorithm can pay an overhead of $\left(1 - \frac{9\pi}{2\sqrt{2}}\bar{\eta}\right)^{-2}$ in the number of samples, relative to the noiseless setting. An interesting feature of this result is that there is an upper threshold $\bar{\eta} < \frac{2\sqrt{2}}{9\pi} \approx 0.1$ above which the algorithm has no success guarantee according to our analysis. This threshold corresponds to a 0.05 deviation in the Hadamard test probabilities. Note that when $\bar{\eta} = 0$, we recover the noiseless case and the bounds of Theorem 3.1 match those of Theorem 2.1.

Next we analyze the physical settings in which the BAN model can be applicable. For a given accuracy, we estimate the required circuit depth as compared to the effective dephasing scale in order for Algorithm 1 to succeed with high probability.

A Physical Setting for the BAN model

Now we discuss a physical setting in which we can apply the BAN model. For this discussion we use the

following change in the classical signal introduced in [48]

$$Pr(c=1|k) = e^{-k/T_2}\left(\frac{1+\cos(k\theta)}{2}\right) + \frac{1-e^{-k/T_2}}{2}. \tag{13}$$

We consider k as a proxy for the time duration of the quantum circuit in this sense: each application of the unitary is composed of gates which take a certain amount of time to execute. Thus the execution of one application of U takes a certain total amount T. Thus the total amount of time taken to execute the whole circuit k times is kT. We can see this time scale to 1 and analyze the problem in these "units". The largest potential deviation from the true probability will occur when k attains its largest value K. Then there are two time scales in this setting: the first is K, the time scale taken to run the deepest quantum circuit and the second is T_2, the time scale, imposed by the dephasing noise. While this physical setting might be unrealistic for the NISQ setting it's definitely plausible for an early fault tolerant device on platforms like ion-traps where physical (decay) times (T_1) are much longer than T_2 times, and can be on the order of billions of seconds. Another example where this simple dephasing model may apply is the case where the logical qubits are stored in cat states [49]. The cat states have been shown to have inherent exponential suppression of bit-flip errors. Measurements on the logical qubits to determine the scale of T_2 can be done [50]. For our discussion, the important quantity is the time scale ratio $\frac{K}{T_2}$. We give an estimate of the time scale ratio for which the algorithm succeeds with high probability. We compute $\eta_{1,k}$ as

$$\begin{aligned}\frac{\eta_{1,k}}{2} &= e^{-k/T_2}\left(\frac{1+\cos(k\theta)}{2}\right) + \frac{1-e^{-k/T_2}}{2} - \left(\frac{1+\cos(k\theta)}{2}\right) \\ &= \frac{e^{-k/T_2}-1}{2}\cos(k\theta),\end{aligned} \tag{14}$$

and similarly for $\eta_{2,k}$, but by replacing the cosine function with sine function. In order to apply the BAN model, we must determine the largest potential magnitude realized by $\eta_{1,k}$ (and $\eta_{2,k}$) such that

$$|\eta_{1,k}| < \frac{2\sqrt{2}}{9\pi}. \tag{15}$$

Then we have

$$\left|\frac{e^{-k/T_2}-1}{2}\cos(k\theta)\right| < \frac{2\sqrt{2}}{9\pi}. \tag{16}$$

In the worst case, $|\cos(k\theta)| = 1$. Therefore we have

$$0 < \frac{k}{T_2} < -\ln\left(\frac{1}{2} - \frac{2\sqrt{2}}{9\pi}\right) \approx 0.916. \tag{17}$$

This gives $0 < \frac{K}{T_2} < -\ln\left(\frac{1}{2} - \frac{2\sqrt{2}}{9\pi}\right)$ as the only physical bound on $\frac{K}{T_2}$. This implies that the circuit run-time is allowed to go to at most 0.916 times the dephasing time before we lose performance guarantee..

In the case that $K << T_2$, Equation 13 can be approximated as

$$Pr(c=1|k) = \frac{1+\cos(k\theta)}{2} + \frac{k}{2T_2}, \tag{18}$$

leading us to identify $\eta_{1,k} = \frac{k}{T_2}$. A similar approximation can be made to obtain $\eta_{2,k} = \frac{k}{T_2}$. In this limit, we imagine approximating the noisy likelihood function as getting an error term linearly related to the ratio of circuit depth and coherence time and therefore one interesting question is " what is the maximum depth of the circuit such that this simple model holds and therefore one is allowed to think of about noise entering even more simply?"

More specifically "for a given accuracy ϵ, what is the highest time scale ratio such that simplified Equation 18 provides performance guarantees"? Requiring that $\epsilon = 0.0004$ with $\bar{\eta} < \frac{2\sqrt{2}}{9\pi}$, we get that the dephasing time should be at least 5 times the algorithmic run-time. This means that for the given accuracy we can approximate (Equation 13) as (Equation 18) without losing the performance guarantee of the algorithm.

3.2 Gaussian noise model

In this subsection we look at a more realistic noise model. We still operate under the assumption that the noise is bounded; however the deviations are chosen from a specific noise model. The physical setting that we model is that in any one run of the algorithm, the particular setting (e.g. different θ or different circuit compilations or different devices or different days of using the same device) is drawn randomly. We model this randomness with Gaussian random variable perturbations of the ideal Hadamard test probabilities. The *Gaussian noise model* assumes that once the algorithm parameters are set, the Hadamard test biases, $\eta_{1,k}, \eta_{2,k}$, are then drawn independently from Gaussian distributions with standard deviation σ and mean $\mu = 0$. When we analyze the success probability of the algorithm, we will assume that in any run of the algorithm the noise parameters are each drawn once and fixed for the remainder of the algorithm. Accordingly, the expectation and variance of the estimator

must include a Gaussian-weighted average over the noise parameters.

Our analysis of the performance of the algorithm under this model will leverage the results from the bounded adversarial noise model. The approach is as follows. Observe that for any choice of $\bar{\eta} < \frac{2\sqrt{2}}{9\pi}$, the Gaussian noise model will satisfy the BAN bound with some probability. In the case where the BAN bound is satisfied, then the probability of success is lower bounded by the results of Theorem 3.1. Using the union bound, the probability of failure in the Gaussian noise setting is upper bounded by the sum of the probability of violating the BAN bound and the probability of the algorithm failing given that the BAN bound is satisfied. We take a strategy that aims to establish a simple expression for the number of measurements required to ensure a success probability greater than $1 - \delta$. This approach compromises the tightness of our results in exchange for simplicity of presentation. Our analysis of the algorithm under the Gaussian noise model is essentially a corollary of the analysis of the algorithm under the BAN model.

We first choose $\bar{\eta}$ of the BAN model such that the Gaussian noise model violates this bound with no more than $\frac{\delta}{2}$ probability. Lemma A.10 gives that we should choose $\bar{\eta}^2 = \frac{\sigma^2}{4K} \ln\left(\frac{8K}{\delta}\right)$, so that the failure probability upper bound is $\frac{\delta}{2}$. With this choice of $\bar{\eta}$, the success of the algorithm under the Gaussian noise model is simply a corollary of Theorem 3.1.

Corollary 3.2 (Gaussian Noise Model Analysis)**.** *Let the Hadamard tests be subject to the Gaussian noise model with variance* $\sigma^2 < \frac{64}{81\pi\epsilon} \ln\left(\frac{16\pi}{\delta\epsilon}\right)$. *Given that the inputs are*

$$M \geq \left\lceil \frac{81\pi^2}{2} \ln\left(\frac{16\pi}{\delta\epsilon}\right) \left(1 - \frac{9\sigma}{8}\sqrt{\frac{\epsilon\pi}{\ln\left(\frac{16\pi}{\delta\epsilon}\right)}}\right)^{-2} \right\rceil \tag{19}$$

and $K \geq \lceil \frac{2\pi}{\epsilon} \rceil$ *Algorithm 1 gives an* ϵ*-accurate estimate of any* $\theta \in [0, \pi]$ *with success probability greater than* $1 - \delta$.

This shows that, under a model of random error with bounded variance, with high probability an accurate estimate can be achieved by increasing the number of samples relative to the noiseless case.

4 Conclusion and Outlook

This work contributes to building a bridge between the noisy intermediate-scale quantum (NISQ) era and the fault-tolerant quantum computing (FTQC) era. We are motivated by the question of how can we predict the performance of quantum algorithms in the era when quantum error correction is costly and algorithms will benefit from a degree of robustness. "Robustness" is used to describe algorithms with reliable performance even under noisy conditions. In this early fault-tolerant quantum computing (EFTQC) era, there is a need for analytical tools that can ensure the robustness of a broad class of quantum algorithms.

To address this need, we develop and analyze a new algorithm for the EFTQC era that exhibits characteristics of many recently-developed quantum algorithms for EFTQC [29, 30, 54, 34, 31]. As illustrated in Figure 3 the tools that we developed are well-suited for analyzing the robustness of such algorithms. Our analysis paves the way to understanding algorithms that are robust to some degree of error. We leave determining the exact levels of errors and thus determining the cost of error-correction for future work.

The randomized Fourier estimation (RFE) algorithm estimates the phase angle θ given the operation c-U and input state $|\psi\rangle$, where $U|\psi\rangle = e^{i\theta}|\psi\rangle$. With a single ancilla qubit, using classical randomization and quantum measurements, the algorithm estimates the discrete Fourier spectrum of the time signal $g(k) = e^{i\theta k}$ to learn the location of the spectral peak that encodes θ. With this prototypical EFTQC quantum algorithm established, we analyze its robustness to error. We develop a methodology for modeling the impact of error on such algorithms. To this end, we introduce an important concept: an *algorithmic noise model*, which lets us describe the pertinent features of circuit-level noise on the functioning of the quantum algorithm. Through this concept we establish rigorous performance guarantees on the RFE algorithm subject to such noise models. This is a step towards understanding the viability of *EFTQC* and designing algorithms suitable for this era.

Like many quantum algorithms for phase estimation, our algorithm has a runtime that scales as $O(\frac{1}{\epsilon}\log\left(\frac{1}{\delta\epsilon}\right))$. We expect that the upper bounds obtained in the analysis are not tight and we expect that more sophisticated algorithms for phase estimation [55, 56, 48] would have better performance. However, the RFE algorithm facilitates an intuitive analysis of the impact of noise on performance. In the most general noise setting, as long as the noise level is bounded by $\bar{\eta}$, our algorithm has a runtime that scales as $O(\frac{1}{\epsilon}\log\left(\frac{1}{\delta\epsilon}\right))$ at the cost of extra mea-

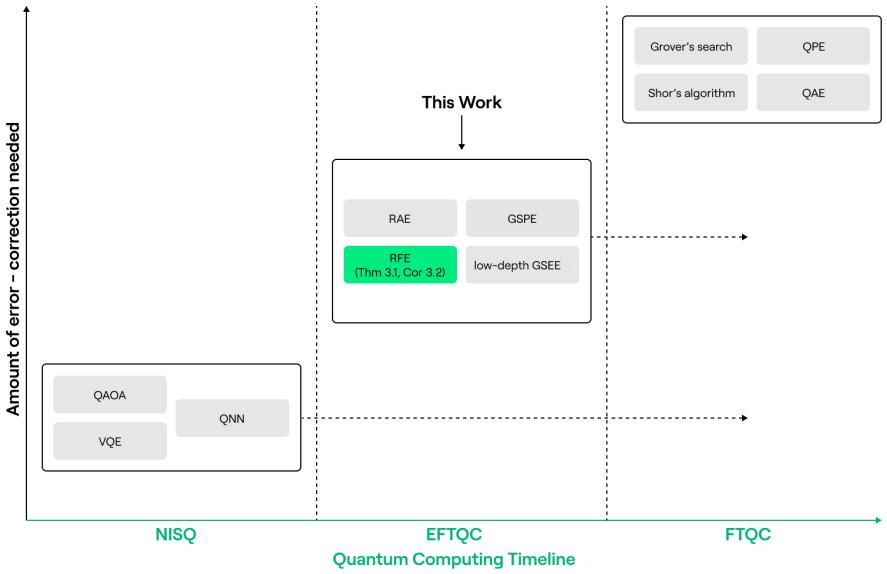

Figure 3: In the EFTQC era, error-correction will be costly and a degree of robustness will benefit quantum algorithms. Theorem 3.1 and Corollary 3.2 quantify degree of robustness to guarantee success of Algorithm 1 under noisy conditions. This figure categorizes quantum algorithms according to the degree of quantum error-correction they require and therefore the era of quantum computing for which they are suited. We have used acronyms to abbreviate following algorithms quantum alternating operator ansatz (QAOA) [13], variational quantum eigensolver (VQE) [12], quantum neural networks (QNN) [51], robust amplitude estimation (RAE) [25], low-depth ground state energy estimation (GSEE)[29, 31], ground state property estimation (GSPE)[30], randomized Fourier estimation (RFE, developed in this work), quantum phase estimation (QPE) [52], and quantum amplitude estimation (QAE) [53].

surements (dependent on $\bar{\eta}$). We also discuss a physical setting that can be related to such a noise model. In the era of costly quantum error correction it will be important to weigh the cost-benefit of varying degree of quantum error correction. Error correction reduces $\bar{\eta}$ and therefore reduces the expected runtime, but it also comes at its own increased runtime of error correction overhead. Our methodology is a tool to enable researchers to explore such trade-offs.

Towards such exploration, an important future direction is the analysis of RFE and related algorithms under more realistic noise models. In particular, it would be valuable to develop algorithmic noise models which are parameterized by quantities determined from the quantum device. This would help to make the overall methodology more predictive. [58] demonstrates a step in this direction.

Our work sets a foundation for future investigations of the robustness of algorithms for EFTQC. Other tasks, such as amplitude estimation, ground state energy estimation, and ground state property estimation, can be viewed as more-involved variants of the phase estimation task. They vary in the nature of the input state $|\psi\rangle$ and its relationship to the unitary U.

For example, ground state energy estimation can be implemented in the setting where $|\psi\rangle$ is a superposition of eigenstates of H and U approximates time evolution under this Hamiltonian [29].

As with RFE, recent early fault-tolerant algorithms developed for these tasks have a structure that is amenable to our robustness analysis [29, 30, 54, 34, 31]. In each of these algorithms, the quantum computer is used to gather time signal data and properties of the signal's spectrum are estimated by post-processing this data. Our methodology may be used to model error in time signal and analyze the robustness of such quantum algorithms. We hope that this work equips researchers with tools to navigate the era of early fault-tolerant quantum computing.

Acknowledgements We thank Yiqing Zhou and Qiyao Liang for many insightful discussions during the course of this work. We thank Archismita Dalal, Max Radin, and Daniel Stilck França for helpful feedback on earlier versions of the manuscript and Zofia Włoczewska for help in generating figures.

A Proofs of Main Results

In this section, we discuss the preliminaries necessary to prove results about performance of Algorithm 1. The following lemma determines how the distance between the Fourier coefficients estimated in the algorithm and the actual Fourier coefficients impacts the distance between the output phase and actual phase.

Lemma A.1. *If $K = \lceil \frac{2\pi}{\epsilon} \rceil \geq 4$ and during Algorithm 1 $|\hat{f}_j - f_j| \leq \frac{4}{9\pi}$ for all $j \in \{0, \ldots, K-1\}$, then Algorithm 1 outputs $\hat{\theta}$ such that $|\hat{\theta} - \theta| \leq \epsilon$.*

Proof. We will show that when the coefficient estimates are in spec, the algorithm is guaranteed to output one of the two "adjacent" discrete frequencies (i.e. those nearest to θ). Choosing $K = \lceil \frac{2\pi}{\epsilon} \rceil$ then ensures that outputting either of these adjacent frequencies gives $|\hat{\theta} - \theta| \leq \epsilon$. It remains to show that when all Fourier coefficients are in spec, the algorithm outputs an adjacent frequency. Our approach will be to lower bound the magnitude of the expected coefficient of two frequencies adjacent to θ and upper bound the magnitude of the expected coefficient of any non-adjacent frequency. These bounds will ensure a gap of $\frac{4}{9\pi}$ between the magnitudes of these expected coefficients. The "in spec" distance of $\frac{4}{9\pi}$ then ensures that if all estimated coefficients are within spec, an adjacent frequency will have the largest magnitude of all frequencies. Consequently, the algorithm will output an adjacent frequency.

The smallest expected magnitude of a "close" frequency (i.e. the adjacent frequency closer to θ) is greater than $\frac{2}{\pi}$ and Corollary A.6 ensures that the expected magnitude of any non-adjacent frequency is less than $\frac{10}{9\pi}$. The minimum distance between the magnitudes of a close frequency and any non-adjacent frequency is greater than $\frac{4}{9\pi}$. So when the estimates are in spec (i.e. deviating by no more than $\frac{4}{9\pi}$ from their expected value), none of the non-adjacent frequencies can have an estimated coefficient of largest magnitude. □

In order to prove Lemma A.3, Corollary A.4, Lemma A.5 and Corollary A.6 we will make use of Lemma A.2 which gives a general upper and lower bounds on the expected coefficients.

Lemma A.2. *In Algorithm 1, the magnitude of the expected Fourier coefficient of discrete frequency label j evaluates to*

$$|f_j| = \left| S_K \left(j - \frac{K\theta}{2\pi} \right) \right|, \tag{20}$$

where $S_J(x) := \frac{\sin \pi x}{J \sin(\frac{\pi x}{J})}$. Note: for large J, $S_J(x)$ approximates a sinc function, $S_J(x) = \frac{\sin(\pi x)}{\pi x} + O(\frac{1}{J^2})$.

Proof. Defining $\omega = \exp\left(\frac{2\pi i}{K}\right)$, we can express the expected Fourier coefficient of discrete frequency label j as

$$\begin{aligned} f_j &= \mathbb{E}_{s,c,k}(c + is)\omega^{jk} && (21)\\ &= \mathbb{E}_k e^{ik\theta}\omega^{jk}. && (22) \end{aligned}$$

Defining $\Theta = K\theta/2\pi$, we have

$$\begin{aligned} f_j &= \mathbb{E}_k \omega^{-\Theta k}\omega^{jk} && (23)\\ &= \mathbb{E}_k \omega^{(-\Theta + j)k} && (24)\\ &= \frac{1}{K}\left(\frac{1 - \omega^{(j-\Theta)K}}{1 - \omega^{(j-\Theta)}}\right) && (25)\\ &= \frac{1}{K}\left(\omega^{(j-\Theta)(K-1)/2}\frac{\sin \pi(j-\Theta)}{\sin \frac{\pi(j-\Theta)}{K}}\right) && (26) \end{aligned}$$

Defining $S_K(x) = \frac{\sin \pi x}{K \sin \frac{\pi x}{K}}$, we can write the magnitude of the Fourier coefficient as

$$|f_j| = |S_K(j - \Theta)| \,. \tag{27}$$

□

Lemma A.3. *For* $|x| \leq 1/2$,

$$\frac{2}{\pi} \leq \left| \frac{\sin \pi x}{K \sin \frac{\pi x}{K}} \right| . \tag{28}$$

Proof. Note that $S_K(x) = \frac{\sin \pi x}{K \sin\left(\frac{\pi x}{K}\right)}$. On the interval $0 \leq x \leq \frac{1}{2}$ the function $|S_K(x)|$ is a monotonically decreasing function and it achieves minimum at $x = \frac{1}{2}$. Therefore

$$|S_K(x)| \geq \left| \frac{\sin\left(\frac{\pi}{2}\right)}{K \sin\left(\frac{\pi}{2K}\right)} \right| . \tag{29}$$

Using that this function is monotonically decreasing with K and for all $K > 1$, $|S_K(x)|$ is lower bounded by the limit of this function as K goes to infinity, that is

$$|S_K(x)| \geq \frac{2}{\pi}. \tag{30}$$

□

Corollary A.4. *We have for* j *satisfying* $|j - \Theta| \leq \frac{1}{2}$,

$$|f_j| \geq \frac{2}{\pi}. \tag{31}$$

In other words, the smallest expected magnitude of a close frequency is greater than $\frac{2}{\pi}$.

Proof. Our starting point is the expression for $|f_j|$ given in Lemma A.2,

$$|f_j| = \left| S_K\left(j - \frac{K\theta}{2\pi}\right) \right| = |S_K(j - \Theta)| \,, \tag{32}$$

where $\Theta = \frac{K\theta}{2\pi}$. Next, the sum $j + \Theta$ must lie between 0 and K. Since j corresponds to a close frequency (i.e. $|j - \Theta| \leq \frac{1}{2}$) and because on the interval $0 \leq |j - \Theta| \leq \frac{1}{2}$, using Lemma A.3 we have

$$|f_j| \geq \frac{2}{\pi}. \tag{33}$$

□

Lemma A.5. *For* $1 \leq |x| \leq K/2$, $K \geq 4$,

$$\left| \frac{\sin \pi x}{K \sin \frac{\pi x}{K}} \right| \leq \frac{10}{9\pi}. \tag{34}$$

Proof. We can use the inequality $S_K(x) = \frac{\sin \pi x}{K \sin\left(\frac{\pi x}{K}\right)} \leq \frac{1}{K \sin\left(\frac{\pi x}{K}\right)}$ to write

$$|S_K(x)| \leq \frac{1}{\left|K \sin\left(\frac{\pi x}{K}\right)\right|}. \tag{35}$$

Since $\left|\sin\left(\frac{\pi x}{K}\right)\right|$ is monotonically increasing on the interval $1 \leq x \leq \frac{K}{2}$, then we have a minima at $x = 1$. Therefore

$$|S_K(x)| \leq \frac{1}{\left|K \sin(\frac{\pi}{K})\right|} \tag{36}$$

Furthermore, when $K \geq 4$ we have $|S_K(x)| \leq \frac{1}{2\sqrt{2}} \leq \frac{10}{9\pi}$. □

Corollary A.6. *We have for j satisfying $1 \leq |j - \Theta| \leq \frac{K}{2}$, $K \geq 4$,*

$$|f_j| \leq \frac{10}{9\pi}. \tag{37}$$

The expected magnitude of a non-adjacent frequency is less than $\frac{10}{9\pi}$.

Proof. As shown in Lemma A.2, the magnitude of the expected value of any coefficient is

$$|f_j| = |S_K(j - \Theta)|\,, \tag{38}$$

where $\Theta = \frac{K\theta}{2\pi}$. Since both j and Θ are bound between 0 and $\frac{K}{2}$, so is their difference $j - \Theta$. Furthermore, assuming that j corresponds to a non-adjacent frequency ensures $|j - \Theta| \geq 1$. For $K \geq 4$, we use Lemma A.5 and have

$$|f_j| \leq \frac{1}{2\sqrt{2}} \leq \frac{10}{9\pi}. \tag{39}$$

□

Lemma A.7 (Likelihood of all coefficients being in spec)**.** *In Algorithm 1, the likelihood of having $|\hat{f}_j - f_j| \leq \frac{4}{9\pi}$. for all $j \in \{0, \ldots, K-1\}$ is greater than $1 - 4K\exp\left(-\frac{2M}{81\pi^2}\right)$.*

Proof. We use a union bound and a concentration inequality to bound the probability that *all* Fourier coefficients are in spec.

Union bound: The algorithm failure probability is less than the probability that at least one Fourier coefficient is out of spec. Let q upper bound the probability that the worst-case Fourier coefficient fails to be in spec. By the union bound, the likelihood that at least one coefficient fails to be in spec is upper bounded as

$$\Pr(\text{any out of spec}) \leq \sum_l \Pr(l\text{th coefficient out of spec}) \leq Kq. \tag{40}$$

Next, we upper-bound q.

Hoeffding bound:

During Algorithm 1, each of the parity samples gives an i.i.d. estimate

$$\begin{aligned}\hat{f}_j^{(i)} &= (c_i + is_i)\exp\left(\frac{2\pi i k_i j}{K}\right) \\ &= (c_i + is_i)\left(\cos\left(\frac{2\pi k_i j}{K}\right) + i\sin\left(\frac{2\pi k_i j}{K}\right)\right),\end{aligned}$$

which are then averaged to obtain $\hat{f}_j = \frac{1}{M}\sum_{i=1}^{M}\hat{f}_j^{(i)}$. Thus, we can use a concentration inequality to upper bound the probability that this quantity deviates from its expected value. The real and imaginary components of a single sample estimate lie within $[-\sqrt{2}, \sqrt{2}]$. The Hoeffding inequality bounds the likelihood that the sum of M such estimates deviates by more than amount t

$$\Pr\left(\left|\sum_{i=1}^{M}\mathrm{Re}(\hat{f}_j^{(i)}) - \mathbb{E}\sum_{i=1}^{M}\mathrm{Re}(\hat{f}_j^{(i)})\right| \geq t\right) \leq 2\exp\left(-\frac{t^2}{4M}\right), \tag{41}$$

where we have used that the random variables are bound by an interval of length $2\sqrt{2}$. Setting $t = \frac{2\sqrt{2}M}{9\pi}$, we can re-express the above statement as

$$\Pr\left(\left|\frac{1}{M}\sum_{i=1}^{M}\mathrm{Re}(\hat{f}_j^{(i)}) - \mathbb{E}\mathrm{Re}(\hat{f}_j)\right| \geq \frac{2\sqrt{2}}{9\pi}\right) \leq 2\exp\left(-\frac{M}{81\pi^2}\right). \tag{42}$$

The same relationship holds for the imaginary components. For any Fourier coefficient, the estimate will be in spec (i.e. distance $\frac{4}{9\pi}$ in the complex plane) if both the real and imaginary components are within $\frac{2\sqrt{2}}{9\pi}$. Using the union bound, the likelihood that at least one of the components (real or imaginary) deviate by more than $\frac{2\sqrt{2}}{9\pi}$ is less than twice the likelihood of any one deviating by more than this amount. Putting this together, we can upper bound the likelihood that a given coefficient is out of spec,

$$\Pr\left(j\text{th coefficient out of spec}\right) = \tag{43}$$

$$\Pr\left(\left|\hat{f}_j - \mathbb{E}\hat{f}_j\right| \geq \frac{4}{9\pi}\right) \leq 4\exp\left(-\frac{2M}{81\pi^2}\right). \tag{44}$$

Given that this bound is independent of j, we have that q, the maximum failure probability over j, is bounded as $q \leq 4\exp\left(-\frac{2M}{81\pi^2}\right)$. Finally, the likelihood that at least one coefficient estimate is out of spec is bounded as

$$\Pr(\text{any out of spec}) \leq Kq \leq 4K\exp\left(-\frac{2M}{81\pi^2}\right). \tag{45}$$

□

Lemma A.8. *Subject to bounded adversarial noise with bound $\bar{\eta} < \frac{2\sqrt{2}}{9\pi}$, in Algorithm 1, the likelihood of having $|\hat{f}_j - f_j| \leq \frac{4}{9\pi}$ for all $j \in \{0, \ldots, K-1\}$ is greater than $1 - 4K\exp\left(-\frac{2M}{81\pi^2}\left(1 - \frac{9\pi}{2\sqrt{2}}\bar{\eta}\right)^2\right)$.*

Proof. We begin by upper bounding the likelihood that any estimated Fourier coefficient is out of spec. When including the BAN model, the violation of in-spec is due to a combination of statistical error and BAN error. The algorithm generates noisy estimates of the Fourier coefficient. We use the minimum distance between these noisy estimates and the actual Fourier coefficients to get an upper bound on the probability that any estimated Fourier coefficient is out of spec, we use the triangular inequality for this purpose. Note that, η_j is the deviation in the jth iteration of the algorithm and the maximum value of $|\eta_j|$ is $\sqrt{2}\bar{\eta}$, since both $\eta_{1,k}$ and $\eta_{2,k}$ have maximum value $\bar{\eta}$ and hence the complex vector has maximum length $\sqrt{2}\bar{\eta}$.

$$\Pr\left(\left|\hat{f}_j - f_j\right| \geq \frac{4}{9\pi}\right) \leq \Pr\left(\left|\hat{f}_j - f_j - \eta_j\right| \geq \frac{4}{9\pi} - \sqrt{2}\bar{\eta}\right) \tag{46}$$

$$\leq \Pr\left(\left|\hat{f}_j - \mathbb{E}\hat{f}_j\right| \geq \frac{4}{9\pi} - \sqrt{2}\bar{\eta}\right) \tag{47}$$

We use a union bound and a concentration inequality to bound the probability that *all* Fourier coefficients are in spec.

Union bound: The algorithm failure probability is less than the probability that at least one Fourier coefficient is out of spec. Let q upper bound the probability that the worst-case Fourier coefficient fails to be in spec. By the union bound, the likelihood that at least one coefficient fails to be in spec is upper bounded as

$$\Pr(\text{any out of spec, noisy}) \leq \tag{48}$$

$$\sum_l \Pr(l\text{th coefficient out of spec, noisy}) \leq Kq. \tag{49}$$

Next, we upper-bound q.

Hoeffding bound: During Algorithm 1, each of the parity samples gives an i.i.d. estimate

$$\hat{f}_j^{(i)} = (c_i + is_i)\exp\left(\frac{2\pi i k_i j}{K}\right)$$

$$= (c_i + is_i)\left(\cos\left(\frac{2\pi k_i j}{K}\right) + i\sin\left(\frac{2\pi k_i j}{K}\right)\right),$$

which are then averaged to obtain $\hat{f}_j = \frac{1}{M}\sum_{i=1}^{M}\hat{f}_j^{(i)}$. Thus, we can use a concentration inequality to upper bound the probability that this quantity deviates from its expected value. The real and imaginary components of a single sample estimate lie within $[-\sqrt{2}, \sqrt{2}]$. The Hoeffding inequality bounds the likelihood that the sum of M such estimates deviates by more than amount t

$$\Pr\left(\left|\sum_{i=1}^{M}\mathrm{Re}(\hat{f}_j^{(i)}) - \mathbb{E}\sum_{i=1}^{M}\mathrm{Re}(\hat{f}_j^{(i)})\right| \geq t\right) \leq 2\exp\left(-\frac{t^2}{4M}\right), \tag{50}$$

where we have used that the random variables are bound by an interval of length $2\sqrt{2}$. The choice of t depends on $\bar{\eta}$. We choose $t = \frac{M}{\sqrt{2}}\left(\frac{4}{9\pi} - \sqrt{2}\bar{\eta}\right)$ according to our definition of "in-spec" and we can re-express the above statement as

$$\Pr\left(\left|\frac{1}{M}\sum_{i=1}^{M}\mathrm{Re}(\hat{f}_j^{(i)}) - \mathbb{E}\mathrm{Re}(\hat{f}_j)\right| \geq \left(\frac{4}{9\pi} - \sqrt{2}\bar{\eta}\right)\frac{1}{\sqrt{2}}\right) \tag{51}$$

$$\leq 2\exp\left(-\frac{2M\left(1 - \frac{9\pi}{2\sqrt{2}}\bar{\eta}\right)^2}{81\pi^2}\right). \tag{52}$$

The same relationship holds for the imaginary components. For any Fourier coefficient, the estimate will be in spec (i.e. distance $\frac{4}{9\pi}$ in the complex plane) if both the real and imaginary components are within $\frac{1}{\sqrt{2}}\left(\frac{4}{9\pi} - \sqrt{2}\bar{\eta}\right)$. Using the union bound, the likelihood that at least one of the components (real or imaginary) deviate by more than $\frac{1}{\sqrt{2}}\left(\frac{4}{9\pi} - \sqrt{2}\bar{\eta}\right)$ is less than twice the likelihood of any one deviating by more than this amount.

Putting this together, we can upper bound the likelihood that a given coefficient is out of spec,

$$\Pr\left(\left|\hat{f}_j - \mathbb{E}\hat{f}_j\right| \geq \frac{4}{9\pi} - \sqrt{2}\bar{\eta}\right) \leq 4\exp\left(-\frac{2M\left(1 - \frac{9\pi}{2\sqrt{2}}\bar{\eta}\right)^2}{81\pi^2}\right). \tag{53}$$

The upper bound on the probability that the jth estimated Fourier coefficient is out-of-spec depends on $\bar{\eta}$ as

$$\Pr(j\text{th coefficient out of spec, noisy}) = \tag{54}$$

$$\Pr\left(\left|\hat{f}_j - f_j\right| \geq \frac{4}{9\pi}\right) \tag{55}$$

$$\leq 4\exp\left(-\frac{2M\left(1 - \frac{9\pi}{2\sqrt{2}}\bar{\eta}\right)^2}{81\pi^2}\right). \tag{56}$$

Given that this bound is independent of j, we have that q, the maximum failure probability over j, is bounded as $q \leq 4\exp\left(-\frac{2M\left(1-\frac{9\pi}{2\sqrt{2}}\bar{\eta}\right)^2}{81\pi^2}\right)$. Finally, the likelihood that at least one coefficient estimate is out of spec is bounded as

$$\Pr(\text{any out of spec, noisy}) \leq Kq \tag{57}$$

$$\leq 4K\exp\left(-\frac{2M\left(1 - \frac{9\pi}{2\sqrt{2}}\bar{\eta}\right)^2}{81\pi^2}\right). \tag{58}$$

□

Lemma A.9. *Let $\hat{z}$ be a random complex variable. The complex variance, $\mathrm{Var}(\hat{z}) = \mathbb{E}(|\hat{z}|^2) - |\mathbb{E}(\hat{z})|^2$, allows us to establish the following concentration inequality*

$$\Pr\left(|\hat{z} - \mathbb{E}(\hat{z})| \geq t\sqrt{\mathrm{Var}(\hat{z})}\right) \leq 4\exp\left(-\frac{t^2}{8}\right). \tag{59}$$

Proof. This inequality is established through the following chain of inequalities:

$$\Pr\left(|\hat{z} - \mathbb{E}(\hat{z})| \geq t\sqrt{\mathrm{Var}(\hat{z})}\right) = \Pr\left(|\hat{z} - \mathbb{E}(\hat{z})|^2 \geq t^2\mathrm{Var}(\hat{z})\right) \tag{60}$$

$$\leq \Pr\left(|\mathrm{Re}(\hat{z} - \mathbb{E}(\hat{z}))| \geq \frac{1}{\sqrt{2}}t\sqrt{\mathrm{Var}(\hat{z})} \text{ or } |\mathrm{Im}(\hat{z} - \mathbb{E}(\hat{z}))| \geq \frac{1}{\sqrt{2}}t\sqrt{\mathrm{Var}(\hat{z})}\right) \tag{61}$$

$$\leq \Pr\left(|\mathrm{Re}(\hat{z} - \mathbb{E}(\hat{z}))| \geq \frac{1}{\sqrt{2}}t\sqrt{\mathrm{Var}(\hat{z})}\right) + \Pr\left(|\mathrm{Im}(\hat{z} - \mathbb{E}(\hat{z}))| \geq \frac{1}{\sqrt{2}}t\sqrt{\mathrm{Var}(\hat{z})}\right) \tag{62}$$

$$\leq \Pr\left(|\mathrm{Re}(\hat{z} - \mathbb{E}(\hat{z}))| \geq \frac{1}{\sqrt{2}}t\sqrt{\mathrm{Var}(\mathrm{Re}(\hat{z}))}\right) + \Pr\left(|\mathrm{Im}(\hat{z} - \mathbb{E}(\hat{z}))| \geq \frac{1}{\sqrt{2}}t\sqrt{\mathrm{Var}(\mathrm{Im}(\hat{z}))}\right) \tag{63}$$

$$\leq 2\exp\left(-\frac{t^2}{8}\right) + 2\exp\left(-\frac{t^2}{8}\right) \tag{64}$$

$$\leq 4\exp\left(-\frac{t^2}{8}\right). \tag{65}$$

Here we establish an upper bound on the probability of a lower bound on the variance of a complex random variable. This would be useful in the next lemma, Lemma A.10. We start with equation (60) as a set up step, equation (61) is obtained using the definition of absolute value of a complex variable. Equation (62) is obtained by using the union bound and equation (63) again uses the definition of absolute value of a complex variable. □

Lemma A.10. *Under the Gaussian noise model with standard deviation σ and $\mu = 0$, the probability that there is at least one j for which $|\hat{\eta}_j| \geq \bar{\eta}$ is less than $4K\exp\left(-\frac{\sigma^2}{4K\bar{\eta}^2}\right)$.*

Proof. Using Lemma A.9, we have $\Pr(|\hat{\eta}_j - \mathbb{E}_k(\hat{\eta}_j)| \geq t\sqrt{\mathrm{Var}(\hat{\eta}_j)}) \leq 4\exp\left(-\frac{t^2}{8}\right)$. A union bound over all j gives $\Pr(\text{any } |\hat{\eta}_j| \geq t\sqrt{\mathrm{Var}(\hat{\eta}_j)}) \leq 4K\exp\left(-\frac{t^2}{8}\right)$. Setting $t^2 = \mathrm{Var}(\hat{\eta}_j)/\bar{\eta}^2$, we find that the failure probability can be upper bounded by evaluating $\mathrm{Var}(\hat{\eta}_j)$.

$$\mathrm{Var}(\hat{\eta}_j) = \mathbb{E}|\hat{\eta}_j|^2 \tag{66}$$

$$= \mathbb{E}|\hat{\eta}_{1,j} + i\hat{\eta}_{2,j}|^2 \tag{67}$$

$$= 2\mathbb{E}|\hat{\eta}_{1,j}|^2 \tag{68}$$

$$= \frac{2}{K^2}\sum_{k=0}^{K-1}\sum_{k'=0}^{K-1}\omega^{j(k-k')}\mathbb{E}\hat{\eta}_{1,k}\hat{\eta}_{1,k'} \tag{69}$$

$$= \frac{2}{K^2}\sum_{k=0}^{K-1}\sigma^2 \tag{70}$$

$$= 2\sigma^2/K, \tag{71}$$

The above equations are obtained using the definition of a complex variable, properties of complex variables, their probabilities and properties of the expectation function. This gives $\Pr(\text{any } |\hat{\eta}_j| \geq t\sqrt{\mathrm{Var}(\hat{\eta}_j)}) \leq 4K\exp\left(-\frac{\sigma^2}{4K\bar{\eta}^2}\right)$. □

Note that, in the above calculation, if σ is increasing with k, which might be a more accurate model of error accumulation in a quantum circuit, we can incorporate this into the calculation of the variance. For example, if $\sigma_k = k\sigma$ for some σ, then we can replace the final expression with $2K(K-1)(2K-1)/6\sigma^2/K^2 \leq 2K\sigma^2/3$.

B Motivation for choice of noise models and relevance to the physical setting

In this section, we look at a generalized noise model that we use to motivate the BAN model. Svore, Hastings and Freeman ([9]) presented the idea of information theoretic phase estimation. They also used ideas from classical signal processing to make their algorithm more efficient. They consider the circuit in Figure 4 for basic measurements.

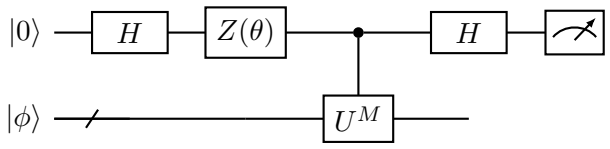

Figure 4: (Figure 1 in [9])The circuit is used as a measurement operator.

B.1 Preliminary remarks and notation

For this more general physical setting we shall assume that the initial state is prepared imperfectly and that it's a mixture of pure states i.e,

$$\rho = \rho^0 \otimes \sum_i a_i \rho_i, \tag{72}$$

where ρ^0 is the ancilla state and $\sum_i \rho_i$ is the initial mixed state. One of the terms will be the eigenstate which we denote as ρ_ϕ, where ϕ is the corresponding phase. We of course have the restriction that

$$\sum_i a_i = 1. \tag{73}$$

Let us work under the condition that there are errors in the unitary gate, U, but crucially we assume the controlled version does not introduce any new noise in the circuit. Then we model the noisy c-U gate as

$$\tilde{U} = \epsilon \circ U,$$

where $\epsilon(\rho) = \sum_\ell \beta_\ell E_\ell \rho E_\ell^\dagger$. Let $\rho^0 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $\rho^1 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$, $\rho^{01} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $\rho^{10} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $\rho^+ = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, $\rho^- = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$, $\rho^{01(H)} = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}$, $\rho^{10(H)} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}$.

B.2 Calculation for ancilla output

Consider the density matrix of the initial state of the noisy input state:

$$\rho = \rho^0 \otimes \sum_i a_i \rho_i. \tag{74}$$

After applying the first Hadamard gate on the ancilla qubit, the density matrix changes as follows:

$$\rho = \rho^+ \otimes \sum_i a_i \rho_i. \tag{75}$$

After applying the $Z(\theta)$ rotation gate on the ancilla qubit, the density matrix changes as follows:

$$\rho = \frac{1}{2}[\rho^0 + \rho^1 + \rho^{01}e^{-i\theta} + \rho^{10}e^{i\theta}] \otimes \sum_i a_i\rho_i. \tag{76}$$

After applying the controlled rotation gate, U^M with the ancilla qubit as control qubit, the density matrix changes as follows:

$$\begin{aligned}\rho &= [\rho^0 \otimes I + \rho^1 \otimes \tilde{U}_L^M] \left(\frac{1}{2}[\rho^0 + \rho^1 + \rho^{01}e^{-i\theta} + \rho^{10}e^{i\theta}] \otimes \sum_i a_i\rho_i\right) [\rho^0 \otimes I + \rho^1 \otimes \tilde{U}_R^{M^\dagger}] \\ &= \frac{1}{2}\left[\rho^0 \otimes \sum_i a_i\rho_i + \rho^{10}e^{i\theta} \otimes \tilde{U}_L^M \sum_i a_i\rho_i + \rho^{01}e^{-i\theta} \otimes \left(\sum_i a_i\rho_i\right)\tilde{U}_R^{M^\dagger} + \rho^1 \otimes \tilde{U}_L^M \left(\sum_i a_i\rho_i\right)\tilde{U}_R^{M^\dagger}\right].\end{aligned} \tag{77}$$

We use the notation $\tilde{U}_L^M$ and $\tilde{U}_R^{M^\dagger}$ to emphasize that we are considering a right and left multiplication of U^M in $\tilde{U}^M$.

After applying the second Hadamard gate on the ancilla qubit, the density matrix changes as follows:

$$\rho = \frac{1}{2}\left[\rho^+ \otimes \sum_i a_i\rho_i + \rho^{10(H)}e^{i\theta} \otimes \tilde{U}_L^M \sum_i a_i\rho_i + \rho^{01(H)}e^{-i\theta} \otimes \left(\sum_i a_i\rho_i\right)\tilde{U}_R^{M^\dagger} + \rho^- \otimes \tilde{U}_L^M \left(\sum_i a_i\rho_i\right)\tilde{U}_R^{M^\dagger}\right]. \tag{78}$$

This can be rewritten as follows:

$$\begin{aligned}\rho = &\frac{\rho^0 \otimes \left(\sum_i a_i\rho_i + e^{i\theta}\tilde{U}_L^M \sum_i a_i\rho_i + e^{-i\theta}\left(\sum_i a_i\rho_i\right)\tilde{U}_R^{M^\dagger} + \tilde{U}_L^M\left(\sum_i a_i\rho_i\right)\tilde{U}_R^{M^\dagger}\right)}{4} \\ &+ \frac{\rho^1 \otimes \left(\sum_i a_i\rho_i - e^{i\theta}\tilde{U}_L^M \sum_i a_i\rho_i - e^{-i\theta}\left(\sum_i a_i\rho_i\right)\tilde{U}_R^{M^\dagger} + \tilde{U}_L^M\left(\sum_i a_i\rho_i\right)\tilde{U}_R^{M^\dagger}\right)}{4} \\ &+ \cdots\end{aligned} \tag{79}$$

Theorem B.1. *If $\tilde{U}_L = \epsilon \circ U$ where $\epsilon(\rho) = \sum_\ell \beta_\ell E_\ell \rho E_\ell^\dagger$ with $\beta_\ell \in \mathbb{R}$, $\sum_j \beta_\ell = 1$, in particular β_{ℓ^*} is probability of no error and $E_\ell \in$ Pauli Group then when $\rho = \rho_\phi$ and $s \in \{3, M\}$*

$$\begin{aligned}\tilde{U}_L^s(\rho_\phi) = {} & e^{2\pi i s\phi_k}\beta_{k_1^*}\beta_{k_2^*}\cdots\beta_{k_s^*}\rho_\phi + e^{2\pi i s\phi_k}\beta_{k_1^*}^*\beta_{k_2^*}\cdots\beta_{k_{s-1}^*}\sum_{k_s \neq k_s^*}\beta_{k_s}E_{k_s}\rho_\phi E_{k_s}^\dagger + \\ & e^{2\pi i\phi_k(s-1)}\beta_{k_1^*}\beta_{k_2^*}\cdots\beta_{k_{s-2}^*}\sum_{k_{s-1}\neq k_{s-1}^*,k_s}\beta_{k_{s-1}}\beta_{k_s}E_{k_s}UE_{k_{s-1}}\rho_\phi E_{k_{s-1}}^\dagger E_{k_s}^\dagger + \ldots \\ & e^{2\pi i\phi_k(s-j)}\beta_{k_1^*}\beta_{k_2^*}\cdots\beta_{k_{s-j}^*}\sum_{k_{s-j}\neq k_{s-j}^*,k_{s-j+1},\ldots k_{s-j}}\beta_{k_{s-j}}\cdots\beta_{k_s}E_{k_s}\cdots UE_{k_{s-j}}\rho_\phi E_{k_{s-j}}^\dagger E_{k_s}^\dagger + \ldots \\ & e^{2\pi i\phi_k}\sum_{k_1\neq k_1^*,k_2\ldots,k_s}\beta_{k_1}\cdots\beta_{k_s}E_{k_s}\cdots UE_{k_1}\rho_\phi E_{k_1}^\dagger\cdots E_{k_s}^\dagger\end{aligned} \tag{80}$$

Proof. Base case: $s = 1$, doesn't establish the pattern for an induction argument.

$$\begin{aligned}\tilde{U}_L\rho_\phi &= (\epsilon \circ U)\rho_\phi = \epsilon \circ (U\rho_\phi) \\ &= \epsilon \circ (e^{2\pi i\phi_k}\rho_\phi) \\ &= e^{2\pi i\phi_k}(\epsilon \circ \rho_\phi) \\ &= e^{2\pi i\phi_k}\sum_{k_1}\beta_{k_1}E_{k_1}\rho_\phi E_{k_1}^\dagger \\ &= e^{2\pi i\phi_k}\left(\beta_{k_1^*}\rho_\phi + \sum_{k_1\neq k_1^*}\beta_{k_1}E_{k_1}\rho_\phi E_{k_1}^\dagger\right)\end{aligned} \tag{81}$$

Base case: $s = 2$, doesn't establish the pattern for an induction argument.

$$\begin{aligned}
\tilde{U}_L^2\rho_\phi &= (\epsilon \circ U)\left(e^{2\pi i\phi_k}\left(\beta_{k_1^*}\rho_\phi + \sum_{k_1\neq k_1^*}\beta_{k_1}E_{k_1}\rho_\phi E_{k_1}^\dagger\right)\right)\\
&= \epsilon\circ\left(e^{2\pi i\phi_k}\left(\beta_{k_1^*}U\rho_\phi + \sum_{k_1\neq k_1^*}\beta_{k_1}UE_{k_1}\rho_\phi E_{k_1}^\dagger\right)\right)\\
&= \epsilon\circ\left(e^{4\pi i\phi_k}\beta_{k_1^*}\rho_\phi + e^{2\pi i\phi_k}\sum_{k_1\neq k_1^*}\beta_{k_1}UE_{k_1}\rho_\phi E_{k_1}^\dagger\right)\\
&= \left(e^{4\pi i\phi_k}\beta_{k_1^*}\sum_{k_2}\beta_{k_2}E_{k_2}\rho_\phi E_{k_2}^\dagger + e^{2\pi i\phi_k}\sum_{k_1\neq k_1^*,k_2}\beta_{k_1}\beta_{k_2}E_{k_2}UE_{k_1}\rho_\phi E_{k_1}^\dagger E_{k_2}^\dagger\right)\\
&= \left(e^{4\pi i\phi_k}\beta_{k_1^*}\beta_{k_2^*}\rho_\phi + e^{4\pi i\phi_k}\sum_{k_2\neq k_2^*}\beta_{k_1^*}\beta_{k_2}E_{k_2}\rho_\phi E_{k_2}^\dagger + e^{2\pi i\phi_k}\sum_{k_1\neq k_1^*,k_2}\beta_{k_1}\beta_{k_2}E_{k_2}UE_{k_1}\rho_\phi E_{k_1}^\dagger E_{k_2}^\dagger\right)
\end{aligned} \tag{82}$$

Base case: $s = 3$

$$\begin{aligned}
\tilde{U}_L^3\rho_\phi &= (\epsilon\circ U)\left(e^{4\pi i\phi_k}\beta_{k_1^*}\beta_{k_2^*}\rho_\phi + e^{4\pi i\phi_k}\sum_{k_2\neq k_2^*}\beta_{k_1^*}\beta_{k_2}E_{k_2}\rho_\phi E_{k_2}^\dagger + e^{2\pi i\phi_k}\sum_{k_1\neq k_1^*,k_2}\beta_{k_1}\beta_{k_2}E_{k_2}UE_{k_1}\rho_\phi E_{k_1}^\dagger E_{k_2}^\dagger\right)\\
&= \epsilon\circ\left(e^{4\pi i\phi_k}\beta_{k_1^*}\beta_{k_2^*}U\rho_\phi + e^{4\pi i\phi_k}\sum_{k_2\neq k_2^*}\beta_{k_1^*}\beta_{k_2}UE_{k_2}\rho_\phi E_{k_2}^\dagger + e^{2\pi i\phi_k}\sum_{k_1\neq k_1^*,k_2}\beta_{k_1}\beta_{k_2}UE_{k_2}UE_{k_1}\rho_\phi E_{k_1}^\dagger E_{k_2}^\dagger\right)\\
&= \epsilon\circ\left(e^{6\pi i\phi_k}\beta_{k_1^*}\beta_{k_2^*}\rho_\phi + e^{4\pi i\phi_k}\sum_{k_2\neq k_2^*}\beta_{k_1^*}\beta_{k_2}UE_{k_2}\rho_\phi E_{k_2}^\dagger + e^{2\pi i\phi_k}\sum_{k_1\neq k_1^*,k_2}\beta_{k_1}\beta_{k_2}UE_{k_2}UE_{k_1}\rho_\phi E_{k_1}^\dagger E_{k_2}^\dagger\right)\\
&= e^{6\pi i\phi_k}\beta_{k_1^*}\beta_{k_2^*}\sum_{k_3\neq k_3^*}E_{k_3}\rho_\phi E_{k_3}^\dagger + e^{4\pi i\phi_k}\sum_{k_2\neq k_2^*,k_3}\beta_{k_1^*}\beta_{k_2}\beta_{k_3}E_{k_3}UE_{k_2}\rho_\phi E_{k_2}^\dagger E_{k_3}^\dagger\\
&+ e^{2\pi i\phi_k}\sum_{k_1\neq k_1^*,k_2,k_3}\beta_{k_1}\beta_{k_2}\beta_{k_3}E_{k_3}UE_{k_2}UE_{k_1}\rho_\phi E_{k_1}^\dagger E_{k_2}^\dagger E_{k_3}^\dagger\\
&= e^{6\pi i\phi_k}\beta_{k_1^*}\beta_{k_2^*}\beta_{k_3^*}\rho_\phi + e^{6\pi i\phi_k}\beta_{k_1^*}\beta_{k_2^*}\sum_{k_3\neq k_3^*}E_{k_3}\rho_\phi E_{k_3}^\dagger + e^{4\pi i\phi_k}\sum_{k_2\neq k_2^*,k_3}\beta_{k_1^*}\beta_{k_2}\beta_{k_3}E_{k_3}UE_{k_2}\rho_\phi E_{k_2}^\dagger E_{k_3}^\dagger\\
&+ e^{2\pi i\phi_k}\sum_{k_1\neq k_1^*,k_2,k_3}\beta_{k_1}\beta_{k_2}\beta_{k_3}E_{k_3}UE_{k_2}UE_{k_1}\rho_\phi E_{k_1}^\dagger E_{k_2}^\dagger E_{k_3}^\dagger
\end{aligned} \tag{83}$$

To prove the general statement we proceed by strong induction we assume it holds for $s \in \{3, M\}$, then by assumption we have that

$$\begin{aligned}
\tilde{U}_L^{s-1}(\rho_\phi) &= e^{2\pi i(s-1)\phi_k}\beta_{k_1^*}\beta_{k_2^*}\cdots\beta_{k_{s-1}^*}\rho_\phi + e^{2\pi i(s-1)\phi_k}\beta_{k_1^*}\beta_{k_2^*}\cdots\beta_{k_{s-2}^*}\sum_{k_{s-1}\neq k_{s-1}^*}\beta_{k_{s-1}}E_{k_{s-1}}\rho_\phi E_{k_{s-1}}^\dagger\\
&+ e^{2\pi i\phi_k(s-2)}\beta_{k_1}^*\beta_{k_2}^*\cdots\beta_{k_{s-3}^*}\sum_{k_{s-2}\neq k_{s-2}^*,k_{s-1}}\beta_{k_{s-2}}\beta_{k_{s-1}}E_{k_{s-1}}UE_{k_{s-2}}\rho_\phi E_{k_{s-2}}^\dagger E_{k_{s-1}}^\dagger + \cdots\\
&+ e^{2\pi i\phi_k(s-1-j)}\beta_{k_1^*}\beta_{k_2^*}\cdots\beta_{k_{s-j-1}^*}\sum_{k_{s-1-j}\neq k_{s-j-1}^*,k_{s-j},\ldots k_{s-1}}\beta_{k_{s-j-1}}\cdots\beta_{k_{s-1}}E_{k_{s-1}}\cdots UE_{k_{s-j-1}}\rho_\phi E_{k_{s-1-j}}^\dagger E_{k_{s-1}}^\dagger + \cdots\\
&+ e^{2\pi i\phi_k}\sum_{k_1\neq k_1^*,k_2\ldots,k_{s-1}}\beta_{k_1}\cdots\beta_{k_{s-1}}E_{k_{s-1}}\cdots UE_{k_1}\rho_\phi E_{k_1}^\dagger\cdots E_{k_{s-1}}^\dagger
\end{aligned} \tag{84}$$

from this we have that

$$
\begin{aligned}
(\tilde{U}_L^s)\rho_\phi &= \tilde{U}_L(\tilde{U}_L^{s-1}\rho_\phi) \\
&= (\epsilon \circ U)\Bigg(e^{2\pi i(s-1)\phi_k}\beta_{k_1^*}\beta_{k_2^*}\dots\beta_{k_{s-1}^*}\rho_\phi + e^{2\pi i(s-1)\phi_k}\beta_{k_1^*}\beta_{k_2^*}\dots\beta_{k_{s-2}^*}\sum_{k_{s-1}\neq k_{s-1}^*}\beta_{k_{s-1}}E_{k_{s-1}}\rho_\phi E_{k_{s-1}}^\dagger \\
&+ e^{2\pi i\phi_k(s-2)}\beta_{k_1^*}\beta_{k_2^*}\dots\beta_{k_{s-3}^*}\sum_{k_{s-2}\neq k_{s-2}^*,k_{s-1}}\beta_{k_{s-2}}\beta_{k_{s-1}}E_{k_{s-1}}UE_{k_{s-2}}\rho_\phi E_{k_{s-2}}^\dagger E_{k_{s-1}}^\dagger + \dots \\
&+ e^{2\pi i\phi_k(s-1-j)}\beta_{k_1^*}\beta_{k_2^*}\dots\beta_{k_{s-j-1}^*}\sum_{k_{s-1-j}\neq k_{s-j-1}^*,k_{s-j},\dots k_{s-1}}\beta_{k_{s-j-1}}\dots\beta_{k_{s-1}}E_{k_{s-1}}\dots UE_{k_{s-j-1}}\rho_\phi E_{k_{s-1-j}}^\dagger E_{k_{s-1}}^\dagger + \dots \\
&+ e^{2\pi i\phi_k}\sum_{k_1\neq k_1^*,k_2\dots,k_{s-1}}\beta_{k_1}\dots\beta_{k_{s-1}}E_{k_{s-1}}\dots UE_{k_1}\rho_\phi E_{k_1}^\dagger\dots E_{k_{s-1}}^\dagger\Bigg) \\
&= e^{2\pi is\phi_k}\beta_{k_1^*}\dots\beta_{k_{s-1}^*}\left(\beta_{k_s^*}\rho_\phi + \sum_{k_s\neq k_{s^*}}\beta_{k_s}E_{k_s}\rho_\phi E_{k_s}^\dagger\right) \\
&+ e^{2\pi i(s-1)\phi_k}\beta_{k_1^*}\dots\beta_{k_{s-2}^*}\sum_{k_{s-1}\neq k_{s-1}^*,k_s}\beta_{k_{s-1}}\beta_{k_s}E_{k_s}E_{k_{s-1}}\rho_\phi E_{k_{s-1}}^\dagger E_{k_s}^\dagger + \dots \\
&+ e^{2\pi(s-1-j)\phi_k}\beta_{k_1^*}\dots\beta_{k_{s-1-j}}\sum_{k_{s-1-j}\neq k_{s-1-j}^*,k_{s-j},\dots k_{s-1},k_s}\beta_{k_{s-1-j}}\dots\beta_{k_{s-1}}\beta_{k_s}E_{k_s}UE_{k_{s-1}}\dots UE_{k_{s-1-j}}\rho_\phi E_{k_{s-1-j}}^\dagger\dots E_{k_{s-1}}^\dagger E_{k_s}^\dagger \\
&+ e^{2\pi i\phi_k}\sum_{k_1\neq k_1^*,k_2\dots,k_{s-1},k_s}\beta_{k_1}\dots\beta_{k_{s-1}}E_{k_{s-1}}\beta_{k_s}E_{k_s}\dots UE_{k_1}\rho_\phi E_{k_1}^\dagger\dots E_{k_{s-1}}^\dagger E_{k_s}^\dagger
\end{aligned}
\tag{85}
$$

Hence, proved. □

From the above theorem we have the following expressions

$$
\begin{aligned}
\tilde{U}_L^M(\rho_\phi) &= e^{2\pi iM\phi_k}\beta_{k_1^*}\beta_{k_2^*}\dots\beta_{k_M^*}\rho_\phi + e^{2\pi iM\phi_k}\beta_{k_1^*}\beta_{k_2^*}\dots\beta_{k_{M-1}^*}\sum_{k_M\neq k_M^*}\beta_{k_M}E_{k_M}\rho E_{k_M}^\dagger + \\
&e^{2\pi i\phi_k(M-1)}\beta_{k_1^*}^*\beta_{k_2}^*\dots\beta_{k_{M-2}^*}^*\sum_{k_{M-1}\neq k_{M-1}^*,k_M}\beta_{k_{M-1}}\beta_{k_M}E_{k_M}UE_{k_{M-1}}\rho E_{k_{M-1}}^\dagger E_{k_M}^\dagger + \\
&e^{2\pi i\phi_k(M-n+1)}\beta_{k_1^*}^*\beta_{k_2}^*\dots\beta_{k_{M-n}^*}^*\sum_{k_{M-n+1}\neq k_{M-n+1}^*,k_{M-n+2},\dots k_M}\beta_{k_{M-1}}\beta_{k_M}E_{k_{M-n+1}}\dots UE_{k_{M-n+1}}\rho E_{k_{M-1}}^\dagger E_{k_{M-n+1}}^\dagger\dots E_{k_M}^\dagger + \\
&e^{2\pi i\phi_k}\sum_{k_1\neq k_1^*,k_2\dots,k_M}\beta_{k_1}\dots\beta_{k_M}E_{k_M}\dots UE_{k_1}\rho E_{k_1}^\dagger\dots E_{k_M}^\dagger
\end{aligned}
\tag{86}
$$

$$
\begin{aligned}
(\rho_\phi)\tilde{U}_R^{M\dagger} &= e^{-2\pi iM\phi_k}\beta_{k_1^*}\beta_{k_2^*}\dots\beta_{k_M^*}\rho_\phi + e^{-2\pi iM\phi_k}\beta_{k_1^*}\beta_{k_2^*}\dots\beta_{k_{M-1}^*}\sum_{k_M\neq k_M^*}\beta_{k_M}E_{k_M}\rho E_{k_M}^\dagger + \\
&e^{-2\pi i\phi_k(M-1)}\beta_{k_1^*}\beta_{k_2^*}\dots\beta_{k_{M-2^*}}\sum_{k_{M-1}\neq k_{M-1}^*,k_M}\beta_{k_{M-1}}\beta_{k_M}E_{k_M}E_{k_{M-1}}\rho E_{k_M}^\dagger U^\dagger E_{k_{M-1}}^\dagger + \\
&e^{-2\pi i\phi_k(M-n+1)}\beta_{k_1^*}\beta_{k_2^*}\dots\beta_{k_{M-n}^*}\sum_{k_{M-n+1}\neq k_{M-n+1}^*,k_{M-n+2},\dots k_M}\beta_{k_{M-n_1}}\dots\beta_{k_M}E_{k_M}\dots E_{k_{M-n+1}}\rho E_{k_{M-n+1}}^\dagger\dots U^\dagger E_{k_M}^\dagger + \\
&e^{-2\pi i\phi_k}\sum_{k_1\neq k_1^*,k_2\dots,k_M}\beta_{k_1}\dots\beta_{k_M}E_{k_1}\dots E_{k_M}\rho E_{k_1}^\dagger\dots U^\dagger E_{k_M}^\dagger +
\end{aligned}
\tag{87}
$$

$$\begin{aligned}\tilde{U}_L^M(\rho_\phi)\tilde{U}_R^{M\dagger} &= \beta_{k_1^*}\beta_{k_2^*}\dots\beta_{k_M^*}\rho_\phi + \beta_{k_1^*}\beta_{k_2^*}\dots\beta_{k_{M-1}^*}\sum_{k_M\neq k_M^*}\beta_{k_M}E_{k_M}\rho E_{k_M}^\dagger + \\ &\beta_{k_1^*}\beta_{k_2^*}\dots\beta_{k_{M-2}*}\sum_{k_{M-1}\neq k_{M-1}^*,k_M}\beta_{k_{M-1}}\beta_{k_M}E_{k_M}UE_{k_{M-1}}\rho E_{k_M}^\dagger U^\dagger E_{k_{M-1}}^\dagger + \\ &\beta_{k_1^*}\beta_{k_2^*}\dots\beta_{k_{M-n}^*}\sum_{k_{M-n+1}\neq k_{M-n+1}^*,k_{M-n+2},\dots k_M}\beta_{k_{M-n_1}}\dots\beta_{k_M}E_{k_M}\dots UE_{k_{M-n+1}}\rho E_{k_{M-n+1}}^\dagger\dots U^\dagger E_{k_M}^\dagger + \\ &\sum_{k_1\neq k_1^*,k_2\dots,k_M}\beta_{k_1}\dots\beta_{k_M}E_{k_1}\dots E_{k_M}\rho E_{k_1}^\dagger\dots U^\dagger E_{k_M}^\dagger\end{aligned} \tag{88}$$

A more straightforward calculation shows that when ρ is not an eigenstate,

$$\tilde{U}_L^M\rho = \sum_{k_1,k_2,\dots,k_M}\gamma_{k_1}\gamma_{k_2}\cdots\gamma_{k_M}E_{k_M}\dots UE_{k_2}UE_{k_1}U\rho E_{k_1}^\dagger E_{k_2}^\dagger\dots E_{k_M}^\dagger, \tag{89}$$

$$\rho\tilde{U}_R^{M\dagger} = \sum_{k_1,k_2,\dots,k_M}\gamma_{k_1}\gamma_{k_2}\cdots\gamma_{k_M}E_{k_M}\dots E_{k_2}E_{k_1}\rho U^\dagger E_{k_1}^\dagger U^\dagger E_{k_2}^\dagger\dots U^\dagger E_{k_M}^\dagger, \tag{90}$$

$$\tilde{U}_L^M\rho\tilde{U}_R^{M\dagger} = \sum_{k_1,k_2,\dots,k_M}\gamma_{k_1}\gamma_{k_2}\cdots\gamma_{k_M}E_{k_M}\dots UE_{k_2}UE_{k_1}U\rho U^\dagger E_{k_1}^\dagger U^\dagger E_{k_2}^\dagger\dots U^\dagger E_{k_M}^\dagger. \tag{91}$$

The probability of measuring 0 on the ancilla is then

$$\begin{aligned}p(0|x) &= \frac{1}{4}Tr\left(\sum_j a_j\rho_j + e^{i\theta}\tilde{U}^M\sum_j a_j\rho_j + e^{-i\theta}\sum_j a_j\rho_j\tilde{U}^{M\dagger} + \tilde{U}^M\sum_j a_j\rho_j\tilde{U}^{M\dagger}\right) \\ &= \frac{1}{4}\left(Tr(a_\phi\rho_\phi) + \sum_{j\neq\phi}a_jTr(\rho_j) + a_\phi Tr(e^{i\theta}\tilde{U}^M\rho_\phi) + \sum_{j\neq\phi}a_jTr(e^{i\theta}\tilde{U}^M\rho_j) + a_\phi Tr(e^{-i\theta}\rho_\phi\tilde{U}^{M\dagger})\right. \\ &\left. + \sum_{j\neq\phi}a_jTr(e^{-i\theta}\rho_j\tilde{U}^{M\dagger}) + a_\phi Tr(\tilde{U}^M\rho_\phi\tilde{U}^{M\dagger}) + \sum_{j\neq\phi}a_jTr(\tilde{U}^M\rho_j\tilde{U}^{M\dagger})\right) \\ &= \frac{1}{4}\left(a_\phi + \sum_{j\neq\phi}a_j + a_\phi + \sum_{j\neq\phi}a_j + 2Re\left[a_\phi + Tr(e^{i\theta\tilde{U}}\rho_\phi) + \sum_{j\neq\phi}a_jr(e^{i\theta\tilde{U}}\rho_i)\right]\right) \\ &= \frac{1}{2}\left(1 + a_\phi\beta_{k_1^*}\dots\beta_{k_{M-1}^*}(\beta_{k_M^*} + \sum_{k_M\neq k_M^*}\beta_{k_M})\cos(2\pi M\phi_k+\theta)\right. \\ &+ a_\phi\beta_{k_1^*}\dots\beta_{k_{M-2}^*}\sum_{k_{M-1}\neq k_{M-1}^*,k_M}\beta_{k_{M-1}}\beta_{k_M}\delta_{k_{M-1},k_M}\cos(2\pi M\phi_k+\alpha_{k_{M-1},k_M}+\theta)+\dots \\ &\left. + a_\phi\sum_{k_1\neq k_1^*,k_2,\dots,k_M}\beta_{k_1}\dots\beta_{k_M}\delta_{k_1,\dots k_M}\cos(2\pi\phi_k+\alpha_{k_1,\dots k_M}+\theta) + \sum_{j\neq\phi}a_j\sum_{k_1,k_2,\dots,k_M}\gamma_{k_1}\dots\gamma_{k_m}\sigma_{k_1,\dots k_M}\cos(2M\phi_k+\eta_{k_1,\dots k_M}+\theta)\right)\end{aligned} \tag{92}$$

In Equation (92), α_i, η_i represent the additional rotation due to noise, and $\delta_i, \gamma_i, \sigma_i$ represent probabilities ; here i denotes various subscripts used in equation (92). The number of "noise" terms in Equation (92) as well as the damping of the term corresponding to the required increase as the depth M of the circuit increases. For simplicity, the damping factor as well as the "additive" noise terms in Equation (92) can be consolidated into a single term, η representing total noise. This in essence gives us the probability of outcome of the real Hadamard test in Equation (8) which corresponds to the BAN model. Similar calculations can be used to compute the outcome probability of the imaginary Hadamard test as given in Equation (9).

References

[1] M. Reiher, N. Wiebe, K.M. Svore, D. Wecker, and M. Troyer. Elucidating reaction mechanisms on quantum computers *Proceedings of the National Academy of Science.* 114, 7555, 2017. https://doi.org/10.1073/pnas.1619152114.

[2] Jérôme F Gonthier, Maxwell D Radin, Corneliu Buda, Eric J Doskocil, Clena M Abuan, and Jhonathan Romero. Measurements as a roadblock to near-term practical quantum advantage in chemistry: resource analysis. *Phys. Rev. Research 4, 033154*, August 26, 2022. https://doi.org/10.1103/PhysRevResearch.4.033154

[3] Yuval R. Sanders, Dominic W. Berry, Pedro C.S. Costa, Louis W. Tessler, Nathan Wiebe, Craig Gidney, Hartmut Neven, and Ryan Babbush. Elucidating reaction mechanisms on quantum computers *PRX Quantum.* 1, 2020. https://doi.org/10.1103/PRXQuantum.1.020312.

[4] Shouvanik Chakrabarti, Rajiv Krishnakumar, Guglielmo Mazzola, Nikitas Stamatopoulos, Stefan Woerner, and William J. Zeng. A Threshold for Quantum Advantage in Derivative Pricing. *Quantum.* 5, 2020. https://doi.org/10.22331/q-2021-06-01-463

[5] Seunghoon Lee, Joonho Lee, Huanchen Zhai, Yu Tong, Alexander M Dalzell, Ashutosh Kumar, Phillip Helms, Johnnie Gray, Zhi-Hao Cui, Wenyuan Liu, et al. Is there evidence for exponential quantum advantage in quantum chemistry? *Nature Communications volume* 14, Article number: 1952, 2023. https://doi.org/10.1038/s41467-023-37587-6

[6] Alán Aspuru-Guzik, Anthony D Dutoi, Peter J Love, and Martin Head-Gordon. Simulated quantum computation of molecular energies. *Science*, 309(5741):1704–1707, 2005. https://doi.org/10.1126/science.1113479

[7] Christoph Durr and Peter Hoyer. A quantum algorithm for finding the minimum. *arXiv preprint quant-ph/9607014*, 1996. https://doi.org/10.48550/arXiv.quant-ph/9607014

[8] Zhiyan Ding, Yulong Dong, Yu Tong and Lin Lin. Robust ground-state energy estimation under depolarizing noise. *arXiv preprint arXiv:2301.11257*, 2023. https://doi.org/10.48550/arXiv.2307.11257

[9] Krysta Marie Svore, Matthew B. Hastings and Michael H. Freedman. Faster phase estimation. *Quantum Inf. Comput.* 14 (2013): 306-328. https://doi.org/10.48550/arXiv.1304.0741

[10] Peter W Shor. Algorithms for quantum computation: discrete logarithms and factoring. In *Proceedings 35th annual symposium on foundations of computer science*, pages 124–134. Ieee, 1994. https://doi.org/10.1109/SFCS.1994.365700

[11] N Cody Jones, Rodney Van Meter, Austin G Fowler, Peter L McMahon, Jungsang Kim, Thaddeus D Ladd, and Yoshihisa Yamamoto. Layered architecture for quantum computing. *Physical Review X*, 2(3):031007, 2012. https://doi.org/10.1103/PhysRevX.2.031007

[12] Alberto Peruzzo, Jarrod McClean, Peter Shadbolt, Man-Hong Yung, Xiao-Qi Zhou, Peter J Love, Alán Aspuru-Guzik, and Jeremy L O'brien. A variational eigenvalue solver on a photonic quantum processor. *Nature communications*, 5(1):1–7, 2014. https://doi.org/10.1038/ncomms5213

[13] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. A quantum approximate optimization algorithm. *arXiv preprint arXiv:1411.4028*, 2014. https://doi.org/10.48550/arXiv.1411.4028

[14] Eric Anschuetz, Jonathan Olson, Alán Aspuru-Guzik, and Yudong Cao. Variational quantum factoring. In *International Workshop on Quantum Technology and Optimization Problems*, pages 74–85. Springer, 2019. https://doi.org/10.48550/arXiv.1808.08927

[15] Peter D Johnson, Alexander A Kunitsa, Jérôme F Gonthier, Maxwell D Radin, Corneliu Buda, Eric J Doskocil, Clena M Abuan, and Jhonathan Romero. Reducing the cost of energy estimation in the variational quantum eigensolver algorithm with robust amplitude estimation. *arXiv preprint arXiv:2203.07275*, 2022. https://doi.org/10.48550/arXiv.2203.07275

[16] Simon J Devitt, William J Munro, and Kae Nemoto. Quantum error correction for beginners. *Reports on Progress in Physics*, 76(7):076001, 2013. https://doi.org/10.1088/0034-4885/76/7/076001

[17] Austin G Fowler, Matteo Mariantoni, John M Martinis, and Andrew N Cleland. Surface codes: Towards practical large-scale quantum computation. *Physical Review A*, 86(3):032324, 2012. https://doi.org/10.1103/PhysRevA.86.032324

[18] Hector Bombin, Isaac H Kim, Daniel Litinski, Naomi Nickerson, Mihir Pant, Fernando Pastawski, Sam Roberts, and Terry Rudolph. Interleaving: Modular architectures for fault-tolerant photonic quantum computing. *arXiv preprint arXiv:2103.08612*, 2021. https://doi.org/10.48550/arXiv.2103.08612

[19] Marco Fellous-Asiani, Jing Hao Chai, Robert S Whitney, Alexia Auffèves, and Hui Khoon Ng. Limitations in quantum computing from resource constraints. *PRX Quantum*, 2(4):040335, 2021. https://doi.org/10.1103/PRXQuantum.2.040335

[20] Rajeev Acharya, Igor Aleiner, Richard Allen, Trond I Andersen, Markus Ansmann, Frank Arute, Kunal Arya, Abraham Asfaw, Juan Atalaya, Ryan Babbush, et al. Suppressing quantum errors by scaling a surface code logical qubit. *arXiv preprint arXiv:2207.06431*, 2022. https://doi.org/10.48550/arXiv.2207.06431

[21] Isaac H Kim, Ye-Hua Liu, Sam Pallister, William Pol, Sam Roberts, and Eunseok Lee. Fault-tolerant resource estimate for quantum chemical simulations: Case study on li-ion battery electrolyte molecules. *Physical Review Research*, 4(2):023019, 2022. https://doi.org/10.1103/PhysRevResearch.4.023019

[22] Simon J Devitt, Austin G Fowler, and Lloyd CL Hollenberg. Robustness of shor's algorithm. *arXiv preprint quant-ph/0408081*, 2004. https://doi.org/10.48550/arXiv.quant-ph/0408081

[23] Oded Regev and Liron Schiff. Impossibility of a quantum speed-up with a faulty oracle. In *International Colloquium on Automata, Languages, and Programming*, pages 773–781. Springer, 2008. https://doi.org/10.48550/arXiv.1202.1027

[24] András Gilyén, Yuan Su, Guang Hao Low, and Nathan Wiebe. Quantum singular value transformation and beyond: exponential improvements for quantum matrix arithmetics [full version]. *arXiv preprint arXiv:1806.01838*, 4, 2018. https://doi.org/10.1145/3313276.3316366

[25] Guoming Wang, Dax Enshan Koh, Peter D Johnson, and Yudong Cao. Minimizing estimation runtime on noisy quantum computers. *PRX Quantum*, 2(1):010346, 2021. https://doi.org/10.1103/PRXQuantum.2.010346

[26] Tomoki Tanaka, Yohichi Suzuki, Shumpei Uno, Rudy Raymond, Tamiya Onodera, and Naoki Yamamoto. Amplitude estimation via maximum likelihood on noisy quantum computer. *Quantum Information Processing*, 20(9):1–29, 2021. https://doi.org/10.1007/s11128-021-03215-9

[27] Tudor Giurgica-Tiron, Sonika Johri, Iordanis Kerenidis, Jason Nguyen, Neal Pisenti, Anupam Prakash, Ksenia Sosnova, Ken Wright, and William Zeng. Low-depth amplitude estimation on a trapped-ion quantum computer. *Physical Review Research*, 4(3):033034, 2022. https://doi.org/10.48550/arXiv.2109.09685

[28] Amara Katabarwa, Alex Kunitsa, Borja Peropadre, and Peter Johnson. Reducing runtime and error in vqe using deeper and noisier quantum circuits. *arXiv preprint arXiv:2110.10664*, 2021. https://doi.org/10.48550/arXiv.2110.10664

[29] Lin Lin and Yu Tong. Heisenberg-limited ground-state energy estimation for early fault-tolerant quantum computers. *PRX Quantum*, 3(1):010318, 2022. https://doi.org/10.1103/PRXQuantum.3.010318

[30] Ruizhe Zhang, Guoming Wang, and Peter Johnson. Computing ground state properties with early fault-tolerant quantum computers. *Quantum*, 6:761, 2022. https://doi.org/10.22331/q-2022-07-11-761

[31] Guoming Wang, Daniel Stilck-França, Ruizhe Zhang, Shuchen Zhu, and Peter D Johnson. Quantum algorithm for ground state energy estimation using circuit depth with exponentially improved dependence on precision. *arXiv preprint arXiv:2209.06811*, 2022. https://doi.org/10.22331/q-2023-11-06-1167

[32] Guoming Wang, Sukin Sim, and Peter D. Johnson. State Preparation Boosters for Early Fault-Tolerant Quantum Computation *Quantum.* 6, 2022. https://doi.org/10.22331/q-2022-10-06-829

[33] Yu Tong. Designing algorithms for estimating ground state properties on early fault-tolerant quantum computers *Quantum.* 6, 2022. https://doi.org/10.22331/qv-2022-07-22-65

[34] Yulong Dong, Lin Lin, and Yu Tong. Ground state preparation and energy estimation on early fault-tolerant quantum computers via quantum eigenvalue transformation of unitary matrices. *arXiv preprint arXiv:2204.05955*, 2022. https://doi.org/10.1103/PRXQuantum.3.040305

[35] Andrew M Childs, Robin Kothari, and Rolando D Somma. Quantum algorithm for systems of linear equations with exponentially improved dependence on precision. *SIAM Journal on Computing*, 46(6):1920–1950, 2017. https://doi.org/10.1137/16M1087072

[36] Hsin-Yuan Huang, Richard Kueng, and John Preskill. Predicting many properties of a quantum system from very few measurements. *Nature Physics*, 16(10):1050–1057, 2020. https://doi.org/10.1038/s41567-020-0932-7

[37] Aram W Harrow, Avinatan Hassidim, and Seth Lloyd. Quantum algorithm for linear systems of equations. *Physical review letters*, 103(15):150502, 2009. https://doi.org/10.1103/PhysRevLett.103.150502

[38] Gilles Brassard, Peter Høyer, and Alain Tapp. Quantum counting. In *Automata, Languages and Programming: 25th International Colloquium, ICALP'98 Aalborg, Denmark, July 13–17, 1998 Proceedings 25*, pages 820–831, 1998. Springer. https://doi.org/10.1007/BFb0055105

[39] Vittorio Giovannetti, Seth Lloyd, and Lorenzo Maccone. Quantum-enhanced measurements: beating the standard quantum limit. *Science*, 306(5700):1330–1336, 2004. American Association for the Advancement of Science. https://doi.org/10.1126/science.1104149

[40] Christopher E Granade, Christopher Ferrie, Nathan Wiebe, and David G Cory. Robust online Hamiltonian learning. *New Journal of Physics*, 14(10):103013, 2012. IOP Publishing. https://doi.org/10.1088/1367-2630/14/10/103013

[41] Christopher Ferrie, Christopher E Granade, and DG Cory. Adaptive Hamiltonian estimation using Bayesian experimental design. In *AIP Conference Proceedings 31st*, volume 1443, number 1, pages 165–173, 2012. American Institute of Physics. https://doi.org/10.1063/1.3703632

[42] Thomas E O'Brien, Brian Tarasinski, and Barbara M Terhal. Quantum phase estimation of multiple eigenvalues for small-scale (noisy) experiments. *New Journal of Physics*, 21(2):023022, 2019. IOP Publishing. https://doi.org/10.1088/1367-2630/aafb8e

[43] Stefano Paesani, Andreas A Gentile, Raffaele Santagati, Jianwei Wang, Nathan Wiebe, David P Tew, Jeremy L O'Brien, and Mark G Thompson. Experimental Bayesian quantum phase estimation on a silicon photonic chip. *Physical review letters*, 118(10):100503, 2017. APS. https://doi.org/10.1103/PhysRevLett.118.100503

[44] Alessandro Lumino, Emanuele Polino, Adil S Rab, Giorgio Milani, Nicolò Spagnolo, Nathan Wiebe, and Fabio Sciarrino. Experimental phase estimation enhanced by machine learning. *Physical Review Applied*, 10(4):044033, 2018. APS. https://doi.org/10.1103/PhysRevApplied.10.044033

[45] Hongkang Ni, Haoya Li, and Lexing Ying. On low-depth algorithms for quantum phase estimation. *arXiv preprint arXiv:2302.02454*, 2023. https://doi.org/10.22331/q-2023-11-06-1165

[46] Haoya Li, Hongkang Ni, and Lexing Ying. Adaptive low-depth quantum algorithms for robust multiple-phase estimation. *Physical Review A*, 108(6):062408, 2023. APS. https://doi.org/10.1103/PhysRevA.108.062408

[47] Qiyao Liang, Yiqing Zhou, Archismita Dalal, and Peter D Johnson. Modeling the Performance of Early Fault-Tolerant Quantum Algorithms. *arXiv preprint arXiv:2306.17235*, 2023. https://doi.org/10.1103/PhysRevResearch.6.023118

[48] Nathan Wiebe and Chris Granade. Efficient bayesian phase estimation. *Physical review letters*, 117(1):010503, 2016. https://doi.org/10.1103/PhysRevLett.117.010503

[49] Christopher Chamberland, Kyungjoo Noh, Patricio Arrangoiz-Arriola, Earl T. Campbell, Connor T. Hann, Joseph Iverson, Harald Putterman, Thomas C. Bohdanowicz, Steven T. Flammia, Andrew Keller, Gil Refael, John Preskill, Liang Jiang, Amir H. Safavi-Naeini, Oskar Painter, and Fernando G.S.L. Brandão. Building a fault-tolerant quantum computer using concatenated cat codes. *PRX Quantum*, 3:010329, 2 2022. https://doi.org/10.1103/PRXQuantum.3.010329

[50] L. Hu, Y. Ma, W. Cai, X. Mu, Y. Xu, W. Wang, Y. Wu, H. Wang, Y. P. Song, C. L. Zou, S. M. Girvin, L. M. Duan, and L. Sun. Quantum error correction and universal gate set operation on a binomial bosonic logical qubit. *Nature Physics 2019 15:5*, 15:503–508, 2 2019. https://doi.org/10.1038/s41567-018-0414-3

[51] Edward Farhi and Hartmut Neven. Classification with quantum neural networks on near term processors. *arXiv preprint arXiv:1802.06002*, 2018. https://doi.org/10.48550/arXiv.1802.06002

[52] A Yu Kitaev. Quantum measurements and the abelian stabilizer problem. *arXiv preprint quant-ph/9511026*, 1995. https://doi.org/10.48550/arXiv.quant-ph/9511026

[53] G Brassard, P Hoyer, M Mosca, and A Tapp. Quantum amplitude amplification and estimation. *Quantum Computation and Quantum Information: A Millennium Volume. AMS Contemporary Mathematics Series*, 2000. https://doi.org/10.1090/conm/305/05215

[54] Kianna Wan, Mario Berta, and Earl T Campbell. Randomized quantum algorithm for statistical phase estimation. *Physical Review Letters*, 129(3):030503, 2022. https://doi.org/10.1103/PhysRevLett.129.030503

[55] Zhengfeng Ji, Guoming Wang, Runyao Duan, Yuan Feng, and Mingsheng Ying. Parameter estimation of quantum channels. *IEEE Transactions on Information Theory*, 54(11):5172–5185, 2008. https://doi.org/10.1109/TIT.2008.929940

[56] Miroslav Dobšíček, Göran Johansson, Vitaly Shumeiko, and Göran Wendin. Arbitrary accuracy iterative quantum phase estimation algorithm using a single ancillary qubit: A two-qubit benchmark. *Physical Review A*, 76(3):030306, 2007. https://doi.org/10.1103/PhysRevA.76.030306

[57] Shelby Kimmel, Guang Hao Low, and Theodore J. Yoder. Robust calibration of a universal single-qubit gate set via robust phase estimation. *Phys. Rev. A*, 92:062315, Dec 2015. https://doi.org/10.1103/PhysRevA.92.062315

[58] Q.Liang, Y.Zhou, A.Dalal and P.D.Johnson, Modeling the Performance of Early Fault-Tolerant Quantum Algorithms. *arXiv:2306.17235 [quant-ph]*, 2023. https://doi.org/10.1103/PhysRevResearch.6.023118