

# QUANTUM MONTE-CARLO INTEGRATION: THE FULL ADVANTAGE IN MINIMAL CIRCUIT DEPTH

STEVEN HERBERT\*

*Cambridge Quantum Computing Ltd, 9a Bridge Street, Cambridge, CB2 1UB, UK  
Department of Computer Science and Technology, University of Cambridge, UK*

**ABSTRACT.** This paper proposes a method of quantum Monte-Carlo integration that retains the full quadratic quantum advantage, without requiring any arithmetic or the quantum Fourier transform to be performed on the quantum computer. No previous proposal for quantum Monte-Carlo integration has achieved all of these at once. The heart of the proposed method is a Fourier series decomposition of the sum that approximates the expectation in Monte-Carlo integration, with each component then estimated individually using quantum amplitude estimation. The main result is presented as theoretical statement of asymptotic advantage, and numerical results are also included to illustrate the practical benefits of the proposed method. The method presented in this paper is the subject of a patent application [Quantum Computing System and Method: Patent application GB2102902.0 and SE2130060-3].

## I. INTRODUCTION

Monte-Carlo integration (MCI) solves the problem of estimating the mean of some probability distribution by averaging samples therefrom. This could either be an average of the samples themselves, or of the samples with some function applied first. It is known that a quadratic quantum speed-up in MCI is available [1] by calling quantum amplitude estimation (QAE) [2] as a sub-routine. Such is the ubiquity of Monte-Carlo methods throughout the physical, biological, data and information sciences, that this has, in turn, spawned a great deal of interest in *quantum Monte-Carlo integration* (QMCI) – most notably in applications related to finance such as *option pricing* [3–12].

However, the theoretical possibility of using quantum computing to perform MCI more efficiently has not masked a number of clear obstacles that are in the way of actually doing so in practice. Most notably the quantum advantage in MCI is an advantage in *query* complexity: the number of uses of a quantum state that encodes the probability distribution in question will be fewer than the number of classical samples from that distribution – but this immediately raises the question of how to prepare such a quantum state. Recently it was shown that there is no advantage when the oft-cited Grover-Rudolph method [13] is used to prepare states encoding probability distribution when a full audit of the *computational* complexity is undertaken [14].

Even should this problem be resolved, there are a number of reasons to be skeptical about the possibility of QMCI yielding a useful quantum advantage in the NISQ era. Most notably, QAE, as originally described, requires use of the quantum Fourier transform (QFT), which is unlikely to be practical on near-term quantum hardware. On this front, though, there are reasons for optimism – it has been shown that the QFT can be omitted in favour of classical post-processing to extract an estimate of the statistical quantity of interest from measurements of the quantum state [15–18]. Moreover, recent efforts have been made to further cut the required circuit depth by interpolating between classical and quantum MCI [19, 20].

To a certain extent, however, these advances have served only to reveal further issues with NISQ compatible QMCI: specifically that the “natural” quantity to estimate on a quantum computer is:

$$\mathbb{E}(\sin^2(X)) = \sum_{x:p(x) \neq 0} p(x) \sin^2(x) \quad (1)$$

where  $X$  are samples from  $p(x)$ . Any other quantity, including the mean of  $p(x)$  itself, will incur a significant amount of arithmetic to be performed on the quantum computer. Indeed, this cost is not unlike that of the QFT – not a problem asymptotically, but almost certainly prohibitive in the NISQ era. The foremost proposal to work around the need for quantum arithmetic is showcased in *Option Pricing using Quantum Computers* [21] in which a technique proposed by Woerner and Egger [9] is used. Specifically, the technique shifts and rescales the support of the probability distribution such that for all  $x : p(x) \neq 0$ ,

---

\*Steven.Herbert@cambridgequantum.com

Method	Computes	MSE	Arithmetic
Classical Monte-Carlo Integration	$\mathbb{E}(f(X)g(Y))$	$\Theta(q^{-1})$	Classical
Quantum Monte-Carlo Integration	$\mathbb{E}(f(X))$	$\Theta(q^{-2})$	Quantum & classical
Rescaled QMCI [9, 21]	$\mathbb{E}(X)$	$\Theta(q^{-4/3})$	Classical only
Fourier QMCI (method proposed here)	$\mathbb{E}(f(X)g(Y))$	$\Theta(q^{-2})$	Classical only

**Table 1.** Comparison of different MCI estimation methods:  $X$  and  $Y$  are the values of different variables sampled from some multivariate probability distribution;  $q$  is the number of samples (classical) or uses of a circuit that prepares a state encoding the probability distribution (quantum); and  $f$  and  $g$  are functions. We can see that our proposed method is the “best of all worlds”. It is capable of computing the product of functions of two random variables, like classical MCI; and it achieves the full quantum speed-up, without requiring arithmetic operations to be performed quantumly – something that none of the existing QMCI algorithms achieve.

$\sin^2(x + \pi/4) \approx x + 1/2$ , thus enabling an estimate of the mean to be extracted. However, this method has an adverse effect on the rate of convergence, which can be seen as a false economy: the requirement of quantum arithmetic has been removed to reduce the circuit depth, but this comes at a cost of reduced convergence rate, and so correspondingly deeper circuits will be required to achieve the specified estimation accuracy.

In this article we propose an alternative solution that still eliminates the need for quantum arithmetic, but also upholds the full quadratic quantum advantage. Instead of *squeezing* the support of the probability distribution to correspond to an approximately linear region of a trigonometric function, instead we propose that the function applied to the samples is *extended* as a periodic, piecewise function. This periodic function can be decomposed as a Fourier series, whose various components have means that can be estimated individually and then recombined. With judicious choices about how to do this Fourier series decomposition, we show that the quadratic quantum advantage is indeed retained.

Furthermore, by performing the QMCI in this way, there is no “favoured” quantity to compute (in the sense of (1)), and computation of the mean of any function applied to the samples is equally easy in principle. This immediately suggests a way in which an actual computational advantage, as opposed a theoretical advantage in query complexity, may be realised using QMCI. Quantum advantage can be realised using the proposed method if: (a) there exists some probability distribution that can be exactly prepared as a quantum state in shallow-depth; (b) a function is then applied to the random samples such that the computed mean is a quantity that cannot be analytically evaluated. This is a clear and useful distinction over, for example, the method proposed by Stamatopoulos *et al* where the distribution of interest

must be directly encoded into the quantum state [21].

Finally, it is worth highlighting that our proposed method of QMCI can also be used to compute the expectation of the product of functions of two correlated random variables, again with the full quadratic quantum advantage and without requiring quantum arithmetic. This may find application in, for example, the computation of correlations and covariances of random variables from multivariate probability distributions.

Table 1 summarises the benefits of the method of QMCI proposed here compared to existing alternatives. Convergence is measured in mean-squared error (MSE) as a function of the number of samples from the probability distribution,  $q$ . The table includes classical MCI for comparison, but it should be noted that there are classical alternatives, for example quadrature and quasi-Monte-Carlo methods. Ostensibly, these appear to have favourable performance (achieving MSE convergence proportional to  $q^{-2}$ ), however they suffer the curse of dimensionality: the complexity grows exponentially with the number of dimensions. Hence it is noteworthy that our proposed method applies to random variables marginalised out from multivariate probability distributions as, for a sufficiently large number of dimensions, classical MCI will be the most appropriate classical algorithm to benchmark against.

The remainder of this paper is organised as follows: in Section II the problem at hand is specified, the relevant definitions given, and precise details of the motivation presented; in Section III the main theoretical result is given; in Section IV the result pertaining to the expectation of products of random variables is given as a Corollary, whose proof is then given in Appendix A; in Section V a simulated example is given as a numerical proof of principle; and in Section VI the results of the paper are discussed. The method

presented in this paper is the subject of a patent application [22].

## II. PRELIMINARIES AND MOTIVATION

Suppose we have the  $d$ -dimensional discrete probability distribution  $p$  (that is a probability distribution over some finite subset of  $\mathbb{R}^d$ ). The points of non-zero probability mass are spaced at constant intervals and for simplicity, we let the number of points along each axis be a power of 2, say  $2^{N_j}$  for the  $j^{\text{th}}$  dimension (we can “zero-pad” so there is no loss of generality in this assumption). The probability distribution is encoded in a quantum state,  $|p\rangle$ , of the form:

$$|p\rangle = \sum_{\substack{x^{(1)} \in \{0,1\}^{N_1} \\ \vdots \\ x^{(d)} \in \{0,1\}^{N_d}}} \sqrt{p(x^{(1)}, \dots, x^{(d)})} |x^{(1)} \dots x^{(d)}\rangle \quad (2)$$

and we assume that the quantum algorithm has access to a circuit,  $P$ , that prepares  $|p\rangle$ , ie,  $|p\rangle = P|0\rangle$  (where  $|0\rangle$  denotes the tensor product of an appropriate number of qubits in the  $|0\rangle$  state).

QMCI estimates the expectation of some function,  $f(x)$ , of samples from the marginal distribution of a single dimension of the multivariate probability distribution. This dimension must therefore be specified – let this be the  $i^{\text{th}}$  dimension – so that QMCI returns an estimate of the mean of the probability distribution implicitly defined by  $f(x)$  applied to samples,  $X$ , from  $p(x^{(i)})$ , that is an estimate of:

$$\mathbb{E}(f(X)) = \sum_{x^{(i)}} f(x^{(i)}) p(x^{(i)}) \quad (3)$$

To achieve this, the support of  $p(x^{(i)})$  is also taken as an input to the QMCI algorithm: this is given as the value that the first point of the distribution corresponds to,  $x_l$  and the spacing interval,  $\Delta$ . It is also useful to express the maximum value of the support of the distribution,  $x_u = x_l + (2^{N_i} - 1)\Delta$ .

In order to obtain a quantum advantage using the QMCI method in this paper, we further require that  $f(x)$  is continuous, has continuous first derivative and bounded piecewise-continuous second and third derivatives over the support of  $p(x^{(i)})$ . QMCI can be achieved by constructing a matrix  $A = R(P \otimes I)$  where

$$\begin{aligned} & R|x^{(1)}x^{(2)} \dots x^{(d)}\rangle|0\rangle \\ &= |x^{(1)}x^{(2)} \dots x^{(d)}\rangle \left( \sqrt{1 - f(x^{(i)})} |0\rangle + \sqrt{f(x^{(i)})} |1\rangle \right) \end{aligned}$$

Thus  $A|0\rangle = \cos \theta_a |\Psi_0\rangle|0\rangle + \sin \theta_a |\Psi_1\rangle|1\rangle$  where  $|\Psi_0\rangle$  and  $|\Psi_1\rangle$  are just some quantum states and  $\sin^2 \theta_a =$

$\sum_{x^{(i)}} f(x^{(i)}) p(x^{(i)})$ , ie, the expectation we are seeking (see Ref. [15] for a more detailed exposition of this), which can be estimated by calling some QAE algorithm as a subroutine.

**Definition 1.** A QAE algorithm takes as an input a circuit  $A$  where  $A|0\rangle = \cos \theta_a |\Psi_0\rangle|0\rangle + \sin \theta_a |\Psi_1\rangle|1\rangle$ , and returns an estimate of  $s = \sin^2 \theta_a$ . The allowed number of uses, denoted “ $q$ ”, of the circuit  $A$  is given as an input to the QAE algorithm.

From this definition, we can write any QAE algorithm as a function in pseudocode form:

$$\hat{s} = \text{QAE}(A, q) \quad (4)$$

where  $\hat{s}$  expresses the fact that this is an estimate of  $s$ . It is also useful to formally define the MSE,  $\hat{\epsilon}^2$ :

$$\hat{\epsilon}^2 = \mathbb{E}((\hat{s} - s)^2) \quad (5)$$

The convergence of any QAE algorithm can be expressed as:

$$\hat{\epsilon}^2(q) \in \Theta(q^{-\lambda}) \quad (6)$$

for some  $\lambda$ , and the method proposed in this paper assumes that the convergence rate of the QAE algorithm as a function of the number of uses of  $A$  is known. That is, there exists a constant  $k_1$  and we know the constant  $\lambda$  such that:

$$\hat{\epsilon}^2(q) \leq k_1 q^{-\lambda} \quad (7)$$

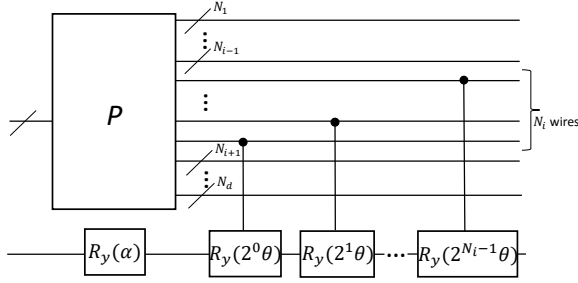
A classical MCI algorithm can be used to perform QAE, in the sense of the above definition, and in this case  $\lambda = 1$ , whereas typical QAE algorithms have  $\lambda = 2$ . However, the method we propose here works for any QAE algorithm, including a proposed method that interpolates between the classical and quantum (so  $1 \leq \lambda \leq 2$ ) [19, 20, 23].

QAE algorithms use the circuit  $A$  (and hence  $P$ ) to build a “Grover iterate” circuit,

$$Q = -AS_0A^\dagger S_\chi \quad (8)$$

where  $S_0$  and  $S_\chi$  do not depend on  $P$  or  $A$ . Further details do not directly concern us here – apart from that circuit depth of QMCI is usually counted in number of sequential uses of  $Q$  [2, 15].

**Problems with existing methods.** We can see that performing QAE when  $A = R(P \otimes I)$  returns an estimate of  $s = \sum f(x)p(x)$ , and thus *does* yield a quantum speed-up. However, this comes at the cost of using the circuit  $R$ , which encodes the function applied to the random samples, and in general this will be prohibitively complex. Even the (classically) trivial case where  $f(x) = x$  (ie, we are just finding the mean of  $p(x)$ ) would entail significant amounts of arithmetic to be performed quantumly. An important exception is the special case where  $f(x) =$



**Fig. 1.** Circuit diagram of  $A(P, i, \beta, n, \omega)$ , where  $\alpha = n\omega x_l - \beta$  and  $\theta = n\omega\Delta$ , in which  $x_l$  and  $\Delta$  are the first point of probability mass and spacing (as defined in Section II) for the  $i^{th}$  dimension of  $P$ .

$\sin^2(mx + c)$  for some constants  $m$  and  $c$ , which can be achieved by a bank of  $R_y$  rotation gates, as shown in Fig. 1.

### III. QMCI OF A SINGLE RANDOM VARIABLE

The main result of the paper requires QAE to be performed on the quantum circuit  $A(P, i, \beta, n, \omega)$  given in Fig. 1.  $P$  is circuit that prepares from  $|0\rangle$  a quantum state encoding a multivariate probability distribution, as in (2), and  $i$  denotes the dimension of interest (we assume that the support of the  $i^{th}$  dimension is known, in terms of the value of the first point of non-zero probability mass,  $x_l$  and spacing between points of probability mass,  $\Delta$ );  $n$  is an integer and  $\omega$  and  $\beta$  real numbers.

**Proposition 2.** *Let  $\text{QAE}(A, q)$  be a QAE algorithm in the sense of Definition 1, which MSE convergence parameterised by  $\lambda$  as in (6).*

i.  $1 - 2\text{QAE}(A(P, i, 0, n, \omega), q)$  is an estimate of:

$$\sum_{x^{(i)}} p(x^{(i)}) \cos(n\omega x^{(i)}) \quad (9)$$

with  $\text{MSE} \in \Theta(q^{-\lambda})$ .

ii.  $1 - 2\text{QAE}(A(P, i, \pi/2, n, \omega), q)$  is an estimate of:

$$\sum_{x^{(i)}} p(x^{(i)}) \sin(n\omega x^{(i)}) \quad (10)$$

with  $\text{MSE} \in \Theta(q^{-\lambda})$ .

*Proof.* The function of circuit  $A$  is to apply an  $R_y$  rotation to the last qubit that is proportional to some constant plus the value of the  $i^{th}$  register of  $P$ . Recall that the circuit  $P$ , when applied to  $|0\rangle$ , encodes a multivariate probability distribution in the computational basis states of a quantum state, as defined

in (2), thus from Fig. 1 we can see that, when  $\beta = 0$ ,  $A|0\rangle$  prepares the state:

$$A|0\rangle = \sum_{x^{(1)} \dots x^{(d)}} \sqrt{p(x^{(1)}, \dots, x^{(d)})} |x^{(1)} \dots x^{(d)}\rangle \left( \cos\left(\frac{n\omega x^{(i)}}{2}\right) |0\rangle + \sin\left(\frac{n\omega x^{(i)}}{2}\right) |1\rangle \right) \quad (11)$$

and thus the amplitude of the  $|1\rangle$  component of the final qubit is:

$$\begin{aligned} & \sum_{x^{(1)} \dots x^{(d)}} p(x^{(1)}, \dots, x^{(d)}) \sin^2(n\omega x^{(i)}/2) \\ &= \sum_{x^{(i)}} p(x^{(i)}) \sin^2(n\omega x^{(i)}/2) \\ &= \sum_{x^{(i)}} \frac{1}{2} p(x^{(i)}) (1 - \cos(n\omega x^{(i)})) \\ &= \frac{1}{2} \left( 1 - \sum_{x^{(i)}} p(x^{(i)}) \cos(n\omega x^{(i)}) \right) \end{aligned} \quad (12)$$

where  $p(x^{(i)})$  is the marginal distribution – note that the marginalisation has been achieved automatically.

Thus  $\text{QAE}(A(P, i, 0, n, \omega), q)$  estimates the value of  $0.5(1 - \sum_{x^{(i)}} p(x^{(i)}) \cos(n\omega x^{(i)}))$  with  $\text{MSE} \in \Theta(q^{-\lambda})$ , by (6). So it follows that  $1 - 2\text{QAE}(A(P, i, 0, n, \omega), q)$  therefore estimates  $\sum_{x^{(i)}} p(x^{(i)}) \cos(n\omega x^{(i)})$ , and as this differs from the original estimate only by being shifted and scaled by a constant factor, the estimation accuracy remains such that  $\text{MSE} \in \Theta(q^{-\lambda})$ .

Turning now to the second part of the Proposition, by the same reasoning as before, when  $\beta = 0$ ,  $A|0\rangle$  prepares the state:

$$A|0\rangle = \sum_{x^{(1)} \dots x^{(d)}} \sqrt{p(x^{(1)}, \dots, x^{(d)})} |x^{(1)} \dots x^{(d)}\rangle \left( \cos\left(\frac{n\omega x^{(i)}}{2} - \frac{\pi}{4}\right) |0\rangle + \sin\left(\frac{n\omega x^{(i)}}{2} - \frac{\pi}{4}\right) |1\rangle \right) \quad (13)$$

and thus the amplitude of the  $|1\rangle$  component of the final qubit is:

$$\begin{aligned} & \sum_{x^{(i)}} p(x^{(i)}) \sin^2\left(\frac{n\omega x^{(i)}}{2} - \frac{\pi}{4}\right) \\ &= \sum_{x^{(i)}} \frac{1}{2} p(x^{(i)}) \left( 1 - \cos\left(\frac{n\omega x^{(i)}}{2} - \frac{\pi}{2}\right) \right) \\ &= \frac{1}{2} \left( 1 - \sum_{x^{(i)}} p(x^{(i)}) \sin(n\omega x^{(i)}) \right) \end{aligned} \quad (14)$$

Thus  $1 - 2\text{QAE}(A(P, i, \pi/2, n, \omega), q)$  gives an estimate of  $\sum_{x^{(i)}} p(x^{(i)}) \sin(n\omega x^{(i)})$  with  $\text{MSE} \in \Theta(q^{-\lambda})$ .  $\square$

Proposition 2 now allows us to proceed to the main result of the paper.

**Theorem 3.** *The quantity,  $\mu = \mathbb{E}(f(X))$ , where  $X \sim p(x^{(i)})$  and  $f$  is a function that is continuous in value and first derivative and whose second and third derivatives are piecewise-continuous and bounded, can be estimated with  $\text{MSE} \in \Theta(q^{-\lambda})$ , where  $q$  is the number of uses of a circuit preparing the quantum state  $|p\rangle$  (ie, as defined in (2)) and  $\lambda$  is the convergence rate of some QAE subroutine (ie, as defined in (7)) which operates on circuits of the form defined in Fig. 1.*

**Remark.** *This result substantiates the “best of all worlds” claim made in Section I: the convergence of the QMCI estimate is equal to that of the underlying QAE algorithm; whilst the circuit in Fig. 1 requires only a single bank of rotation gates to construct  $A$  from  $P$ , which is minimal. That is, if the probability distribution of interest is over  $2^{N_i}$  points then at least  $N_i$  two-qubit gates are required in order that the amplitude of a further qubit is equal to the mean of some function the probability distribution (ignoring unnatural cases where some of the data-points of the probability distribution are omitted altogether).*

*Proof.* We prove the theorem by giving an explicit construction. The proof strategy is to first build a periodic function,  $f(x)$ , from  $f(x)$ ; then to show that the expectation can be approximated by performing QAE on the components of a Fourier series decomposition of  $f(x)$  using Proposition 2; and finally to show that the total number of uses of  $P$  can be distributed amongst the Fourier series components such that the claimed overall convergence rate holds.

First, we build a piecewise periodic function  $f(x)$  such that  $f(x) = f(x)$  over the support of  $p(x)$ ;  $f(x)$  is also constructed such that it is continuous in value and first derivative, and has second and third derivatives that are piecewise-continuous and bounded. For example, some  $f(x)$  of the following form, which repeats with period  $x_{\bar{u}} - x_l$ , is suitable in general:

$$f(x) = \begin{cases} f(x) & \text{if } x_l \leq x < x_u \\ \tilde{f}(x) & \text{if } x_u \leq x < x_{\bar{u}} \end{cases} \quad (15)$$

where  $x_{\bar{u}} > x_u$  and  $\tilde{f}(x)$  is itself sufficiently smooth (that is, continuous in value and first derivative, and has second and third derivatives that are piecewise-continuous and bounded) and chosen such  $f(x_l) = \tilde{f}(x_{\bar{u}})$ ,  $f'(x_l) = \tilde{f}'(x_{\bar{u}})$ ,  $f(x_u) = \tilde{f}(x_u)$  and  $f'(x_u) = \tilde{f}'(x_u)$  (where  $f'$ ,  $\tilde{f}'$  denotes differentiation w.r.t.  $x$ ). It is easy to see that a cubic will always suffice to fit such a  $\tilde{f}(x)$ , and on occasion a simpler function may

do. As  $f(x)$  is periodic, it has a Fourier series:

$$f(x) = c + \sum_{n=1}^{\infty} a_n \cos(n\omega x) + b_n \sin(n\omega x) \quad (16)$$

where  $\omega = 2\pi/T$  and  $T$  is the period of the piecewise periodic function. By construction,  $f(x)$  is continuous in value and first derivative and has second and third derivatives that are piecewise-continuous and bounded, and so the Fourier series coefficients decay as  $1/n^3$  or faster<sup>1</sup> and thus we can write:

$$f(x) = c + \sum_{n=1}^{\infty} \frac{1}{n^3} \left( \tilde{a}_n \cos(n\omega x) + \tilde{b}_n \sin(n\omega x) \right) \quad (17)$$

where  $|\tilde{a}_n|, |\tilde{b}_n| \in \mathcal{O}(1)$ . As  $f(x) = f(x)$  whenever  $p(x^{(i)}) \neq 0$ , we can express:

$$\begin{aligned} \mu &= \sum_{x^{(i)}} p(x^{(i)}) f(x^{(i)}) \\ &= \sum_{x^{(i)}} p(x^{(i)}) f(x^{(i)}) \\ &= \sum_{x^{(i)}} p(x^{(i)}) \left( \sum_{n=1}^{\infty} \frac{1}{n^3} \left( \tilde{a}_n \cos(n\omega x^{(i)}) + \tilde{b}_n \sin(n\omega x^{(i)}) \right) + c \right) \\ &= c + \sum_{n=1}^{\infty} \frac{\tilde{a}_n}{n^3} \left( \sum_{x^{(i)}} p(x^{(i)}) \cos(n\omega x^{(i)}) \right) \\ &\quad + \sum_{n=1}^{\infty} \frac{\tilde{b}_n}{n^3} \left( \sum_{x^{(i)}} p(x^{(i)}) \sin(n\omega x^{(i)}) \right) \end{aligned} \quad (18)$$

To estimate  $\mu$ , we can thus estimate each of the parenthesised sums in final line of (18) individually. These can be estimated using QAE according to the procedure described in Proposition 2. To analyse the overall accuracy of the consequent estimate of  $\mu$ , we first define:

$$\mu_n^{(a)} = \sum_{x^{(i)}} p(x^{(i)}) \cos(n\omega x^{(i)}) \quad (19)$$

$$\mu_n^{(b)} = \sum_{x^{(i)}} p(x^{(i)}) \sin(n\omega x^{(i)}) \quad (20)$$

and further define  $\hat{\mu}_n^{(a)}$  and  $\hat{\mu}_n^{(b)}$  as the estimates of  $\mu_n^{(a)}$  and  $\mu_n^{(b)}$  respectively such that:

$$\mu_n^{(a)} = \hat{\mu}_n^{(a)} + \nu_n^{(a)} + \epsilon_n^{(a)} \quad (21)$$

$$\mu_n^{(b)} = \hat{\mu}_n^{(b)} + \nu_n^{(b)} + \epsilon_n^{(b)} \quad (22)$$

where  $\nu_n^{(a)}$  and  $\nu_n^{(b)}$  are biases (noting that the possibility of biased estimation has not been precluded)

<sup>1</sup>This is a relatively commonly-used result, however there does not appear to be a standard citation – so a theorem to this effect is given in Appendix B

and  $\epsilon_n^{(a)}$  and  $\epsilon_n^{(b)}$  are zero-mean random variables. The inaccuracy of estimating  $\mu$  in this way does not solely arise because of the inaccuracy of the estimates of each term in (18), we must also address the fact that the Fourier series is an infinite sum, and so obviously we cannot estimate all of the terms in practise. Therefore we truncate at some  $n_{max}$ . That is, our estimate,  $\hat{\mu}$ , of  $\mu$  is:

$$\hat{\mu} = c + \sum_{n=1}^{n_{max}} \frac{\tilde{a}_n}{n^3} \hat{\mu}_n^{(a)} + \frac{\tilde{b}_n}{n^3} \hat{\mu}_n^{(b)} \quad (23)$$

and to keep track of the error incurred by this truncation, we define the value of the truncated part of the sum:

$$\eta = \sum_{n=n_{max}+1}^{\infty} \frac{\tilde{a}_n}{n^3} \left( \sum_{x^{(i)}} \cos(n\omega x^{(i)}) p(x^{(i)}) \right) + \frac{\tilde{b}_n}{n^3} \left( \sum_{x^{(i)}} \sin(n\omega x^{(i)}) p(x^{(i)}) \right) \quad (24)$$

and it is also convenient to define:

$$\nu^{(a)} = \sum_{n=1}^{n_{max}} \frac{\tilde{a}_n}{n^3} \nu_n^{(a)} \quad (25)$$

$$\nu^{(b)} = \sum_{n=1}^{n_{max}} \frac{\tilde{b}_n}{n^3} \nu_n^{(b)} \quad (26)$$

Substituting (19) – (24) into (18) and also using the definitions (25) and (26) we can express the actual mean,  $\mu$ , in terms of our estimate of the mean,  $\hat{\mu}$ :

$$\begin{aligned} \mu &= \hat{\mu} + \sum_{n=1}^{n_{max}} \frac{\tilde{a}_n}{n^3} \left( \nu_n^{(a)} + \epsilon_n^{(a)} \right) + \frac{\tilde{b}_n}{n^3} \left( \nu_n^{(b)} + \epsilon_n^{(b)} \right) + \eta \\ &= \hat{\mu} + \nu^{(a)} + \nu^{(b)} + \sum_{n=1}^{n_{max}} \frac{\tilde{a}_n}{n^3} \epsilon_n^{(a)} + \frac{\tilde{b}_n}{n^3} \epsilon_n^{(b)} + \eta \end{aligned} \quad (27)$$

from which we can express the MSE:

$$\begin{aligned} \hat{\epsilon}^2 &= \mathbb{E}((\mu - \hat{\mu})^2) \\ &= \mathbb{E} \left( \left( \nu^{(a)} + \nu^{(b)} + \eta + \sum_{n=1}^{n_{max}} \frac{\tilde{a}_n}{n^3} \epsilon_n^{(a)} + \frac{\tilde{b}_n}{n^3} \epsilon_n^{(b)} \right)^2 \right) \\ &= (\nu^{(a)} + \nu^{(b)})^2 + \eta^2 + (\nu^{(a)} + \nu^{(b)})\eta \\ &\quad + \sum_{n=1}^{n_{max}} \left( \frac{\tilde{a}_n}{n^3} \right)^2 \mathbb{E}((\epsilon_n^{(a)})^2) + \left( \frac{\tilde{b}_n}{n^3} \right)^2 \mathbb{E}((\epsilon_n^{(b)})^2) \end{aligned} \quad (28)$$

using the fact that  $\epsilon_n^{(a)}$  and  $\epsilon_n^{(b)}$  are zero-mean by design, and also that estimates corresponding to the various terms in the Fourier series are independent.

To prove the theorem it is now incumbent upon us to show that the number of uses of  $P$  assigned to the QAE estimation of each of the terms in (18) and

$n_{max}$  can be chosen such that the claimed convergence holds. To do this, we first let  $q_n$  be the number of uses of  $P$  allowed in the QAE estimate of *each* of the sine and cosine components of the  $n^{th}$  term in the Fourier series such that:

$$q_n = q_0 n^{-\kappa} \quad (29)$$

for some constants  $q_0$  and  $\kappa$  the latter of which we will later give numerical value. Recalling that  $|\tilde{a}_n|, |\tilde{b}_n| \in \mathcal{O}(1)$ , we can define a constant  $k_2$  such that  $k_2 \geq |\tilde{a}_n|, |\tilde{b}_n|$  (in the following analysis we will need to introduce further new constants, and will henceforth do so without explicit announcement by using  $k$  with subscript incremented relative to the most recently introduced constant). We now address the terms in (28). First, we note that the magnitude of the bias,  $\nu_n^{(a)}$ , is upper-bounded by the square root of the MSE, which has been bounded in (7), so (starting with (25) and also using (29), and  $k_2 \geq |\tilde{a}_n|$  as defined above):

$$\begin{aligned} |\nu^{(a)}| &\leq k_2 \sqrt{k_1} \sum_{n=1}^{n_{max}} \frac{1}{n^3} q_n^{-\lambda/2} \\ &\leq k_2 \sqrt{k_1} q_0^{-\lambda/2} \sum_{n=1}^{\infty} n^{\kappa\lambda/2-3} \\ &= k_3 q_0^{-\lambda/2} \end{aligned} \quad (30)$$

so long as  $\kappa$  is chosen such that  $\kappa < 4/\lambda$  and the sum therefore converges. In exactly the same way we get:

$$|\nu^{(b)}| \leq k_3 q_0^{-\lambda/2} \quad (31)$$

Turning now to the term (of (28)),  $\mathbb{E}((\epsilon_n^{(a)})^2)$ , which is the mean-squared error of the unbiased part of the estimate of  $\mu_n^{(a)}$ , and is therefore upper-bounded by  $k_1 q_n^{-\lambda}$  (this would hold with equality for an unbiased estimator and be strictly less for an estimator with non-zero bias). So we get (again using (7)):

$$\begin{aligned} \sum_{n=1}^{n_{max}} \left( \frac{\tilde{a}_n}{n^3} \right)^2 \mathbb{E}((\epsilon_n^{(a)})^2) &\leq k_1 \sum_{n=1}^{n_{max}} \left( \frac{\tilde{a}_n}{n^3} \right)^2 q_n^{-\lambda} \\ &\leq k_1 k_2^2 q_0^{-\lambda} \sum_{n=1}^{\infty} n^{\kappa\lambda-6} \\ &= k_4 q_0^{-\lambda} \end{aligned} \quad (32)$$

where the previously specified condition,  $\kappa < 4/\lambda$ , suffices to ensure that this sum also converges. Likewise we get

$$\sum_{n=1}^{n_{max}} \left( \frac{\tilde{b}_n}{n^3} \right)^2 \mathbb{E}((\epsilon_n^{(b)})^2) \leq k_4 q_0^{-\lambda} \quad (33)$$

The final quantity in (28) that we must address is the value of the truncated term,  $\eta$ , defined in (24).

The magnitudes of  $\cos(n\omega x^{(i)})$  and  $\sin(n\omega x^{(i)})$  are bounded above by 1 for all  $x^{(i)}$ , and so

$$-1 \leq \sum_{x^{(i)}} \cos(n\omega x^{(i)}) p(x^{(i)}) \leq 1 \quad (34)$$

$$-1 \leq \sum_{x^{(i)}} \sin(n\omega x^{(i)}) p(x^{(i)}) \leq 1 \quad (35)$$

and thus from (24), using  $k_2 \geq |\tilde{a}_n|, |\tilde{b}_n|$  (by definition) we have

$$\begin{aligned} |\eta| &\leq \sum_{n=n_{max}+1}^{\infty} \frac{|\tilde{a}_n| + |\tilde{b}_n|}{n^3} \\ &\leq 2k_2 \sum_{n=n_{max}+1}^{\infty} \frac{1}{n^3} \\ &\leq 2k_2 \int_{n_{max}}^{\infty} \frac{1}{x^3} dx \\ &= \frac{k_2}{n_{max}^2} \end{aligned} \quad (36)$$

Putting this all together by substituting (30), (31), (32), (33) and (36) into (28) we get:

$$\hat{\epsilon}^2 \leq 4k_3^2 q_0^{-\lambda} + \frac{k_2^2}{n_{max}^4} + \frac{2k_2 k_3 q_0^{-\lambda/2}}{n_{max}^2} + 2k_4 q_0^{-\lambda} \quad (37)$$

From which we can see that in order to achieve  $\hat{\epsilon}^2 \propto q_0^{-\lambda}$  it suffices to choose  $n_{max}$  such that:

$$n_{max} = \lceil q_0^{\lambda/4} \rceil \quad (38)$$

We must now count up the total number of uses of  $P$ . Recall that we have assigned  $q_n$  uses of  $P$  to *each* of the sine and cosine components of the  $n^{th}$  term in the Fourier series, thus we have  $\sum_{n=1}^{n_{max}} 2q_n$  uses of  $P$ . However, in the above analysis we have allowed  $q_n$  to take non-integer values, but in practise we must round to an integer. To ensure that the claim of convergence holds, we round up in each case. If all  $n_{max}$  Fourier components require  $q_n$  to be rounded up by an amount approaching 1 for each of the sine and cosine components then we will require an extra  $2n_{max}$  uses of  $P$  in the worst case. Thus we can evaluate the total number of uses of  $P$ :

$$\begin{aligned} q_{tot} &= \sum_{n=1}^{n_{max}} 2q_n + 2n_{max} \\ &= 2q_0 \sum_{n=1}^{n_{max}} n^{-\kappa} + 2\lceil q_0^{\lambda/4} \rceil \\ &\leq 2q_0 \sum_{n=1}^{\infty} n^{-\kappa} + 2\lceil q_0^{\lambda/4} \rceil \end{aligned} \quad (39)$$

From (30) we require that  $\kappa < 4/\lambda$ , and we also have that in practise  $1 < \lambda \leq 2$  (ie, at  $\lambda \leq 1$  there is no quantum advantage, and  $\lambda = 2$  is the Heisenberg limit). Thus choosing any  $\kappa < 2$  always suffices, and

so we let  $\kappa = 2 - \delta$  for some  $0 < \delta < 1$ . As  $\lambda \leq 2$ , we have that  $\lambda/4 \leq 1/2$ , and thus we get:

$$\begin{aligned} q_{tot} &\leq 2q_0 \sum_{n=1}^{\infty} n^{-2+\delta} + 2q_0^{1/2} \\ &\in \Theta(q_0) \end{aligned} \quad (40)$$

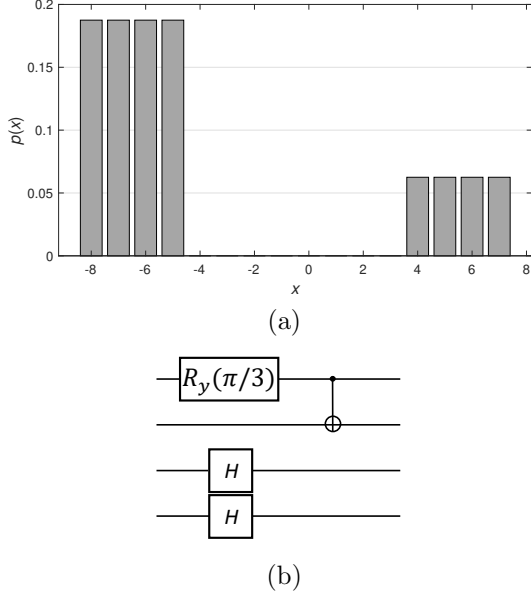
because the sum converges when  $0 < \delta < 1$ . Thus we have shown that the total number of queries is proportional to  $q_0$ . Putting this together with  $\hat{\epsilon}^2 \propto q_0$  from (37) and (38) we can thus see that  $\hat{\epsilon}^2 \in \Theta(q_{tot}^{-\lambda})$ , proving the theorem.  $\square$

It is important to be careful about the evaluation of the Fourier coefficients. If these can only be computed using numerical integration, then we have essentially shifted the computational load, but not reduced the overall complexity. For many common functions of interest, such as the mean ( $f(x) = x$ ), second moment ( $f(x) = x^2$ ) and indeed any higher-order moments, then it is easy to see that the Fourier coefficients can be calculated symbolically, and so concerns on this front can be allayed. Furthermore, if there are functions whose Fourier coefficients cannot be found symbolically, but nevertheless are commonly-used, then it may be reasonable to assume that these have been pre-computed to high accuracy and stored, and thus should not count against the complexity of any individual run. However, more generally the Fourier coefficients should be symbolically computable in order that the method attains the advertised speed-up once all of the complexity is counted.

#### IV. QMCI OF THE PRODUCT OF TWO CORRELATED RANDOM VARIABLES

As the proposed method of QMCI applies to quantum states that encode *multivariate* probability distributions, a natural next step is to apply the proposed method to functions of multiple random variables. For the sake of definiteness, and for simplicity, we restrict the following analysis to computing the product of functions of two random variables. Moreover, this setting covers computation of correlation and covariance, and indeed most quantities that may be of interest in practise. However, in principle the following approach of using trigonometric identities to break the Monte-Carlo integral down into terms that can be estimated using QAE also applies to the product of (functions of) three or more random variables.

**Corollary 4.** *The quantity,  $\mu = \mathbb{E}(f(X)g(Y))$ , where  $(X, Y) \sim p(x^{(i)}, x^{(j)})$  and  $f$  and  $g$  are functions that are continuous in value and first derivative and whose second and third derivatives are piecewise-continuous and bounded, can be estimated with  $MSE \in \Theta(q^{-\lambda})$ ,*



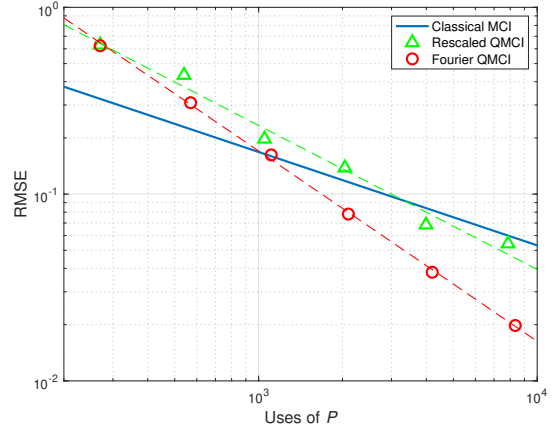
**Fig. 2.** (a) The probability distribution used for the numerical results; (b) a circuit that prepares a quantum state encoding this distribution.

where  $q$  is the number of uses of a circuit preparing the quantum state  $|p\rangle$  (ie, as defined in (2)) and  $\lambda$  is the convergence rate of some QAE subroutine (ie, as defined in (7)) which operates on circuits of the form defined in Fig. 4.

*Proof (sketch).* The idea is to build piecewise periodic functions for both  $f$  and  $g$ , and thus expand each as a Fourier series. This results in a double sum each term of which contains products of sines and cosines. In particular, these will always be either  $\cos(\rho)\cos(\sigma)$ ,  $\sin(\rho)\sin(\sigma)$  or  $\cos(\rho)\sin(\sigma)$  for some  $\rho, \sigma$ . The next step is to use trigonometric identities to express these terms as functions of single sines and cosines (with correspondingly altered angles). Thus we can decompose the sum representing the expectation we need to compute into a double sum over sines and cosines. The analysis then follows along the same lines as that of Theorem 1, and so to avoid tedious repetition we defer the full proof to Appendix A.  $\square$

## V. NUMERICAL RESULTS

To illustrate the practical utility of our method of QMCI, we present some simulated results. So that we can compare to both classical MCI and the rescaled QMCI method of Stamatopoulos *et al* [21] we consider the case where  $f(x) = x$ , that is we simply compute the mean of the sampled distribution without applying any other function. We compare for a 16-point univariate probability distribution supported



**Fig. 3.** A comparison of classical MCI; rescaled QMCI; and Fourier QMCI for the probability distribution shown in Fig. 2(a).

on  $\{-8, -7, \dots, 7\}$ , shown in Fig. 2(a), which is encoded in a quantum state that can be prepared using the circuit shown in Fig. 2(b). This choice of probability distribution and state preparation circuit has been made largely for practical reasons: there is little to be gained (in terms of proving the principle) by addressing a multivariate distribution, and this would significantly slow down the simulations; and on a similar note, it is desirable to use some probability distribution whose encoding state can be prepared in a very shallow circuit. We have avoided a distribution whose mean is the median of the support, as in this case the rescaled QMCI method can perform artificially well.

The results are shown in Fig. 3, where the simulated results for both “Fourier QMCI” (the method we propose in this paper) and “Rescaled QMCI” (the method of Stamatopoulos *et al*) were averaged over 500 runs. We can see that for sufficiently small desired root-MSE (RMSE), Fourier QMCI outperforms both Classical MCI and Rescaled QMCI, as expected. Indeed, the gradient of the line-of-best-fit for Fourier QMCI is  $-1.02$ , which is very close to the theoretical value of  $-1$  (note that this is *root*-MSE, hence  $-1$  not  $-2$  as for MSE), and for rescaled QMCI is  $-0.771$ , clearly inferior and not too far off the theoretical value of  $0.667$ . It is also worth highlighting that the “cross-over” point between Fourier QMCI and classical MCI occurs very close to the third data-point of the former. This represents 1100 uses of  $P$  in total, but more significantly a circuit depth of only 8 sequential uses of  $Q$  in any single quantum circuit.



## VI. DISCUSSION

In this paper we have proposed a method for performing QMCI by expanding the Monte-Carlo integral as a Fourier series and estimating each component individually using amplitude estimation. Our primary motivation for this was to achieve the full quadratic quantum advantage without incurring the need to perform arithmetic operations on the quantum computer, however an added bonus is that our method of QMCI has an innate flexibility that allows it to be readily adapted to a wide variety of settings of real-world relevance. Specifically, we have shown that QMCI can be performed when any function is to be applied to the random samples (with only a mild technical restriction on the smoothness of that function). Moreover, even the expectation of functions of multiple variables can be computed – allowing the quadratic advantage to be achieved when estimating, for example, the covariance and correlation of some data.

**Will Fourier QMCI yield a useful quantum advantage in the NISQ era?** To answer this question, we first must clarify what, exactly, we mean by “the NISQ era”. It is fast becoming clear that the dichotomy between the NISQ era and the Full-Scale Fault-Tolerant (FSFT) era is too simplistic – at least when viewing NISQs in terms of their original definition by Preskill [24]. In these strict terms, the quantum community has very much been focussed on variational algorithms as the only route to NISQ era advantage, however this disposition is somewhat myopic. If the predictions in scaling-up of qubit number (if not fidelity) that hardware manufacturers are making (see eg. Ref [25]) come to fruition, then it is reasonable to look forward to a time when some qubits will be allocated as ancillas to enable some degree of error-correction (if not full fault-tolerance) and others as ancillas to perform entanglement distribution to reduce the depth of otherwise costly qubit routing subcircuits (see eg. Ref. [26,27]), thus effectively facilitating the simulation of a quantum computer with lower qubit-number but higher quantum-volume on the hardware. In such a case, it will in turn be reasonable to think not only of running quantum algorithms with variationally-optimised ansätze, but also those with explicitly-constructed circuits, such as QAE, so long as they can adequately handle the in-clement noise. Removing the need to perform depth-consuming arithmetic operations on the quantum computer, as we have done in this paper, is an important step towards this.

**Will Fourier QMCI still be useful beyond the NISQ era?** One of the most enduring ideas in quantum algorithm design is that, because quantum computers will not offer useful computational speed-ups for all problems, the focus should therefore be on using the quantum computer only when it offers a computational advantage. Moreover, even within a single task those parts of the computation that can be handed off to a classical computer without adversely impacting overall performance should be. This essential paradigm is unlikely to be restricted solely to the NISQ era, especially as the transition from the NISQ era to FSFT era is unlikely to be particularly sharp (as outlined above). In this paper, we have shown how the task of Monte-Carlo integration can be sped-up by using a quantum computer to perform amplitude estimation (where there is a quantum advantage), but not arithmetic (where there is no quantum advantage). We therefore expect that the techniques we have proposed will find useful application for the foreseeable future.

## ACKNOWLEDGEMENT

Thanks to Roland Guichard for writing the code used for the numerical results, and to Ross Duncan for reviewing the paper. Also, special thanks to Zen Harper for pointing out the additional requirement of piecewise-continuity and boundedness on the second and third derivatives that I had implicitly assumed in the first version, and for providing the more rigorous proof of the Fourier series convergence.

## REFERENCES

- [1] A. Montanaro, “Quantum speedup of Monte Carlo methods,” *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 471, no. 2181, p. 20150301, 2015. [Online]. Available: <https://royalsocietypublishing.org/doi/abs/10.1098/rspa.2015.0301>
- [2] G. Brassard, P. Hoyer, M. Mosca, and A. Tapp, “Quantum amplitude amplification and estimation,” 2000.
- [3] D. An, N. Linden, J.-P. Liu, A. Montanaro, C. Shao, and J. Wang, “Quantum-accelerated multilevel monte carlo methods for stochastic differential equations in mathematical finance,” 2020.
- [4] R. Orús, S. Mugel, and E. Lizaso, “Quantum computing for finance: Overview and prospects,” *Reviews in Physics*, vol. 4, p. 100028, 2019. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S2405428318300571>
- [5] D. J. Egger, R. G. Gutiérrez, J. C. Mestre, and S. Woerner, “Credit risk analysis using quantum computers,” 2019.
- [6] S. Chakrabarti, R. Krishnakumar, G. Mazzola, N. Stamatopoulos, S. Woerner, and W. J. Zeng, “A threshold for quantum advantage in derivative pricing,” 2020.
- [7] P. Rebentrost and S. Lloyd, “Quantum computational finance: quantum algorithm for portfolio optimization,” 2018.

- [8] K. Kaneko, K. Miyamoto, N. Takeda, and K. Yoshino, "Quantum pricing with a smile: Implementation of local volatility model on quantum computer," 2020.
- [9] S. Woerner and D. J. Egger, "Quantum risk analysis," *npj Quantum Information*, vol. 5, no. 1, Feb 2019. [Online]. Available: <http://dx.doi.org/10.1038/s41534-019-0130-6>
- [10] P. Rebentrost, B. Gupt, and T. R. Bromley, "Quantum computational finance: Monte carlo pricing of financial derivatives," *Physical Review A*, vol. 98, no. 2, Aug 2018. [Online]. Available: <http://dx.doi.org/10.1103/PhysRevA.98.022321>
- [11] D. J. Egger, C. Gambella, J. Marecek, S. McFaddin, M. Mevissen, R. Raymond, A. Simonetto, S. Woerner, and E. Yndurain, "Quantum computing for finance: State-of-the-art and future prospects," *IEEE Transactions on Quantum Engineering*, vol. 1, pp. 1–24, 2020.
- [12] K. Miyamoto and K. Shiohara, "Reduction of qubits in a quantum algorithm for monte carlo simulation by a pseudo-random-number generator," *Physical Review A*, vol. 102, no. 2, Aug 2020. [Online]. Available: <http://dx.doi.org/10.1103/PhysRevA.102.022424>
- [13] L. Grover and T. Rudolph, "Creating superpositions that correspond to efficiently integrable probability distributions," 2002.
- [14] S. Herbert, "The problem with grover-rudolph state preparation for quantum monte-carlo," 2021.
- [15] Y. Suzuki, S. Uno, R. Raymond, T. Tanaka, T. Onodera, and N. Yamamoto, "Amplitude estimation without phase estimation," *Quantum Information Processing*, vol. 19, no. 2, Jan 2020. [Online]. Available: <http://dx.doi.org/10.1007/s11128-019-2565-2>
- [16] D. Grinko, J. Gacon, C. Zoufal, and S. Woerner, "Iterative quantum amplitude estimation," 2019.
- [17] S. Aaronson and P. Rall, "Quantum approximate counting, simplified," *Symposium on Simplicity in Algorithms*, p. 24–32, Jan 2020. [Online]. Available: <http://dx.doi.org/10.1137/1.9781611976014.5>
- [18] K. Nakaji, "Faster amplitude estimation," 2020.
- [19] I. Kerenidis and A. Prakash, "A method for amplitude estimation with noisy intermediate-scale quantum computers. U.S. Patent Application No. 16/892,229," 2020.
- [20] T. Giurgica-Tiron, I. Kerenidis, F. Labib, A. Prakash, and W. Zeng, "Low depth algorithms for quantum amplitude estimation," 2020.
- [21] N. Stamatopoulos, D. J. Egger, Y. Sun, C. Zoufal, R. Iten, N. Shen, and S. Woerner, "Option pricing using quantum computers," *Quantum*, vol. 4, p. 291, Jul 2020. [Online]. Available: <http://dx.doi.org/10.22331/q-2020-07-06-291>
- [22] S. Herbert, "Quantum Computing System and Method: Patent application GB2102902.0 and SE2130060-3," 2021.
- [23] A. Bouland, W. van Dam, H. Joorati, I. Kerenidis, and A. Prakash, "Prospects and challenges of quantum finance," 2020.
- [24] J. Preskill, "Quantum computing in the nisq era and beyond," *Quantum*, vol. 2, p. 79, Aug 2018. [Online]. Available: <http://dx.doi.org/10.22331/q-2018-08-06-79>
- [25] "IBM quantum roadmap." [Online]. Available: <https://www.ibm.com/blogs/research/2021/02/quantum-development-roadmap>
- [26] N. d. Beaudrap and S. Herbert, "Quantum linear network coding for entanglement distribution in restricted architectures," *Quantum*, vol. 4, p. 356, Nov 2020. [Online]. Available: <http://dx.doi.org/10.22331/q-2020-11-01-356>
- [27] N. d. Beaudrap and S. Herbert, "Method of Operating a Quantum Information Processing System. U.S. Patent Application No. 17/064,980," 2020.

# APPENDIX A. PROOF OF COROLLARY 4

**Corollary 4.** *The quantity,  $\mu = \mathbb{E}(f(X)g(Y))$ , where  $(X, Y) \sim p(x^{(i)}, x^{(j)})$  and  $f$  and  $g$  are functions that are continuous in value and first derivative and whose second and third derivatives are piecewise-continuous and bounded, can be estimated with  $MSE \in \Theta(q^{-\lambda})$ , where  $q$  is the number of uses of a circuit preparing the quantum state  $|p\rangle$  (ie, as defined in (2)) and  $\lambda$  is the convergence rate of some QAE subroutine (ie, as defined in (7)) which operates on circuits of the form defined in Fig. 4.*

*Proof.* The proof follows similar lines to the univariate case. First we construct periodic piecewise functions to extend the two functions of interest, then show that it is possible to truncate and estimate each Fourier component individually to achieve the claimed convergence. To start with, we let  $f$  and  $g$  be periodic piecewise functions (continuous in value and first derivative, and whose second and third derivatives are piecewise-continuous and bounded) for  $f$  and  $g$  respectively – so that  $f(x) = f(x)$  whenever  $p(x)$  is non-zero (and likewise for  $g(x)$ ). Thus we can write the Fourier series:

$$f(x) = c^{(1)} + \sum_{n=1}^{\infty} \frac{1}{n^3} \left( \tilde{a}_n^{(1)} \cos(n\omega^{(1)}x) + \tilde{b}_n^{(1)} \sin(n\omega^{(1)}x) \right) \quad (41)$$

$$g(x) = c^{(2)} + \sum_{m=1}^{\infty} \frac{1}{m^3} \left( \tilde{a}_m^{(2)} \cos(m\omega^{(2)}x) + \tilde{b}_m^{(2)} \sin(m\omega^{(2)}x) \right) \quad (42)$$

where  $|\tilde{a}_n^{(1)}|, |\tilde{a}_m^{(2)}|, |\tilde{b}_n^{(1)}|, |\tilde{b}_m^{(2)}| \in \mathcal{O}(1)$ . Thus we can write the quantity of interest,  $\mu$ , in terms of the Fourier decompositions:

$$\begin{aligned} \mu &= \mathbb{E}(f(X)g(Y)) \\ &= \sum_{x^{(i)}, x^{(j)}} \left( c^{(1)} + \sum_{n=1}^{\infty} \frac{1}{n^3} \left( \tilde{a}_n^{(1)} \cos(n\omega^{(1)}x^{(i)}) + \tilde{b}_n^{(1)} \sin(n\omega^{(1)}x^{(i)}) \right) \right) \\ &\quad \left( c^{(2)} + \sum_{m=1}^{\infty} \frac{1}{m^3} \left( \tilde{a}_m^{(2)} \cos(m\omega^{(2)}x^{(j)}) + \tilde{b}_m^{(2)} \sin(m\omega^{(2)}x^{(j)}) \right) \right) p(x^{(i)}, x^{(j)}) \\ &= c^{(1)} + c^{(2)} + \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \frac{1}{(nm)^3} \left( \sum_{x^{(i)}, x^{(j)}} \left( \tilde{a}_n^{(1)} \cos(n\omega^{(1)}x^{(i)}) + \tilde{b}_n^{(1)} \sin(n\omega^{(1)}x^{(i)}) \right) \right. \\ &\quad \left. \left( \tilde{a}_m^{(2)} \cos(m\omega^{(2)}x^{(j)}) + \tilde{b}_m^{(2)} \sin(m\omega^{(2)}x^{(j)}) \right) p(x^{(i)}, x^{(j)}) \right) \end{aligned} \quad (43)$$

Once again, we assume that the infinite sums are truncated at  $\lceil q_0^{\lambda/4} \rceil$ , thus we have:

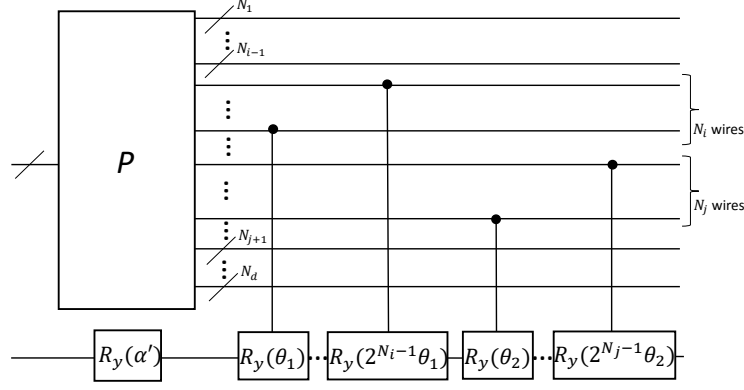
$$\begin{aligned} \mu &= c^{(1)} + c^{(2)} + \sum_{n=1}^{\lceil q_0^{\lambda/4} \rceil} \sum_{m=1}^{\lceil q_0^{\lambda/4} \rceil} \frac{1}{(nm)^3} \left( \sum_{x^{(i)}, x^{(j)}} \left( \tilde{a}_n^{(1)} \cos(n\omega^{(1)}x^{(i)}) + \tilde{b}_n^{(1)} \sin(n\omega^{(1)}x^{(i)}) \right) \right. \\ &\quad \left. \left( \tilde{a}_m^{(2)} \cos(m\omega^{(2)}x^{(j)}) + \tilde{b}_m^{(2)} \sin(m\omega^{(2)}x^{(j)}) \right) p(x^{(i)}, x^{(j)}) \right) + \eta \end{aligned} \quad (44)$$

where

$$\begin{aligned}
\eta &= \sum_{n=\lceil q_0^{\lambda/4} \rceil + 1}^{\infty} \sum_{m=1}^{\infty} \frac{1}{(nm)^3} \left( \sum_{x^{(i)}, x^{(j)}} \left( \tilde{a}_n^{(1)} \cos(n\omega^{(1)} x^{(i)}) + \tilde{b}_n^{(1)} \sin(n\omega^{(1)} x^{(i)}) \right) \right. \\
&\quad \left. \left( \tilde{a}_m^{(2)} \cos(m\omega^{(2)} x^{(j)}) + \tilde{b}_m^{(2)} \sin(m\omega^{(2)} x^{(j)}) \right) p(x^{(i)}, x^{(j)}) \right) \\
&+ \sum_{n=1}^{\infty} \sum_{m=\lceil q_0^{\lambda/4} \rceil + 1}^{\infty} \frac{1}{(nm)^3} \left( \sum_{x^{(i)}, x^{(j)}} \left( \tilde{a}_n^{(1)} \cos(n\omega^{(1)} x^{(i)}) + \tilde{b}_n^{(1)} \sin(n\omega^{(1)} x^{(i)}) \right) \right. \\
&\quad \left. \left( \tilde{a}_m^{(2)} \cos(m\omega^{(2)} x^{(j)}) + \tilde{b}_m^{(2)} \sin(m\omega^{(2)} x^{(j)}) \right) p(x^{(i)}, x^{(j)}) \right) \\
&- \sum_{n=\lceil q_0^{\lambda/4} \rceil + 1}^{\infty} \sum_{m=\lceil q_0^{\lambda/4} \rceil + 1}^{\infty} \frac{1}{(nm)^3} \left( \sum_{x^{(i)}, x^{(j)}} \left( \tilde{a}_n^{(1)} \cos(n\omega^{(1)} x^{(i)}) + \tilde{b}_n^{(1)} \sin(n\omega^{(1)} x^{(i)}) \right) \right. \\
&\quad \left. \left( \tilde{a}_m^{(2)} \cos(m\omega^{(2)} x^{(j)}) + \tilde{b}_m^{(2)} \sin(m\omega^{(2)} x^{(j)}) \right) p(x^{(i)}, x^{(j)}) \right) \\
\Rightarrow |\eta| &\leq \left| \sum_{n=\lceil q_0^{\lambda/4} \rceil + 1}^{\infty} \sum_{m=1}^{\infty} \frac{1}{(nm)^3} \left( \sum_{x^{(i)}, x^{(j)}} \left( \tilde{a}_n^{(1)} \cos(n\omega^{(1)} x^{(i)}) + \tilde{b}_n^{(1)} \sin(n\omega^{(1)} x^{(i)}) \right) \right. \right. \\
&\quad \left. \left. \left( \tilde{a}_m^{(2)} \cos(m\omega^{(2)} x^{(j)}) + \tilde{b}_m^{(2)} \sin(m\omega^{(2)} x^{(j)}) \right) p(x^{(i)}, x^{(j)}) \right) \right| \\
&+ \left| \sum_{n=1}^{\infty} \sum_{m=\lceil q_0^{\lambda/4} \rceil + 1}^{\infty} \frac{1}{(nm)^3} \left( \sum_{x^{(i)}, x^{(j)}} \left( \tilde{a}_n^{(1)} \cos(n\omega^{(1)} x^{(i)}) + \tilde{b}_n^{(1)} \sin(n\omega^{(1)} x^{(i)}) \right) \right. \right. \\
&\quad \left. \left. \left( \tilde{a}_m^{(2)} \cos(m\omega^{(2)} x^{(j)}) + \tilde{b}_m^{(2)} \sin(m\omega^{(2)} x^{(j)}) \right) p(x^{(i)}, x^{(j)}) \right) \right| \\
&+ \left| \sum_{n=\lceil q_0^{\lambda/4} \rceil + 1}^{\infty} \sum_{m=\lceil q_0^{\lambda/4} \rceil + 1}^{\infty} \frac{1}{(nm)^3} \left( \sum_{x^{(i)}, x^{(j)}} \left( \tilde{a}_n^{(1)} \cos(n\omega^{(1)} x^{(i)}) + \tilde{b}_n^{(1)} \sin(n\omega^{(1)} x^{(i)}) \right) \right. \right. \\
&\quad \left. \left. \left( \tilde{a}_m^{(2)} \cos(m\omega^{(2)} x^{(j)}) + \tilde{b}_m^{(2)} \sin(m\omega^{(2)} x^{(j)}) \right) p(x^{(i)}, x^{(j)}) \right) \right| \\
&\leq \sum_{n=\lceil q_0^{\lambda/4} \rceil + 1}^{\infty} \sum_{m=1}^{\infty} \frac{1}{(nm)^3} 4k_2^2 + \sum_{n=1}^{\infty} \sum_{m=\lceil q_0^{\lambda/4} \rceil + 1}^{\infty} \frac{1}{(nm)^3} 4k_2^2 + \sum_{n=\lceil q_0^{\lambda/4} \rceil + 1}^{\infty} \sum_{m=\lceil q_0^{\lambda/4} \rceil + 1}^{\infty} \frac{1}{(nm)^3} 4k_2^2 \\
&\leq 12k_2^2 \sum_{n=\lceil q_0^{\lambda/4} \rceil + 1}^{\infty} \sum_{m=1}^{\infty} \frac{1}{(nm)^3} \\
&\leq k_5 \left( \lceil q_0^{\lambda/4} \rceil \right)^{-2} \\
&\leq k_6 q_0^{-\lambda/2}
\end{aligned} \tag{45}$$

We now address the other terms in (44):

$$\begin{aligned}
& \sum_{x^{(i)}, x^{(j)}} \left( \tilde{a}_n^{(1)} \cos(n\omega^{(1)} x^{(i)}) + \tilde{b}_n^{(1)} \sin(n\omega^{(1)} x^{(i)}) \right) \left( \tilde{a}_m^{(2)} \cos(m\omega^{(2)} x^{(j)}) + \tilde{b}_m^{(2)} \sin(m\omega^{(2)} x^{(j)}) \right) p(x^{(i)}, x^{(j)}) \\
&= \tilde{a}_n^{(1)} \tilde{a}_m^{(2)} \sum_{x^{(i)}, x^{(j)}} \cos(n\omega^{(1)} x^{(i)}) \cos(m\omega^{(2)} x^{(j)}) p(x^{(i)}, x^{(j)}) \\
&\quad + \tilde{a}_n^{(1)} \tilde{b}_m^{(2)} \sum_{x^{(i)}, x^{(j)}} \cos(n\omega^{(1)} x^{(i)}) \sin(m\omega^{(2)} x^{(j)}) p(x^{(i)}, x^{(j)}) \\
&\quad + \tilde{b}_n^{(1)} \tilde{a}_m^{(2)} \sum_{x^{(i)}, x^{(j)}} \sin(n\omega^{(1)} x^{(i)}) \cos(m\omega^{(2)} x^{(j)}) p(x^{(i)}, x^{(j)}) \\
&\quad + \tilde{b}_n^{(1)} \tilde{b}_m^{(2)} \sum_{x^{(i)}, x^{(j)}} \sin(n\omega^{(1)} x^{(i)}) \sin(m\omega^{(2)} x^{(j)}) p(x^{(i)}, x^{(j)}) \\
&= 0.5 \tilde{a}_n^{(1)} \tilde{a}_m^{(2)} \sum_{x^{(i)}, x^{(j)}} \cos(n\omega^{(1)} x^{(i)} - m\omega^{(2)} x^{(j)}) p(x^{(i)}, x^{(j)}) \\
&\quad + 0.5 \tilde{a}_n^{(1)} \tilde{a}_m^{(2)} \sum_{x^{(i)}, x^{(j)}} \cos(n\omega^{(1)} x^{(i)} + m\omega^{(2)} x^{(j)}) p(x^{(i)}, x^{(j)}) \\
&\quad + 0.5 \tilde{a}_n^{(1)} \tilde{b}_m^{(2)} \sum_{x^{(i)}, x^{(j)}} \sin(n\omega^{(1)} x^{(i)} + m\omega^{(2)} x^{(j)}) p(x^{(i)}, x^{(j)}) \\
&\quad - 0.5 \tilde{a}_n^{(1)} \tilde{b}_m^{(2)} \sum_{x^{(i)}, x^{(j)}} \sin(n\omega^{(1)} x^{(i)} - m\omega^{(2)} x^{(j)}) p(x^{(i)}, x^{(j)}) \\
&\quad + 0.5 \tilde{b}_n^{(1)} \tilde{a}_m^{(2)} \sum_{x^{(i)}, x^{(j)}} \sin(n\omega^{(1)} x^{(i)} + m\omega^{(2)} x^{(j)}) p(x^{(i)}, x^{(j)}) \\
&\quad + 0.5 \tilde{b}_n^{(1)} \tilde{a}_m^{(2)} \sum_{x^{(i)}, x^{(j)}} \sin(n\omega^{(1)} x^{(i)} - m\omega^{(2)} x^{(j)}) p(x^{(i)}, x^{(j)}) \\
&\quad + 0.5 \tilde{b}_n^{(1)} \tilde{b}_m^{(2)} \sum_{x^{(i)}, x^{(j)}} \cos(n\omega^{(1)} x^{(i)} - m\omega^{(2)} x^{(j)}) p(x^{(i)}, x^{(j)}) \\
&\quad - 0.5 \tilde{b}_n^{(1)} \tilde{b}_m^{(2)} \sum_{x^{(i)}, x^{(j)}} \cos(n\omega^{(1)} x^{(i)} + m\omega^{(2)} x^{(j)}) p(x^{(i)}, x^{(j)}) \\
&= 0.5 (\tilde{a}_n^{(1)} \tilde{a}_m^{(2)} + \tilde{b}_n^{(1)} \tilde{b}_m^{(2)}) \sum_{x^{(i)}, x^{(j)}} \cos(n\omega^{(1)} x^{(i)} - m\omega^{(2)} x^{(j)}) p(x^{(i)}, x^{(j)}) \\
&\quad + 0.5 (\tilde{a}_n^{(1)} \tilde{a}_m^{(2)} - \tilde{b}_n^{(1)} \tilde{b}_m^{(2)}) \sum_{x^{(i)}, x^{(j)}} \cos(n\omega^{(1)} x^{(i)} + m\omega^{(2)} x^{(j)}) p(x^{(i)}, x^{(j)}) \\
&\quad + 0.5 (\tilde{a}_n^{(1)} \tilde{b}_m^{(2)} + \tilde{b}_n^{(1)} \tilde{a}_m^{(2)}) \sum_{x^{(i)}, x^{(j)}} \sin(n\omega^{(1)} x^{(i)} + m\omega^{(2)} x^{(j)}) p(x^{(i)}, x^{(j)}) \\
&\quad + 0.5 (\tilde{a}_n^{(1)} \tilde{b}_m^{(2)} - \tilde{b}_n^{(1)} \tilde{a}_m^{(2)}) \sum_{x^{(i)}, x^{(j)}} \sin(n\omega^{(1)} x^{(i)} - m\omega^{(2)} x^{(j)}) p(x^{(i)}, x^{(j)}) \tag{46}
\end{aligned}$$



**Fig. 4.** Circuit diagram of  $A'(P, i, j, \beta, n, m, \omega^{(1)}, \omega^{(2)})$  where:  $\alpha' = n\omega^{(1)}x_l^{(i)} + m\omega^{(2)}x_l^{(j)} - \beta$ ,  $\theta_1 = n\omega^{(1)}\Delta^{(i)}$  and  $\theta_2 = m\omega^{(2)}\Delta^{(j)}$ .

We can now use QAE to find an approximation of each of the four terms in the final line of (46). To do this, we define:

$$\mu_{nm}^{(a)} = \sum_{x^{(i)}, x^{(j)}} \cos(n\omega^{(1)}x^{(i)} - m\omega^{(2)}x^{(j)})p(x^{(i)}, x^{(j)}) \quad (47)$$

$$\mu_{nm}^{(b)} = \sum_{x^{(i)}, x^{(j)}} \cos(n\omega^{(1)}x^{(i)} + m\omega^{(2)}x^{(j)})p(x^{(i)}, x^{(j)}) \quad (48)$$

$$\mu_{nm}^{(c)} = \sum_{x^{(i)}, x^{(j)}} \sin(n\omega^{(1)}x^{(i)} + m\omega^{(2)}x^{(j)})p(x^{(i)}, x^{(j)}) \quad (49)$$

$$\mu_{nm}^{(d)} = \sum_{x^{(i)}, x^{(j)}} \sin(n\omega^{(1)}x^{(i)} - m\omega^{(2)}x^{(j)})p(x^{(i)}, x^{(j)}) \quad (50)$$

These can each be estimated using QAE. In particular we define a family of parameterised circuits,  $A'(P, i, j, \beta, n, m, \omega^{(1)}, \omega^{(2)})$  shown in Fig. 4, where the support of the  $i^{th}$  dimension of the multivariate probability distribution is defined in terms of  $x_l^{(i)}$  and  $\Delta^{(i)}$  (and analogously for the  $j^{th}$  dimension). Once again  $\beta$  is set to  $\pi/2$  when the sinusoidal components are being estimated, and in the case of  $\mu_{nm}^{(a)}$  and  $\mu_{nm}^{(d)}$  we set  $m$  to  $-m$  to achieve the negation. We let each term have  $q_n = q_0(nm)^{2-\delta}$  uses of  $P$  (for some  $0 < \delta < 1$ ), and thus achieve the estimates:

$$\mu_{nm}^{(a)} = \hat{\mu}_{nm}^{(a)} + \nu_{nm}^{(a)} + \epsilon_{nm}^{(a)} \quad (51)$$

$$\mu_{nm}^{(b)} = \hat{\mu}_{nm}^{(b)} + \nu_{nm}^{(b)} + \epsilon_{nm}^{(b)} \quad (52)$$

$$\mu_{nm}^{(c)} = \hat{\mu}_{nm}^{(c)} + \nu_{nm}^{(c)} + \epsilon_{nm}^{(c)} \quad (53)$$

$$\mu_{nm}^{(d)} = \hat{\mu}_{nm}^{(d)} + \nu_{nm}^{(d)} + \epsilon_{nm}^{(d)} \quad (54)$$

where again  $\hat{\mu}_{nm}^{(a)}$  is the estimate,  $\nu_{nm}^{(a)}$  is the bias of the estimator and  $\epsilon_{nm}^{(a)}$  is the zero-mean error of the estimator (and likewise for the other terms).

Making the appropriate substitutions in (44) we get

$$\begin{aligned} \mu - \hat{\mu} = 0.5 \sum_{n=1}^{\lceil q_0^{\lambda/4} \rceil} \sum_{m=1}^{\lceil q_0^{\lambda/4} \rceil} \frac{1}{(nm)^3} & \left( (\tilde{a}_n^{(1)}\tilde{a}_m^{(2)} + \tilde{b}_n^{(1)}\tilde{b}_m^{(2)})(\nu_{nm}^{(a)} + \epsilon_{nm}^{(a)}) + (\tilde{a}_n^{(1)}\tilde{a}_m^{(2)} - \tilde{b}_n^{(1)}\tilde{b}_m^{(2)})(\nu_{nm}^{(b)} + \epsilon_{nm}^{(b)}) \right. \\ & \left. + (\tilde{a}_n^{(1)}\tilde{b}_m^{(2)} + \tilde{b}_n^{(1)}\tilde{a}_m^{(2)})(\nu_{nm}^{(c)} + \epsilon_{nm}^{(c)}) + (\tilde{a}_n^{(1)}\tilde{b}_m^{(2)} - \tilde{b}_n^{(1)}\tilde{a}_m^{(2)})(\nu_{nm}^{(d)} + \epsilon_{nm}^{(d)}) \right) + \eta \end{aligned} \quad (55)$$

Once again, using the fact that  $|\tilde{a}_n^{(1)}|, |\tilde{a}_m^{(2)}|, |\tilde{b}_n^{(1)}|, |\tilde{b}_m^{(2)}| \in \mathcal{O}(1)$  we use the constant  $k_2$  to upper-bound the magnitudes thereof, and get:

$$|\mu - \hat{\mu}| \leq k_2^2 \left| \sum_{n=1}^{\lceil q_0^{\lambda/4} \rceil} \sum_{m=1}^{\lceil q_0^{\lambda/4} \rceil} \frac{1}{(nm)^3} \left( \nu_{nm}^{(a)} + \epsilon_{nm}^{(a)} + \nu_{nm}^{(b)} + \epsilon_{nm}^{(b)} + \nu_{nm}^{(c)} + \epsilon_{nm}^{(c)} + \nu_{nm}^{(d)} + \epsilon_{nm}^{(d)} \right) + \eta \right| \quad (56)$$

We also define:

$$\nu = k_2^2 \sum_{n=1}^{\lceil q_0^{\lambda/4} \rceil} \sum_{m=1}^{\lceil q_0^{\lambda/4} \rceil} \frac{1}{(nm)^3} \nu_{nm}^{(a)} + \nu_{nm}^{(b)} + \nu_{nm}^{(c)} + \nu_{nm}^{(d)} \quad (57)$$

Which thus allows to express:

$$\begin{aligned} \mathbb{E}((\mu - \hat{\mu})^2) &\leq \mathbb{E} \left( \left( \nu + k_2^2 \sum_{n=1}^{\lceil q_0^{\lambda/4} \rceil} \sum_{m=1}^{\lceil q_0^{\lambda/4} \rceil} \frac{1}{(nm)^3} (\epsilon_{nm}^{(a)} + \epsilon_{nm}^{(b)} + \epsilon_{nm}^{(c)} + \epsilon_{nm}^{(d)}) + \eta \right)^2 \right) \\ &= \nu^2 + \eta^2 + 2\nu\eta + k_2^2 \sum_{n=1}^{\lceil q_0^{\lambda/4} \rceil} \sum_{m=1}^{\lceil q_0^{\lambda/4} \rceil} \frac{1}{(nm)^6} \left( \mathbb{E}((\epsilon_{nm}^{(a)})^2) + \mathbb{E}((\epsilon_{nm}^{(b)})^2) + \mathbb{E}((\epsilon_{nm}^{(c)})^2) + \mathbb{E}((\epsilon_{nm}^{(d)})^2) \right) \end{aligned} \quad (58)$$

We have already upper-bounded  $|\eta| \leq k_6 q_0^{-\lambda/2}$  in (45), and we now bound the other terms, recalling that  $q_{nm} = q_0(nm)^{-2+\delta}$  uses of  $P$  for each QAE subroutine. Starting with  $\nu$ , and noting that  $|\nu_{nm}^{(a)}|$  is upper-bounded by the square root of the MSE (and likewise for  $|\nu_{nm}^{(b)}|, |\nu_{nm}^{(c)}|$  and  $|\nu_{nm}^{(d)}|$ ), we get:

$$\begin{aligned} |\nu| &\leq k_2^2 \sum_{n=1}^{\lceil q_0^{\lambda/4} \rceil} \sum_{m=1}^{\lceil q_0^{\lambda/4} \rceil} \frac{4}{(nm)^3} \sqrt{k_1 q_n^{-\lambda}} \\ &= 4k_2^2 \sum_{n=1}^{\lceil q_0^{\lambda/4} \rceil} \sum_{m=1}^{\lceil q_0^{\lambda/4} \rceil} \frac{1}{(nm)^3} k_1^{1/2} (q_0(nm)^{-2+\delta})^{-\lambda/2} \\ &\leq 4k_1^{1/2} k_2^2 q_0^{-\lambda/2} \sum_{n=1}^{\infty} n^{\lambda(1-\delta/2)-3} \sum_{m=1}^{\infty} m^{\lambda(1-\delta/2)-3} \\ &= k_7 q_0^{-\lambda/2} \end{aligned} \quad (59)$$

by noting that  $\lambda \leq 2$ . Similarly, again we have that:

$$\mathbb{E}((\epsilon_{nm}^{(a)})^2), \mathbb{E}((\epsilon_{nm}^{(b)})^2), \mathbb{E}((\epsilon_{nm}^{(c)})^2), \mathbb{E}((\epsilon_{nm}^{(d)})^2) \leq k_1 q_n^{-\lambda} \quad (60)$$

with equality for unbiased estimates. Thus we get:

$$\begin{aligned} \sum_{n=1}^{\lceil q_0^{\lambda/4} \rceil} \sum_{m=1}^{\lceil q_0^{\lambda/4} \rceil} \frac{1}{(nm)^6} \left( \mathbb{E}((\epsilon_{nm}^{(a)})^2) + \mathbb{E}((\epsilon_{nm}^{(b)})^2) + \mathbb{E}((\epsilon_{nm}^{(c)})^2) + \mathbb{E}((\epsilon_{nm}^{(d)})^2) \right) &\leq \sum_{n=1}^{\lceil q_0^{\lambda/4} \rceil} \sum_{m=1}^{\lceil q_0^{\lambda/4} \rceil} \frac{1}{(nm)^6} k_1 (q_0(nm)^{-2+\delta})^{-\lambda} \\ &= k_1 q_0^{-\lambda} \sum_{n=1}^{\lceil q_0^{\lambda/4} \rceil} n^{\lambda(2-\delta)-6} \sum_{m=1}^{\lceil q_0^{\lambda/4} \rceil} m^{\lambda(2-\delta)-6} \\ &= k_1 q_0^{-\lambda} \sum_{n=1}^{\infty} n^{\lambda(2-\delta)-6} \sum_{m=1}^{\infty} m^{\lambda(2-\delta)-6} \\ &= k_8 q_0^{-\lambda} \end{aligned} \quad (61)$$

again using  $\lambda \leq 2$ . Substituting (45), (59) and (61) into (58) we get:

$$\begin{aligned} \mathbb{E}((\mu - \hat{\mu})^2) &\leq (k_2^2 k_8 + (k_6 + k_7)^2) q_0^{-\lambda} \\ &\in \mathcal{O}(q_0^{-\lambda}) \end{aligned} \quad (62)$$

Even though we haven't shown the lower bound, it is clear that the method presented here won't allow us to do better than  $\Omega(q_0^{-\lambda})$ , so we can conclude that the complexity is  $\Theta(q_0^{-\lambda})$ . Finally, we must count up

the number of uses of  $P$ . Once again, we must include the possibility of needing to round up – we include a total of  $(\lceil q_0^{\lambda/4} \rceil)^2$  Fourier components (that is, because both random variables being multiplied have  $\lceil q_0^{\lambda/4} \rceil$  terms), and each of these requires four uses of QAE, thus we get that the total number of uses of  $P$  can be counted up as:

$$\begin{aligned}
q_{tot} &= \sum_{n=1}^{\lceil q_0^{\lambda/4} \rceil} \sum_{m=1}^{\lceil q_0^{\lambda/4} \rceil} 4k_1 q_0 (nm)^{-2+\delta} + 4(\lceil q_0^{\lambda/4} \rceil)^2 \\
&\leq 4k_1 q_0 \sum_{n=1}^{\infty} n^{-2+\delta} \sum_{m=1}^{\infty} m^{-2+\delta} + 4(\lceil q_0^{\lambda/4} \rceil)^2 \\
&\in \Theta(q_0)
\end{aligned} \tag{63}$$

using the fact that  $\delta < 1$ . Thus we have shown that the total number of uses of  $P$  is just a constant multiple of  $q_0$ , and from (62) we have shown that the MSE decays as  $q_0^{-\lambda}$ , thus proving the Corollary.  $\square$



## APPENDIX B. CONVERGENCE OF FOURIER SERIES

Let  $g(x)$  be a function that is continuous in value and first derivative and piecewise-continuous and bounded in second and third derivatives, which we differentiate twice to give a function,  $f(x)$  that is piecewise-continuous and bounded in value and first derivative.

**Theorem 5** (Harper). *Let  $f : \mathbb{R} \rightarrow \mathbb{C}$  be periodic of period  $T > 0$ , and let  $\{t_j\}$  be a finite dissection of  $[0, T]$ :*

$$0 = t_0 < t_1 < t_2 < \dots < t_N = T.$$

*Suppose that  $f, f'$  are continuous at all points of  $[0, T] \setminus \{t_j\}$  (where  $f'$  denotes the differential w.r.t  $x$ ), and that*

$$\int_0^T |f'(x)| dx = \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} |f'(x)| dx < \infty. \quad (64)$$

*(Note that this makes sense even though  $f'$  need not exist at  $t_j$ . Also note that this easily implies that  $\lim_{x \rightarrow t_j^+} f(x)$  and  $\lim_{x \rightarrow t_{j+1}^-} f(x)$  exist).*

*Then the Fourier coefficients of  $f$  for the interval  $[0, T]$  decay like  $1/|n|$ .*

*Proof.* For convenience we use the exponential form of the Fourier series. Let  $d_n$  be the coefficient of the  $n^{\text{th}}$  component of the exponential Fourier series, which is defined:

$$d_n = \frac{1}{T} \int_0^T f(x) \exp(-in\omega x) dx \quad (65)$$

Starting with the integral between  $t_j$  and  $t_{j+1}$ , which we can integrate by parts:

$$\int_{t_j}^{t_{j+1}} f(x) \exp(-in\omega x) dx = -\frac{1}{in\omega} [f(x) \exp(-in\omega x)]_{t_j}^{t_{j+1}} + \frac{1}{in\omega} \int_{t_j}^{t_{j+1}} f'(x) \exp(-in\omega x) dx, \quad (66)$$

so that

$$\left| \int_{t_j}^{t_{j+1}} f(x) \exp(-in\omega x) dx \right| \leq \frac{1}{|n\omega|} \left( |f(t_j^+)| + |f(t_{j+1}^-)| + \int_{t_j}^{t_{j+1}} |f'(x)| dx \right). \quad (67)$$

**Note:** we need the one-sided limits  $\lim_{x \rightarrow t_j^+} f(x)$  and  $\lim_{x \rightarrow t_{j+1}^-} f(x)$  to exist, which they easily do because of (64), as remarked above.

There are *finitely many* such terms, so summing gives

$$\left| \frac{1}{T} \int_0^T f(x) \exp(-in\omega x) dx \right| \leq \frac{1}{|n\omega|T} \left( A + \int_0^T |f'(x)| dx \right) = \frac{B}{|n\omega|} \quad (68)$$

for some  $A, B > 0$  (which of course depend on  $f$ ). Referring back to (65) we can see that we have shown that  $|d_n|$  indeed decay like  $1/|n|$ .  $\square$

If we now integrate the Fourier series of  $f(x)$  twice w.r.t  $x$  to get the Fourier series of our original function,  $g(x)$ , then we can see that we get Fourier coefficients decaying as  $1/|n|^3$ . That is because the  $n^{\text{th}}$  term in the Fourier series is  $d_n \exp(in\omega x)$ , and  $\int d_n \exp(in\omega x) dx = \frac{d_n}{in\omega} \exp(in\omega x)$  (and likewise for the second integration).