Cambridge Quantum

Variational Quantum Amplitude Estimation

by Kirill Plekhanov, Matthias Rosenkranz, Mattia Fiorentini, Michael Lubasch

Cambridge Quantum
KIRIL PLEKHANOV
kirill.plekhanov@cambridgequantum.com
MATTHIAS ROSENKRANZ
matthias.rosenkranz@cambridgequantum.com
MATTIA FIORENTINI
mattia.fiorentini@cambridgequantum.com
MICHAEL LUBASCH
michael.lubasch@cambridgequantum.com

Variational quantum amplitude estimation

Kirill Plekhanov,* Matthias Rosenkranz, Mattia Fiorentini, and Michael Lubasch

Cambridge Quantum Computing Limited, SW1E 6DR London, United Kingdom

(Dated: September 9, 2021)

We propose to perform amplitude estimation with the help of constant-depth quantum circuits that variationally approximate states during amplitude amplification. In the context of Monte Carlo (MC) integration, we numerically show that shallow circuits can accurately approximate many amplitude amplification steps. We combine the variational approach with maximum likelihood amplitude estimation [Y. Suzuki *et al.*, Quantum Inf. Process. **19**, 75 (2020)] in variational quantum amplitude estimation (VQAE). VQAE can exhibit a cubic quantum speedup over classical MC sampling if the variational cost is ignored. If this cost is taken into account, VQAE typically has larger computational requirements than classical MC sampling. To reduce the variational cost, we propose adaptive VQAE and numerically show that it can outperform classical MC sampling.

I. INTRODUCTION

Amplitude estimation [1] is a powerful algorithm that can achieve a quadratic quantum speedup over classical Monte Carlo (MC) methods [2]. It has a wide range of applications, e.g. in quantum chemistry [3, 4], machine learning [5–7], and finance [8–10] where it can help with tasks such as risk analysis [11, 12] and the pricing of financial derivatives [13, 14].

The original amplitude estimation procedure [1] has hardware requirements that are challenging for current quantum devices and, therefore, reducing these requirements is currently an active area of research. Crucial breakthroughs were obtained in recent proposals which succeeded in replacing the hardware-intensive components of traditional amplitude estimation – controlled multi-qubit gates and quantum Fourier transform – by classical post-processing [15–18]. Alternatively, one can systematically reduce the circuit depth by interpolating between classical MC methods and amplitude estimation [19]. Additionally, classical pre-processing can replace costly quantum arithmetic [20].

In this article, we address the question whether the quantum computational requirements for amplitude estimation can be further decreased by making use of variational quantum algorithms [21–23]. We present variational quantum amplitude estimation (VQAE) in which the depth of the entire quantum circuit is always kept below a desired maximum value by means of variational optimization. VQAE is based on maximum likelihood amplitude estimation (MLAE) [15]. We present a naïve and an adaptive VQAE algorithm. Adaptive VQAE rescales the amplitude to reduce the cost of the variational optimization. The advantage of VQAE over MLAE is that the maximum circuit depth of VQAE is independent of the total number of MLAE steps, whereas in MLAE this depth grows linearly with the number of MLAE steps. The advantage of VQAE over classical MC sampling is that VQAE can have a lower computational cost.
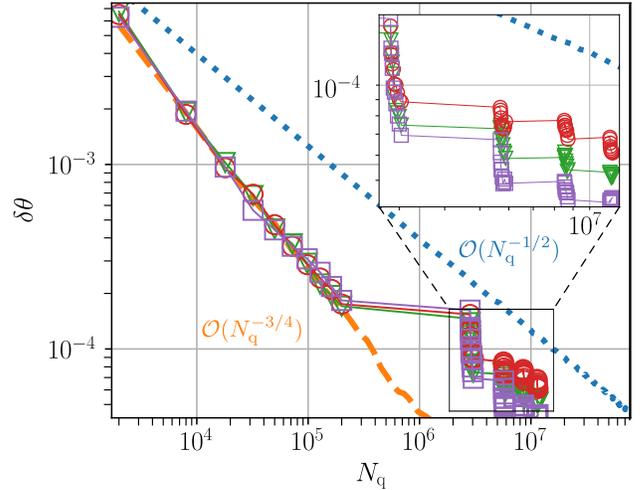

FIG. 1. Amplitude estimation error $\delta\theta$ as a function of the computational cost, i.e. the number of queries N_q. We compare adaptive VQAE (symbols with lines guide to the eye) to MLAE (dashed orange) and classical MC sampling (dotted blue). We calculate the rescaled mean value of a shifted Lorentzian (red circles), Gaussian (green triangles), and log-normal (purple squares) probability distribution. In VQAE, the first 10 amplitude estimates are computed via MLAE, then one step of the variational optimization is performed, which is followed by the next iteration of 10 MLAE steps. This procedure of variational approximation followed by MLAE is repeated three more times, resulting in a final error $\delta\theta \approx 6 \cdot 10^{-5}$. We see that, for this error, N_q is up to an order of magnitude smaller in VQAE than in classical MC sampling. Throughout this calculation, VQAE's circuit depth is the depth of the initial state plus at most 10 times the depth of the query operator, whereas MLAE's circuits have the depth of the initial state plus, at the end, 50 times the depth of the query operator.

Figure 1 shows that, for the problems considered here, VQAE outperforms classical MC sampling and additionally keeps the overall circuit depth below a fixed value.

This article is organized as follows. Firstly, in Section II, we define the problems considered here and explain the original quantum algorithm for amplitude esti-

* kirill.plekhanov@cambridgequantum.com

mation as well as the classical MC approach. Then, in Section III we present our variational methods, study variational errors of constant-depth quantum circuits, and develop naïve and adaptive VQAE. We conclude this article and discuss potential next steps in Section IV.

II. BACKGROUND

In this section, we first define the problem that we are interested in. Next, we explain quantum amplitude estimation and classical MC sampling.

A. Problem definition

Throughout this article, we focus on the calculation of expectation values

$$\mathbb{E}_p[f] = \sum_x p(x) f(x) \tag{1}$$

where the sum runs over 2^n equidistant values of $x \in [0, 1)$, $p(x)$ represents a probability distribution and $f(x)$ a real-valued function. Here n is the qubit count of the wave function that encodes $p(x)$ and $f(x)$ in its amplitudes. We consider three probability distributions: a Gaussian

$$p_{\text{G}}(x) = \frac{1}{\mathcal{N}_{\text{G}}} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right), \tag{2}$$

Cauchy-Lorentz

$$p_{\text{C-L}}(x) = \frac{1}{\mathcal{N}_{\text{C-L}}} \frac{\sigma}{(x - \mu)^2 + \sigma^2}, \tag{3}$$

and log-normal distribution

$$p_{\text{l-n}}(x) = \frac{1}{\mathcal{N}_{\text{l-n}}(c_0 + c_1 x)} \exp\left(-\frac{(\ln(c_0 + c_1 x) - \mu)^2}{2\sigma^2}\right). \tag{4}$$

The normalization constants $\mathcal{N}_{\text{G}}$, $\mathcal{N}_{\text{C-L}}$ and $\mathcal{N}_{\text{l-n}}$ are chosen so that $\sum_x p(x) = 1$.

We choose the following parameters for our analysis. We fix the total number of qubits encoding $f(x)$ and $p(x)$ to $n = 5$. In our calculations with the Gaussian and Cauchy-Lorentz distribution, we use $\mu = 0.5$ and $\sigma = 0.1$. In our calculations with the log-normal distribution, we use $c_0 = 0$, $c_1 = 10$, $\mu = 1.5$, and $\sigma = 0.2$. For the function $f(x)$, we use

$$f(x) = Cx \tag{5}$$

with some $C > 0$. For this choice of parameters, the expectation value (1) is approximately $\mathbb{E}_p[f] \approx 0.5\,C$ for all distributions.

B. Quantum amplitude estimation

Let us present a way to encode the solution to (1) on a quantum computer. We assume $f(x)$ and $p(x)$ are functions that map $[0, 1)$ to $[0, 1]$. We consider real numbers $x \in [0, 1)$ that satisfy

$$x = \sum_i \frac{x_i}{2^{n-i}}, \quad x_i \in \{0, 1\} \tag{6}$$

and that we identify with n-bit strings $\{x_i,\ i = 1, 2, \ldots, n\}$. Each bit string shall correspond to a quantum state $|x\rangle = |x_1,\ x_2,\ \ldots,\ x_n\rangle$ in the computational basis of a n-qubit register. Additionally we have a quantum circuit $\mathcal{A}$ that acts on a register of $n + 1$ qubits and produces a state $|\chi_0\rangle_{n+1} = \mathcal{A} |0\rangle_{n+1}$ such that

$$|\chi_0\rangle_{n+1} = \sqrt{1 - a} |\psi_{\text{bad}}\rangle_n |0\rangle + \sqrt{a} |\psi_{\text{good}}\rangle_n |1\rangle. \tag{7}$$

Here $|\psi_{\text{bad}}\rangle_n$ and $|\psi_{\text{good}}\rangle_n$ are two normalized quantum states of a n-qubit register which is connected to one additional ancilla qubit. We define the good state

$$|\psi_{\text{good}}\rangle_n = \frac{1}{\sqrt{a}} \sum_x \sqrt{p(x) f(x)} |x\rangle_n \tag{8}$$

so that $a = \mathbb{E}_p[f]$ of Eq. (1) coincides with the probability of measuring the ancilla qubit in the state $|1\rangle$.

To determine a, the amplitude estimation algorithm uses the Grover operator $\mathcal{Q} = -\mathcal{R}_\chi \mathcal{R}_{\text{good}}$ [1, 24, 25] where

$$\mathcal{R}_\chi = \mathbb{1} - 2 |\chi_0\rangle_{n+1} \langle\chi_0|_{n+1} = \mathbb{1} - 2\mathcal{A} |0\rangle_{n+1} \langle 0|_{n+1} \mathcal{A}^\dagger$$
$$\mathcal{R}_{\text{good}} = \mathbb{1} - 2 |\psi_{\text{good}}\rangle_n |1\rangle \langle\psi_{\text{good}}|_n \langle 1| \tag{9}$$

are reflections in a two-dimensional subspace $\mathcal{H}_\chi$ spanned by states $|\psi_{\text{bad}}\rangle_n |0\rangle$ and $|\psi_{\text{good}}\rangle_n |1\rangle$. We define $a = \sin^2(\theta)$ and explicitly write out the action of Q:

$$|\chi_m\rangle_{n+1} = Q^m |\chi_0\rangle_{n+1} = \cos[(2m + 1)\theta] |\psi_{\text{bad}}\rangle_n |0\rangle$$
$$+ \sin[(2m + 1)\theta] |\psi_{\text{good}}\rangle_n |1\rangle. \tag{10}$$

Therefore, the subspace $\mathcal{H}_\chi$ is stable under the action of $\mathcal{Q}$ and the only effect of $\mathcal{Q}$ is to rotate by an angle of 2θ. The original amplitude estimation algorithm then uses quantum phase estimation to find the eigenvalues of $\mathcal{Q}$ equal to $\exp(\pm 2i\theta)$ and provides an estimate of a with an error

$$\epsilon \leq 2\pi \frac{\sqrt{a(1 - a)}}{N_{\text{q}}} + \frac{\pi^2}{N_{\text{q}}^2}, \tag{11}$$

with a probability of at least $8/\pi^2$ [1].

Traditional amplitude estimation has high requirements on quantum hardware because it uses quantum phase estimation. This algorithm needs the quantum Fourier transform and multiple controlled $\mathcal{Q}^m$ operations where $\{m = 1, 2, 4, \ldots, 2^M\}$. The depth of the

corresponding quantum circuit is mostly determined by the depth of the last controlled $\mathcal{Q}^m$ operator for which $m = 2^M$. In general, the total circuit depth scales like the total number of queries $\mathcal{O}(N_q) \sim \mathcal{O}(1/\epsilon)$ inversely proportional to the desired error ϵ.

To avoid these deep quantum circuits, several recent articles propose new ways to carry out amplitude estimation circumventing quantum phase estimation [15, 17, 18] and circuits of depth $\mathcal{O}(1/\epsilon)$ [19]. One proposal is MLAE [15] in which one combines measurements of the states $|\chi_m\rangle_{n+1}$ with a maximum likelihood estimation of a. For an exponential schedule $\{m = 1, 2, 4, \ldots, 2^M\}$, this algorithm has the query cost $N_q \sim \mathcal{O}(1/\epsilon)$. A linear schedule $\{m = 1, 2, 3, \ldots, M\}$ increases the query cost to $N_q \sim \mathcal{O}(\epsilon^{-4/3})$. Note that in this case N_q scales quadratically with the maximum circuit depth M. Following the same idea of reducing the hardware requirements, the authors of Ref. [19] present two algorithms with computational cost $N_q \sim \mathcal{O}(1/\epsilon^{\beta+1})$ for quantum circuits of reduced depth $\mathcal{O}(1/\epsilon^{1-\beta})$. These algorithms are controlled by an external parameter β which allows one to interpolate between the quantum regime at $\beta = 0$ and the classical MC regime at $\beta = 1$.

C. Classical MC sampling

We perform classical MC sampling in the following way. We sample from the state $|\chi_0\rangle_{n+1}$ of Eq. (7) and measure the ancilla qubit. We compute a as the relative frequency of measuring the ancilla qubit in the state $|1\rangle$. This calculation of a has the error $\epsilon = \sqrt{a(1-a)/N_q}$ so that the total number of queries required for a certain error ϵ is $N_q \sim \mathcal{O}(1/\epsilon^2)$ [26].

Comparing this query cost with the previous ones, we find that traditional amplitude estimation as well as MLAE with exponential schedule achieve a quadratic quantum speedup over classical MC sampling. Both MLAE with linear schedule and the algorithms in [19] obtain a reduced quantum speedup.

Note that, throughout this article, the query complexity is defined in terms of $\mathcal{A}$ operators, with two applications of $\mathcal{A}$ required per application of $\mathcal{Q}$, see Eq. (9). Also, the depth of quantum circuits is measured in units of $\mathcal{A}$.

III. VARIATIONAL ALGORITHMS

Here we present our variational algorithms. We first explain the general VQAE formalism, then our naïve implementation, and finally the adaptive VQAE approach.

A. General formalism

The VQAE algorithm is based on the maximum likelihood framework of Ref. [15] with linearly incremental

sequence $\{m = 1, 2, 3, \ldots, M\}$. In this framework, the depth of the quantum circuit implementing the state $|\chi_m\rangle_{n+1} = \mathcal{Q}^m |\chi_0\rangle_{n+1}$ increases linearly with m. To prevent the circuit depth from increasing indefinitely, we add to this framework a variational step during which the quantum state $|\chi_m\rangle_{n+1}$ of depth $\mathcal{O}(m)$ is approximated by a quantum state of a constant depth $\mathcal{O}(1)$. We perform the variational approximation every k-th power of $\mathcal{Q}$, with $0 < k < M$. For all the other iterations, we simply apply the corresponding power of $\mathcal{Q}$ to the variational state. This results in Algorithm 1 presented below.

Algorithm 1 Variational quantum amplitude estimation

Require: $|\phi_{i=0}\rangle_{n+1} = |\chi_0\rangle_{n+1}$, $\mathcal{Q}$
  for $0 \leq i \leq \lfloor M/k \rfloor$ **do**
    for $0 \leq j < k$ **do**
      Sample the circuit $\mathcal{Q}^j |\phi_i\rangle_{n+1}$ and collect h samples. Save the number of times the ancilla qubit is $|1\rangle$ in a variable h_m, with $m = i \cdot k + j$
    end for
    if $i \neq \lfloor M/k \rfloor$ **then**
      Perform the variational approximation
$$|\phi_{i+1}\rangle_{n+1} \approx \mathcal{Q}^k |\phi_i\rangle_{n+1}$$
    end if
  end for
  Use $\{h_m\}$ to carry out the maximum likelihood estimation

---

The maximum likelihood post-processing [15] consists in maximizing the likelihood function $L(\{h_m\}, x) = \prod_m L_m(h_m, x)$ with

$$L_m(h_m, x) = [\sin^2((2m+1)x)]^{h_m} [\cos^2((2m+1)x)]^{h-h_m}, \tag{12}$$

so that the estimate of the phase θ becomes

$$\hat{\theta} = \arg\max_x \Big(\ln L(\{h_m\}, x) \Big). \tag{13}$$

Our implementations of the maximum likelihood estimation use $h = 2 \times 10^3$ samples. The minimization of $L(\{h_m\}, x)$ is accomplished by means of a brute-force search algorithm that uses 5×10^3 grid points.

We variationally approximate states $\mathcal{Q}^k |\phi_i\rangle_{n+1}$ by minimizing $|| |\phi_{\text{var}}(\boldsymbol{\lambda})\rangle_{n+1} - \mathcal{Q}^k |\phi_i\rangle_{n+1} ||^2$ which is equivalent to maximizing the objective function

$$\mathcal{F}(\boldsymbol{\lambda}) = \mathfrak{Re}\Big(\langle\phi_{\text{var}}(\boldsymbol{\lambda})|_{n+1} \mathcal{Q}^k |\phi_i\rangle_{n+1} \Big) \tag{14}$$

with respect to the variational parameters $\boldsymbol{\lambda}$. The depth of the quantum circuit for $\mathcal{F}(\boldsymbol{\lambda})$ is $2k + 2$. In general, the variational quantum state $|\phi_{\text{var}}(\boldsymbol{\lambda})\rangle_{n+1}$ is a parameterized quantum circuit (PQC)

$$|\phi_{\text{var}}(\boldsymbol{\lambda})\rangle_{n+1} = \mathcal{U}_{\text{var}}(\boldsymbol{\lambda}) |\phi_{\text{init}}\rangle_{n+1} = \prod_j e^{-i\lambda_j G_j} |\phi_{\text{init}}\rangle_{n+1}, \tag{15}$$

FIG. 2.   Variational PQC ansatz of depth d for 6 qubits. The variational parameters λ_j are in single-qubit rotation gates $R_y(\lambda_j) = \exp(-i\lambda_j\sigma_y/2)$ where σ_y denotes the y Pauli matrix. In naïve VQAE, $|\phi_0\rangle_{n+1} = |0\rangle_{n+1}$ and the circuit structure inside the dashed box is repeated d times. In adaptive VQAE, $|\phi_0\rangle_{n+1} = |\chi_0\rangle_{n+1}$ and only the dark blue gates with adjacent CNOTs form our variational ansatz.

where G_j are Hermitian operators acting on the $(n+1)$-qubit register and $|\phi_{\text{init}}\rangle_{n+1}$ is some initial state. For our purposes, we are interested in hardware-efficient quantum circuits that produce real-valued quantum states. We use the PQC shown in Fig. 2 that is composed of d layers with 15 parameterized single-qubit rotation gates and 10 CNOT gates per layer.

One single variational update of a PQC consists of n_s sweeps over all circuit parameters, during which all parameters are updated simultaneously. To perform the optimization, it is convenient to introduce a coordinate-wise version of Eq. (14) for the j-th parameter

$$f_j(x) = \mathcal{F}(\lambda_1, \ \lambda_2, \ \ldots, \ \lambda_{j-1}, \ x, \ \lambda_{j+1}, \ \ldots). \quad (16)$$

The optimization of the parameterized state in Eq. (15) can then be performed via a particle swarm approach [27, 28], the coordinate-wise update [29–33], or gradient based methods with the parameter-shift rule [34–39]

$$\frac{\mathrm{d}f_j(\lambda_j)}{\mathrm{d}\lambda_j} = f_j(\lambda_j + \pi/4) - f_j(\lambda_j - \pi/4). \quad (17)$$

We obtained the best results using the gradient based approach with the Adam optimizer [40]. Therefore this technique is being used throughout this article for the computation of all results. Each gradient calculation requires two evaluations of the coordinate-wise objective function $f_j(\lambda_j \pm \pi/4)$. On a quantum computer, f_j can be determined via the Hadamard test [41]. In our numerical simulations, we emulate the measurement of the Hadamard circuit by first evaluating the exact value of f_j and then sampling it using a binomial distribution with the probability $(1 + f_j)/2$ and n_f independent Bernoulli trials [26].

The variational approximation step significantly affects the total number of queries N_q used by VQAE. In MLAE with a linearly incremental sequence, the total number of queries is equal to

$$N_q = \sum_{m=1}^{M} h(2m + 1) = hM(M + 2), \quad (18)$$

where $2m + 1$ is the depth of the quantum circuit encoding $|\chi_m\rangle_{n+1} = \mathcal{Q}^m |\chi_0\rangle_{n+1}$. In VQAE, the total number of queries is composed of two separate contributions. The first one accounts for the sampling of the quantum circuits $|\chi_m\rangle_{n+1}$ and we denote it by N_{samp}. The second contribution corresponds to the variational approximation cost, which we denote by N_{var}. We assume that the number of queries required per variational approximation is independent of the iteration number m and changes only as a function of the desired variational error as well as the depth of the circuit for the objective function, which is equal to $2k + 2$. For convenience, we denote the cost of a single variational update for $k = 1$ by $N_{\text{var}/1}$, so that the total number of variational queries reads

$$N_{\text{var}} = N_{\text{var}/1}(2k + 2)\lfloor M/k \rfloor \sim \mathcal{O}(k\lfloor M/k \rfloor). \quad (19)$$

The number of sampling queries is equal to

$$N_{\text{samp}} = hk(k + 2)\lfloor M/k \rfloor + h(M\%k)(M\%k + 2), \quad (20)$$

where $\lfloor \cdot \rfloor$ denotes the floor function and $\%$ is the modulo operation. In the limit $M \gg k$, $N_{\text{var}} \sim \mathcal{O}(M)$ and $N_{\text{samp}} \sim \mathcal{O}(kM)$. Note that both contributions scale like $\mathcal{O}(M)$ which is quadratically better than the scaling $\mathcal{O}(M^2)$ of MLAE in Eq. (18).

B.   Naïve VQAE

In our naïve implementation of the VQAE algorithm, the initial state of the PQC in Eq. (15) and Fig. 2 is $|\phi_{\text{init}}\rangle_{n+1} = |0\rangle_{n+1}$.

Let us first explore the expressive power of the corresponding variational state $|\phi_{\text{var}}(\boldsymbol{\lambda})\rangle$. To this end, we perform amplitude amplifications followed by variational approximations of the resulting state with $k = 1$ and $M = 50$. To evaluate the quality of the variational approximation, we calculate the infidelity

$$\mathcal{I}_m = 1 - \langle\phi_i|_{n+1} Q^j |\chi_m\rangle_{n+1}, \ m = i \cdot k + j, \quad (21)$$
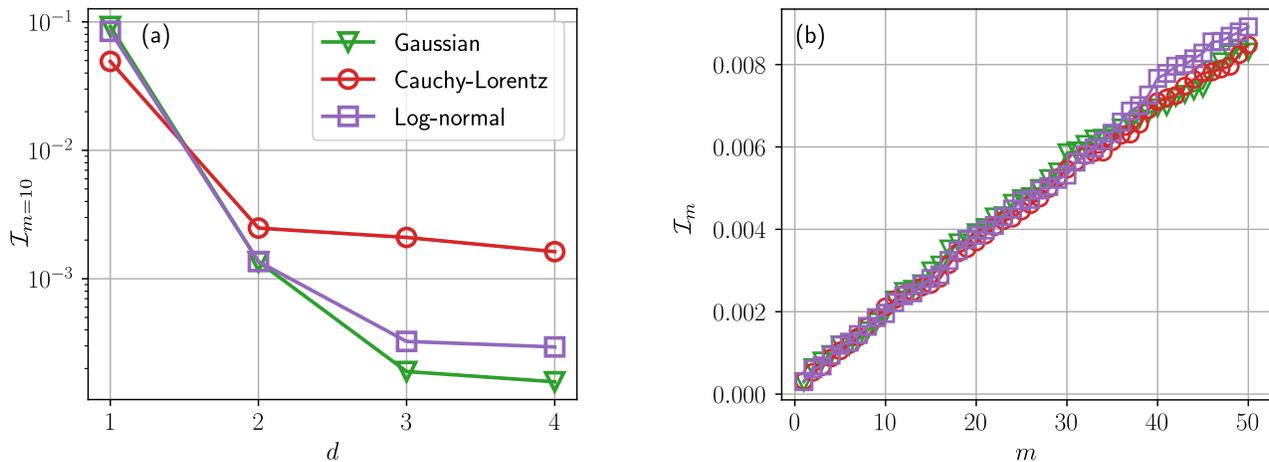
FIG. 3. Infidelity $\mathcal{I}_m$ in Eq. (21) for the probability distributions of Eqs. (2)-(4) (see legend) shown (a) as a function of d for $m = 10$ and (b) as a function of m for $d = 4$. We see that the infidelity decreases with increasing d, due to the corresponding increase of the expressive power of the PQC. The infidelity increases linearly as a function of m slowly with a slope of ≈ 0.00017 that is approximately the same for the three distributions. These results are obtained via naïve VQAE with $k = 1$, $M = 50$, and $n_s = 1000$, using the numerically exact gradient without sampling and Adam with the initial learning rate $\beta = 0.1$. We consider 100 randomly initialized PQC and (a) shows one example calculation and (b) the average over all 100 calculations.

where for $k = 1$ we have $j = 0$ and $i = m$. Figure 3(a) shows the results of such calculations performed for different depths d of the PQC for $m = 10$ and Fig. 3(b) shows the infidelity as a function of m for $d = 4$. We observe that the accuracy of the variational ansatz increases with the depth and saturates at $d \approx 4$. The infidelity increases linearly with m. This behaviour is seen for all probability distributions considered.

Next, we present the amplitude estimation results of naïve VQAE. Figure 4 shows the convergence of $\delta\theta$ as a function of N_q, under the assumption that $N_{\text{var}/1} = 0$ and $k = 1$. The resulting error is compared with the one of classical MC sampling which scales like $\delta\theta \sim \mathcal{O}(N_q^{-1/2})$ and the one of MLAE which scales like $\delta\theta \sim \mathcal{O}(N_q^{-3/4})$. Interestingly, we find that the convergence of $\delta\theta$ changes as a function of M. For small values of M, it follows the ideal VQAE scaling $\delta\theta \sim \mathcal{O}(N_q^{-3/2})$ as if the variational approximation is performed without error. We emphasize that this scaling is cubically better than the one of classical MC sampling. The second convergence regime is observed for larger values of M. In this regime, the error follows the MC scaling with $\delta\theta \sim \mathcal{O}(N_q^{-1/2})$. One can understand this new scaling in terms of a one-dimensional classical random walk with an average step of length 2θ and some finite variance induced by variational approximations. Provided that the variance is much smaller than θ, after $M \sim N_q$ steps the total distance of the random walk converges to the value $2M\theta$ with a relative error scaling as $\mathcal{O}(M^{-1/2})$.

Finally, we take into account the cost of the variational approximation, to obtain a more complete assessment of the algorithmic performance of naïve VQAE. To estimate the cost of a single variational update, we use that $N_{\text{var}/1} = 2n_f n_s n_p$, where n_p is the number of pa-

rameters of a PQC and the factor 2 comes from the two evaluations of the objective function in Eq. (17). For the PQC in Fig. 2 with $d = 4$, the number of parameters is $n_p = 60$. Additionally, we choose $n_f \sim n_p \sim 100$ so that $N_{\text{var}/1} \sim 1.2 \times 10^6$ and $N_{\text{var}} \sim 4.8 \times 10^6 M$. This large variational cost is the dominant part in the calculation of the total number of queries N_q. Ultimately, it leads to a performance of naïve VQAE that is worse than the one of classical MC sampling. Reducing any of n_f, n_s, or n_p decreases the variational cost but also increases the variational error which then leads to a worse final amplitude estimation error.

C. Adaptive VQAE

To reduce the variational cost of VQAE, in this section we present the adaptive VQAE algorithm. This algorithm makes use of adaptive rescaling of the function f, based on importance sampling by scaling, which is well known in statistics and commonly used in MC simulations [42].

To introduce the adaptive VQAE algorithm, we first note that the amplitude $a = \mathbb{E}_p[f]$ is linear in f, meaning that rescaling the function f with a proportionality constant r also rescales the amplitude a:

$$a' = \mathbb{E}_p[f'] = \mathbb{E}_p[rf] = ra. \tag{22}$$

The rescaled function f' can then be used to encode a new quantum state $|\chi'_0\rangle_{n+1}$, provided that $f'(x) \leq 1$ for all x, which is required for the successful state preparation via Eq. (7). Next, a new Grover operator $\mathcal{Q}'$ can be defined and amplitude estimation can be performed to estimate the amplitude a'. To proceed further, we make
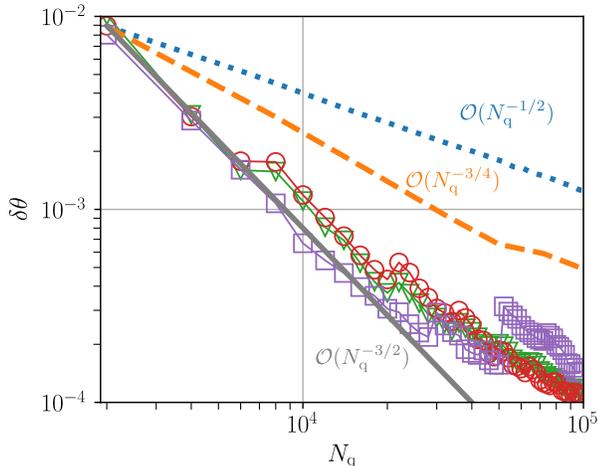
FIG. 4. Amplitude estimation error $\delta\theta$ as a function of the number of queries N_q obtained using naïve VQAE with $k = 1$ under the assumption of zero variational cost $N_{var/1} = 0$. We observe that for small M, the error follows the ideal VQAE scaling $\delta\theta \sim \mathcal{O}(N_q^{-3/2})$ (solid gray line). For larger values of M, the scaling changes to the MC scaling $\delta\theta \sim \mathcal{O}(N_q^{-1/2})$ (dotted blue line). The result is also compared to the MLAE scaling $\delta\theta \sim \mathcal{O}(N_q^{-3/4})$ (dashed orange line). The legend as well as the simulation parameters are the same as in Fig. 3.

the observation that after k iterations, the overlap between the states $|\chi'_0\rangle_{n+1}$ and $|\chi'_k\rangle_{n+1} = \mathcal{Q}'^k |\chi'_0\rangle_{n+1}$ is equal to $\cos(2k\theta')$. Hence, one can always find a commensurate amplitude $0 \leq a' \leq 1$ and a phase θ' so that

$$a' = \sin^2(\theta'), \quad \theta' = \pi l/k, \quad l \in \mathbb{Z}, \tag{23}$$

and the overlap between the corresponding states $|\chi'_0\rangle_{n+1}$ and $|\chi'_k\rangle_{n+1}$ is equal to one. Combining these two observations, we conclude that for a proper choice of the renormalization factor r satisfying $rf(x) \leq 1$ for all x, it is possible to rescale the function f such that the k-th power of the corresponding Grover operator becomes the identity operator $\mathcal{Q}'^k = \mathcal{I}$.

Finding the exact renormalization factor r requires exact knowledge of the initial amplitude a. However, as we show in the following, a loose estimate a_l, obtained from a moderate number of MC samples of the initial state $|\chi_0\rangle_{n+1}$, is sufficient. Assuming that such a loose amplitude estimate is provided, the renormalization factor can then be approximately expressed as $r_l = a'/a_l$, with a' defined as in Eq. (23). As a result, the overlap between the states $|\chi'_0\rangle_{n+1}$ and $|\chi'_k\rangle_{n+1}$ becomes

$$\langle\chi'_0|_{n+1} |\chi'_k\rangle_{n+1} = \cos(\theta' - k\delta\theta') \approx 1 - \frac{(k\delta\theta')^2}{2}, \tag{24}$$

where $\delta\theta' = \theta' - \arcsin\sqrt{r_l a_l}$. We see that for small deviations, the overlap decreases quadratically with k.

Next, we use the VQAE algorithm to estimate the amplitude a' by means of the Grover operator $\mathcal{Q}'$ and

the initial state $|\chi'_0\rangle_{n+1}$. The variational approximation is performed at every k-th step, when the overlap between $|\chi'_0\rangle_{n+1}$ and $|\chi'_k\rangle_{n+1}$ reaches maximum. Additionally, we assume that the PQC has the initial state $|\phi_{init}\rangle_{n+1} = |\chi'_0\rangle_{n+1}$ so that the variational quantum state of Eq. (15) reads

$$|\phi_{var}(\boldsymbol{\lambda})\rangle_{n+1} = \mathcal{U}_{var}(\boldsymbol{\lambda}) |\chi'_0\rangle_{n+1}. \tag{25}$$

Having the PQC initialized to identity at the beginning of each optimization step, the only role of the variational quantum circuit is to correct the deviation of Eq. (24) originating from an imprecise value of the renormalization constant r_l. Finally, at the end of the calculation, an estimation of the phase $\hat{\theta}'$ corresponding to the renormalized amplitude $\hat{a}' = \sin^2 \hat{\theta}'$ is obtained. To go back to the original formulation of the problem and compare the results, we use the inverse transformation

$$\hat{\theta} = \arcsin\sqrt{\hat{a}}, \quad \hat{a} = \hat{a}'/r_l, \tag{26}$$

where the renormalization constant r_l has to be exactly the same as the one used for the function rescaling. This last step concludes the adaptive VQAE algorithm.

We analyze the performance of the adaptive VQAE algorithm with a simplified variational ansatz consisting of only six single-qubit rotation gates and four CNOT gates, as shown in Fig. 2 in dark blue color. This simplified ansatz has only six parameters in total, which significantly reduces the number of variational queries as well as the effects of the noise due to finite sampling. We determine the loose estimate of the amplitude a_l via 5×10^5 MC samples. As a result, much smaller values of infidelity are achieved for n_f being an order of magnitude smaller than in our naïve VQAE computations. We also note that for smaller values of a, the initial MC estimation of a' gets worse and, as a consequence, more sweeps are required to ensure the convergence of the variational ansatz.

Our results for adaptive VQAE are presented in Fig. 1, where we show the convergence of $\delta\theta$ as a function of N_q for $k = 10$. The simulations use Adam with the initial learning rate $\beta = 10^{-3}$, $n_f = 100$, $n_s = 100$, and $n_p = 6$, resulting in $N_{var/1} = 2n_f n_s n_p = 1.2 \times 10^5$. As in Fig. 4, we compare with the classical MC scaling $\delta\theta \sim \mathcal{O}(N_q^{-1/2})$ and the MLAE scaling $\delta\theta \sim \mathcal{O}(N_q^{-3/4})$. The major difference of the adaptive VQAE as compared to all previously studied methods is a large starting cost which corresponds to the amount of MC samples required for the evaluation of a_l. This starting cost, however, represents only an additive contribution to N_q and, hence, is insignificant in the regime of our interest when the number of queries gets large. Additionally, we find that, thanks to a significant improvement of the number of query calls and the overall precision of the variational state, the resulting final error $\delta\theta$ of the adaptive VQAE algorithm surpasses the classical MC error, therefore leading to a genuine quantum advantage.

Interestingly, we observe that in the regime of small k, the performance of the adaptive VQAE algorithm decreases. This has several reasons. Firstly, the precision of the maximum likelihood estimation decreases when the angle $\theta' = \pi l/k$ (where $l = 1$ in our case) becomes larger than $\pi/4$, *i.e.* for $k \leq 4$. Hence, to perform an estimation with such small values of k, a different statistical inference technique has to be considered. Secondly, for small values of k, the rescaling factor can become much larger than one and then leads to more efficient classical MC sampling. In calculations with $C = 0.1$, this leads to a loss of quantum advantage for $k = 5$, corresponding to $r \approx 7.508$. The quantum advantage of the adaptive VQAE algorithm holds true for other cases with $k > 5$.

IV. DISCUSSION

In this article, we provide numerical evidence that variational quantum algorithms and constant-depth quantum circuits can lead to a quantum advantage over classical MC sampling in the context of amplitude estimation. Our results are proof of concept that it is feasible, in principle, for current quantum devices to achieve a quantum advantage by performing amplitude estimation. The quantum circuits used for our numerical demonstrations, however, are still challenging for this generation of gate-based quantum computers. Therefore, an exciting next step is to find other problems and applications for which VQAE has low quantum hardware requirements and can be realized on actual quantum devices.

We can imagine future applications for VQAE in several areas, including combinatorial optimization, quantum machine learning, and quantum chemistry. In the context of combinatorial optimization, VQAE enables

us to use constant-depth quantum circuits to carry out Grover search, which can find the optimal solution with a quadratic quantum speedup over brute-force search. Here it is also enticing to study whether such a variational Grover search algorithm can benefit from filtering operators [43]. In relation to quantum machine learning, VQAE has the potential to make it possible for current quantum devices to accelerate inference in Bayesian networks [44], which can then be compared with state-of-the-art variational quantum algorithms for inference [45]. With regards to quantum chemistry, the concept of VQAE can be combined with variational quantum phase estimation (VQPE) [46–48] to realize VQPE with shallow circuits on actual quantum hardware. In this context, one interesting application is to use accurate quantum chemistry results obtained with a quantum computer to train an ansatz for the exchange-correlation energy in density functional theory by means of machine learning [49–51].

We anticipate that the efficiency of the VQAE algorithm can be further increased. Firstly, it would be interesting to analyse whether a local cost function – which can help mitigate the negative effect of barren plateaus [52] – improves the variational optimization and reduces the required number of variational queries. Secondly, the performance of our variational algorithm crucially depends on the maximum likelihood estimation procedure. It would be interesting to investigate whether alternative approaches perform better, e.g. iterative QAE [18] or QoPrime AE [19].

[1] G. Brassard, P. Hoyer, M. Mosca, and A. Tapp, Quantum amplitude amplification and estimation, Contemp. Math. **305**, 53 (2002).

[2] A. Montanaro, Quantum speedup of Monte Carlo methods, Proc. Math. Phys. Eng. Sci. **471** (2015).

[3] E. Knill, G. Ortiz, and R. D. Somma, Optimal quantum measurements of expectation values of observables, Phys. Rev. A **75**, 012328 (2007).

[4] I. Kassal, S. P. Jordan, P. J. Love, M. Mohseni, and A. Aspuru-Guzik, Polynomial-time quantum algorithm for the simulation of chemical dynamics, Proc. Natl. Acad. Sci. USA **105**, 18681 (2008).

[5] N. Wiebe, A. Kapoor, and K. M. Svore, Quantum Algorithms for Nearest-Neighbor Methods for Supervised and Unsupervised Learning, Quantum Info. Comput. **15**, 316 (2015).

[6] N. Wiebe, A. Kapoor, and K. M. Svore, Quantum Deep Learning, Quantum Info. Comput. **16**, 541 (2016).

[7] I. Kerenidis, J. Landman, A. Luongo, and A. Prakash, q-means: A quantum algorithm for unsupervised machine learning, Adv. Neural Inf. Process. Syst. **32** (2019).

[8] R. Orús, S. Mugel, and E. Lizaso, Quantum computing for finance: Overview and prospects, Rev. Phys. **4**, 100028 (2019).

[9] 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 Trans. Quantum Eng. **1**, 1 (2020).

[10] A. Bouland, W. van Dam, H. Joorati, I. Kerenidis, and A. Prakash, Prospects and challenges of quantum finance, arXiv:2011.06492 (2020).

[11] S. Woerner and D. J. Egger, Quantum risk analysis, npj Quantum Inf. **5**, 15 (2019).

[12] M. C. Braun, T. Decker, N. Hegemann, S. F. Kerstan, and C. Schäfer, A Quantum Algorithm for the Sensitivity Analysis of Business Risks, arXiv:2103.05475 (2021).

[13] P. Rebentrost, B. Gupt, and T. R. Bromley, Quantum computational finance: Monte Carlo pricing of financial derivatives, Phys. Rev. A **98**, 022321 (2018).

[14] N. Stamatopoulos, D. J. Egger, Y. Sun, C. Zoufal, R. Iten, N. Shen, and S. Woerner, Option Pricing using Quantum Computers, Quantum **4**, 291 (2020).

[15] Y. Suzuki, S. Uno, R. Raymond, T. Tanaka, T. Onodera, and N. Yamamoto, Amplitude estimation without phase estimation, Quantum Inf. Process. **19**, 75 (2020).

[16] T. Tanaka, Y. Suzuki, S. Uno, R. Raymond, T. Onodera, and N. Yamamoto, Amplitude estimation via maximum likelihood on noisy quantum computer, arXiv:2006.16223 (2020).

[17] S. Aaronson and P. Rall, Quantum approximate counting, simplified, in *Symposium on Simplicity in Algorithms* (SIAM, 2020) pp. 24–32.

[18] D. Grinko, J. Gacon, C. Zoufal, and S. Woerner, Iterative quantum amplitude estimation, npj Quantum Inf. **7**, 1 (2021).

[19] T. Giurgica-Tiron, I. Kerenidis, F. Labib, A. Prakash, and W. Zeng, Low depth algorithms for quantum amplitude estimation, arXiv:2012.03348 (2020).

[20] S. Herbert, Quantum Monte-Carlo Integration: The Full Advantage in Minimal Circuit Depth, arXiv:2105.09100 (2021).

[21] M. Benedetti, E. Lloyd, S. Sack, and M. Fiorentini, Parameterized quantum circuits as machine learning models, Quantum Sci. Technol. **4**, 043001 (2019).

[22] M. Cerezo, A. Arrasmith, R. Babbush, S. C. Benjamin, S. Endo, K. Fujii, J. R. McClean, K. Mitarai, X. Yuan, L. Cincio, and P. J. Coles, Variational Quantum Algorithms, arXiv:2012.09265 (2020).

[23] K. Bharti, A. Cervera-Lierta, T. H. Kyaw, T. Haug, S. Alperin-Lea, A. Anand, M. Degroote, H. Heimonen, J. S. Kottmann, T. Menke, W.-K. Mok, S. Sim, L.-C. Kwek, and A. Aspuru-Guzik, Noisy intermediate-scale quantum (NISQ) algorithms, arXiv:2101.08448 (2021).

[24] G. Brassard and P. Hoyer, An exact quantum polynomial-time algorithm for Simon's problem, in *Proceedings of the Fifth Israeli Symposium on Theory of Computing and Systems* (IEEE, 1997) pp. 12–23.

[25] L. K. Grover, Quantum computers can search rapidly by using almost any transformation, Phys. Rev. Lett. **80**, 4329 (1998).

[26] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C* (Cambridge University Press, New York, USA, 1992).

[27] J. Kennedy and R. Eberhart, Particle swarm optimization, in *Proceedings of ICNN'95 - International Conference on Neural Networks*, Vol. 4 (1995) pp. 1942–1948.

[28] M. R. Bonyadi and Z. Michalewicz, Particle Swarm Optimization for Single Objective Continuous Space Problems: A Review, Evol. Comput. **25**, 1 (2017).

[29] J. Gil Vidal and D. O. Theis, Calculus on parameterized quantum circuits, arXiv:1812.06323 (2018).

[30] R. M. Parrish, J. T. Iosue, A. Ozaeta, and P. L. McMahon, A Jacobi Diagonalization and Anderson Acceleration Algorithm For Variational Quantum Algorithm Parameter Optimization, arXiv:1904.03206 (2019).

[31] K. M. Nakanishi, K. Fujii, and S. Todo, Sequential minimal optimization for quantum-classical hybrid algorithms, Phys. Rev. Research **2**, 043158 (2020).

[32] M. Ostaszewski, E. Grant, and M. Benedetti, Structure optimization for parameterized quantum circuits, Quantum **5**, 391 (2021).

[33] M. Benedetti, M. Fiorentini, and M. Lubasch, Hardware-efficient variational quantum algorithms for time evolu-

tion, Phys. Rev. Research **3**, 033083 (2021).

[34] E. Hazan, K. Y. Levy, and S. Shalev-Shwartz, Beyond convexity: Stochastic quasi-convex optimization, arXiv:1507.02030 (2015).

[35] Y. Suzuki, H. Yano, R. Raymond, and N. Yamamoto, Normalized Gradient Descent for Variational Quantum Algorithms, arXiv:2106.10981 (2021).

[36] J. Li, X. Yang, X. Peng, and C.-P. Sun, Hybrid Quantum-Classical Approach to Quantum Optimal Control, Phys. Rev. Lett. **118**, 150503 (2017).

[37] K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii, Quantum circuit learning, Phys. Rev. A **98**, 032309 (2018).

[38] M. Schuld, V. Bergholm, C. Gogolin, J. Izaac, and N. Killoran, Evaluating analytic gradients on quantum hardware, Phys. Rev. A **99**, 032331 (2019).

[39] L. Banchi and G. E. Crooks, Measuring analytic gradients of general quantum evolution with the stochastic parameter shift rule, Quantum **5**, 386 (2021).

[40] D. P. Kingma and J. Ba, Adam: A Method for Stochastic Optimization, arXiv:1412.6980 (2014).

[41] D. Aharonov, V. Jones, and Z. Landau, A polynomial quantum algorithm for approximating the Jones polynomial, Algorithmica **55**, 395 (2009).

[42] B. Arouna, Adaptive Monte Carlo Method, A Variance Reduction Technique, Monte Carlo Methods Appl. **10**, 1 (2004).

[43] D. Amaro, C. Modica, M. Rosenkranz, M. Fiorentini, M. Benedetti, and M. Lubasch, Filtering variational quantum algorithms for combinatorial optimization, arXiv:2106.10055 (2021).

[44] G. H. Low, T. J. Yoder, and I. L. Chuang, Quantum inference on Bayesian networks, Phys. Rev. A **89**, 062315 (2014).

[45] M. Benedetti, B. Coyle, M. Fiorentini, M. Lubasch, and M. Rosenkranz, Variational inference with a quantum computer, arXiv:2103.06720 (2021).

[46] R. M. Parrish and P. L. McMahon, Quantum Filter Diagonalization: Quantum Eigendecomposition without Full Quantum Phase Estimation, arXiv:1909.08925 (2019).

[47] N. H. Stair, R. Huang, and F. A. Evangelista, A Multireference Quantum Krylov Algorithm for Strongly Correlated Electrons, J. Chem. Theory Comput. **16**, 2236 (2020).

[48] K. Klymko, C. Mejuto-Zaera, S. J. Cotton, F. Wudarski, M. Urbanek, D. Hait, M. Head-Gordon, K. B. Whaley, J. Moussa, N. Wiebe, W. A. de Jong, and N. M. Tubman, Real time evolution for ultracompact Hamiltonian eigenstates on quantum hardware, arXiv:2103.08563 (2021).

[49] J. C. Snyder, M. Rupp, K. Hansen, K.-R. Müller, and K. Burke, Finding Density Functionals with Machine Learning, Phys. Rev. Lett. **108**, 253002 (2012).

[50] M. Lubasch, J. I. Fuks, H. Appel, A. Rubio, J. I. Cirac, and M.-C. Bañuls, Systematic construction of density functionals based on matrix product state computations, New J. Phys. **18**, 083039 (2016).

[51] S. Dick and M. Fernandez-Serra, Machine learning accurate exchange and correlation functionals of the electronic density, Nat. Commun. **11**, 3509 (2020).

[52] M. Cerezo, A. Sone, T. Volkoff, L. Cincio, and P. J. Coles, Cost function dependent barren plateaus in shallow parameterized quantum circuits, Nat. Commun. **12**, 1791 (2021).