Harnessing Bayesian Statistics to Accelerate Iterative Quantum Amplitude Estimation

Qilin Li[*1], Atharva Vidwans[*2], Yazhen Wang[1], and Micheline B. Soley[†2,3,4]

[1]Department of Statistics, University of Wisconsin-Madison, 1205 University Ave., Madison, WI 53706, USA
[2]Department of Chemistry, University of Wisconsin-Madison, 1101 University Ave., Madison, WI 53706, USA
[3]Department of Physics, University of Wisconsin-Madison, 1150 University Ave., Madison, WI 53706, USA
[4]Data Science Institute, University of Wisconsin-Madison, 1205 University Ave., Madison, WI 53706, USA

arXiv:2507.23074v2 [quant-ph] 7 Jan 2026

We establish a unified statistical framework that underscores the crucial role statistical inference plays in Quantum Amplitude Estimation (QAE), a task essential to fields ranging from chemistry to finance and machine learning. We use this framework to harness Bayesian statistics for improved measurement efficiency with rigorous interval estimates at all iterations of Iterative Quantum Amplitude Estimation. We demonstrate the resulting method, Bayesian Iterative Quantum Amplitude Estimation (BIQAE), accurately and efficiently estimates both quantum amplitudes and molecular ground-state energies to high accuracy, and show in analytic and numerical sample complexity analyses that BIQAE outperforms all other QAE approaches considered. Both rigorous mathematical proofs and numerical simulations conclusively indicate Bayesian statistics is the source of this advantage, a finding that invites further inquiry into the power of statistics to expedite the search for quantum utility.

1 Introduction

Reduction of the measurement cost of quantum amplitude estimation is essential to accelerate the utility of quantum computers. The inherently probabilistic nature of quantum computing frequently entails a large number of quantum circuit measurements to definitively determine the proportion of a quantum state in a specific target state [1–13], and finding such a proportion is inherent to fundamental mathematical tasks on quantum computers, from integration to optimization [14–16]. These tasks in turn play a foundational role in fields ranging from chemistry [17–21] to particle physics [22–28], condensed matter [29–33] to fluid mechanics [34–37], and finance [38–50] to machine learning [51, 52]. It is therefore expected that advancements to QAE would have broad impact.

A Bayesian approach that harnesses past information to predict future outcomes is emerging as a force in quantum algorithm development. Bayesian variants of standard frequentist approaches to quantum computing have been proposed in a wide array of quantum computing algorithms, including QAE [53–55], the quantum approximate optimization algorithm [56–61], phase estimation [62–75, 72], phase difference estimation [76–78], the variational quantum eigensolver [79–83, 20, 84–89, 61], and quantum annealing [90, 91]. In QAE specifically, Bayesian statistics has been found to be associated with key developments in driving down computational costs [53–55]. The Bayesian approach of Robust Amplitude Estimation (RAE), which employs fixed-length quantum circuits with Bayesian updated parameters, has been found to effectively reduce the impact of so-called estimation dead spots [53]. Likewise, the Bayesian approach of Bayesian Amplitude Estimation (BAE), which employs adaptive circuits with a Bayesian-informed scheduler, has successfully yielded the lowest number of oracle calls to reach a given target accuracy (quantum sample complexity) to date [55].

However, the impact of Bayesian statistics itself in QAE measurement cost reduction has his-

[*] These authors contributed equally to this work.
[†] Corresponding Author: msoley@wisc.edu

torically been difficult to isolate due to changes in the underlying QAE implementation required to inject Bayesian statistics in a noisy quantum environment. Implementation of Bayesian enhanced sampling in RAE coincides with abbreviation of the quantum circuit to avoid quantum decoherence effects [53], and similarly BAE fuses a Bayesian approach with a dynamic grid expansion scheduling strategy [55]. There is therefore a pressing need for a controlled experiment based on a clear-cut QAE algorithm capable of remaining unchanged except for the injection of Bayesian statistics in order to showcase the true power of what Bayesian inference offers QAE.

We aim to clearly demonstrate the direct impact that Bayesian inference offers to reduce measurement cost in QAE to provide a proof of the quantum speedup offered by Bayesian statistics and to provide conclusive numerical evidence of such a speedup. To do so, we establish a statistical framework for three key QAE methods—Classical Quantum Amplitude Estimation (Classical QAE) [1], Amplified Amplitude Estimation (AAE) [92], and Iterative Quantum Amplitude Estimation (IQAE) [93]—in the statistical perspective of asymptotic mean squared error (AMSE) and the normal approximation, and we use the resulting framework to introduce "Bayesian Iterative Quantum Amplitude Estimation (BIQAE)" based on the injection of Bayesian statistics into IQAE. We show the resulting statistical framework quantifies a double-digit percentage speedup of BIQAE relative to IQAE that can be completely attributed to Bayesian inference with rigorous mathematical proofs, and that BIQAE outperforms all other QAE methods considered—including benchmark IQAE and state-of-the-art BAE—over many orders of magnitude in the target accuracy for estimation of arbitrary quantum amplitudes. BIQAE also lowers the quantum sample complexity of ground-state energy approximation relative to benchmark IQAE for all molecular test cases considered over a broad range of equilibrium and nonequilibrium bond distances. These findings set the stage for development of further schemes to capitalize on Bayesian inference in QAE and to develop new approaches to reduce measurement cost in quantum mechanics with a Bayesian rather than a frequentist perspective.

2 Statistical Framework for QAE

2.1 Preliminaries

We define the QAE problem in terms of the action of a quantum oracle $\mathcal{A}$ on an $(n+1)$-qubit zero state $|0\rangle_n |0\rangle$. The resulting output state may be expressed in terms of two normalized states $|\psi_0\rangle_n$ and $|\psi_1\rangle_n$, which need not be orthogonal, as

$$\mathcal{A}|0\rangle_n |0\rangle = \sqrt{1-a}\,|\psi_0\rangle_n |0\rangle + \sqrt{a}\,|\psi_1\rangle_n |1\rangle\,,$$

where the oracle $\mathcal{A}$ may be chosen freely without loss of generality with the states $|\psi_0\rangle_n$ and $|\psi_1\rangle_n$ defined accordingly. The objective of QAE is then to estimate the amplitude $a \in [0,1]$ of the quantum state $|\psi_1\rangle_n |1\rangle$.

As a benchmark, Classical QAE (or the "standard-sampling" approach) approximates the quantum amplitude a as the sample mean of n measurement outcomes of the quantum circuit associated with oracle $\mathcal{A}$, where the ancilla states $|0\rangle$ and $|1\rangle$ are mapped to the outcomes 0 and 1, respectively. Where the *quantum sample complexity* is defined by the total number of accesses to the oracle, the quantum sample complexity required to achieve a target accuracy of ε with Classical QAE follows as on the order of $\mathcal{O}(1/\varepsilon^2)$ [1]. Restatement of the sample complexity in terms of asymptotic mean-square error (AMSE), as derived in Appendix A.1, then yields a sample complexity required to achieve a mean squared error (MSE) ε^2 of

$$N_{\text{oracle}} = \frac{1}{\varepsilon^2} a(1-a). \tag{1}$$

The relatively high measurement cost of Classical QAE is addressed via amplified estimation (or the "enhanced-sampling" approach), which yields quadratic improvement in the sample complexity [1]. Let the amplitude a be parametrized as

$$a = \sin^2\theta,$$

where $\theta \in [0, \frac{\pi}{2}]$, and let the Grover operator $\mathcal{Q}$ be defined as

$$\mathcal{Q} \coloneqq \mathcal{A}\mathcal{S}_0\mathcal{A}^\dagger\mathcal{S}_{\psi_0},$$

where $\mathcal{S}_0 \coloneqq \mathbb{I}_{n+1} - 2|0\rangle_{n+1}\langle 0|_{n+1}$ and $\mathcal{S}_{\psi_0} \coloneqq \mathbb{I}_{n+1} - 2|\psi_0\rangle_n \langle\psi_0|_n \otimes |0\rangle\langle 0|$. Repeated application of the Grover operator $\mathcal{Q}$ k times then results in the amplified state [1]

$$\begin{aligned}\mathcal{Q}^k\mathcal{A}|0\rangle_n |0\rangle = &\cos\left((2k+1)\theta\right)|\psi_0\rangle_n |0\rangle \\ &+ \sin\left((2k+1)\theta\right)|\psi_1\rangle_n |1\rangle\,,\end{aligned} \tag{2}$$

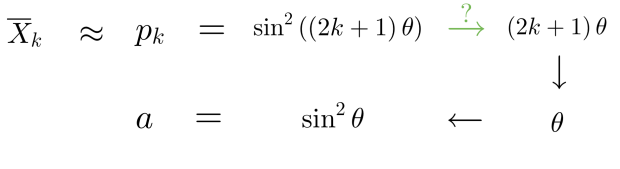


Figure 1: Pipeline of enhanced-sampling amplified estimation with identifiability challenge highlighted in green.

such that the *amplified target probability* is

$$p_k \coloneqq \sin^2((2k+1)\theta). \tag{3}$$

Given that the probability p_k is amplified relative to the original probability a, where enhanced sampling estimates p_k with the same accuracy as a and subsequently rescales based on inference of θ, p_k results in a higher-accuracy estimate of a than direct measurement of a itself (see Appendix A.2 Example 1).

Abstraction of the quantum amplitude a in amplified estimation from the amplified state therefore follows an analogous procedure to abstraction of the quantum amplitude a from the original state in Classical QAE—with an added identifiability challenge. For a fixed number of Grover operators k, the circuit $\mathcal{Q}^k\mathcal{A}$ applied to the $(n+1)$-qubit zero state is measured N_k times. Let $X_{i,k}$ represent the outcome of the i-th application. The sample mean is calculated as $\overline{X}_k = \frac{1}{N_k}\sum_{i=1}^{N_k} X_{i,k}$. This mean approximates p_k, which for known k yields θ, which is in turn used to calculate a according to the pipeline schematic in Fig. 1. For each value of p_k there exist $2k+1$ valid solutions of Eq. (3)

$$(2k+1)\theta = \begin{cases} \arcsin\left(\sqrt{p_k}\right) + l\cdot\frac{\pi}{2} & \text{if } l \text{ is even,} \\ \arccos\left(\sqrt{p_k}\right) + l\cdot\frac{\pi}{2} & \text{if } l \text{ is odd,} \end{cases} \tag{4}$$

where $l = 0, 1, 2, \ldots, 2k$ (see Appendix A.3). Each choice of the l value and its corresponding possible solution in Eq. (4) leads to a distinct estimate for a, each of which is well-reasoned; yet, based solely on information about a single p_k $(k > 0)$, it is impossible to determine which possible solution is the true solution. As a result, without further information, a would remain unidentifiable, and additional information is required to resolve this ambiguity, thereby identifying the unique *amplified angle* $(2k+1)\theta$ that corresponds to the true l value for a given k. This *quadrant index* $l(k)$ may be readily interpreted geometrically as the quadrant index of the amplified angle

$$l(k) = \left\lfloor \frac{(2k+1)\theta}{(\pi/2)} \right\rfloor, \tag{5}$$

where $\lfloor\cdot\rfloor$ denotes the floor function.

2.2 Amplified Amplitude Estimation

The recently developed Amplified Amplitude Estimation (AAE) method [92] resolves the identifiability problem in amplified estimation via the small-angle assumption and can be readily formulated in the statistical framework of AMSE (thus providing a conceptual stepping stone to IQAE and BIQAE), as follows:

To begin, we specify that the AAE approach considers cases in which the amplitude a is sufficiently small that there exists only one solution to the identifiability challenge Eq. (4).

Assumption 1. *For a given k, a is sufficiently small that $(2k+1)\theta$ does not exceed $\frac{\pi}{2}$.*

With Assumption 1, the quadrant index $l(k)$ in Eq. (4) must be zero, such that there exists only one valid solution.

According to the proof outlined in Appendix C.1, the quantum sample complexity of the approach can then be stated in the statistical framework of AMSE according to Theorem 1:

Theorem 1. *Suppose Assumption 1 holds with a certain k. Then,*

$$\hat{a}_0 \coloneqq \sin^2\left(\frac{1}{2k+1}\arcsin\sqrt{\bar{X}_k}\right) \tag{6}$$

is the Maximum Likelihood Estimator (MLE) for a. The corresponding asymptotic MSE is

$$AMSE_{\hat{a}_0}(a) = \frac{1}{(2k+1)^2}\frac{1}{N_k}a(1-a),$$

and to achieve a target accuracy ε, the quantum sample complexity is

$$N_{oracle} = (2k+1)N_k = \frac{1}{2k+1}\frac{1}{\varepsilon^2}a(1-a). \tag{7}$$

Comparison of the quantum sample complexity of AAE Eq. (7) to that of Classical QAE Eq. (1) then confirms AAE decreases the quantum sample complexity by a factor $\frac{1}{2k+1}$, thereby demonstrating its advantage over Classical QAE

and emphasizing that the greater the number of applied $\mathcal{Q}$ operators, the greater the efficiency of the amplified estimator, provided Assumption 1 holds. Furthermore, in the case where $a = O(\varepsilon)$, the quantum sample complexity becomes $N_{\text{oracle}} = O\left(\frac{1}{\sqrt{\varepsilon}}\right)$, a conclusion that can be directly derived from Theorem 1 where the order of k is chosen to be $\Theta\left(\sqrt{\varepsilon}\right)$. The complexity $N_{\text{oracle}} = O\left(\frac{1}{\sqrt{\varepsilon}}\right)$ exceeds Heisenberg-limited scaling in this context but does not refute the Heisenberg limit, as it relies on prior knowledge.

2.3 IQAE

We proceed by formulating IQAE in the same statistical framework of AMSE with the additional introduction of the normal approximation, which transparently demonstrates IQAE's estimator for a achieves a quadratic advantage, up to a $\log\log(1/\varepsilon)$ factor, without the need for the small-angle assumption. Note the key insight exploited by IQAE here is that one can obtain the same AMSE and quantum sample complexity as AAE, provided that the quadrant index is known (see Appendix A.4 for a formal statement), thereby replacing the need for the small-angle assumption.

To begin the statistical derivation of IQAE, we consider that, since $\theta \in \left[0, \frac{\pi}{2}\right]$ implies the initial quadrant index must be zero ($l(0) = 0$) per the definition of the quadrant index Eq. (5), we may initiate the AAE process with zero Grover operators.

Progressively increasing the number k of Grover operators while carefully tracking the evolution of the quadrant index then provides the means to increase the efficiency of the estimator while preserving identifiability, as follows: Formally, let $0 = k_0 < k_1 < k_2 < \cdots < k_t < \cdots < k_{T-1}$ be the sequence of T k-values processed in pursuit of the desired accuracy. Beginning with $t = 0$, for each t, the amplitude is amplified with k_t Grover operators and N_t measurements are performed on the corresponding quantum circuit, which constitutes a k_t-dependent amplification and measurement process we term *stage* t. With the definition $K_t \coloneqq 2k_t + 1$, which represents the number of accesses to the oracle $\mathcal{A}$ in a single execution of the circuit with k_t applications of the Grover operator, we organize these stages into a sequence $\{(K_t, N_t) : t = 0, 1, \ldots, T-1\}$ that we term the *schedule* that is designed to ensure the feasibility of tracking the quadrant index at all stages.

To develop a schedule that is both efficient and facilitates statistical analysis, we consider the distinct advantages of schedules previously employed in Maximum-Likelihood QAE (MLQAE) [7] and Iterative QAE (IQAE) [93]. MLQAE puts forward both linear and exponential K-schedules with preset values of N_t for all t. Prior numerical results indicate such an exponential schedule with fixed N_t values achieves a quadratic advantage at the expense of the lack of a theoretical guarantee of its feasibility. IQAE, in contrast, adaptively selects both K_t and N_t such that the quadrant index of each stage l_t is estimated based on measurements carried out in the preceding stage, which continuously refines a confidence interval at confidence level α for a and guarantees that its radius eventually becomes smaller than a prescribed accuracy threshold ε. A rigorous theoretical analysis indicates IQAE achieves optimal complexity up to a multiplicative $\log\log(1/\varepsilon)$ factor; however, derivation of this complexity upper bound is relatively intricate due to IQAE's fully adaptive algorithmic structure, which requires accounting for many possible algorithmic behaviors.

Based on these points, we develop two hybrid approaches that combine the analytical advantages of MLQAE's exponential K-schedule with IQAE's beneficial confidence intervals.

2.3.1 Base-3 Hybrid-Scheduled IQAE

We first introduce a hybrid (K, N)-schedule for IQAE to combine the benefits of exponential K-schedules and confidence intervals: a base-3 exponential K-schedule

$$K_t = 3^t, \quad t = 0, 1, \ldots, T-1$$

with an adaptive N-schedule in which the algorithm incrementally performs batches of N_{incre} measurements (termed *iterations*) until the confidence interval for θ is narrow enough that $l(k_{t+1})$ is determined with a desired degree of high confidence α_t. We set $\alpha_t = \alpha/T$ such that, by the union bound, the choice of α_t ensures that the final confidence interval for a achieves the overall confidence level α. As derived in Appendix C.2, to obtain the desired confidence interval for a with a radius no more than ε, the proposed sched-

ule requires the number of stages to be just

$$T = \left\lceil \log_3 \left(\frac{\pi}{2} \frac{1}{\varepsilon} \sqrt{a(1-a)} \right) \right\rceil . \tag{8}$$

To aid analysis, the resulting hybrid schedule admits both ternary-expansion and geometric interpretations: Combined with the formula for the quadrant index Eq. (5), imposition of the hybrid (K, N)-schedule ensures $l(k_t)$ represents the t-th digit in the ternary expansion of $\theta/(\pi/2)$. At stage t, the task of identifying the quadrant index is then equivalent to determining the $(t+1)$-th digit in this ternary expansion given its first t digits. Alternatively, from a geometric perspective, the task of identifying the quadrant index is equivalent to determining the *reference interval* to which the amplified angle $3^t\theta$ belongs, where the *reference interval* at stage t is defined as one of the three equal-angle divisions of the quadrant containing $3^t\theta$

$$l_t \cdot \frac{\pi}{2} + j \cdot \frac{\pi}{6} + \left[0, \frac{\pi}{6}\right], \quad j = 0, 1, 2,$$

where l_t is used as an abbreviation of $l(k_t)$, the quadrant index at stage t.

To ascertain the quantum sample complexity, we consider the simplest case in which the confidence interval for the amplified angle is relatively broad at each stage relative to the reference intervals:

Assumption 2. *Let ε_t be the radius of the confidence interval for θ obtained at stage t. Assume that*

$$\varepsilon_t \geq \frac{\gamma}{2} \frac{1}{3^{t+1}} \frac{\pi}{2}$$

for some constant $\gamma \in (0, 1]$, and for all $t = 0, 1, \ldots, T-2$.

The hybrid schedule and normal approximation then lead to the following upper bound for the quantum sample complexity (see detailed proof in Appendix C.3):

Theorem 2. *Under Assumption 2, the accumulated quantum sample complexity of the proposed base-3 hybrid schedule to obtain a confidence interval for a of radius ε with an overall confidence level of α (≤ 0.1) is upper bounded by*

$$N_{oracle} \leq \frac{6}{\pi} \left(\frac{3}{\gamma^2} + 2 \right) \frac{1}{\varepsilon} \log \left(\frac{2}{\alpha} \left\lceil \log_3 \left(\frac{\pi}{4\varepsilon} \right) \right\rceil \right) \times \sqrt{a(1-a)}.$$

The proposed base-3 upper bound thereby achieves the same order quantum sample complexity $\mathcal{O}\left(\frac{1}{\varepsilon} \log \log \frac{1}{\varepsilon}\right)$ as that of standard-scheduled IQAE [93], and thus achieves the optimal order up to a double logarithmic factor. Note the use of the normal approximation in this derivation provides the conceptual and mathematical basis for the creation of Normal-BIQAE, introduced in Sec. 3.3.1.

2.3.2 Base-3/Base-5 Hybrid-Scheduled IQAE

To observe the same ε-dependence of the upper bound of the sample complexity with hybrid-scheduled IQAE in extreme cases in which Assumption 2 does not hold, we introduce a second hybrid (K, N)-schedule for IQAE that includes another base in its K-schedule, as follows:

In base 5, the reference intervals are

$$l_t \cdot \frac{\pi}{2} + j \cdot \frac{\pi}{10} + \left[0, \frac{\pi}{10}\right], \quad j = 0, 1, \ldots, 4,$$

Simultaneous monitoring of both sets of reference intervals for base 3 and base 5 to determine if either set completely contains the confidence interval therefore guarantees that

$$\varepsilon_t \geq \frac{1}{2} \frac{1}{K_t} \frac{1}{15} \frac{\pi}{2}, \quad \text{for } t = 0, 1, \ldots, T-2, \tag{9}$$

which reflects the fact that, where the amplified angle falls close to the boundaries of the reference intervals defined for base 3 and thus violates Assumption 2, the amplified angle nonetheless rests sufficiently far from the reference interval boundaries in base 5 to avoid entrapment.

Replacement of Assumption 2 with Eq. (9) then yields the accumulated complexity bound for the base-3/base-5 schedule of

$$N_{\text{oracle}} \leq \frac{227}{\pi} \frac{1}{\varepsilon} \log \left(\frac{2}{\alpha} \left\lceil \log_3 \left(\frac{\pi}{4\varepsilon} \right) \right\rceil \right) \sqrt{a(1-a)}, \tag{10}$$

as formally shown in Appendix A.5.

The resulting bound thus retains the same order upper bound as base-3 hybrid-scheduled IQAE in Theorem 2, and thus the same order as in standard-scheduled IQAE [93], with a difference only in the constant factor.

2.3.3 IQAE Workflow

We execute IQAE according to the workflow depicted in Fig. 2a, which forms the structural scaffolding for BIQAE. Repeated sets of measurements of the quantum circuit are fed into a classical `ComputeCI` module to compute the confidence interval for θ based on all measurements at the current stage, where confidence intervals are computed as normal approximated intervals for hybrid-scheduled IQAE or as Chernoff-Hoeffding or Clopper-Pearson confidence intervals for standard-scheduled IQAE (see Appendix A.9). The `FindNextK` module then determines the following $K := 2k+1$, as outlined in the aforementioned strategy for hybrid-scheduled IQAE or via a linear search over a candidate set of K values for standard-scheduled IQAE. The updated K-value is then employed both to update the form of the quantum circuit and to inform the classical `ComputeCI` module, and the process continues such that the confidence interval for a continuously shrinks until its radius is smaller than the target accuracy ε. Note this workflow only permits IQAE to retain memory of the confidence interval determined in the most recent stage, without preference within said confidence interval.

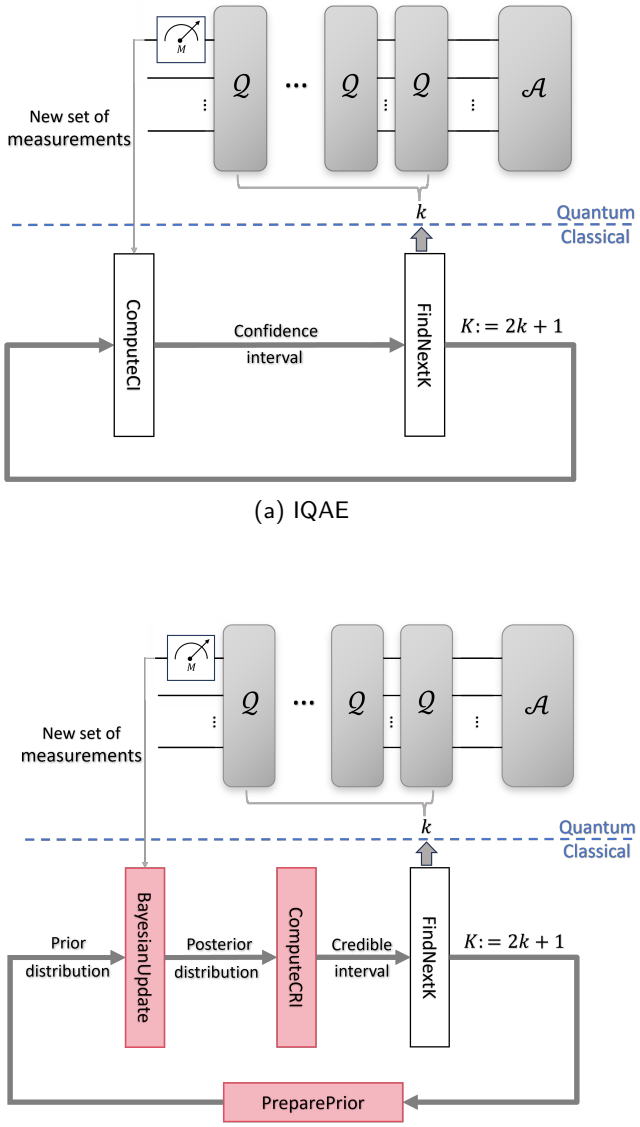


Figure 2: Schematic illustrations of IQAE and BIQAE algorithms. (a) In IQAE, measurements of amplified-amplitude quantum circuits (top) are fed to a classical computer (bottom), which both iteratively improves the estimate of the confidence interval of θ (`ComputeCI`) and increments the number of Grover operators k of the quantum circuit (`FindNextK`) until the quantum amplitude is estimated with the desired accuracy. (b) BIQAE introduces three new modules to improve IQAE's efficiency via Bayesian inference (highlighted in red): `BayesianUpdate` and `PreparePrior`, which maintain and refine the prior and posterior distribution of the amplified target probability; and `ComputeCRI`, the Bayesian substitute to `ComputeCI`, which calculates a credible interval for θ based on the posterior distribution in place of confidence interval.

3 BIQAE Method

3.1 BIQAE Framework

To harness and refine not only the interval estimate but also the *posterior distribution* of the amplitude, we incorporate Bayesian inference into IQAE as follows:

For each stage t, we consider a prior distribution over p_{k_t} that models understanding of p_{k_t} thus far as

$$p_{k_t} \sim \pi_t^{\text{prior}}(p_{k_t}), \tag{11}$$

where $\pi_t^{\text{prior}}(\cdot)$ denotes a prior distribution whose form remains to be specified. We hold this prior distribution fixed from the conclusion of stage $t-1$ throughout the remainder stage t.

In each iteration within stage t, we increment N_t by N_{incre} to enhance the accuracy of the estimate. Let N_t denote the total number of measurements gathered at stage t. We map the measurement outcomes to 0 and 1. Then, the distri-

bution of the sample average $\bar{X}_t$ is given by

$$\bar{X}_t | p_{k_t} \sim \frac{1}{N_t} \mathrm{Bin}\left(N_t, p_{k_t}\right). \tag{12}$$

We refer to this distribution of the sample average $\bar{X}_t$ given p_{k_t} as the likelihood model and denote the density function by $\mathcal{L}_t\left(\bar{X}_t \middle| p_{k_t}\right)$.

A Bayesian update then integrates the N_t measurements collected at stage t with the prior distribution to compute a posterior distribution that reflects an updated understanding about p_{k_t}. This computation is carried out using Bayes' rule as

$$\pi_t^{\mathrm{post}}\left(p_{k_t} \middle| \bar{X}_t\right) \propto \pi_t^{\mathrm{prior}}(p_{k_t}) \cdot \mathcal{L}_t\left(\bar{X}_t \middle| p_{k_t}\right), \tag{13}$$

where $\pi_t^{\mathrm{post}}\left(p_{k_t} \middle| \bar{X}_t\right)$ denotes the posterior distribution.

An interval estimate for the angle θ is then obtained by first computing an interval estimate for p_{k_t} and then translating this interval to θ. The interval estimate for p_{k_t} is derived from the posterior distribution of p_{k_t} by taking its upper and lower $\alpha_t/2$-quantiles

$$\left[p_{k_t}^l, p_{k_t}^u\right] = \left[\left[\Pi_t^{\mathrm{post}}\right]^{-1}\left(\frac{\alpha_t}{2}\right), \left[\Pi_t^{\mathrm{post}}\right]^{-1}\left(1-\frac{\alpha_t}{2}\right)\right], \tag{14}$$

where Π_t^{post} represents the cumulative distribution function of the posterior distribution. This approach, known as *credible intervals* in Bayesian statistics, replaces the confidence intervals in IQAE, as the latter is incompatible with the posterior distribution.

To obtain the θ location information required to determine the next k, the resulting credible interval for p_{k_t} is then transformed by

$$\left[\theta_t^l, \theta_t^u\right] = f_t\left(\left[p_{k_t}^l, p_{k_t}^u\right]\right), \tag{15}$$

where $f_t(\cdot)$ denotes the deterministic functional relationship between θ and p_{k_t}

$$f_t(x) := \begin{cases} \frac{1}{K_t}\left(\arcsin\left(\sqrt{x}\right) + l_t\frac{\pi}{2}\right) & \text{if } l_t \text{ is even,} \\ \frac{1}{K_t}\left(\arccos\left(\sqrt{x}\right) + l_t\frac{\pi}{2}\right) & \text{if } l_t \text{ is odd.} \end{cases}$$

To track the evolution of the bases such that l_t is known, let $\left[\theta_{t-1}^l, \theta_{t-1}^u\right]$ denote the interval estimate for θ from the previous stage. Then, l_t is readily computed as

$$l_t = \left\lfloor \frac{K_t \theta_{t-1}^l}{\pi/2} \right\rfloor.$$

Based on the interval estimate obtained from Eq. (15), a linear search is performed to find the largest feasible k-value as in standard-scheduled IQAE. Denote the k-value output as k_{next}. If $k_{\mathrm{next}} = k_t$, the k-value remains identical, and the algorithm continues with a new iteration. If instead $k_{\mathrm{next}} > k_t$, the number of $\mathcal{Q}$ operators in the associated quantum circuit is increased to $k_{t+1} := k_{\mathrm{next}}$ to improve measurement efficiency. A new stage $t+1$ then begins, which requires preparation of a new prior distribution for $p_{k_{t+1}}$ before resumption of the iteration process.

To prepare the new prior upon identification of a greater feasible k-value, the posterior distribution for p_{k_t} is transformed into a prior distribution of the target probability in the upcoming stage as

$$p_{k_{t+1}} = \sin^2\left(K_{t+1} \cdot f_t(p_{k_t})\right), \tag{16}$$

namely

$$p_{k_t} \sim \pi_t^{\mathrm{post}} \xrightarrow{\sin^2(K_{t+1}\cdot f_t(\cdot))} p_{k_{t+1}} \sim \pi_{t+1}^{\mathrm{prior}}. \tag{17}$$

In practice, this transformation is implemented approximately based on the chosen form of prior distribution.

3.2 BIQAE Workflow

We execute BIQAE according to the workflow diagram provided in Fig. 2b, which differs from the workflow of IQAE in (i) the addition of a `BayesianUpdate` module, which updates the posterior distribution based on previous measurements and the prior distribution according to Bayes' rule, (ii) the insertion of a `PreparePrior` module, which prepares the prior based on the posterior distribution of the preceding stage, and (iii) the substitution of IQAE's `ComputeCI` module, which computes frequentist confidence intervals, with BIQAE's `ComputeCRI` module, which computes Bayesian credible intervals in BIQAE. Note that whereas IQAE feeds measurements into the `ComputeCI` module to compute the confidence interval, BIQAE inputs new measurements into the `BayesianUpdate` module to refine the posterior distribution of the amplitude a (thereby integrating prior knowledge with new measurements) and subsequently passes the posterior distribution on to `ComputeCRI` to compute the credible interval. This flow of information ensures the algorithm learns from both the previous interval

estimate and the shape of the distribution within this interval to further increase the measurement efficiency of the quantum amplitude estimate.

3.3 Implementation of BIQAE

We introduce two approaches to implement the aforementioned Bayesian framework: (i) BIQAE with a normal prior (Normal-BIQAE) for theoretical insights into quantification of the improvement brought about by injection of Bayesian inference into IQAE and (ii) BIQAE with a beta prior (Beta-BIQAE) for superior practical performance.

3.3.1 Normal-BIQAE

Prior. To model the probability distributions in BIQAE according to a normal prior distribution, we adopt as the prior Eq. (11) a distribution of the form

$$p_{k_t} \sim \mathcal{N}(\mu_{0,t}, \Delta_t^2) \tag{18}$$

for normal distribution $\mathcal{N}$, prior mean $\mu_{0,t}$, and prior variance Δ_t^2 at stage t.

`BayesianUpdate`. To leverage the associated conjugate normal prior, we approximate the likelihood model in Eq. (12) in `BayesianUpdate` by the normal distribution

$$\bar{X}_t | p_{k_t} \sim \mathcal{N}\left(p_{k_t}, \frac{1}{N_t} p_{k_t}(1-p_{k_t})\right). \tag{19}$$

Combining Eqs. (13), (18), and (19), the posterior distribution follows as

$$\begin{aligned} p_{k_t} | \bar{X}_t &\sim \mathcal{N}\left(\mu_{\text{post},t}, \sigma^2_{\text{post},t}\right), \\ \mu_{\text{post},t} &= \frac{\frac{\mu_{0,t}}{\Delta_t^2} + \frac{\bar{X}_t}{\hat{\sigma}_t^2/N_t}}{\frac{1}{\Delta_t^2} + \frac{1}{\hat{\sigma}_t^2/N_t}}, \\ \sigma^2_{\text{post},t} &= \left(\frac{1}{\Delta_t^2} + \frac{1}{\hat{\sigma}_t^2/N_t}\right)^{-1}, \end{aligned} \tag{20}$$

where $\hat{\sigma}_t^2 := \bar{X}_t(1-\bar{X}_t)$ is an asymptotic unbiased estimator for the variance.

`ComputeCRI`. Since the resulting posterior distribution is normal, the credible interval for p_{k_t} Eq. (14) in the `ComputeCRI` module becomes

$$\left[p_{k_t}^l, p_{k_t}^u\right] = \text{trunc}_{[0,1]}(\mu_{\text{post},t} \pm z_{\alpha_t/2} \cdot \sigma_{\text{post},t}), \tag{21}$$

where $z_{\alpha_t/2}$ is the $(1-\alpha_t/2)$-quantile of the standard normal distribution.

`PreparePrior`. An approximation is then required to prepare a normal prior from the normal posterior distribution Eq. (20) since, although the posterior distribution is normal and the posterior-to-prior transformation Eq. (17) is smooth, the exact prior distribution after transformation is no longer a normal distribution. Assuming N_t is sufficiently large that the normality in the central limit theorem applies, we thus approximate the prior distribution for $p_{k_{t+1}}$ in the `PreparePrior` module by the normal distribution according to the delta method (see Appendix A.6), which yields

$$\begin{aligned} p_{k_{t+1}} &\sim \mathcal{N}\left(\mu_{0,t+1}, \Delta_{t+1}^2\right), \\ \mu_{0,t+1} &= \sin^2\left(K_{t+1} \cdot f_t(\mu_{\text{post},t})\right), \\ \Delta_{t+1}^2 &= \left(\frac{K_{t+1}}{K_t}\right)^2 \frac{\mu_{0,t+1}(1-\mu_{0,t+1})}{\mu_{\text{post},t}(1-\mu_{\text{post},t})} \cdot \sigma^2_{\text{post},t}. \end{aligned}$$

Quantum Sample Complexity. Where N_t is sufficiently large that the normality in central limit theorem applies, we derive for the resulting Normal-BIQAE implementation an upper bound on the quantum sample complexity at stage t required to obtain an interval estimate for θ with radius ε_t of

$$N_{\text{oracle},t} \leq \frac{z^2_{\alpha_t/2}}{4K_t}\left(\frac{1}{\varepsilon_t^2} - \frac{1}{\varepsilon_{t-1}^2}\right), \tag{22}$$

where ε_{t-1} is the radius of the interval estimate for θ obtained at stage $t-1$ (see Appendix A.7). Compared to the upper bound of Normal-IQAE (see Appendix C.3)

$$N_{\text{oracle},t} = \frac{z^2_{\alpha_t/2}}{4K_t}\frac{1}{\varepsilon_t^2}, \tag{23}$$

Normal-BIQAE therefore reduces the complexity by a factor of $\varepsilon_t^2/\varepsilon_{t-1}^2$.

3.3.2 Beta-BIQAE

Prior. In Beta-BIQAE, we assume a prior Eq. (11) in the form of a beta distribution

$$p_{k_t} \sim \text{Beta}(a_{0,t}, b_{0,t}), \tag{24}$$

where $a_{0,t}, b_{0,t}$ are the shape parameters.

`BayesianUpdate`. A beta-distributed prior allows us to directly use the binomial distribution Eq. (12) to derive a beta-distributed posterior, which avoids the need for normal approximation. Using Eqs. (12), (13), and (24); the posterior distribution is

$$\begin{aligned} p_{k_t}|\bar{X}_t &\sim \text{Beta}(a_{\text{post},t}, b_{\text{post},t}), \\ a_{\text{post},t} &= a_{0,t} + N_t\bar{X}_t, \\ b_{\text{post},t} &= b_{0,t} + N_t\left(1-\bar{X}_t\right). \end{aligned} \tag{25}$$

`ComputeCRI`. The interval estimate Eq. (14) in the `ComputeCRI` module then follows as

$$\left[p^l_{k_t}, p^u_{k_t}\right] = \left[\text{BetaCDF}^{-1}\left(\frac{\alpha_t}{2}, a_{\text{post},t}, b_{\text{post},t}\right), \text{BetaCDF}^{-1}\left(1-\frac{\alpha_t}{2}, a_{\text{post},t}, b_{\text{post},t}\right)\right]$$

to accommodate the beta-distributed posterior, where $\text{BetaCDF}^{-1}(\cdot, a, b)$ is the inverse of the cumulative density function of the beta distribution with shape parameters a and b.

`PreparePrior`. Analogously to Normal-BIQAE, given that the exact prior distribution of $p_{k_{t+1}}$ after the transformation Eq. (17) is no longer a beta distribution, an appropriate beta distribution is required to approximate the true distribution. However, unlike Normal-BIQAE, Beta-BIQAE does not admit use of the delta method to enable the derivation of a closed-form approximation. We therefore employ a sampling procedure in `PreparePrior` to obtain an asymptotically optimal beta approximation of the exact prior distribution, where optimality is defined as minimization of the distance (Kullback–Leibler divergence) between the chosen approximation and the exact prior (see Appendix A.8 for a rigorous establishment of this optimality result).

Specifically, given the posterior distribution derived in Eq. (25), we prepare the prior as follows:

1. Draw a sufficiently large sample size R (for example, $R = 1000$) of independent and identically distributed (i.i.d.) values $\{y_1, y_2, \ldots, y_R\}$ from the beta posterior in Eq. (25) to accurately approximate the underlying distribution.

2. Map each value y_i to the parameter space for $p_{k_{t+1}}$ through the transformation described in Eq. (17) to obtain

$$\tilde{y}_i \coloneqq \sin^2\left(K_{t+1} \cdot f_t(y_i)\right)$$

where $\{\tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_R\}$ is identical to a sample drawn from the exact prior distribution of $p_{k_{t+1}}$.

3. Treat $\{\tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_R\}$ as if they were drawn from a beta distribution and use the Maximum Likelihood Estimator (MLE) to infer its shape parameters.

4. Use the maximum likelihood estimates as the shape parameters $(a_{0,t+1}, b_{0,t+1})$ of the beta distribution Eq. (24).

The chosen sampling procedure is performed only at transitions between stages, namely, a total of $T-1$ times where $T = \mathcal{O}\left(\log(1/\varepsilon)\right)$ and requires less than a second of classical computing time per application, thereby circumventing the high classical cost associated with Monte Carlo techniques that are required where the form of the prior is not specified to allow conjugacy [55].

In practice, the per-stage sample size N_t employed in implementations of IQAE and BIQAE is typically quite small (often $N_t < 100$). Whereas Normal-BIQAE depends on a normal approximation, such that as N_t decreases the performance of Normal-BIQAE deteriorates (namely, both the quantum sample complexity and the loss of nominal coverage at level $1-\alpha$ increase, see Appendix B.5); Beta-BIQAE makes no such approximation, such that Beta-BIQAE is capable of handling a broader range of sample sizes, including the practical small sample regime (see Section 4.1 and Appendix B.5).

3.3.3 BIQAE with a Noninformative Prior

We further recognize that, although not originally derived according to a Bayesian perspective, IQAE constitutes implementation of BIQAE with a noninformative prior, as derived in detail in Appendix A.9. Precisely, IQAE with Chernoff-Hoeffding confidence intervals (termed Normal-IQAE or IQAE-CH) corresponds to BIQAE with a noninformative normal prior, and IQAE with Clopper-Pearson confidence intervals (termed Beta-IQAE or IQAE-CP) corresponds to BIQAE with a noninformative beta prior.

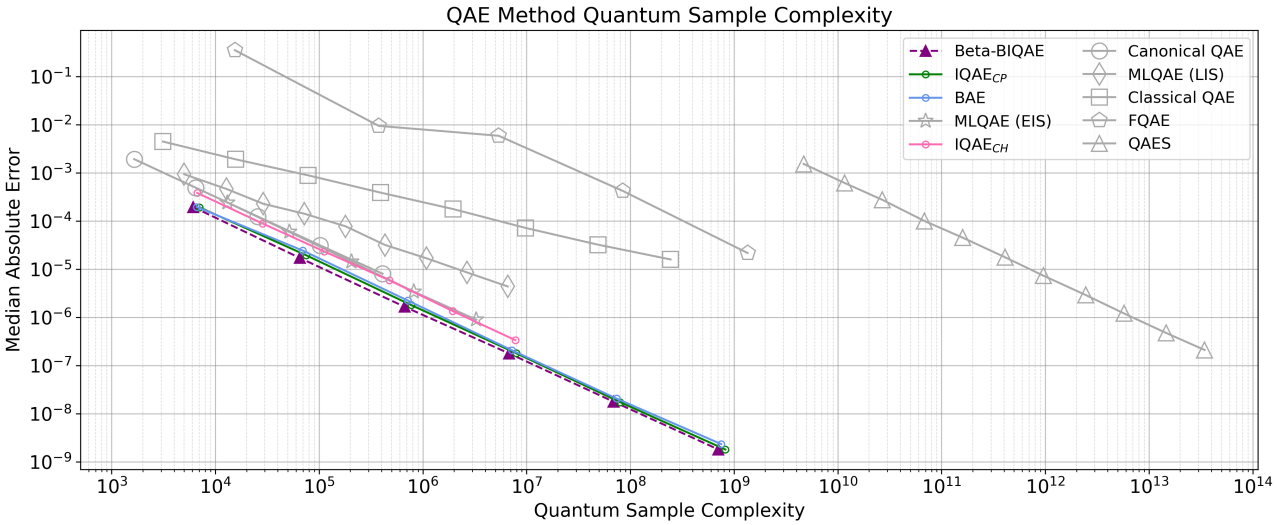


Figure 3: Quantum sample complexity comparison of BIQAE to benchmark QAE approaches for quantum amplitude $a = 0.5$ over 200 repetitions. Legend is ordered from highest to lowest performance, with Bayesian and IQAE methods emphasized in color (BIQAE, purple triangles; IQAE-CP, green circles; BAE, blue circles; and IQAE-CH, pink circles).

3.4 Computational Methods

To evaluate the broad applicability and measurement efficiency of BIQAE, we conduct numerical experiments on the ability of BIQAE to estimate both amplitudes of specified quantum states and ground-state energies of diatomic and triatomic molecules. We have made code for BIQAE and all computational tests readily accessible via Github at `http://www.github.com/Kirin0570/BIQAE` [94]. For benchmark methods, we employ the open-source code [95] that accompanies ref. [55] with minor modifications for output formatting and storage.

To examine the performance of BIQAE for quantum state amplitude estimation relative to existing QAE techniques, we compare the quantum sample complexity of Beta-BIQAE, Canonical QAE [96], Classical QAE [1], IQAE [93], BAE [55], MLQAE [7], FQAE (Faster Quantum Amplitude Estimation) [6], and QAES (QAE, Simplified) [5]. Experiments are performed with the standardized quantum circuit design presented in Appendix A.10 at a variety of target accuracies, with each experiment independently repeated with a stated number of *repetitions* to compute summary statistics of the performance metrics. Given that BIQAE, IQAE, FQAE and Classical QAE are able to produce interval estimates whereas BAE, MLQAE, and Canonical QAE provide only point estimates; we simulate all approaches based on confidence intervals with a confidence level of $\alpha = 0.05$ during execution of the algorithm and subsequently compare all QAE methods as a whole from a point estimation perspective in which we reduce the confidence interval to a point estimate by taking its center. To further aid direct comparison with literature results for IQAE [93], (i) we adopt as the weight of each shot of the quantum circuit the number of the applications of the Grover operator (k_t) instead of the number of accesses to the oracle $\mathcal{A}$ ($K_t = 2k_t + 1$) and (ii) in cases where the quantum amplitude is prespecified, we consider a value of $a = 0.5$—an amplitude that is simultaneously expected to yield the worst-case sample complexity because all derived bounds exhibit dependence on $a(1-a)$ either directly or through its square root. A hyperparameter of incremental shot batch size $N_{\text{incre}} = 10$ is employed for all relevant experiments with IQAE and BIQAE.

To assess the accuracy and efficiency of BIQAE for molecular ground-state energy approximation, we employ BIQAE to approximate the energy of the ground-state wavefunction of four molecules – H_2, LiH, HF, and BeH_2. For each molecule, the electronic Hamiltonian is generated at the Hartree-Fock level of electronic structure theory in the Born-Oppenheimer approximation with frozen core orbitals, Jordan-Wigner mapping, and qubit tapering [97–102]. The equilibrium bond length is assumed unless otherwise stated. The ground-state wavefunction is then prepared using the variational quantum eigensolver, as de-

tailed in Appendix A.11, prior to QAE for expectation value estimation. Where the accuracy is fixed, we compare the quantum sample complexity of Classical QAE, IQAE-CP, and Beta-BIQAE. Where the quantum sample complexity is fixed, we compare Classical QAE and Beta-BIQAE using the same number of shots in Classical QAE as that needed to reach the desired target accuracy for each Pauli string $\varepsilon = 10^{-3}$ with Beta-BIQAE. We further restrict the target accuracy to $\varepsilon = 10^{-4}$ for HF given that the error entailed by energy estimation grows as the number and coefficient magnitude of Pauli terms increases. All molecular results represent averages over 200 repetitions.

4 Numerical Simulations

4.1 Quantum Sample Complexity Scaling Analysis

Quantum sample complexity analysis indicates Beta-BIQAE outperforms all other QAE methods considered for estimation of the benchmark quantum amplitude $a = 0.5$. As shown in Fig. 3, Beta-BIQAE requires a lower quantum sample complexity to estimate the quantum amplitude than BAE, IQAE (-CP and -CH), Canonical QAE, MLQAE (EIS [Exponentially Incremental Schedule] and LIS [Linearly Incremental Schedule]), Classical QAE, FQAE, and QAES; an advantage that persists over five orders of magnitude in the specified median absolute error. Linear regression analysis detailed in Appendix B.1 further indicates Beta-BIQAE's scaling (slope -1.0088, intercept 0.0211, fit $R^2 = 0.9999$) robustly offers the expected quadratic improvement over Classical QAE's scaling (slope -1.9821, intercept -1.1693, fit $R^2 = 0.996$). Beta-BIQAE's scaling also exhibits an intercept that is a multiplicative factor of five lower than the leading alternative, IQAE-CP. [1]

The superior scaling of Beta-BIQAE relative to IQAE-CP is found to be robust to different target accuracies. As shown in Fig. 4, Beta-BIQAE improves upon the measurement cost of IQAE-CP by a near-constant factor of approximately 14% across six orders of magnitude in the target accuracy (see also interval radius and coverage rate comparison of Beta-BIQAE, IQAE-CP, and BAE in Appendix B.2). This constant factor improvement is consistent with the predicted upper bound on the sample complexity Eq. (22) for Normal-BIQAE (in the absence of an analytical result for Beta-BIQAE), as the average interval radius ratio $\varepsilon_{t-1}/\varepsilon_t$ observed for the majority of the Beta-BIQAE process remains approximately 2.5 in Fig. 5, which corresponds to a 16% improvement.[2]

The double-digit percentage improvement of Beta-BIQAE relative to IQAE-CP is also found to persist across the full range of possible quantum amplitudes $a \in [0, 1]$, as evidenced by Fig. 6 for a target accuracy of $\varepsilon = 10^{-8}$. A consistent improvement in the quantum sample complexity of Beta-BIQAE relative to IQAE-CP of approximately 12% to 15% is visible for all a values considered. Additionally, close fitting of the quantum sample complexity as a function of a to the model $N_{\text{oracle}} = \beta\sqrt{a(1-a)}$ (where β is a constant depending on ε) reveals the observed complexities are consistent with the theoretical expectation detailed in Eqs. (10) and (22) that the complexities of IQAE and BIQAE follow the decomposition

$$N_{\text{oracle}} = C \cdot \frac{1}{\varepsilon}(\log\frac{1}{\alpha} + \log\log\frac{1}{\varepsilon}) \cdot \sqrt{a(1-a)}, \tag{26}$$

where C is a constant independent of a that is lower in BIQAE than IQAE and where the effect of differing input amplitude values is entirely captured in the multiplicative factor $\sqrt{a(1-a)}$.

[1] Note that whereas Beta-BIQAE outperforms all other QAE methods considered in terms of quantum sample complexity, Normal-BIQAE outperforms all other QAE methods considered with the exception of IQAE-CP. The more favorable performance of Beta- over Normal-BIQAE is consistent with the fact that the normal approximation (and the accuracy of variance estimation) that underlies Normal-BIQAE may become inaccurate for small sample sizes.

[2] We note that the observed interval radius ratio remains relatively unchanged until the algorithm's final stages, in which the ratio falls. The change observed in the final stages suggests that initial progression with a linear search strategy fluctuates about an exponential K-schedule (in which a base-b hybrid schedule implies an accuracy improvement rate per stage and concomitant radius ratio of b) and final stages satisfy the stopping criterion (namely, the final stage of each repetition is terminated early before the interval estimate becomes sufficiently small to allow advancement to the next larger K-value, which contributes a smaller ratio value to the average).

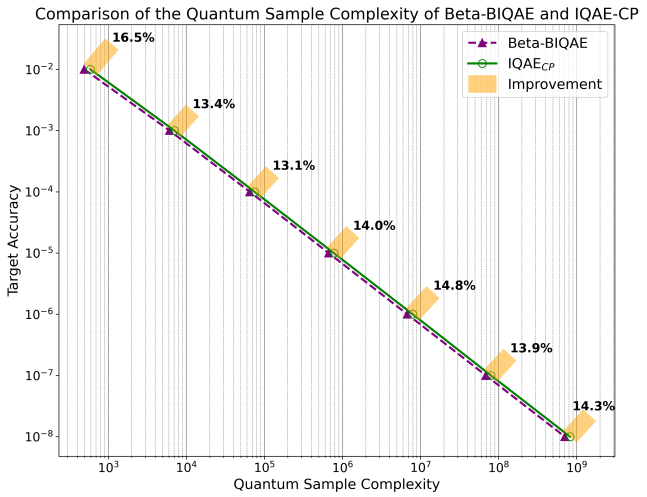


Figure 4: Quantum sample complexity comparison between Beta-BIQAE (dashed purple line with triangles) and IQAE-CP (solid green line with circles) across six orders of magnitude of target accuracy ε (percentage improvements of Beta-BIQAE over IQAE-CP accentuated in text and as yellow bars).

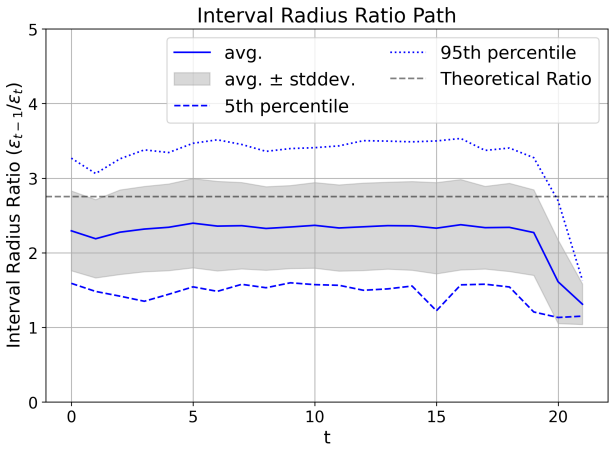


Figure 5: Radius ratio progression across stages in Beta-BIQAE (average value, solid blue line; values within one standard deviation, shaded gray area; 5^{th} percentile, dashed blue line; 95^{th} percentile, dotted blue line). Results are aggregated from 1000 repetitions with a target accuracy of $\varepsilon = 10^{-8}$. A reference line (dashed gray line) is included to mark the theoretically expected ratio from Eq. (22) corresponding to the observed 14.3% improvement in the quantum sample complexity analysis.

In Appendix B.3, we demonstrate that this favorable scaling of Beta-BIQAE relative to IQAE-CP for arbitrary a values holds over six orders of magnitude $\varepsilon \in [10^{-2}, 10^{-7}]$.

4.2 Molecular Ground-State Energy Estimation

As shown in Fig. 7, Beta-BIQAE estimates molecular ground-state energies with significantly greater accuracy than Classical QAE for all molecules at all bond lengths given a fixed quantum sample complexity. The interval estimates are significantly smaller for Beta-BIQAE than Classical QAE by an average factor of $\times 32.84$ for HF, $\times 14.2$ by H_2, $\times 12.16$ for BeH_2, and $\times 12.13$ for LiH. Additionally, in the majority of cases considered, the interval centers for Beta-BIQAE lie closer to the exact diagonalization value than Classical QAE, on average by a difference of 0.137 Hartree, 0.020 Hartree, 0.004 Hartree, and 0.002 Hartree for HF, H_2, LiH, and BeH_2, respectively. Note all of these improvements exceed chemical accuracy.

Examination of the quantum sample complexity required to yield a given target accuracy likewise indicates Beta-BIQAE significantly reduces measurement cost for energy estimation relative to Classical QAE and IQAE-CP: Fig. 8 indicates Beta-BIQAE requires a sample complexity two

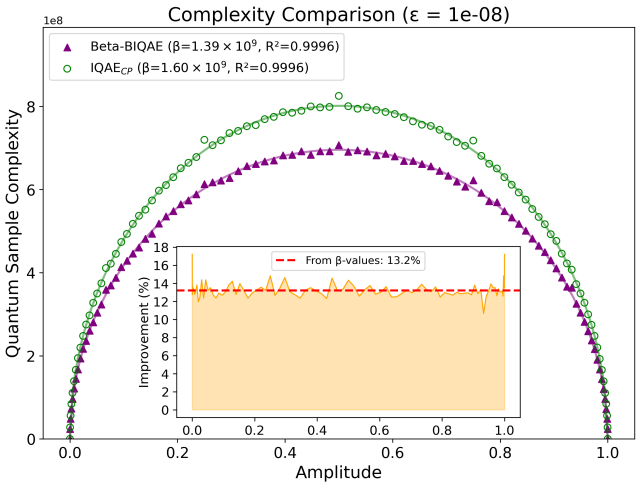


Figure 6: Comparison of the quantum sample complexities of Beta-BIQAE (purple triangles with fit solid line) and IQAE-CP (green circles with fit solid line) across quantum amplitudes $a \in [0, 1]$ at a fixed target accuracy $\varepsilon = 10^{-8}$, averaged over 1000 repetitions. Insets highlight Beta-BIQAE's consistent percentage improvement over IQAE-CP (yellow solid line values with red dotted line from fitted β).

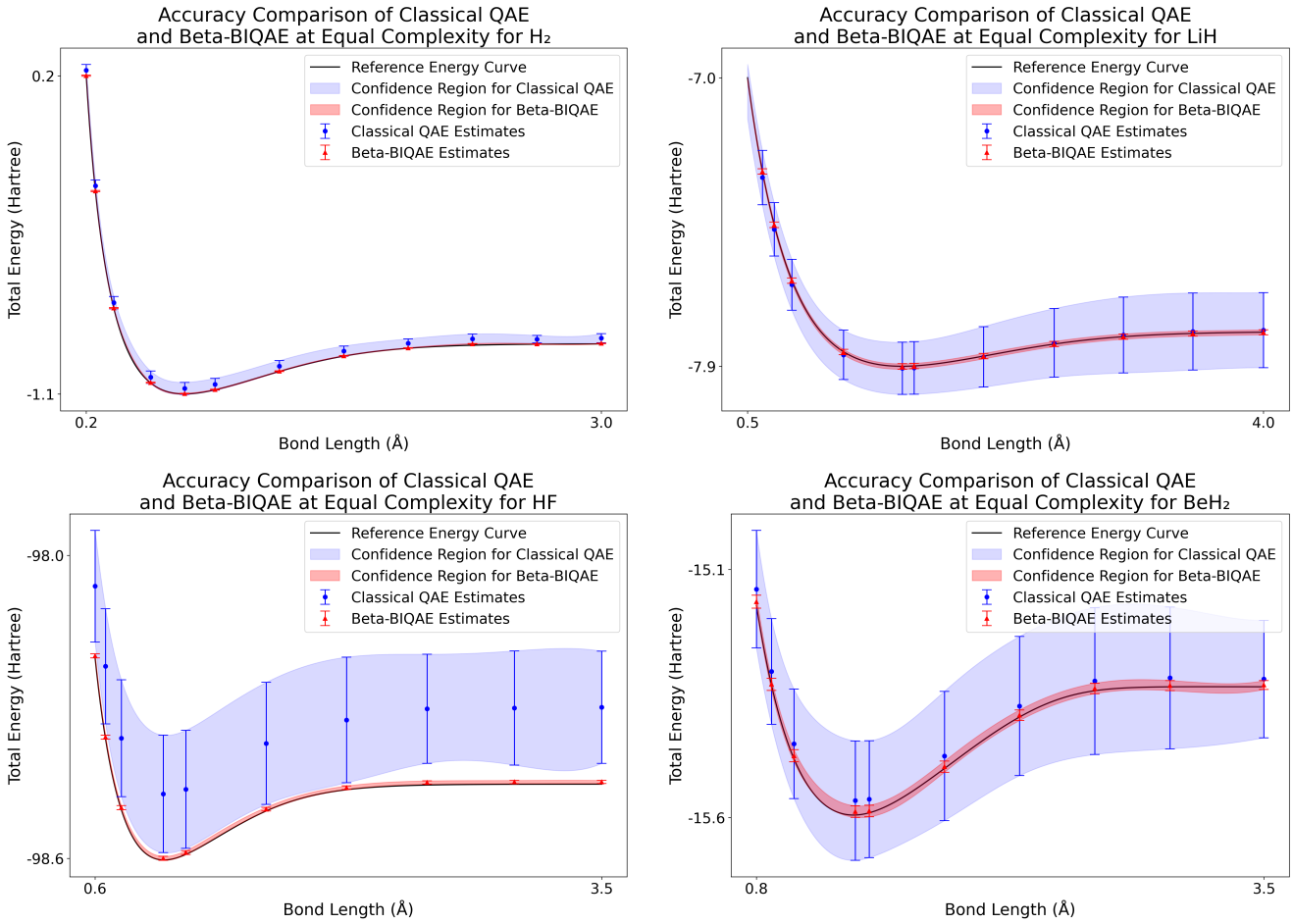


Figure 7: Comparison of Beta-BIQAE (red) and Classical QAE (blue) accuracy for H_2, LiH, HF, and BeH_2 ground-state energies at equal quantum sample complexity for Classical QAE and Beta-BIQAE. Error bars represent interval estimates, points indicate interval centers, and shaded areas depict the confidence region, with the energy curve from exact diagonalization of the Hamiltonian included for reference (black). Results between sampled bond lengths are interpolated with quartic splines.

to three orders of magnitude lower than Classical QAE to reach a target accuracy of $\varepsilon = 10^{-3}$ (three in the $\varepsilon = 10^{-4}$ case of HF) across all bond lengths considered. Furthermore, Beta-BIQAE consistently requires a sample complexity 8% to 13% lower than IQAE-CP across all molecular tests.

The quantum sample complexities to compute expectation values of individual Pauli terms to the specified target accuracies similarly demonstrate the lower measurement cost entailed by BIQAE relative to Classical QAE and IQAE-CP (see Fig. 9). The average quantum sample complexity of Beta-BIQAE is lower than that of Classical QAE for all Pauli terms by up to three orders of magnitude for $\varepsilon = 10^{-3}$ and up to four orders of magnitude for $\varepsilon = 10^{-4}$. Likewise, the average quantum sample complexity of Beta-BIQAE is lower than that of IQAE-CP by up to a double-digit percentage for all Pauli terms for all molecules considered. In particular, this advantage over IQAE-CP averages a percent improvement of 10.16% for H_2, 10.5% for BeH_2, 10.72% for LiH, and 11.4% for HF, in agreement with the expected scaling result for arbitrary quantum amplitudes of $\sim 10 - 16\%$. Moreover, a trend is clearly visible in which the sample complexity is lowest for extreme Pauli string indices with a plateau at intermediate Pauli string index values. Since the Pauli string indices are ordered in descending order of expectation value magnitude, this plateau is consistent with the prediction that higher quantum sample complexities are needed to estimate intermediate quantum amplitudes per the predicted quantum sample complexity dependence on a of Beta-BIQAE, Classical QAE, and

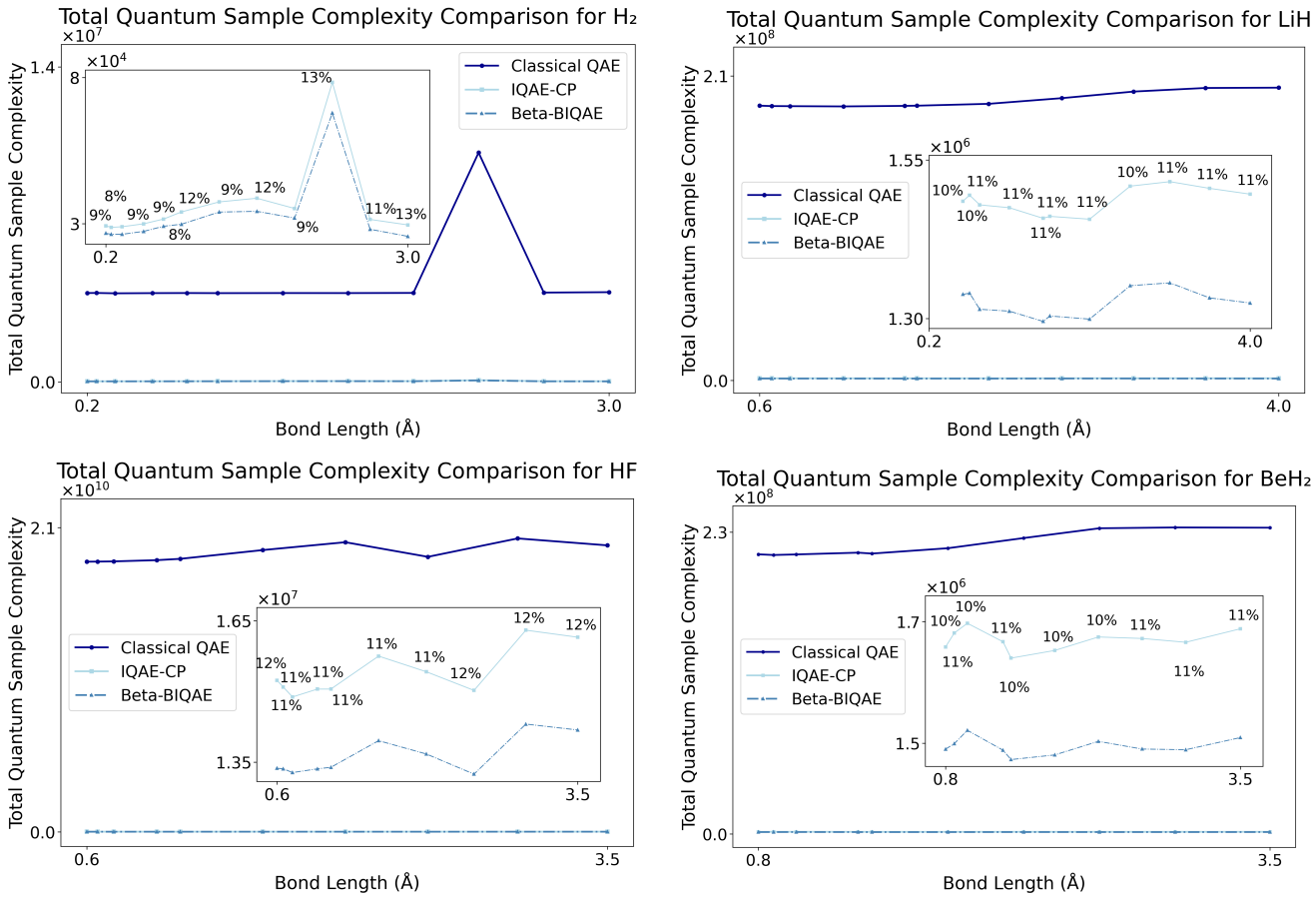


Figure 8: Comparison of quantum sample complexities for estimation of molecular ground-state energies as a function of interatomic bond length for Classical QAE (dark blue circles), IQAE-CP (light blue squares), and Beta-BIQAE (medium blue triangles) with inset illustration of percentage improvement achieved by Beta-BIQAE over IQAE-CP.

IQAE-CP $N_{\text{oracle}} \propto \sqrt{a(1-a)}$ in Eqs. (1) and (26). Note Beta-BIQAE's measurement cost reduction only requires a number of Grover operators comparable to that of IQAE-CP, and thus is not associated with a significant increase in circuit depth, as demonstrated in the extended results presented in Appendix B.4.

5 Discussion

The success of BIQAE for quantum amplitude estimation and molecular ground-state energy approximation shows conclusively the value of Bayesian inference as a way to reduce the cost of QAE and quantum computing algorithms generally. Since QAE is widely used as a subroutine in quantum computing tasks, it is therefore natural to explore BIQAE's implementation to accelerate a variety of calculations currently sought after as targets for quantum utility, including in tandem with supplementary cost reduction techniques that would further accelerate QAE. For example, the successful identification of molecular ground states with BIQAE demonstrated here invites the use of BIQAE to facilitate identification of molecular excited states in conjunction with quantum tomography techniques [103–105]. Consequently, the methods developed in this work are expected to have the potential to impact a wide range of quantum algorithms, thereby underscoring the broader significance of the results.

Given the measurement cost reduction observed here for implementation of Bayesian inference in IQAE at the level of interval estimates, the next step for BIQAE would be to inject Bayesian inference at the level of hyperparameters or (K, N)-scheduling decisions. Bayesian optimization [106–109] could be employed to optimize BIQAE's hyperparameters (such as the sample size N_t and incremental shot batch size N_{incre}), which is expected to improve the constant prefactor associated with the quantum sam-

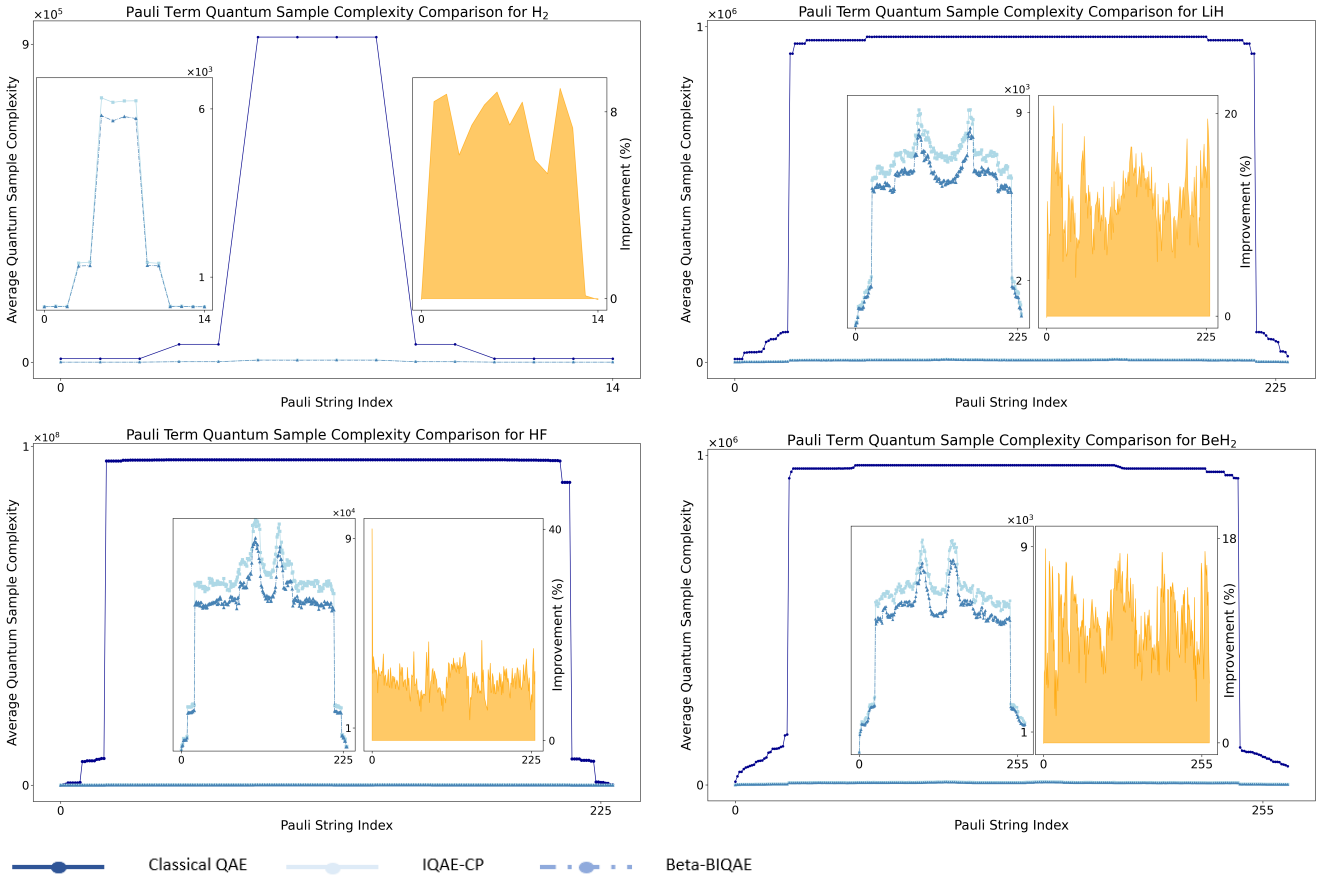


Figure 9: Quantum sample complexity comparison for H_2, LiH, HF, and BeH_2 ground-state energies with Classical QAE (dark blue circles), IQAE-CP (light blue squares), and Beta-BIQAE (medium blue triangles). Insets indicate Beta-BIQAE outperforms IQAE-CP via magnification of average quantum sample complexity results (left) and percentage improvement of Beta-BIQAE over IQAE-CP (right). Pauli strings are sorted in descending order of magnitude of the corresponding expectation value.

ple complexity at fixed target accuracy. Likewise, the schedule employed here to facilitate direct comparison between BIQAE and IQAE literature results could be substituted with a Bayesian-informed schedule (such as that of BAE [55]) to make the most of limited measurement information.

In order to realize BIQAE's gains, a key step would be to modify the algorithm for implementation on current and future quantum computers. The far-term method presented here could be readily modified for today's near-term intermediate scale quantum computers. Importantly, practical implementation of BIQAE with today's quantum computers will require consideration of current realistic noise conditions. Under the noise model considered in BAE [55], measurement outcome probabilities decay exponentially toward $p_k = \frac{1}{2}$ as the number of Grover operators increases $k \to \infty$, such that there exists a fundamental tradeoff that accompanies the addition of Grover operators between stochastic noise reduction and biased noise amplification in the absence of error correction. This tradeoff presently prevents IQAE from fully achieving its predicted efficiency gains, and would therefore be likely to similarly impact a naive implementation of BIQAE in the high Grover operator number limit on near-term quantum computers. Noise-aware schedules and restrictions on the number of Grover operators that have increased the robustness of prior QAE variants [53, 55] provide a possible route to circumvent such obstacles in BIQAE, a direction that we leave as an open question for future research. More generally, strategies may be imported from QAE methods recently successfully implemented on both superconducting and trapped-ion quantum computers [110–117, 21]. These strategies may straightforwardly be used in conjunction with the most recent NISQ accom-

modations of RAE [53] and BAE [55]. Ultimately, the power of BIQAE could be demonstrated without modification on far-term quantum computers if sufficiently reliable logical qubits and low-error gates are developed. Based on the results, it would be fruitful to identify new avenues for Bayesian inference to accelerate the frequentist tasks ubiquitous in quantum approaches to mathematical and scientific applications [14–27, 29–52].

6 Acknowledgments

M. B. S. acknowledges support for this research provided by the Office of the Vice Chancellor for Research and Graduate Education (OVCRGE) at the University of Wisconsin-Madison with funding from the Wisconsin Alumni Research Foundation (WARF). Y. W. acknowledges partial support for this research by OVCRGE from WARF and NSF Grant DMS-2514240. AI was used to refine early drafts of the manuscript.

References

[1] Gilles Brassard, Peter Høyer, Michele Mosca, and Alain Tapp. "Quantum amplitude amplification and estimation". In Samuel J Lomonaco, Jr. and Howard E Brandt, editors, Quantum Computation and Information (Washington, DC, 2000). Volume 305, pages 53–74. AMS Contemporary Mathematics (2002).

[2] Lov K. Grover. "A framework for fast quantum mechanical algorithms". In Proceedings of the Thirtieth Annual ACM Symposium on Theory of Computing. Pages 53–62. Association for Computing Machinery (1998).

[3] Daniel S. Abrams and Colin P. Williams. "Fast quantum algorithms for numerical integrals and stochastic processes" (1999). arXiv:quant-ph/9908083.

[4] Chu-Ryang Wie. "Simpler quantum counting". Quantum Information and Computation **19**, 967–983 (2019).

[5] Scott Aaronson and Patrick Rall. "Quantum approximate counting, simplified". In Symposium on Simplicity in Algorithms. Pages 24–32. SIAM (2020).

[6] Kouhei Nakaji. "Faster amplitude estimation". Quantum Information and Computation **20**, 1109–1123 (2020).

[7] Yohichi Suzuki, Shumpei Uno, Rudy Raymond, Tomoki Tanaka, Tamiya Onodera, and Naoki Yamamoto. "Amplitude estimation without phase estimation". Quantum Information Processing **19**, 1–17 (2020).

[8] Aram W. Harrow and Annie Y. Wei. "Adaptive quantum simulated annealing for Bayesian inference and estimating partition functions". In Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms. Pages 193–212. SIAM (2020).

[9] Srinivasan Arunachalam, Vojtech Havlicek, Giacomo Nannicini, Kristan Temme, and Pawel Wocjan. "Simpler (classical) and faster (quantum) algorithms for Gibbs partition functions". Quantum **6**, 789 (2022).

[10] Patrick Rall and Bryce Fuller. "Amplitude estimation from quantum signal processing". Quantum **7**, 937 (2023).

[11] Kirill Plekhanov, Matthias Rosenkranz, Mattia Fiorentini, and Michael Lubasch. "Variational quantum amplitude estimation". Quantum **6**, 670 (2022).

[12] Tudor Giurgica-Tiron, Iordanis Kerenidis, Farrokh Labib, Anupam Prakash, and William Zeng. "Low depth algorithms for quantum amplitude estimation". Quantum **6**, 745 (2022).

[13] Adam Callison and Dan E. Browne. "Improved maximum-likelihood quantum amplitude estimation" (2023). arXiv:2209.03321.

[14] Ashley Montanaro. "Quantum speedup of Monte Carlo methods". Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences **471**, 20150301 (2015).

[15] Kwangmin Yu, Hyunkyung Lim, and Pooja Rao. "Practical numerical integration on NISQ devices". In Quantum Information Science, Sensing, and Computation XII. Volume 11391, page 1139106. SPIE (2020).

[16] Pooja Rao, Kwangmin Yu, Hyunkyung Lim, Dasol Jin, and Deokkyu Choi. "Amplitude estimation algorithms and implementations on noisy intermediate-scale quantum devices". In OSA Quantum 2.0 Con-

ference. Page QW6A.16. Optica Publishing Group (2020).

[17] Ivan Kassal, Stephen P. Jordan, Peter J. Love, Masoud Mohseni, and Alán Aspuru-Guzik. "Polynomial-time quantum algorithm for the simulation of chemical dynamics". Proceedings of the National Academy of Sciences **105**, 18681–18686 (2008).

[18] Jakob Günther, Alberto Baiardi, Markus Reiher, and Matthias Christandl. "More quantum chemistry with fewer qubits". Physical Review Research **6**, 043021 (2024).

[19] Thomas E. Baker and David Poulin. "Density functionals and Kohn-Sham potentials with minimal wavefunction preparations on a quantum computer". Physical Review Research **2**, 043238 (2020).

[20] Peter D. Johnson, Alexander A. Kunitsa, Jérôme F. Gonthier, Maxwell D. Radin, Corneliu Buda, Eric J. Doskocil, Clena M. Abuan, and Jhonathan Romero. "Reducing the cost of energy estimation in the variational quantum eigensolver algorithm with robust amplitude estimation" (2022). arXiv:2203.07275.

[21] Alexander Kunitsa, Nicole Bellonzi, Shangjie Guo, Jérôme F. Gonthier, Corneliu Buda, Clena M. Abuan, and Jhonathan Romero. "Experimental demonstration of robust amplitude estimation on near-term quantum devices for chemistry applications" (2024). arXiv:2410.00686.

[22] Gabriele Agliardi, Michele Grossi, Mathieu Pellen, and Enrico Prati. "Quantum integration of elementary particle processes". Physics Letters B **832**, 137228 (2022).

[23] Jorge J. Martínez de Lejarza, Michele Grossi, Leandro Cieri, and Germán Rodrigo. "Quantum Fourier iterative amplitude estimation". In 2023 IEEE International Conference on Quantum Computing and Engineering (QCE). Volume 01, pages 571–579. IEEE (2023).

[24] Koichi Miyamoto, Soichiro Yamazaki, Fumio Uchida, Kotaro Fujisawa, and Naoki Yoshida. "Quantum algorithm for the vlasov simulation of the large-scale structure formation with massive neutrinos". Physical Review Research **6**, 013200 (2024).

[25] Jorge J. Martínez de Lejarza, David F. Rentería-Estrada, Michele Grossi, and Germán Rodrigo. "Quantum integration of decay rates at second order in perturbation theory". Quantum Science and Technology **10**, 025026 (2025).

[26] Jorge J. Martínez de Lejarza, Leandro Cieri, Michele Grossi, Sofia Vallecorsa, and Germán Rodrigo. "Loop Feynman integration on a quantum computer". Physical Review D **110**, 074031 (2024).

[27] Euimin Lee, Sangmin Lee, and Shiho Kim. "Quantum walk based Monte Carlo simulation for photon interaction cross sections". Physical Review D **111**, 116001 (2025).

[28] Ifan Williams and Mathieu Pellen. "A general approach to quantum integration of cross sections in high-energy physics". Quantum Science and Technology **10**, 045017 (2025).

[29] Sonika Johri, Damian S. Steiger, and Matthias Troyer. "Entanglement spectroscopy on a quantum computer". Physical Review B **96**, 195136 (2017).

[30] Patrick Rall. "Quantum algorithms for estimating physical quantities using block encodings". Physical Review A **102**, 022408 (2020).

[31] Yulong Dong, Lin Lin, and Yu Tong. "Ground-state preparation and energy estimation on early fault-tolerant quantum computers via quantum eigenvalue transformation of unitary matrices". PRX Quantum **3**, 040305 (2022).

[32] Lorenzo Piroli, Georgios Styliaris, and J. Ignacio Cirac. "Approximating many-body quantum states with quantum circuits and measurements". Physical Review Letters **133**, 230401 (2024).

[33] Anjali A. Agrawal, Joshua Job, Tyler L. Wilson, S. N. Saadatmand, Mark J. Hodson, Josh Y. Mutus, Athena Caesura, Peter D. Johnson, Justin E. Elenewski, Kaitlyn J. Morrell, and Alexander F. Kemper. "Quantifying fault tolerant simulation of strongly correlated systems using the Fermi-Hubbard model" (2024). arXiv:2406.06511.

[34] Frank Gaitan. "Finding flows of a Navier–Stokes fluid through quantum computing". npj Quantum Information **6**, 61 (2020).

[35] Sachin S. Bharadwaj and Katepalli R.

Sreenivasan. "Hybrid quantum algorithms for flow problems". Proceedings of the National Academy of Sciences **120**, e2311014120 (2023).

[36] John Penuel, Amara Katabarwa, Peter D. Johnson, Parker Kuklinski, Benjamin Rempfer, Collin Farquhar, Yudong Cao, and Michael C. Garrett. "Detailed assessment of calculating drag force with quantum computers: Explicit time-evolution precludes exponential advantage for nonlinear differential equations" (2024). arXiv:2406.06323.

[37] Frank Gaitan. "Circuit implementation of oracles used in a quantum algorithm for solving nonlinear partial differential equations". Physical Review A **109**, 032604 (2024).

[38] Nikitas Stamatopoulos, Daniel J. Egger, Yue Sun, Christa Zoufal, Raban Iten, Ning Shen, and Stefan Woerner. "Option pricing using quantum computers". Quantum **4**, 291 (2020).

[39] Adam Bouland, Wim van Dam, Hamed Joorati, Iordanis Kerenidis, and Anupam Prakash. "Prospects and challenges of quantum finance" (2020). arXiv:2011.06492.

[40] Nikitas Stamatopoulos, Guglielmo Mazzola, Stefan Woerner, and William J. Zeng. "Towards quantum advantage in financial market risk using quantum gradient algorithms". Quantum **6**, 770 (2022).

[41] Javier Alcazar, Andrea Cadarso, Amara Katabarwa, Marta Mauri, Borja Peropadre, Guoming Wang, and Yudong Cao. "Quantum algorithm for credit valuation adjustments". New Journal of Physics **24**, 023036 (2022).

[42] Dylan Herman, Cody Googin, Xiaoyuan Liu, Yue Sun, Alexey Galda, Ilya Safro, Marco Pistoia, and Yuri Alexeev. "Quantum computing for finance". Nature Reviews Physics **5**, 450–465 (2023).

[43] Guoming Wang and Angus Kan. "Option pricing under stochastic volatility on a quantum computer". Quantum **8**, 1504 (2024).

[44] Stefan Woerner and Daniel J. Egger. "Quantum risk analysis". npj Quantum Information **5**, 15 (2019).

[45] Andrés Gómez, Álvaro Leitao, Alberto Manzano, Daniele Musso, María R. Nogueiras, Gustavo Ordóñez, and Carlos Vázquez. "A survey on quantum computational finance for derivatives pricing and VaR". Archives of Computational Methods in Engineering **29**, 4137–4163 (2022).

[46] Patrick Rebentrost, Brajesh Gupt, and Thomas R. Bromley. "Quantum computational finance: Monte Carlo pricing of financial derivatives". Physical Review A **98**, 022321 (2018).

[47] Koichi Miyamoto. "Bermudan option pricing by quantum amplitude estimation and Chebyshev interpolation". EPJ Quantum Technology **9**, 1–27 (2022).

[48] Daniel J. Egger, Ricardo García Gutiérrez, Jordi Cahué Mestre, and Stefan Woerner. "Credit risk analysis using quantum computers". IEEE Transactions on Computers **70**, 2136–2145 (2021).

[49] Román Orús, Samuel Mugel, and Enrique Lizaso. "Quantum computing for finance: Overview and prospects". Reviews in Physics **4**, 100028 (2019).

[50] M. C. Braun, T. Decker, N. Hegemann, S. F. Kerstan, and C. Schäfer. "A quantum algorithm for the sensitivity analysis of business risks" (2021). arXiv:2103.05475.

[51] Nathan Wiebe, Ashish Kapoor, and Krysta M. Svore. "Quantum deep learning". Quantum Information and Computation **16**, 541–587 (2016).

[52] Simon Wiedemann, Daniel Hein, Steffen Udluft, and Christian B. Mendl. "Quantum policy iteration via amplitude estimation and Grover search—towards quantum advantage for reinforcement learning". Transactions on Machine Learning Research (2023). url: https://openreview.net/forum?id=HG11PAmwQ6.

[53] Guoming Wang, Dax Enshan Koh, Peter D. Johnson, and Yudong Cao. "Minimizing estimation runtime on noisy quantum computers". PRX Quantum **2**, 010346 (2021).

[54] Dax Enshan Koh, Guoming Wang, Peter D. Johnson, and Yudong Cao. "Foundations for Bayesian inference with engineered likelihood functions for robust amplitude estimation". Journal of Mathematical Physics **63**, 052202 (2022).

[55] Alexandra Ramôa and Luis Paulo Santos.

"Bayesian quantum amplitude estimation". Quantum **9**, 1856 (2025).

[56] Yan Wang. "A quantum approximate Bayesian optimization algorithm for continuous problems". In Proceedings of the IISE Annual Conference and Expo. Pages 235–240. Institute of Industrial and Systems Engineers (2021). url: https://www.proceedings.com/content/061/061116webtoc.pdf.

[57] Samuel Stein, Nathan Wiebe, Yufei Ding, James Ang, and Ang Li. "Quantum Bayesian error mitigation employing poisson modelling over the Hamming spectrum for quantum error mitigation" (2022). arXiv:2207.07237.

[58] Simone Tibaldi, Davide Vodola, Edoardo Tignone, and Elisa Ercolessi. "Bayesian optimization for QAOA". IEEE Transactions on Quantum Engineering **4**, 1–11 (2023).

[59] Jungin E. Kim and Yan Wang. "Quantum approximate Bayesian optimization algorithms with two mixers and uncertainty quantification". IEEE Transactions on Quantum Engineering **4**, 1–17 (2023).

[60] Yanqi Song, Yusen Wu, Sujuan Qin, Qiaoyan Wen, Jingbo B. Wang, and Fei Gao. "Trainability analysis of quantum optimization algorithms from a Bayesian lens" (2023). arXiv:2310.06270.

[61] Zhimin He, Hongxiang Chen, Yan Zhou, Haozhen Situ, Yongyao Li, and Lvzhou Li. "Self-supervised representation learning for Bayesian quantum architecture search". Physical Review A **111**, 032403 (2025).

[62] Nathan Wiebe and Chris Granade. "Efficient Bayesian phase estimation". Physical Review Letters **117**, 010503 (2016).

[63] Stefano Paesani, Andreas A. Gentile, Raffaele Santagati, Jianwei Wang, Nathan Wiebe, David P. Tew, Jeremy L. O'Brien, and Mark G. Thompson. "Experimental Bayesian quantum phase estimation on a silicon photonic chip". Physical Review Letters **118**, 100503 (2017).

[64] Yan Li, Luca Pezzè, Manuel Gessner, Zhihong Ren, Weidong Li, and Augusto Smerzi. "Frequentist and Bayesian quantum phase estimation". Entropy **20**, 628 (2018).

[65] Thomas E. O'Brien, Brian Tarasinski, and Barbara M. Terhal. "Quantum phase estimation of multiple eigenvalues for small-scale (noisy) experiments". New Journal of Physics **21**, 023022 (2019).

[66] Fernando Martínez-García, Davide Vodola, and Markus Müller. "Adaptive Bayesian phase estimation for quantum error correcting codes". New Journal of Physics **21**, 123027 (2019).

[67] Yuxiang Qiu, Min Zhuang, Jiahao Huang, and Chaohong Lee. "Bayesian phase estimation via active learning" (2021). arXiv:2107.00196.

[68] Ewout van den Berg. "Efficient Bayesian phase estimation using mixed priors". Quantum **5**, 469 (2021).

[69] Valentin Gebhart, Augusto Smerzi, and Luca Pezzè. "Bayesian quantum multiphase estimation algorithm". Physical Review Applied **16**, 014035 (2021).

[70] Yuxiang Qiu, Min Zhuang, Jiahao Huang, and Chaohong Lee. "Efficient Bayesian phase estimation via entropy-based sampling". Quantum Science and Technology **7**, 035022 (2022).

[71] Joseph G. Smith, Crispin H. W. Barnes, and David R. M. Arvidsson-Shukur. "Adaptive Bayesian quantum algorithm for phase estimation". Physical Review A **109**, 042412 (2024).

[72] Kentaro Yamamoto, Samuel Duffield, Yuta Kikuchi, and David Muñoz Ramo. "Demonstrating Bayesian quantum phase estimation with quantum error detection". Physical Review Research **6**, 013221 (2024).

[73] Travis Hurant, Ke Sun, Zhubing Jia, Jungsang Kim, and Kenneth R. Brown. "Few-shot, robust calibration of single qubit gates using Bayesian robust phase estimation". In 2024 IEEE International Conference on Quantum Computing and Engineering (QCE). Volume 1, pages 1244–1253. IEEE (2024).

[74] Boyu Zhou, Saikat Guha, and Christos N. Gagatsos. "Bayesian quantum phase estimation with fixed photon states". Quantum Information Processing **23**, 366 (2024).

[75] Su Direkci, Ran Finkelstein, Manuel Endres, and Tuvia Gefen. "Heisenberg-limited Bayesian phase estimation with

low-depth digital quantum circuits" (2024). arXiv:2407.06006.

[76] Kenji Sugisaki, Chikako Sakai, Kazuo Toyota, Kazunobu Sato, Daisuke Shiomi, and Takeji Takui. "Bayesian phase difference estimation: a general quantum algorithm for the direct calculation of energy gaps". Physical Chemistry Chemical Physics **23**, 20152–20162 (2021).

[77] Kenji Sugisaki, Hiroyuki Wakimoto, Kazuo Toyota, Kazunobu Sato, Daisuke Shiomi, and Takeji Takui. "Quantum algorithm for numerical energy gradient calculations at the full configuration interaction level of theory". Journal of Physical Chemistry Letters **13**, 11105–11111 (2022).

[78] Kenji Sugisaki. "Projective measurement-based quantum phase difference estimation algorithm for the direct computation of eigenenergy differences on a quantum computer". Journal of Chemical Theory and Computation **19**, 7617–7625 (2023).

[79] Jarrod R. McClean, Jonathan Romero, Ryan Babbush, and Alán Aspuru-Guzik. "The theory of variational hybrid quantum-classical algorithms". New Journal of Physics **18**, 023023 (2016).

[80] Daochen Wang, Oscar Higgott, and Stephen Brierley. "Accelerated variational quantum eigensolver". Physical Review Letters **122**, 140504 (2019).

[81] Hikaru Wakaura, Takao Tomono, and Shoya Yasuda. "Evaluation on genetic algorithms as an optimizer of variational quantum eigensolver (VQE) method" (2021). arXiv:2110.07441.

[82] Giovanni Iannelli and Karl Jansen. "Noisy Bayesian optimization for variational quantum eigensolvers" (2021). arXiv:2112.00426.

[83] Chris N. Self, Kiran E. Khosla, Alistair W. R. Smith, Frédéric Sauvage, Peter D. Haynes, Johannes Knolle, Florian Mintert, and M. S. Kim. "Variational quantum algorithm with information sharing". npj Quantum Information **7**, 116 (2021).

[84] Kim Nicoli, Christopher J. Anders, Lena Funcke, Tobias Hartung, Karl Jansen, Stefan Kühn, Klaus-Robert Müller, Paolo Stornati, Pan Kessel, and Shinichi Nakajima. "Physics-informed Bayesian optimization of variational quantum circuits". In A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine, editors, Advances in Neural Information Processing Systems. Volume 36, pages 18341–18376. Curran Associates, Inc. (2023). url: https://proceedings.neurips.cc/paper_files/paper/2023/file/3adb85a348a18cdd74ce99fbbab20301-Paper-Conference.pdf.

[85] Farshud Sorourifar, Mohamed Taha Rouabah, Nacer Eddine Belaloui, Mohamed Messaoud Louamri, Diana Chamaki, Erik J. Gustafson, Norm M. Tubman, Joel A. Paulson, and David E. Bernal Neira. "Toward efficient quantum computation of molecular ground-state energies". AIChE Journal **71**, e18887 (2025).

[86] Milena Röhrs, Alexey Bochkarev, and Arcesio C. Medina. "Bayesian optimisation with improved information sharing for the variational quantum eigensolver" (2024). arXiv:2405.14353.

[87] Tao Jiang, John Rogers, Marius S. Frank, Ove Christiansen, Yong-Xin Yao, and Nicola Lanatà. "Error mitigation in variational quantum eigensolvers using tailored probabilistic machine learning". Physical Review Research **6**, 033069 (2024).

[88] Trung Huynh, Gwangil An, Minsu Kim, Yu-Seong Jeon, and Jinhyoung Lee. "A variational quantum algorithm by Bayesian inference with von Mises-Fisher distribution" (2024). arXiv:2410.03130.

[89] Samuele Pedrielli, Christopher J. Anders, Lena Funcke, Karl Jansen, Kim A. Nicoli, and Shinichi Nakajima. "Bayesian parameter shift rule in variational quantum eigensolvers" (2025). arXiv:2502.02625.

[90] Elijah Pelofske, Georg Hahn, and Hristo N. Djidjev. "Advanced anneal paths for improved quantum annealing". In 2020 IEEE International Conference on Quantum Computing and Engineering (QCE). Pages 256–266. IEEE (2020).

[91] Jernej Rudi Finžgar, Martin J. A. Schuetz, J. Kyle Brubaker, Hidetoshi Nishimori, and Helmut G. Katzgraber. "Designing quantum annealing schedules using Bayesian optimization". Physical Review Research **6**, 023063 (2024).

[92] Sophia Simon, Matthias Degroote, Nikolaj Moll, Raffaele Santagati, Michael Streif, and Nathan Wiebe. "Amplified amplitude estimation: Exploiting prior knowledge to improve estimates of expectation values" (2024). arXiv:2402.14791.

[93] Dmitry Grinko, Julien Gacon, Christa Zoufal, and Stefan Woerner. "Iterative quantum amplitude estimation". npj Quantum Information **7**, 52 (2021).

[94] Qilin Li, Atharva Vidwans, Yazhen Wang, and Micheline B. Soley (2025). Code: Kirin0570/BIQAE.

[95] Alexandra Ramôa (2025). Code: alexandrafrca/BAE.

[96] Michael A. Nielsen and Isaac L. Chuang. "Quantum computation and quantum information". Cambridge University Press. (2010). 10th anniversary edition edition.

[97] P. Jordan and E. Wigner. "Über das Paulische Äquivalenzverbot". Zeitschrift fur Physik **47**, 631–651 (1928).

[98] Sergey Bravyi, Jay M. Gambetta, Antonio Mezzacapo, and Kristan Temme. "Tapering off qubits to simulate fermionic hamiltonians" (2017). arXiv:1701.08213.

[99] Ilya G. Ryabinkin, Scott N. Genin, and Artur F. Izmaylov. "Constrained variational quantum eigensolver: Quantum computer search engine in the Fock space". Journal of Chemical Theory and Computation **15**, 249–255 (2019).

[100] Qiming Sun, Timothy C. Berkelbach, Nick S. Blunt, George H. Booth, Sheng Guo, Zhendong Li, Junzi Liu, James D. McClain, Elvira R. Sayfutyarova, Sandeep Sharma, Sebastian Wouters, and Garnet Kin-Lic Chan. "PySCF: the Python-based simulations of chemistry framework". WIREs Computational Molecular Science **8**, e1340 (2018).

[101] Sam McArdle, Suguru Endo, Alán Aspuru-Guzik, Simon C. Benjamin, and Xiao Yuan. "Quantum computational chemistry". Reviews of Modern Physics **92**, 015003 (2020).

[102] Ali Javadi-Abhari, Matthew Treinish, Kevin Krsulich, Christopher J. Wood, Jake Lishman, Julien Gacon, Simon Martiel, Paul D. Nation, Lev S. Bishop, Andrew W. Cross, Blake R. Johnson, and Jay M. Gambetta. "Quantum computing with Qiskit" (2024). arXiv:2405.08810.

[103] Irma Avdic and David A. Mazziotti. "Enhanced shadow tomography of molecular excited states via the enforcement of n-representability conditions by semidefinite programming". Physical Review A **110**, 052407 (2024).

[104] Tony Cai, Donggyu Kim, Yazhen Wang, Ming Yuan, and Harrison H. Zhou. "Optimal large-scale quantum state tomography with Pauli measurements". Annals of Statistics **44**, 682–712 (2016).

[105] Yazhen Wang. "Asymptotic equivalence of quantum state tomography and noisy matrix completion". Annals of Statistics **41**, 2462–2504 (2013).

[106] William R. Thompson. "On the likelihood that one unknown probability exceeds another in view of the evidence of two samples". Biometrika **25**, 285–294 (1933).

[107] Harold J. Kushner. "A new method of locating the maximum point of an arbitrary multipeak curve in the presence of noise". Journal of Basic Engineering **86**, 97–106 (1964).

[108] Jonas Močkus. "On Bayesian methods for seeking the extremum". In IFIP Technical Conference on Optimization Techniques. Pages 400–404. Springer (1974).

[109] Bobak Shahriari, Kevin Swersky, Ziyu Wang, Ryan P. Adams, and Nando De Freitas. "Taking the human out of the loop: A review of Bayesian optimization". Proceedings of the IEEE **104**, 148–175 (2015).

[110] Pooja Rao, Kwangmin Yu, Hyunkyung Lim, Dasol Jin, and Deokkyu Choi. "Quantum amplitude estimation algorithms on IBM quantum devices". In Quantum Communications and Quantum Imaging XVIII. Volume 11507, page 115070O. SPIE (2020).

[111] Almudena Carrera Vazquez and Stefan Woerner. "Efficient state preparation for quantum amplitude estimation". Physical Review Applied **15**, 034027 (2021).

[112] Tomoki Tanaka, Yohichi Suzuki, Shumpei Uno, Rudy Raymond, Tamiya Onodera, and Naoki Yamamoto. "Amplitude estimation via maximum likelihood on noisy quantum computer". Quantum Information Processing **20**, 293 (2021).

[113] Tomoki Tanaka, Shumpei Uno, Tamiya

Onodera, Naoki Yamamoto, and Yohichi Suzuki. "Noisy quantum amplitude estimation without noise estimation". Physical Review A **105**, 012411 (2022).

[114] Salvatore Certo, Anh Dung Pham, and Daniel Beaulieu. "Benchmarking amplitude estimation on a superconducting quantum computer" (2022). arXiv:2201.06987.

[115] Tudor Giurgica-Tiron, Sonika Johri, Iordanis Kerenidis, Jason Nguyen, Neal Pisenti, Anupam Prakash, Ksenia Sosnova, Ken Wright, and William Zeng. "Low-depth amplitude estimation on a trapped-ion quantum computer". Physical Review Research **4**, 033034 (2022).

[116] Archismita Dalal and Amara Katabarwa. "Noise tailoring for robust amplitude estimation". New Journal of Physics **25**, 023015 (2023).

[117] Steven Herbert, Ifan Williams, Roland Guichard, and Darren Ng. "Noise-aware quantum amplitude estimation". IEEE Transactions on Quantum Engineering **5**, 1–23 (2024).

[118] Yunpeng Zhao, Haiyan Wang, Kuai Xu, Yue Wang, Ji Zhu, and Feng Wang. "Adaptive algorithm for quantum amplitude estimation" (2022). arXiv:2206.08449.

[119] Jacob T. Seeley, Martin J. Richard, and Peter J. Love. "The Bravyi-Kitaev transformation for quantum computation of electronic structure". Journal of Chemical Physics **137**, 224109 (2012).

[120] Alberto Peruzzo, Jarrod McClean, Peter Shadbolt, Man-Hong Yung, Xiao-Qi Zhou, Peter J. Love, Alán Aspuru-Guzik, and Jeremy L. O'Brien. "A variational eigenvalue solver on a photonic quantum processor". Nature Communications **5**, 4213 (2014).

[121] Yudong Cao, Jonathan Romero, Jonathan P. Olson, Matthias Degroote, Peter D. Johnson, Mária Kieferová, Ian D. Kivlichan, Tim Menke, Borja Peropadre, Nicolas P. D. Sawaya, Sukin Kim, Libor Veis, and Alán Aspuru-Guzik. "Quantum chemistry in the age of quantum computing". Chemical Reviews **119**, 10856–10915 (2019).

[122] Halbert White. "Maximum likelihood estimation of misspecified models". Econometrica **50**, 1–25 (1982).

A Algorithmic and Computational Details

A.1 Quantum Sample Complexity of Classical QAE

As derived in Appendix C.1, the quantum sample complexity of Classical QAE is formally stated in a statistical framework as follows:

Theorem 3. *Consider n instances of the quantum circuit associated with oracle $\mathcal{A}$. Let X_i denote the outcome of the i^{th} instance, where the state $|0\rangle$ is mapped to the outcome 0 and $|1\rangle$ is mapped to 1. Then, the outcomes $X_1, \ldots, X_n$ are independent and identically distributed (i.i.d.) Bernoulli random variables with success probability $p = a$.*

The sample mean $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ serves as the Uniformly Minimum Variance Unbiased Estimator (UMVUE) for a. The mean squared error (MSE) of $\bar{X}$ with respect to a is

$$MSE_{\bar{X}}(a) = \frac{1}{n}a(1-a).$$

To achieve a target accuracy ε, the required number of shots is

$$N_{shots} = \frac{1}{\varepsilon^2}a(1-a).$$

Since each shot requires one query to the oracle $\mathcal{A}$, the total number of accesses to the oracle is then

$$N_{oracle} = N_{shots} = \frac{1}{\varepsilon^2}a(1-a).$$

A.2 Mechanistic Foundations of Amplitude Estimation

To aid readers new to the field, we provide the following brief, simplified example of how amplitude amplification improves measurement efficiency:

Example 1. *Suppose one seeks to estimate the head probability of a coin while in possession of (i) the ability to estimate the head probability of a biased coin with accuracy 0.02 and (ii) a technique that amplifies the head probability by a factor of $\times 4$. If the unknown head probability is 0.15 and one tosses the coin repeatedly, one will obtain an interval estimate of $[0.13, 0.17]$ without amplification or $[0.58, 0.62]$ with amplification.*

At first glance, the interval lengths appear to be equal for both approaches. However, the interval in the amplified case must be rescaled by the inverse of the original factor (namely, $\times 1/4$) to reorient to the original head probability. The interval after rescaling is $[0.145, 0.155]$, which is four times *narrower than the interval obtained without amplification. Thus, using the amplified approach improves accuracy in the sample case by a factor of four.*

A.3 Identifiability Challenge of Amplitude Estimation

Fig. 10 provides a visual demonstration of the need for additional information to identify the unique amplitude to which a measured probability corresponds in amplitude-amplification-based QAE methods. Given the relationship between the amplified angle and the amplified target probability Eq. (3), there exist multiple angles that correspond to each observed probability Eq. (4). However, these individual solutions may be distinguished uniquely by specification of their base or quadrant index Eq. (5).

A.4 Resolution of the Identifiability Challenge with Arbitrary Quadrant Index

To demonstrate that the quadrant index $l(k)$ need not be zero to uniquely estimate the quantum amplitude a (a requirement of Theorem 1), we relax the small-angle assumption of Assumption 1 to mere knowledge of the value of $l(k)$, as follows:

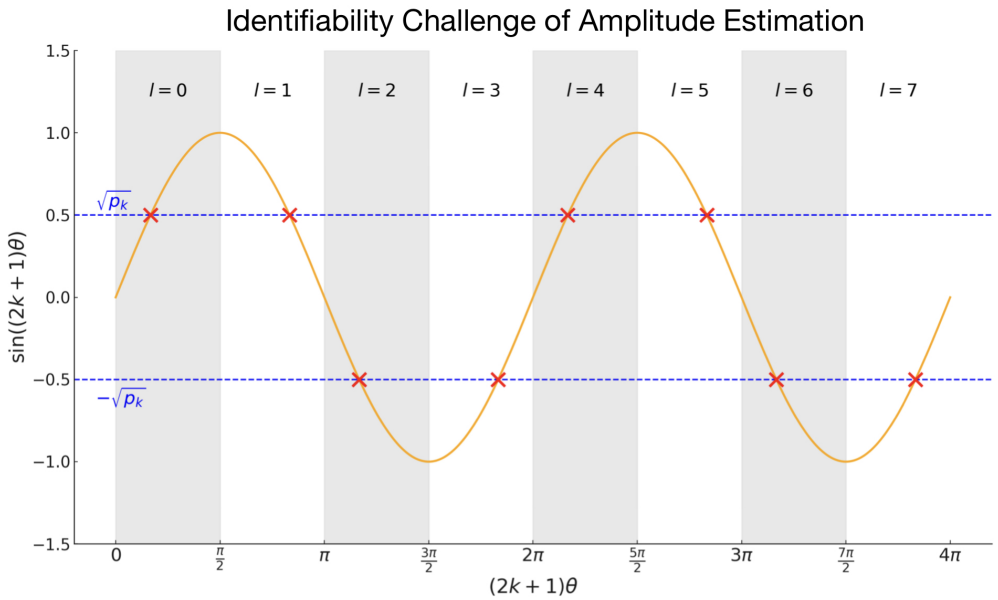


Figure 10: Illustration of the identifiability challenge in amplitude estimation, in which the relationship between the amplified angle and amplified target probability (yellow solid line) entails multiple valid solutions (red points) for each measured amplitude ($p_k = 0.5$ shown, blue dashed line), which are distinguished uniquely by their base or quadrant index (from left to right, $l = 0, 1, 2, 3, \ldots$, alternating shaded gray and white vertical areas).

Assumption 3. *For a given k, the corresponding quadrant index $l(k)$ is known.*

The proof provided in Appendix C.1 then indicates that Theorem 1 extends to

Theorem 4. *Suppose Assumption 3 holds with a certain k. Then,*

$$\hat{a} := \begin{cases} \sin^2\left[\frac{1}{2k+1}\left(\arcsin\left(\sqrt{\bar{X}_k}\right) + l(k)\cdot\frac{\pi}{2}\right)\right] & \text{if } l(k) \text{ is even,} \\ \sin^2\left[\frac{1}{2k+1}\left(\arccos\left(\sqrt{\bar{X}_k}\right) + l(k)\cdot\frac{\pi}{2}\right)\right] & \text{if } l(k) \text{ is odd} \end{cases}$$

is the Maximum Likelihood Estimator (MLE) for a. The corresponding asymptotic MSE is

$$AMSE_{\hat{a}}(a) = \frac{1}{(2k+1)^2}\frac{1}{N_k}a(1-a),$$

and the required number of accesses to the oracle $\mathcal{A}$ to achieve a target accuracy ε is

$$N_{oracle} = \frac{1}{2k+1}\frac{1}{\varepsilon^2}a(1-a).$$

Note this theorem not only establishes that amplified estimation is feasible for arbitrary amplitudes given prior knowledge of the quadrant index, but also that the estimator increases in efficiency as k grows—insights implicit in the design of IQAE [93] that are broadly applicable to the wider class of amplified-estimation-based QAE methods.

A.5 Pictorial Representation and Accumulated Complexity Bound of Base-3/Base-5 Hybrid-Scheduled IQAE

To overcome the problem that base-3 hybrid-scheduled IQAE struggles where the amplified angle $K_t\theta$ falls too close to the reference interval boundary (due to the increasing difficulty of constructing an interval estimate fully contained in a single reference interval as the amplified angle nears said boundaries), we introduce a second set of reference intervals in base 5, as depicted in Fig. 11. Introduction

of the second set of reference intervals ensures that the interval estimate must be fully contained in a single reference interval either in base 3 or 5 whenever its length is less than the smallest separation between reference interval boundaries (specifically, 6 degrees). Replacement of the Assumption 2 that underpins base-3 hybrid-scheduled IQAE with the lower bound on the radius of the confidence interval radius Eq. (9) then results in the following accumulated complexity bound, as derived and proven in Appendix C.4:

Theorem 5. *Using the base-3/base-5 K-schedule, the accumulated sample complexity to achieve a accuracy of ε with an overall confidence level of α (≤ 0.1) is upper bounded by Eq. (10).*

Note this strategy requires no additional ancilla, in contrast to the strategy of ref. [118].

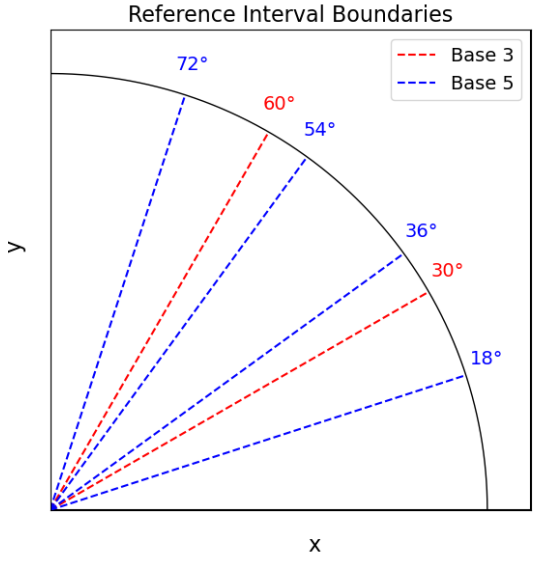


Figure 11: Reference interval boundaries in base 3 (red dashed lines) and base 5 (blue dashed lines) in base-3/base-5 hybrid-scheduled IQAE.

A.6 Pseudocode for Normal-BIQAE

Normal-BIQAE proceeds according to the workflow illustrated in Fig. 12 and modules provided in Algorithms 1–3, in which normal conjugate priors ensure all prior and posterior distributions remain normal throughout the process and only the mean and variance evolve as the method progresses.

Algorithm 1 BayesianUpdate

Require: Prior mean $\mu_{0,t}$ and variance Δ_t^2 for p_{k_t}, number of $\mathcal{Q}$ operators k_t, sample size N_t
Ensure: Posterior mean $\mu_{\text{post},t}$ and variance $\sigma^2_{\text{post},t}$ for p_{k_t}

1: Construct k_t $\mathcal{Q}$ operators into the circuit $\mathcal{A}$
2: Sample shots until a total of N_t shots (including existing shots at stage t) is collected. Store the sample average as $\bar{X}_t$
3: Estimate the population variance:

$$\hat{\sigma}_t^2 = \bar{X}_t(1 - \bar{X}_t)$$

4: Compute the posterior mean and variance:

$$\mu_{\text{post},t} = \frac{\frac{\mu_{0,t}}{\Delta_t^2} + \frac{\bar{X}_t}{\hat{\sigma}_t^2/N_t}}{\frac{1}{\Delta_t^2} + \frac{1}{\hat{\sigma}_t^2/N_t}}, \quad \sigma^2_{\text{post},t} = \left(\frac{1}{\Delta_t^2} + \frac{1}{\hat{\sigma}_t^2/N_t}\right)^{-1}$$

5: **Return** $(\mu_{\text{post},t}, \sigma^2_{\text{post},t})$

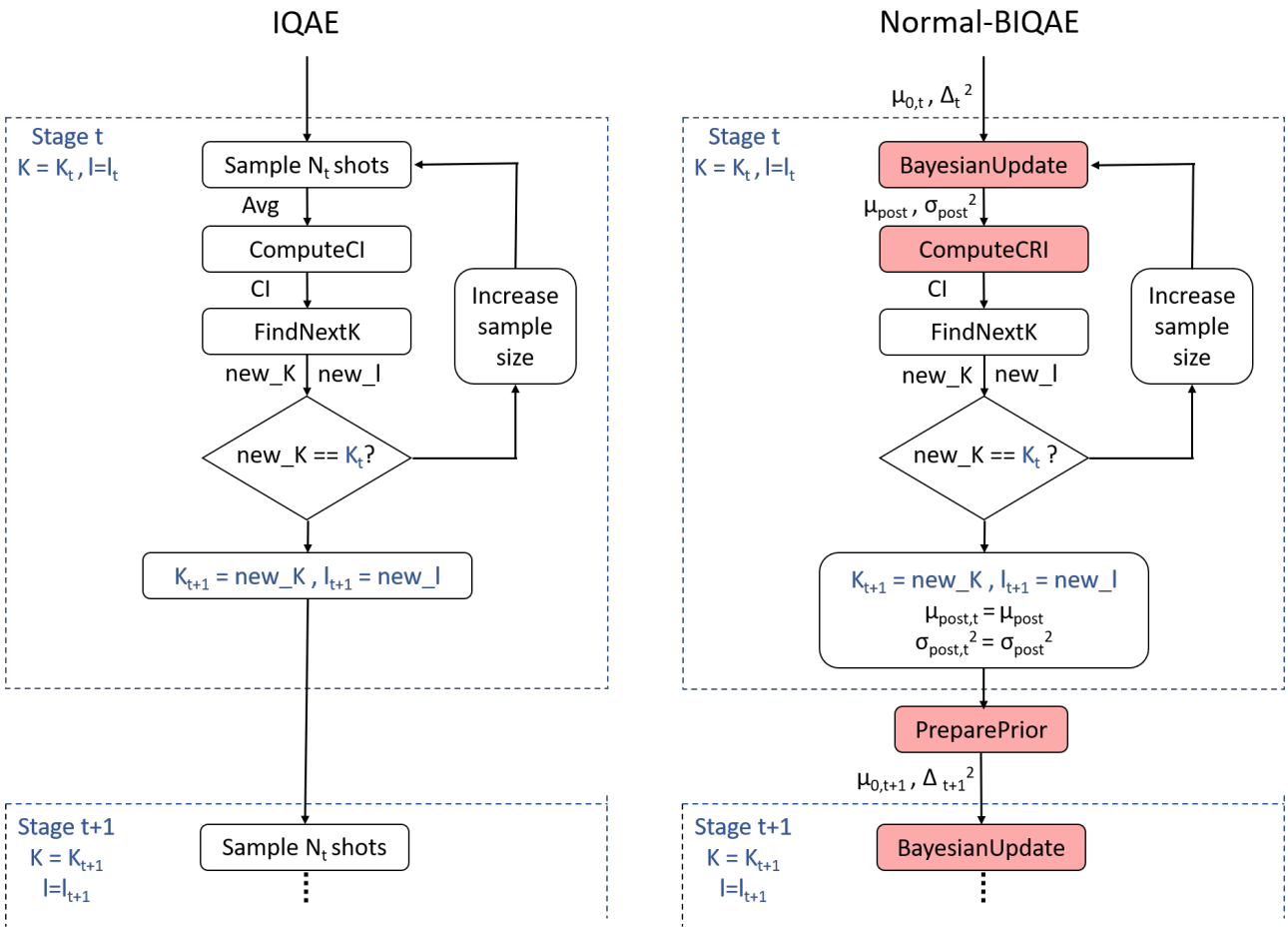


Figure 12: Detailed workflow of Normal-BIQAE (right), juxtaposed with IQAE for comparison (left). The three modules that distinguish BIQAE from IQAE are emphasized in red.

Algorithm 2 ComputeCRI

Require: Posterior $(\mu_{\text{post},t}, \sigma^2_{\text{post},t})$ for p_{k_t}, amplification factor K_t, quadrant index l_t, confidence level α_t

Ensure: An interval estimate $(\theta_{l,t}, \theta_{u,t})$ for θ at confidence level α_t

1: Compute the interval estimate for p_{k_t}:

$$\left[p^l_{k_t}, p^u_{k_t}\right] = \text{trunc}_{[0,1]}(\mu_{\text{post},t} \pm z_{\alpha_t/2} \cdot \sigma_{\text{post},t})$$

2: Map the interval estimate for p_{k_t} to an interval estimate for θ using f_t:

$$\left[\theta^l_t, \theta^u_t\right] = f_t([p^l_{k_t}, p^u_{k_t}])$$

where

$$f_t(x) := \begin{cases} \frac{1}{K_t}\left(\arcsin\left(\sqrt{x}\right) + l_t \cdot \frac{\pi}{2}\right) & \text{if } l_t \text{ is even,} \\ \frac{1}{K_t}\left(\arccos\left(\sqrt{x}\right) + l_t \cdot \frac{\pi}{2}\right) & \text{if } l_t \text{ is odd} \end{cases}$$

3: **Return** $\left[\theta^l_t, \theta^u_t\right]$

Algorithm 3 PreparePrior

Require: Posterior mean and variance $(\mu_{\text{post},t}, \sigma^2_{\text{post},t})$ for p_{k_t} obtained at the end of stage t, current and next amplification factors K_t, K_{t+1}, current quadrant index l_t

Ensure: Prior mean $\mu_{0,t+1}$ and variance Δ^2_{t+1} for $p_{k_{t+1}}$

1: Compute the prior mean for $p_{k_{t+1}}$:

$$\mu_{0,t+1} = \sin^2\left(K_{t+1} \cdot f_t(\mu_{\text{post},t})\right)$$

2: Compute the prior variance for $p_{k_{t+1}}$:

$$\Delta^2_{t+1} = \left(\frac{K_{t+1}}{K_t}\right)^2 \frac{\mu_{0,t+1}(1-\mu_{0,t+1})}{\mu_{\text{post},t}(1-\mu_{\text{post},t})} \cdot \sigma^2_{\text{post},t}$$

3: **Return** $(\mu_{0,t+1}, \Delta^2_{t+1})$

Theorem 6, which is proven in Appendix C.5, provides the theoretical guarantee that underlies Algorithm 3:

Theorem 6. *Let $\mu_{post,t} \in (0,1), \sigma^2_{post,t}$ denote the posterior mean and variance of p_{k_t} obtained at the end of stage t. Then, the prior distribution of $p_{k_{t+1}}$ can be approximated by a normal distribution with mean*

$$\mu_{0,t+1} = \sin^2\left(K_{t+1} \cdot f_t(\mu_{post,t})\right)$$

and variance

$$\Delta^2_{t+1} = \left(\frac{K_{t+1}}{K_t}\right)^2 \frac{\mu_{0,t+1}(1-\mu_{0,t+1})}{\mu_{post,t}(1-\mu_{post,t})} \cdot \sigma^2_{post,t}.$$

A.7 Quantum Sample Complexity Analysis for Normal-BIQAE

To determine the constant-factor reduction of the complexity bound of Normal-BIQAE relative to Normal-IQAE, we derive the sample complexity of a given stage (in place of the aforementioned accumulated sample complexity, see proof in Appendix C.6):

Theorem 7. *At stage t, the quantum sample complexity required to obtain an interval estimate for θ with a radius of ε_t is upper bounded by Eq. (22).*

Theorem 7 indicates a reduction in the upper bound of the quantum sample complexity of Normal-IQAE Eq. (23) of $\frac{z^2_{\alpha_t/2}}{4K_t}\frac{1}{\varepsilon^2_{t-1}}$, which is consistent with a Bayesian perspective in two respects: (i) The measurement cost reduction depends on the accuracy achieved in the previous stage, as expected given that BIQAE leverages prior information from earlier stages to enhance the efficiency of amplitude estimation at a current stage, and (ii) in the case that the interval estimate in the previous stage is sufficiently accurate that its confidence interval radius exceeds that of the current stage $\varepsilon_{t-1} \leq \varepsilon_t$, the upper bound becomes less than or equal to zero, which implies that no additional sampling is required and the algorithm may proceed directly to the next stage.

A.8 Pseudocode for Beta-BIQAE

In contrast to Normal-BIQAE, which updates the mean and variance of a normal distribution between iterations; Beta-BIQAE updates the shape parameters a and b of a beta distribution between iterations according to the workflow in Fig. 13 and modules in Algorithms 4–6.

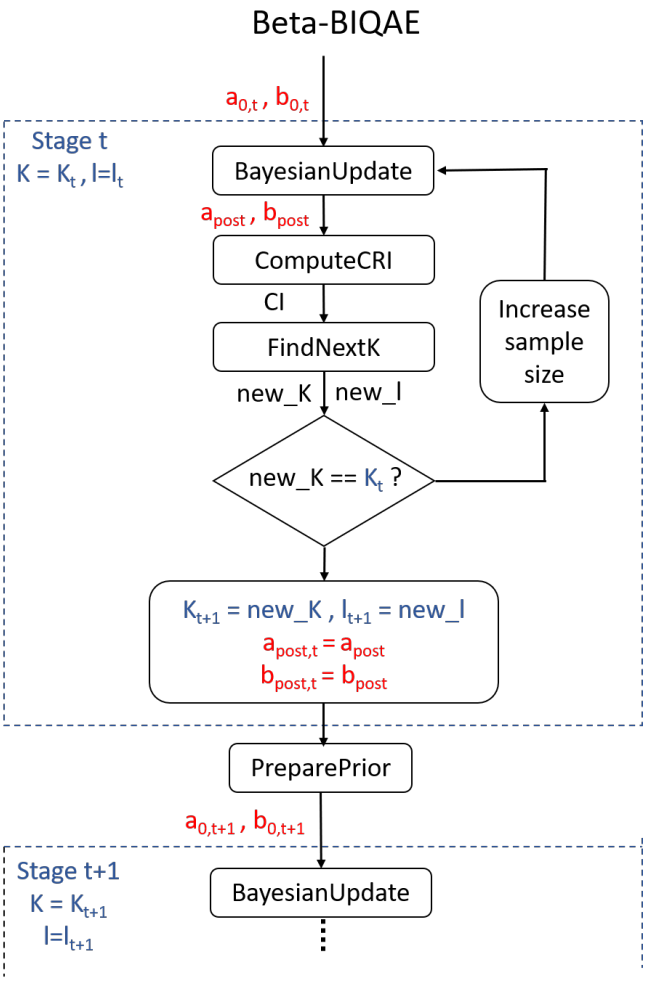


Figure 13: Detailed workflow of Beta-BIQAE, where parameter updates are emphasized in red to contrast with Normal-BIQAE.

Algorithm 4 BayesianUpdateBeta

Require: Prior shape parameters $a_{0,t}$ and $b_{0,t}$ for p_{k_t}, number of $\mathcal{Q}$ operators k_t, sample size N_t
Ensure: Posterior shape parameters $a_{\text{post},t}$ and variance $b_{\text{post},t}$ for p_{k_t}

1: Construct k_t $\mathcal{Q}$ operators into the circuit $\mathcal{A}$
2: Sample shots until a total of N_t shots (including existing shots at stage t) is collected. Store the sample mean as $\bar{X}_t$
3: Compute the posterior shape parameters:

$$a_{\text{post},t} = a_{0,t} + N_t \bar{X}_t,$$

$$b_{\text{post},t} = b_{0,t} + N_t \left(1 - \bar{X}_t\right)$$

4: **Return** $(a_{\text{post},t}, b_{\text{post},t})$

Algorithm 5 ComputeCIBeta

Require: Posterior shape parameters $(a_{\text{post},t}, b_{\text{post},t})$ for p_{k_t}, amplification factor K_t, quadrant index l_t, confidence level α_t

Ensure: An interval estimate $(\theta_{l,t}, \theta_{u,t})$ for θ at confidence level α_t

1: Compute the credible interval for p_{k_t} using the quantiles of the beta distribution:

$$\left[p_{k_t}^l, p_{k_t}^u\right] = \left[\text{BetaCDF}^{-1}\left(\frac{\alpha_t}{2}, a_{\text{post},t}, b_{\text{post},t}\right), \text{BetaCDF}^{-1}\left(1-\frac{\alpha_t}{2}, a_{\text{post},t}, b_{\text{post},t}\right)\right]$$

where BetaCDF^{-1} denotes the inverse cumulative distribution function of the beta distribution.

2: Map the interval estimate for p_{k_t} to an interval estimate for θ using f_t:

$$\left[\theta_t^l, \theta_t^u\right] = f_t([p_{k_t}^l, p_{k_t}^u])$$

where

$$f_t(x) \coloneqq \begin{cases} \frac{1}{K_t}\left(\arcsin\left(\sqrt{x}\right) + l_t \cdot \frac{\pi}{2}\right) & \text{if } l_t \text{ is even,} \\ \frac{1}{K_t}\left(\arccos\left(\sqrt{x}\right) + l_t \cdot \frac{\pi}{2}\right) & \text{if } l_t \text{ is odd} \end{cases}$$

3: **Return** $\left[\theta_t^l, \theta_t^u\right]$

Algorithm 6 PreparePriorBeta

Require: Posterior shape parameters $(a_{\text{post},t}, b_{\text{post},t})$ for p_{k_t} obtained at the end of stage t, current and next amplification factors K_t, K_{t+1}, current quadrant index l_t

Ensure: Shape parameters $(a_{0,t+1}, b_{0,t+1})$ for the beta prior of $p_{k_{t+1}}$

1: Sample $R = 1000$ i.i.d. values $\{y_1, y_2, \ldots, y_R\}$ from the beta posterior $\text{Beta}(a_{\text{post},t}, b_{\text{post},t})$

2: Map each sample y_i to the parameter space for $p_{k_{t+1}}$ using the transformation:

$$\tilde{y}_i \coloneqq \sin^2\left(K_{t+1} \cdot f_t(y_i)\right),$$

where f_t is the mapping defined in Algorithm 5

3: Fit a beta distribution to the transformed samples $\{\tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_R\}$ using maximum likelihood estimation (MLE) to obtain shape parameters $(a_{0,t+1}, b_{0,t+1})$

4: **Return** $(a_{0,t+1}, b_{0,t+1})$

Formally, Theorem 8 establishes that the beta prior fitted by Algorithm 6 is the asymptotically optimal beta approximation to the exact prior distribution for $p_{k_{t+1}}$ that minimizes the Kullback-Leibler (KL) divergence between the exact prior and the fitted beta distribution. The proof is provided in Appendix C.7.

Theorem 8. *Let g denote the density function of the exact prior distribution for $p_{k_{t+1}}$ derived from transforming the posterior distribution for p_{k_t}, and let $\{q_{\alpha,\beta} : (\alpha,\beta) \in \Theta \coloneqq [\alpha_{\min}, \alpha_{\max}] \times [\beta_{\min}, \beta_{\max}]\}$ be a parametric family of beta distributions indexed by (α,β). Suppose R i.i.d. samples $\tilde{Y}_1, \ldots, \tilde{Y}_R \sim g$ are generated by mapping samples from the beta posterior of p_{k_t} (as described in Algorithm 6), and let $(\hat{\alpha}_R, \hat{\beta}_R)$ be the MLE estimates obtained by maximizing the average log-likelihood over the parameter space Θ*

$$(\hat{\alpha}_R, \hat{\beta}_R) = \arg\max_{(\alpha,\beta)\in\Theta} \frac{1}{R}\sum_{i=1}^{R} \log q_{\alpha,\beta}(\tilde{Y}_i).$$

Define the optimal shape parameters (α^*, β^*) as

$$(\alpha^*, \beta^*) = \arg\min_{(\alpha,\beta)\in\Theta} D_{\text{KL}}(g \| q_{\alpha,\beta}),$$

where D_{KL} represents the Kullback-Leibler (KL) divergence.

Then, $(\hat{\alpha}_R, \hat{\beta}_R)$ converges almost surely to (α^*, β^*) as $R \to \infty$.

Note that, although the beta prior fitted in Algorithm 6 is the best approximation within the beta family, it is not guaranteed to be close to the exact prior. In principle, it is possible that the beta family is not expressive enough, and that even the optimal beta approximation cannot closely represent the exact prior. However, this phenomenon is unlikely to occur in practice for BIQAE: Informally, for any smooth and unimodal distribution on $[0, 1]$, there exists a beta distribution that can approximate it closely. We leave the rigorous theoretical justification of this claim to future work. Fig. 14 shows the quality of this approximation with an example.

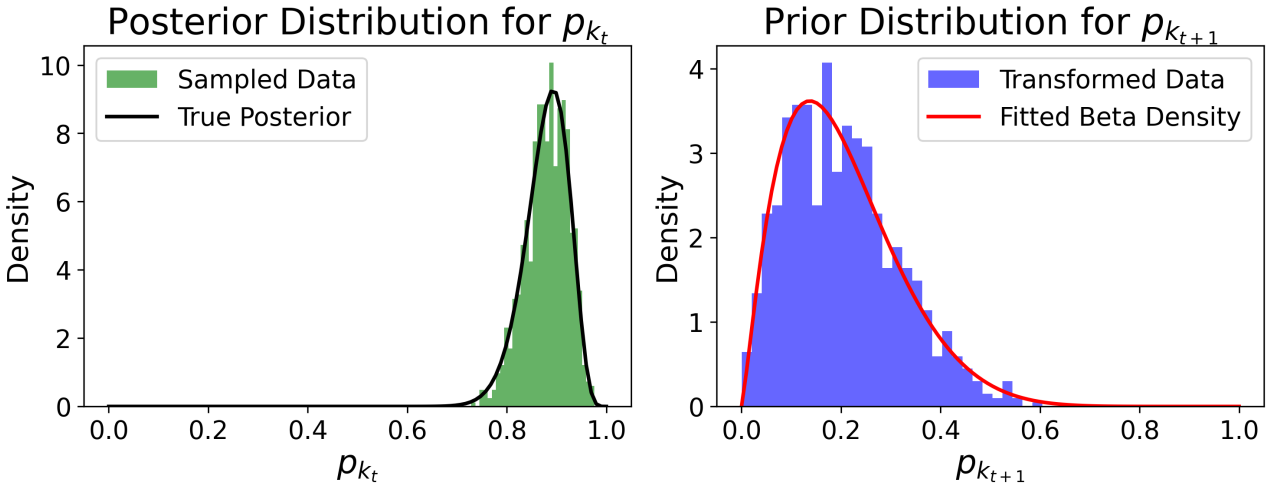


Figure 14: Beta approximation to the exact prior for $p_{k_{t+1}}$. We first generate (left) sampled data (green bars) from the true posterior (black line) and then transform the sampled data as $\sin^2(K_{t+1} \cdot f_t(\cdot))$ to obtain (right) the transformed data (blue bars). Finally, we fit a beta distribution (red line) to the transformed data. The close alignment between the red line and the blue bars demonstrates the accuracy of the approximation.

A.9 Equivalence of Interval Estimates

To demonstrate that IQAE can be derived as an implementation of BIQAE with a noninformative prior, we establish the similitude of the interval estimates in IQAE-CH and Normal-BIQAE and in IQAE-CP and Beta-BIQAE, as follows:

Consider Normal-BIQAE with a noninformative prior $\Delta_t^2 \to \infty$. A Bayesian update according to Algorithm 1 yields a posterior distribution characterized by

$$\mu_{\mathrm{post},t} \to \bar{X}_t, \quad \sigma_{\mathrm{post},t} \to \sqrt{\frac{\bar{X}_t(1-\bar{X}_t)}{N_t}}$$

regardless of $\mu_{0,t}$. Approximation of the z-value as

$$z_{\alpha_t/2} \approx \sqrt{2\log\left(\frac{2}{\alpha_t}\right)}$$

then yields the credible interval

$$\mu_{\mathrm{post},t} \pm z_{\alpha_t/2} \cdot \sigma_{\mathrm{post},t} \approx \bar{X}_t \pm \sqrt{\frac{2\bar{X}_t(1-\bar{X}_t)}{N_t}\log\left(\frac{2}{\alpha_t}\right)}, \tag{27}$$

which is directly comparable to the Chernoff-Hoeffding confidence interval for p_{k_t}

$$\bar{X}_t \pm \sqrt{\frac{1}{2N_t}\log\left(\frac{2}{\alpha_t}\right)}, \tag{28}$$

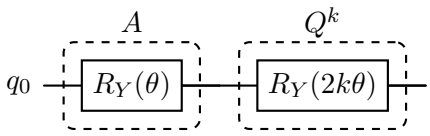


Figure 15: Circuit diagram for the quantum sample complexity scaling analysis.

where the truncation function $\mathrm{trunc}_{[0,1]}(\cdot)$ is omitted for simplicity. The credible interval of Normal-BIQAE Eq. (27) and the confidence interval of IQAE-CP Eq. (28) are thus structurally similar, possessing identical centers with a discrepancy only in first factor under the square root that enables the Chernoff-Hoeffding confidence interval to be slightly more conservative since $\frac{1}{4}$ is always no less than $\bar{X}_t(1-\bar{X}_t)$, such that construction of interval estimates with the two methods is highly similar; we thus do not distinguish linguistically between the Normal-IQAE method corresponding to the approach entailed by Eq. (27) and the IQAE-CH method corresponding to the approach entailed by Eq. (28) in this paper.

Likewise, the credible interval of Beta-BIQAE according to Algorithm 5

$$\left[p_{k_t}^l, p_{k_t}^u\right] = \left[\mathrm{BetaCDF}^{-1}\left(\frac{\alpha_t}{2}, a_{\mathrm{post},t}, b_{\mathrm{post},t}\right), \mathrm{BetaCDF}^{-1}\left(1-\frac{\alpha_t}{2}, a_{\mathrm{post},t}, b_{\mathrm{post},t}\right)\right] \tag{29}$$

is directly comparable to the Clopper-Pearson confidence interval used in IQAE-CP

$$\left[p_{k_t}^l, p_{k_t}^u\right] = \left[\mathrm{BetaCDF}^{-1}\left(\frac{\alpha_t}{2}, N_t\bar{X}_t, 1+N_t\left(1-\bar{X}_t\right)\right), \mathrm{BetaCDF}^{-1}\left(1-\frac{\alpha_t}{2}, 1+N_t\bar{X}_t, N_t\left(1-\bar{X}_t\right)\right)\right] \tag{30}$$

for a noninformative prior $a_{0,t} = b_{0,t} = 0.5$, in particular for relatively large sample sizes N_t (namely, for *Jeffreys prior* with corresponding credible interval *Jeffreys interval*). Given the absence of significant differences between practical applications of the two approaches, we do not distinguish linguistically between the IQAE-CP approach using Clopper-Pearson confidence intervals and the Beta-IQAE approach using Jeffreys intervals in this paper.

A.10 Settings for Quantum Sample Complexity Scaling Analysis

As a proof-of-concept, numerical simulations of the quantum sample complexity for arbitrary quantum states are performed using a simplified circuit design with mock $\mathcal{Q}$-operators that eliminate the impact of circuit depth on classical simulation runtime—a simplification that becomes essential as the target accuracy increases, implying increasingly prohibitively time-consuming numbers of $\mathcal{Q}$-operators. The circuit, shown in Fig. 15, employs a single $R_Y(2k\theta)$ gate to emulate $\mathcal{Q}^k$, in which the behavior of the circuit is constructed with knowledge θ, but the algorithm is executed as if θ were unknown.

Note that although this mock $\mathcal{Q}$ operator is used to facilitate classical simulations, all presented circuit depths and quantum sample complexities accurately account for the internal structure of the $\mathcal{Q}$ operator. Specifically, the total circuit depth is calculated as

$$\text{total circuit depth} = \text{circuit depth of } \mathcal{A} + k \cdot \text{circuit depth of } \mathcal{Q},$$

and, since each $\mathcal{Q}$ operator contains two $\mathcal{A}$ operators, the quantum sample complexity weights each set of N_k shots of the quantum circuit by either k (by the IQAE convention) or $K = 2k+1$ (by the rigorous count of oracle calls).

A.11 Settings for Molecular Ground-State Energy Estimation

To identify the molecular ground-state wavefunction, the electronic Hamiltonian for each molecule and bond length considered is generated in Qiskit Nature [102] via the PySCF driver [100] in parity mapping [119]. Frozen core orbitals and qubit tapering [98] are then employed to reduce the required

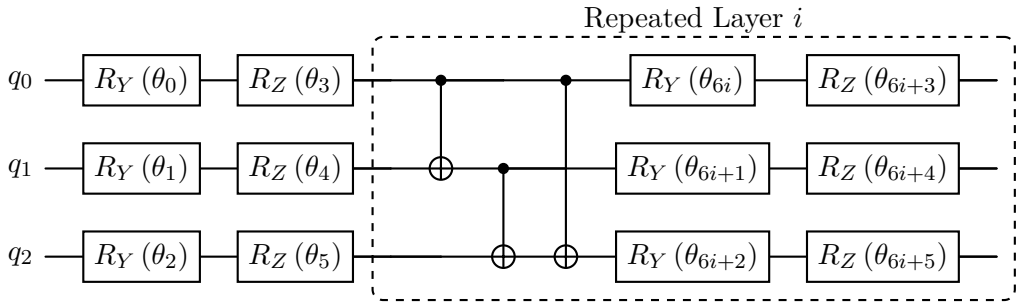


Figure 16: Efficient SU(2) Ansatz, as defined by a single base layer and repeated layers $i = 1, 2, \ldots, L$. The Ansatz for a single repeated layer $L = 1$ is shown.

Molecule	Qubit Number		Ansatz Layer Number	Max. Iter. Number
	Before Taper.	After Taper.		
H_2	4	N/A	12	300
LiH	10	6	12	300
HF	10	6	12	300
BeH_2	12	7	15	500

Table 1: Number of qubits (before and after qubit tapering), layers, and maximum iterations used to determine the electronic ground-state wavefunction for each molecule considered. The number of qubits is fixed according to the form of the mapped Hamiltonian, and the number of Ansatz layers and maximum iterations is chosen to converge the ground-state wavefunction.

number of qubits to represent the Hamiltonian.[3] The ground-state wavefunction is then determined via the Variational Quantum Eigensolver (VQE) [120, 121] for the efficient SU(2) Ansatz with full entanglement, as depicted in Fig. 16; the parameters of the Ansatz are optimized classically with Sequential Least Squares Programming (SLSQP); and the expectation value corresponding to each parameter set is computed classically via the Qiskit Runtime Estimator primitive service with zero shots (see additional details in Table 1). In all molecular ground-state energy calculations, the electronic energy is supplemented by the classically calculated nuclear repulsion energy and core electron energy as determined with PySCF.

B Extended Results

B.1 Linear Regression Analysis

As detailed in Table 2, regression of the quantum sample complexity as a function of the median absolute error in Fig. 3 suggests Beta-BIQAE is robust, improves upon Classical QAE quadratically, and features the lowest quantum sample complexity of all QAE methods considered. Linear regression of the log-log data to produce slope, intercept, and R^2 values suggests all of these QAE methods (with the possible exception of FQAE) reliably obey the expected power law scaling, with R^2 values exceeding 0.99. The slopes indicate that methods based on amplitude amplification offer a quadratic advantage over Classical QAE, as, for example, Beta-BIQAE's slope of -1.0 and Classical QAE's slope of -2.0 correspond to quantum sample complexities of $N_{\text{oracle}} \in \mathcal{O}(\varepsilon^{-1})$ and $\mathcal{O}(\varepsilon^{-2})$, respectively. Similarly, the intercept, which represents the logarithm of the constant overhead, implies that, of all of the QAE methods considered that achieve the quadratic advantage, Beta-BIQAE has the lowest constant factor

[3] Note Jordan-Wigner mapping [97] without qubit tapering is employed for H_2.

Table 2: Linear regression results for the log-log relationship between the quantum sample complexity and the estimation error ($\log_{10} N_{\text{oracles}} \sim \log_{10} \varepsilon$).

Method	**Slope**	**Intercept**	R^2
Beta-BIQAE	−1.0088	0.0211	0.9999
IQAE-CP	−1.0044	0.1253	0.9999
BAE	−1.0137	0.1089	0.9996
Canonical QAE	−1.0058	0.4847	1.0000
Classical QAE	−1.9821	−1.1693	0.9996
IQAE_CH	−1.0056	0.3930	0.9999
MLQAE_EIS	−0.9820	0.5626	0.9999
MLQAE_LIS	−1.3375	−0.3395	0.9979
FQAE	−1.2091	3.6295	0.9677
QAES	−0.9973	6.8670	0.9999

of the quantum sample complexity.

B.2 Interval Radii and Coverage Rates

Beta-BIQAE is also found to provide more accurate and reliable quantum amplitude estimates than IQAE-CP and BAE from an interval radius and coverage rate perspective, where the radius of the interval estimate for BAE is approximated as the product of the critical value v and standard deviation and where the *coverage rate* is defined as the proportion of repeated independent experiments in which the interval contains the true parameter value. As shown in Fig. 17 (left panel), for the critical value $v = 1.96$, which corresponds to 95% coverage under the assumption that the sampling distribution of the point estimate is normal, both Beta-BIQAE and IQAE-CP have a higher coverage rate for all target accuracies considered than BAE, which provides the narrowest intervals (and thus the highest accuracy) but exhibits a coverage rate of only $\leq 87\%$. Likewise, as shown in Fig. 17 (right panel), for $v = 10$, which corresponds to coverage of all but $1.523971 \times 10^{-21}\%$ for a normal distribution, Beta-BIQAE and IQAE-CP have higher coverage rates than BAE, which features coverage rates as low as 87.5% and the largest intervals. Note in the $v = 10$ case Beta-BIQAE yields the smallest intervals and frequently the highest coverage rates for all cases considered, which is indicative of an advantageous balance between a narrow interval radius and a high coverage rate (*i.e.*, accuracy and reliability). The observed performance difference between BIQAE and BAE is also consistent with BIQAE and BAE's distinct methodological approaches. Namely, although BIQAE and BAE both employ Bayesian updates within an iterative QAE framework to extract information from measurement outcomes, (i) BIQAE adopts the K-scheduling strategy of IQAE to provide interval estimates for the amplitude in place of BAE's Bayesian experimental design strategy that schedules K by maximizing the utility function to obtain a point estimate with its standard deviation and (ii) BIQAE utilizes conjugate priors to perform Bayesian updates (in the case of the beta distribution, with the advantage of small-sample efficiency enhancement) in place of BAE's Sequential Monte Carlo.

B.3 Robustness of Beta-BIQAE Advantage Over IQAE-CP

Fig. 18 shows that the near-constant-factor improvement in the sample complexity of Beta-BIQAE over IQAE-CP for all quantum amplitude values a shown for $\varepsilon = 10^{-8}$ in Fig. 6 persists for target accuracies $\varepsilon \in \{10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}, 10^{-7}\}$. The cost reduction of Beta-BIQAE over IQAE-CP is consistently 10.5–12.8% and increases for higher target accuracies. Note deviations from the scaling laws for both Beta-BIQAE and IQAE-CP often appear at the same a values across target accuracies, an interesting area for further study.

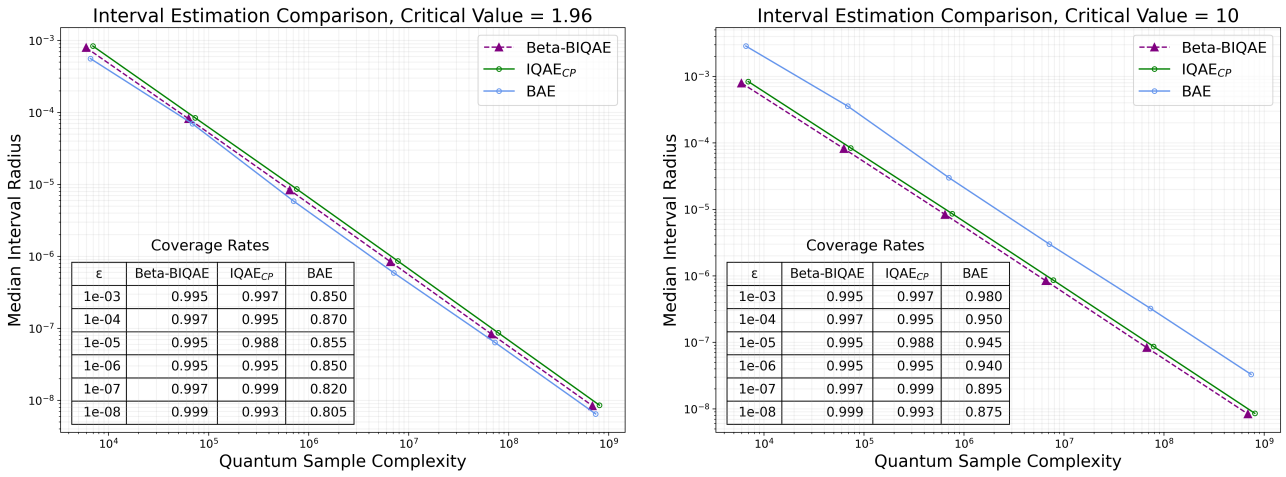


Figure 17: Comparison of the median radius and the coverage rate for BAE and IQAE-based estimates for $a = 0.5$ from an interval estimation perspective. The critical value for BAE is selected as (left) 1.96 or (right) 10.

B.4 Circuit Depth Comparison

Consideration of the quantum circuits required for molecular ground-state energy estimation suggests that the average number of Grover $\mathcal{Q}$ operators needed to estimate the expectation value of Pauli terms is similar for Beta-BIQAE and IQAE-CP for all molecules considered. Given that the number of such operators is the major contributor to the quantum circuit depth in both methods, these results suggest Beta-BIQAE's near-constant-factor improvement over IQAE-CP is largely accomplished without growth of the cost of quantum computation.

B.5 Comparison of BIQAE and IQAE with beta and normal priors

Fig. 20 supports the predicted superiority of the quantum sample complexity and coverage rate of beta-based methods (Beta-BIQAE and IQAE-CP, which employ informative and non-informative beta priors, respectively) over normal-based methods (Normal-BIQAE, Normal-IQAE, and IQAE-CH, which employ an informative normal prior, a non-informative normal prior, and the Chernoff–Hoeffding bound, respectively, where the Chernoff–Hoeffding bound is asymptotically equivalent to the noninformative normal prior up to a more conservative constant factor, as shown in Appendix A.9) in the small per-stage sample size limit. For all target accuracy values considered, the beta-based methods exhibit lower complexity than the normal-based methods. The coverage rate for beta-based methods is likewise superior to that of normal-based methods, notwithstanding the imposition in the Normal-IQAE and Normal-BIQAE implementations of a minimum sample size of at least 30 shots per stage to avoid zero-variance estimates associated with diminished coverage rates and algorithmic failure. For all target accuracies considered, beta-based Beta-BIQAE and IQAE-CP—but not normal-based Normal-BIQAE and Normal-IQAE—achieve the target coverage rate of 0.95. This distinction between the two classes of methods can be attributed to the normal-based methods' variance estimator $\hat{\sigma}_t^2 = \bar{X}_t(1-\bar{X}_t)$, which requires a substantially larger per-stage sample size to have a comparable accuracy to that of the beta-based IQAE or BIQAE settings. IQAE-CH mitigates this issue by replacing the variance estimator with the upper bound $\bar{X}_t(1-\bar{X}_t) \leq \frac{1}{4}$, which stabilizes the coverage rate at the cost of increased complexity due to its conservativeness, in keeping with its coverage intermediate between the remaining normal- and beta-based methods considered. By contrast, the use of beta priors in Beta-BIQAE and IQAE-CP eliminates the small-sample instability entirely, which enables both low quantum sample complexities and high coverage rates.

Note a further observation from Fig. 20 is that Normal-BIQAE consistently improves upon Normal-IQAE, as expected.

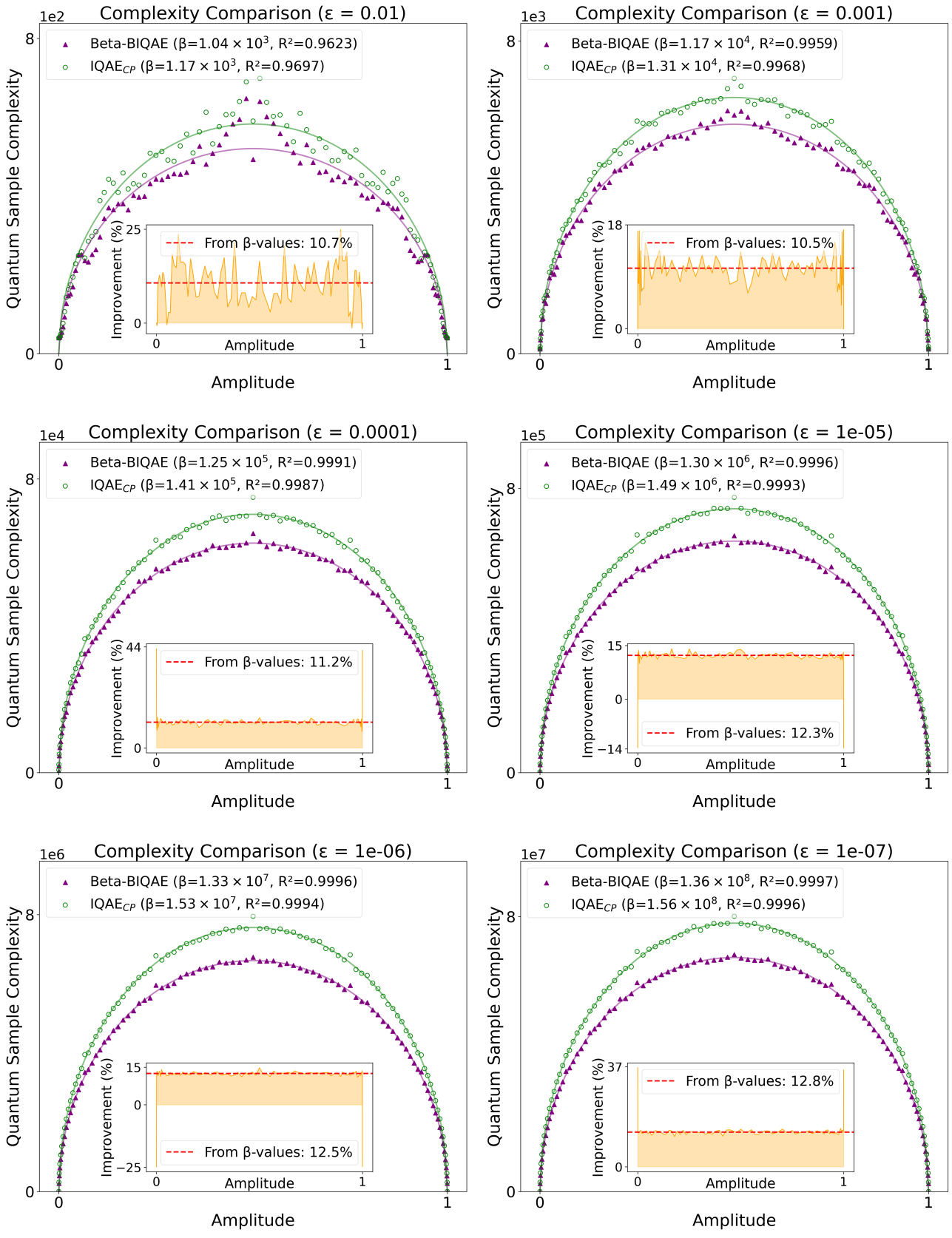


Figure 18: Comparison of the quantum sample complexities of Beta-BIQAE and IQAE-CP as a function of the quantum amplitude for target accuracies $\varepsilon \in \{10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}, 10^{-7}\}$, as in Fig. 6.

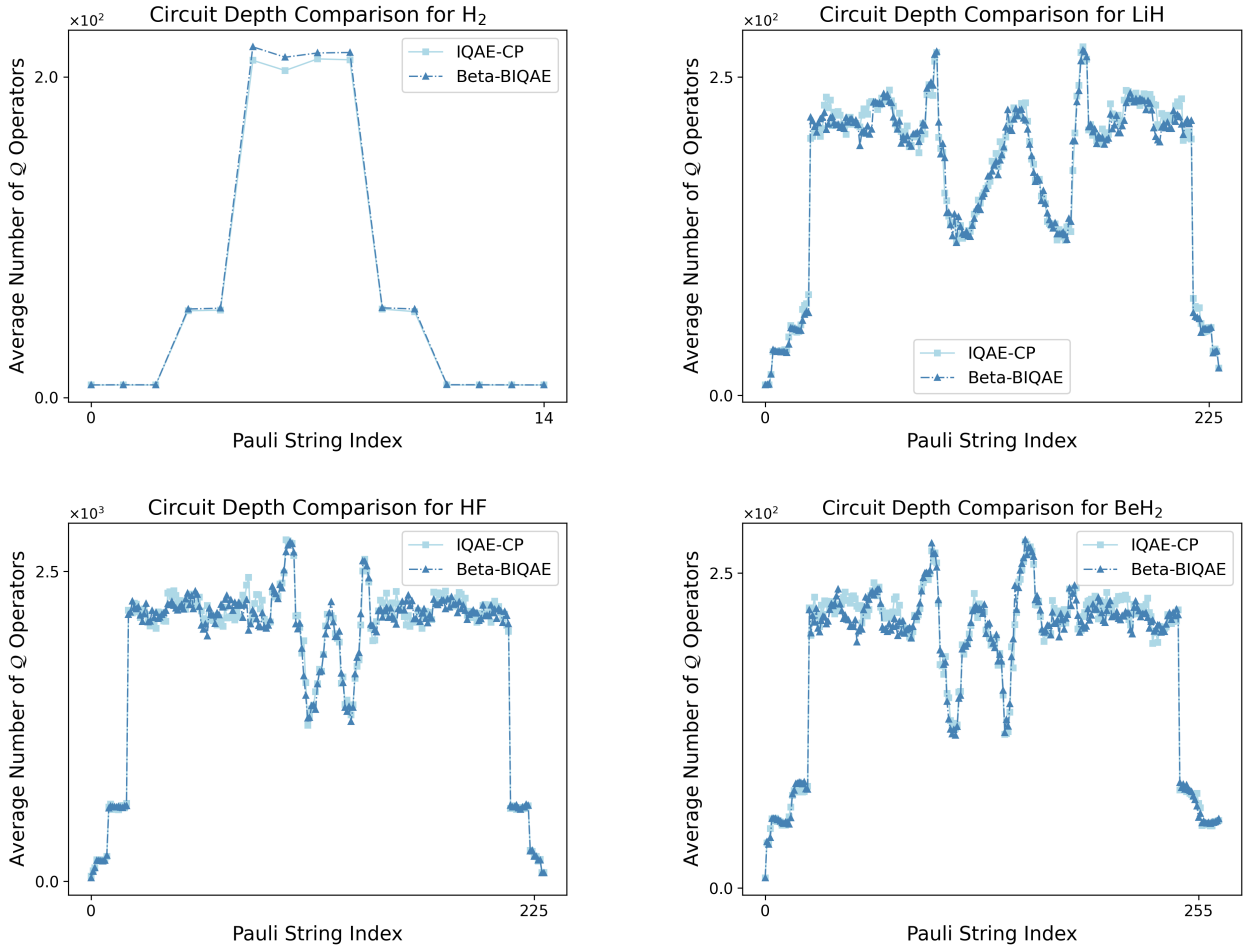


Figure 19: Comparison of the number of Grover operators in the last stage required to estimate Pauli string expectation values with Beta-BIQAE (medium blue triangles) and IQAE-CP (light blue squares). The number is averaged over 200 repetitions.

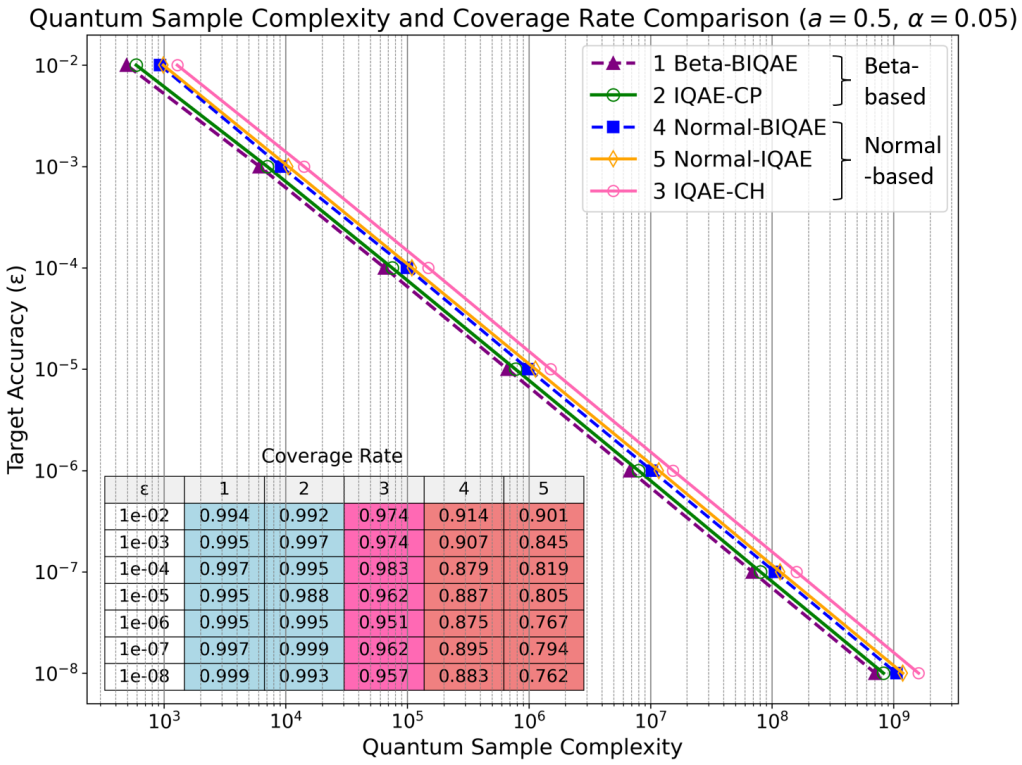


Figure 20: Comparison of quantum sample complexity and coverage rate between beta-based methods (Beta-BIQAE and IQAE-CP) and normal-based methods (Normal-BIQAE, Normal-IQAE, and IQAE-CH). Legends are ordered from the lowest to the highest quantum sample complexity as a function of target accuracy. Inset table cells are shaded to indicate low coverage associated with failure to achieve the target coverage rate of 0.95 (red), intermediate coverage between 0.95 and 0.99 (pink), and high coverage above 0.99 (blue).

C Proofs

C.1 MLE and AMSE of QAE

Proofs of Theorem 4 and Appendix Theorems 3 and 4. Given that $\bar{X}_k$ is the MLE of p_k and the quantum amplitude is $a = \sin^2 f_k(p_k)$ given by Eq. (4), where $f_k(\cdot)$ is defined as

$$f_k(x) := \begin{cases} \frac{1}{2k+1}\left(\arcsin\left(\sqrt{x}\right) + l(k)\cdot\frac{\pi}{2}\right) & \text{if } l(k) \text{ is even,} \\ \frac{1}{2k+1}\left(\arccos\left(\sqrt{x}\right) + l(k)\cdot\frac{\pi}{2}\right) & \text{if } l(k) \text{ is odd,} \end{cases} \tag{31}$$

the MLE of a is $\hat{a} = \sin^2 f_k(\bar{X}_k)$ according to the invariance principle.

Likewise, $\bar{X}_k$ is an unbiased estimator of p_k, and its MSE is readily calculated as

$$\mathrm{MSE}_{\bar{X}_k}(p_k) = \frac{1}{N_k} p_k(1-p_k).$$

The AMSE of $\hat{a}$ can be derived by the delta method as follows:

$$\begin{aligned} \mathrm{AMSE}_{\hat{a}}(a) &= \frac{1}{N_k} p_k(1-p_k)[(\sin^2 f_k(p_k))']^2 \\ &= \frac{1}{N_k} p_k(1-p_k)\cdot 4\sin^2 f_k(p_k)\cos^2 f_k(p_k)[f_k'(p_k)]^2 \\ &= \frac{1}{N_k} p_k(1-p_k)\cdot 4a(1-a)[f_k'(p_k)]^2. \end{aligned}$$

Since the derivative of $f_k(\cdot)$ is given by

$$f_k'(x) = \begin{cases} \frac{1}{2(2k+1)}\frac{1}{\sqrt{x(1-x)}} & \text{if } l(k) \text{ is even,} \\ -\frac{1}{2(2k+1)}\frac{1}{\sqrt{x(1-x)}} & \text{if } l(k) \text{ is odd,} \end{cases} \tag{32}$$

we obtain the AMSE expression given in Theorem 4, namely,

$$\mathrm{AMSE}_{\hat{a}}(a) = \frac{1}{(2k+1)^2}\frac{1}{N_k} a(1-a).$$

The quantum sample complexities of Classical QAE (Appendix Theorem 3) and Amplified Amplitude Estimation (Appendix Theorem 4) then follow as special cases.

C.2 Schedule Number T of Base-3 Hybrid Scheduled IQAE

Proof of Eq. (8). Let ε_T denote the radius of the confidence interval for θ at the end of the algorithm, where the confidence interval must be fully contained within one of the reference intervals at Stage $T-1$. Since each reference interval at Stage $T-1$ has length $\frac{1}{3^T}\frac{\pi}{2}$, we have

$$2\varepsilon_t \le \frac{1}{3^T}\frac{\pi}{2}.$$

The Taylor expansion and definition $a = \sin^2\theta$ imply that the radius ε of the final confidence interval for a satisfies

$$\varepsilon = \left(2\sqrt{a(1-a)} + o(1)\right)\varepsilon_t.$$

The combination of the inequality for ε_t and the equation for ε then yields

$$T \le \log_3\left(\frac{\pi}{2}\frac{1}{\varepsilon}\left(\sqrt{a(1-a)} + o(1)\right)\right).$$

Thus, the choice

$$T = \left\lceil \log_3\left(\frac{\pi}{2}\frac{1}{\varepsilon}\sqrt{a(1-a)}\right)\right\rceil$$

is sufficient to achieve the target accuracy ε.

C.3 Accumulated Quantum Sample Complexity of Base-3 Hybrid Scheduled IQAE

Lemma 1. *Assume the same conditions for the quantum sample complexity of QAE as in Theorem 4. We can construct the following confidence intervals for θ and a:*

1. *An asymptotic level-α confidence interval for θ can be constructed as*

$$[\theta_l, \theta_u] = f_k\left(\mathrm{trunc}_{[0,1]}\left(\bar{X}_k \pm \frac{1}{\sqrt{N_k}}\sqrt{\bar{X}_k(1-\bar{X}_k)}z_{\frac{\alpha}{2}}\right)\right),$$

   where $f_k(\cdot)$ is defined in Eq. (31) and $z_{\frac{\alpha}{2}}$ is the upper $\frac{\alpha}{2}$ quantile of the standard normal distribution. The approximate number of accesses to the oracle required for the radius to be ε is

$$N_{oracle} = \frac{z^2_{\frac{\alpha}{2}}}{4(2k+1)\varepsilon^2}.$$

2. *An asymptotic level-α confidence interval for a can be constructed as*

$$[a_l, a_u] = \sin^2([\theta_l, \theta_u]).$$

   The approximate number of accesses to the oracle required for the radius to be ε is

$$N_{oracle} = \frac{z^2_{\frac{\alpha}{2}}}{(2k+1)\varepsilon^2}a(1-a).$$

Proof of Lemma 1. Note that the distribution of the sample average $\bar{X}_k$ is

$$\bar{X}_k \sim \frac{1}{N_k}\mathrm{Bin}\left(N_k, p_k\right).$$

When N is large, this distribution can be approximated by the following normal distribution:

$$\bar{X}_k \sim \mathcal{N}\left(p_k, \frac{1}{N_k}p_k(1-p_k)\right),$$

and the normal approximation allows us to construct an asymptotic level-α confidence interval for p_k as follows:

$$\mathrm{trunc}_{[0,1]}\left(\bar{X}_k \pm \frac{1}{\sqrt{N_k}}\sqrt{\bar{X}_k(1-\bar{X}_k)}z_{\frac{\alpha}{2}}\right).$$

Consequently, using $\theta = f_k(p_k)$ given by Eq. (4), we obtain an asymptotic level-α confidence interval for θ

$$[\theta_l, \theta_u] = f_k\left(\mathrm{trunc}_{[0,1]}\left(\bar{X}_k \pm \frac{1}{\sqrt{N_k}}\sqrt{\bar{X}_k(1-\bar{X}_k)}z_{\frac{\alpha}{2}}\right)\right).$$

An application of the Taylor expansion leads to the following radius of the confidence interval for θ:

$$\begin{aligned}\frac{\theta_u - \theta_l}{2} &\approx \frac{1}{\sqrt{N_k}}\sqrt{\bar{X}_k(1-\bar{X}_k)}z_{\frac{\alpha}{2}} \cdot |f_k'(\bar{X}_k)| \\ &= \frac{z_{\frac{\alpha}{2}}}{2(2k+1)\sqrt{N_k}}.\end{aligned}$$

Hence, when the radius is equal to ε, the corresponding quantum sample complexity is given by

$$N_{\mathrm{oracle}} = (2k+1)N_k = (2k+1)\frac{z^2_{\frac{\alpha}{2}}}{4(2k+1)^2\varepsilon^2} = \frac{z^2_{\frac{\alpha}{2}}}{4(2k+1)\varepsilon^2}.$$

Given the constructed confidence interval $[\theta_l, \theta_u]$ for θ, we have a natural asymptotic level-α confidence interval for $a = \sin^2\theta$

$$[a_l, a_u] = \sin^2([\theta_l, \theta_u]),$$

and its radius

$$\begin{aligned}\frac{a_u - a_l}{2} &\approx \frac{\theta_u - \theta_l}{2} \cdot 2\sin\theta\cos\theta \\ &= \frac{z_{\frac{\alpha}{2}}}{(2k+1)\sqrt{N_k}}\sqrt{a(1-a)},\end{aligned}$$

where the Taylor expansion is used again. If we make the radius equal to ε, the quantum sample complexity is given by

$$N_{\text{oracle}} = (2k+1)N_k = (2k+1)\frac{z_{\frac{\alpha}{2}}^2}{(2k+1)^2\varepsilon^2}a(1-a) = \frac{z_{\frac{\alpha}{2}}^2}{(2k+1)\varepsilon^2}a(1-a).$$

Proof of Theorem 2. To specify the required confidence at each stage, we set all the confidence levels α_t equal to $\frac{\alpha}{T}$, which ensures that the overall confidence level is α by the union bound. By Lemma 1, the accumulated sample complexity is then given by

$$\sum_{t=0}^{T-1} N_{\text{oracle},t} = \sum_{t=0}^{T-1} \frac{z_{\alpha_t/2}^2}{4K_t\varepsilon_t^2} = \frac{1}{4}z_{\frac{\alpha}{2T}}^2\left(\sum_{t=0}^{T-2}\frac{1}{K_t}\frac{1}{\varepsilon_t^2} + \frac{1}{K_{T-1}}\frac{1}{\varepsilon_{T-1}^2}\right), \tag{33}$$

where ε_t is the radius of the interval estimate for θ obtained at stage t.

To bound the first term in the parentheses of (33), we assume that the radius of the confidence interval is sufficiently large compared to the reference intervals (Assumption 2), which indicates

$$\begin{aligned}\sum_{t=0}^{T-2}\frac{1}{K_t}\frac{1}{\varepsilon_t^2} &\le \sum_{t=0}^{T-2}\frac{1}{3^t}\frac{16}{\gamma^2\pi^2}3^{2t+2} \\ &= \frac{144}{\gamma^2\pi^2}\sum_{t=0}^{T-2}3^t \\ &= \frac{72}{\gamma^2\pi^2}(3^{T-1}-1) \\ &\le \frac{36}{\gamma^2\pi}\frac{1}{\varepsilon}\sqrt{a(1-a)}.\end{aligned}$$

The definition of T thus yields

$$\begin{aligned}T = \left\lceil \log_3\left(\frac{\pi}{2}\frac{1}{\varepsilon}\sqrt{a(1-a)}\right)\right\rceil &\Rightarrow T \le \log_3\left(\frac{\pi}{2}\frac{1}{\varepsilon}\sqrt{a(1-a)}\right) + 1 \\ &\Rightarrow 3^{T-1} \le \frac{\pi}{2}\cdot\frac{1}{\varepsilon}\sqrt{a(1-a)}.\end{aligned}$$

To bound the second term in the parentheses of (33), we derive

$$\begin{aligned}\frac{1}{K_{T-1}}\frac{1}{\varepsilon_{T-1}^2} &= \frac{1}{3^{T-1}}\frac{1}{\varepsilon_{T-1}^2} \\ &\le \frac{6}{\pi\sqrt{a(1-a)}}\frac{\varepsilon}{\varepsilon_{T-1}^2} \\ &= \frac{24}{\pi}\frac{1}{\varepsilon}\sqrt{a(1-a)},\end{aligned}$$

where, again, the second to the last inequality follows from the definition of T as follows:

$$T = \left\lceil \log_3 \left(\frac{\pi}{2} \frac{1}{\varepsilon} \sqrt{a(1-a)} \right) \right\rceil \Rightarrow 3^T \geq \frac{\pi}{2} \frac{1}{\varepsilon} \sqrt{a(1-a)} \Rightarrow \frac{1}{3^{T-1}} \leq \frac{6\varepsilon}{\pi} \frac{1}{\sqrt{a(1-a)}}.$$

An application of the radius of interval estimates for θ and a in Lemma 1 leads to

$$\varepsilon_{T-1} = \frac{\varepsilon}{2\sqrt{a(1-a)}}.$$

Hence, we have

$$\begin{aligned} \sum_{t=0}^{T-1} N_{\mathrm{oracle},t} &\leq \frac{1}{4} z^2_{\frac{\alpha}{2T}} \left(\frac{36}{\gamma^2 \pi} + \frac{24}{\pi} \right) \frac{1}{\varepsilon} \sqrt{a(1-a)} \\ &= \frac{3}{\pi} \left(\frac{3}{\gamma^2} + 2 \right) z^2_{\frac{\alpha}{2T}} \frac{1}{\varepsilon} \sqrt{a(1-a)}. \end{aligned}$$

To evaluate the prefactor, by the upper and lower bounds for the Mills ratio, we approximate z_α as follows:

$$z_\alpha \approx \sqrt{-2 \log \alpha},$$

or, more rigorously,

$$z_\alpha = \sqrt{-(2 + o(1)) \log \alpha},$$

where $o(1)$ represents a term that vanishes as α approaches zero. In fact, for $\alpha \leq 0.1$, this approximation is an upper bound

$$z_\alpha < \sqrt{-2 \log \alpha}.$$

Hence, we obtain

$$z^2_{\frac{\alpha}{2T}} < -2 \log \frac{\alpha}{2T} \leq 2 \log \left(\frac{2}{\alpha} \left\lceil \log_3 \left(\frac{\pi}{4\varepsilon} \right) \right\rceil \right).$$

Finally, we conclude that the accumulated sample complexity is

$$\sum_{t=0}^{T-1} N_{\mathrm{oracle},t} \leq \frac{6}{\pi} \left(\frac{3}{\gamma^2} + 2 \right) \frac{1}{\varepsilon} \log \left(\frac{2}{\alpha} \left\lceil \log_3 \left(\frac{\pi}{4\varepsilon} \right) \right\rceil \right) \sqrt{a(1-a)}.$$

C.4 Accumulated Quantum Sample Complexity of Base-3/Base-5 Hybrid Scheduled IQAE

Proof of Appendix Theorem 5. Given the decomposition of the accumulated complexity Eq. (33), we bound the first term in the parentheses by replacing the assumption that the confidence interval is sufficiently large relative to the base-3 reference intervals (Assumption 2) with the requirement that the confidence interval is contained in the base-3 or base-5 reference intervals [Eq. (9)]. This replacement yields the inequality

$$\varepsilon_t \geq \frac{1}{2} \frac{1}{K_t} \frac{1}{15} \frac{\pi}{2}, \quad \text{for } t = 0, 1, \ldots, T-2$$

and thus the bound

$$\sum_{t=0}^{T-2} \frac{1}{K_t} \frac{1}{\varepsilon_t^2} \leq \frac{3600}{\pi^2} \sum_{t=0}^{T-2} K_t.$$

Note that $\{K_t\}$ is no longer a geometric sequence as in base-3 hybrid schedule IQAE, and recognize that the ratio $\frac{K_{t+1}}{K_t}$ can only be 3 or 5. We have

$$K_t \leq \frac{1}{3} K_{t+1}, \quad \text{for all } t = 0, \ldots, T-2,$$

$$K_t \leq \frac{1}{3^{T-1-t}} K_{T-1}, \quad \text{for all } t = 0, 1, \ldots, T-1.$$

and

$$\sum_{t=0}^{T-2} K_t \leq \sum_{t=0}^{T-2} \frac{1}{3^{T-1-t}} K_{T-1} = K_{T-1} \sum_{t=1}^{T-1} \frac{1}{3^t} \leq \frac{1}{2} K_{T-1}.$$

Therefore, we conclude

$$\begin{aligned} \sum_{t=0}^{T-1} N_{\text{oracle},t} &\leq \frac{1}{4} z^2_{\frac{\alpha}{2T}} \left(\frac{1800}{\pi^2} K_{T-1} + \frac{1}{K_{T-1}} \frac{1}{\varepsilon^2_{T-1}} \right) \\ &= \frac{1}{4} z^2_{\frac{\alpha}{2T}} \left(\frac{1800}{\pi^2} K_{T-1}\varepsilon_{T-1} + \frac{1}{K_{T-1}\varepsilon_{T-1}} \right) \frac{1}{\varepsilon_{T-1}}. \end{aligned}$$

To express the above accumulated quantum sample complexity in terms of the quantum amplitude and determine $K_{T-1}\varepsilon_{T-1}$, we consider that (i) since the interval estimate for the amplified angle $K_{T-1}\theta$ must lie within a single quadrant, such that $K_{T-1}\varepsilon_{T-1} \leq \frac{1}{2}\frac{\pi}{2}$ and (ii) at the last stage, the interval estimate for the amplified angle cannot be fully contained within any reference interval, otherwise the algorithm would proceed to stage T to achieve a more accurate interval estimate. This argument yields the bound

$$K_{T-1}\varepsilon_{T-1} \geq \frac{1}{2}\frac{1}{15}\frac{\pi}{2}.$$

Due to the monotonicity of $\frac{1800}{\pi^2} K_{T-1}\varepsilon_{T-1} + \frac{1}{K_{T-1}\varepsilon_{T-1}}$ with respect to $K_{T-1}\varepsilon_{T-1}$, the accumulated quantum sample complexity is upper bounded by

$$\sum_{t=0}^{T-1} N_{\text{oracle},t} \leq \frac{227}{2\pi} z^2_{\frac{\alpha}{2T}} \frac{1}{\varepsilon_{T-1}}.$$

Replacing $z^2_{\frac{\alpha}{2T}}$ by its upper bound and ε_{T-1} by ε as in Appendix C.3, we derive the accumulated sample complexity[4]

$$\sum_{t=0}^{T-1} N_{\text{oracle},t} \leq \frac{227}{\pi} \frac{1}{\varepsilon} \log\left(\frac{2}{\alpha} \left\lceil \log_3 \left(\frac{\pi}{4\varepsilon} \right) \right\rceil \right) \sqrt{a(1-a)}.$$

C.5 Preparation of the Prior in Normal-BIQAE

Proof of Appendix Theorem 6. Note the relationship between $p_{k_{t+1}}$ and p_{k_t} in Eq. (16)

$$p_{k_{t+1}} = \varphi(p_{k_t}),$$

where φ is defined as

$$\varphi(x) \coloneqq \sin^2\left(K_{t+1} \cdot f_t(x)\right). \tag{34}$$

[4]Note in base-3/base-5 hybrid scheduled IQAE, T is upper bounded by $\left\lceil \log_3 \left(\frac{\pi}{2\varepsilon} \sqrt{a(1-a)} \right) \right\rceil$ because the slowest schedule always selects three as the multiplier.

The posterior distribution of p_{k_t} is approximately normal with mean $\mu_{\text{post},t}$ and variance $\sigma^2_{\text{post},t}$, and, for $x \in (0,1)$,

$$|f_t'(x)| = \frac{1}{2K_t\sqrt{x(1-x)}}. \tag{35}$$

An application of the delta method yields that $p_{k_{t+1}}$ approximately follows a normal distribution with mean

$$\mu_{0,t+1} = \varphi(\mu_{\text{post},t}) = \sin^2\left(K_{t+1} \cdot f_t(\mu_{\text{post},t})\right)$$

and variance

$$\begin{aligned}
\Delta^2_{t+1} &= \left[\varphi'(\mu_{\text{post},t})\right]^2 \cdot \sigma^2_{\text{post},t} \\
&= \left\{2\sin\left[K_{t+1}f_t(\mu_{\text{post},t})\right]\cos\left[K_{t+1}f_t(\mu_{\text{post},t})\right] \cdot K_{t+1}f_t'(\mu_{\text{post},t})\right\}^2 \cdot \sigma^2_{\text{post},t} \\
&= \left(\frac{K_{t+1}}{K_t}\right)^2 \sin^2\left[K_{t+1}f_t(\mu_{\text{post},t})\right]\cos^2\left[K_{t+1}f_t(\mu_{\text{post},t})\right] \cdot \frac{1}{\mu_{\text{post},t}(1-\mu_{\text{post},t})} \cdot \sigma^2_{\text{post},t} \\
&= \left(\frac{K_{t+1}}{K_t}\right)^2 \frac{\mu_{0,t+1}(1-\mu_{0,t+1})}{\mu_{\text{post},t}(1-\mu_{\text{post},t})} \cdot \sigma^2_{\text{post},t}.
\end{aligned}$$

Thus, the prior distribution of $p_{k_{t+1}}$ is approximately normal with mean $\mu_{0,t+1}$ and variance Δ^2_{t+1}.

C.6 Quantum Sample Complexity of Stage t in Normal-BIQAE

Proof of Appendix Theorem 7. Without loss of generality, we assume that the interval estimate for p_{k_t} in Normal-BIQAE as presented in Algorithm 2 is contained in $[0,1]$. We will derive an expression for the required sample size N_t to achieve the stopping criterion at stage t.

At stage t, when the radius of the interval estimate for θ reaches ε_t, we have

$$z_{\alpha_t/2} \cdot \sigma_{\text{post},t} \cdot \left|f_t'(\mu_{\text{post},t})\right| = \varepsilon_t,$$

which upon substitution of $|f_t'(\mu_{\text{post},t})|$ given by Eq. (35) becomes

$$\frac{z^2_{\alpha_t/2}}{4K_t^2\varepsilon_t^2}\frac{1}{\mu_{\text{post},t}(1-\mu_{\text{post},t})} = \sigma^{-2}_{\text{post},t}. \tag{36}$$

Substituting $\sigma^2_{\text{post},t}$ by the expression as detailed in Algorithm 1 yields

$$\frac{z^2_{\alpha_t/2}}{4K_t^2\varepsilon_t^2}\frac{1}{\mu_{\text{post},t}(1-\mu_{\text{post},t})} = \frac{1}{\Delta_t^2} + \frac{N_t}{\bar{X}_t\left(1-\bar{X}_t\right)},$$

and a rearrangement of terms leads to the number of shots N_t as follows:

$$N_t = \frac{\bar{X}_t\left(1-\bar{X}_t\right)}{\mu_{\text{post},t}(1-\mu_{\text{post},t})}\left(\frac{z^2_{\alpha_t/2}}{4K_t^2\varepsilon_t^2} - \frac{\mu_{\text{post},t}(1-\mu_{\text{post},t})}{\Delta_t^2}\right).$$

We derive the sample complexity at stage t by considering the corresponding number of accesses to the oracle $\mathcal{A}$, namely,

$$N_{\text{oracle},t} = K_tN_t = \frac{\bar{X}_t\left(1-\bar{X}_t\right)}{\mu_{\text{post},t}(1-\mu_{\text{post},t})}\left(\frac{z^2_{\alpha_t/2}}{4K_t\varepsilon_t^2} - \frac{\mu_{\text{post},t}(1-\mu_{\text{post},t})}{\Delta_t^2}K_t\right).$$

Note in the above expression (i) the first term inside the parentheses is the sample complexity of Normal-IQAE according to Lemma 1, (ii) the second term inside the parentheses is a positive quantity

that reduces the sample complexity, and (iii) the multiplicative factor outside the parentheses accounts for the impact on the interval length when the center is shifted from the sample mean to the posterior mean.

To show that the sample complexity of Normal-BIQAE is lower than that of IQAE with the standard-schedule, we prove that the added computational cost of the multiplicative factor is outweighed by the reduction in sample complexity brought about by the second term in the parentheses. Specifically, the separation of the first term inside the parentheses and comparison of the multiplicative factor to one lead to the reformulation as follows:

$$N_{\mathrm{oracle},t} = \frac{z^2_{\alpha_t/2}}{4K_t}\frac{1}{\varepsilon_t^2} + \left(\frac{\bar{X}_t\left(1-\bar{X}_t\right)}{\mu_{\mathrm{post},t}\left(1-\mu_{\mathrm{post},t}\right)} - 1\right)\frac{z^2_{\alpha_t/2}}{4K_t}\frac{1}{\varepsilon_t^2} - \frac{\bar{X}_t\left(1-\bar{X}_t\right)}{\Delta_t^2}K_t. \tag{37}$$

To express this quantum sample complexity in terms of the target accuracy, we first factorize the difference between the additional multiplicative factor and one to give

$$\begin{aligned}
\frac{\bar{X}_t(1-\bar{X}_t)}{\mu_{\mathrm{post},t}(1-\mu_{\mathrm{post},t})} - 1 &= \frac{(\bar{X}_t-\mu_{\mathrm{post},t})(1-\bar{X}_t-\mu_{\mathrm{post},t})}{\mu_{\mathrm{post},t}(1-\mu_{\mathrm{post},t})} \\
&= \frac{1}{\mu_{\mathrm{post},t}(1-\mu_{\mathrm{post},t})}\left(\bar{X}_t - \frac{\frac{\mu_{0,t}}{\Delta_t^2}+\frac{\bar{X}_t}{\hat{\sigma}_t^2/N_t}}{\frac{1}{\Delta_t^2}+\frac{1}{\hat{\sigma}_t^2/N_t}}\right)(1-\bar{X}_t-\mu_{\mathrm{post},t}) \\
&= \frac{1}{\mu_{\mathrm{post},t}(1-\mu_{\mathrm{post},t})}\frac{\frac{\bar{X}_t-\mu_{0,t}}{\Delta_t^2}}{\frac{1}{\Delta_t^2}+\frac{1}{\hat{\sigma}_t^2/N_t}}(1-\bar{X}_t-\mu_{\mathrm{post},t}) \\
&= \underbrace{\frac{1}{\mu_{\mathrm{post},t}(1-\mu_{\mathrm{post},t})}\frac{\sigma^2_{\mathrm{post},t}}{\Delta_t^2}}_{\text{First term}}\underbrace{(\bar{X}_t-\mu_{0,t})\cdot(1-\bar{X}_t-\mu_{\mathrm{post},t})}_{\text{Second term}},
\end{aligned}$$

where we have used in the first step the algebraic identity

$$a(1-a)-b(1-b) = (a-b)(1-a-b) \tag{38}$$

and in the second and fourth steps the Bayesian updates $\mu_{\mathrm{post},t}$ and $\sigma^2_{\mathrm{post},t}$ as detailed in Algorithm 1. By the expression for $\sigma_{\mathrm{post},t}$ derived from the stopping criterion Eq. (36), we obtain the first term

$$\frac{\sigma^2_{\mathrm{post},t}}{\mu_{\mathrm{post},t}(1-\mu_{\mathrm{post},t})\Delta_t^2} = \frac{4K_t^2\varepsilon_t^2}{z^2_{\alpha_t/2}\Delta_t^2}.$$

Since $\mu_{\mathrm{post},t}$ is a weighted average of $\bar{X}_t$ and $\mu_{0,t}$, the second term is bounded by[5]

$$(\bar{X}_t-\mu_{0,t})\cdot(1-\bar{X}_t-\mu_{0,t}) = \bar{X}_t(1-\bar{X}_t)-\mu_{0,t}(1-\mu_{0,t}),$$

where the algebraic identity Eq. (38) is used again. Combining the two terms together yields

$$\frac{\bar{X}_t(1-\bar{X}_t)}{\mu_{\mathrm{post},t}(1-\mu_{\mathrm{post},t})} - 1 \le \left[\bar{X}_t(1-\bar{X}_t)-\mu_{0,t}(1-\mu_{0,t})\right]\cdot\frac{1}{\Delta_t^2}\frac{4K_t^2\varepsilon_t^2}{z^2_{\alpha_t/2}}.$$

Substituting these into the sample complexity at stage t in Eq. (37), we obtain

$$\begin{aligned}
N_{\mathrm{oracle},t} &\le \frac{z^2_{\alpha_t/2}}{4K_t}\frac{1}{\varepsilon_t^2} + \left[\bar{X}_t(1-\bar{X}_t)-\mu_{0,t}(1-\mu_{0,t})\right]\cdot\frac{K_t}{\Delta_t^2} - \frac{\bar{X}_t\left(1-\bar{X}_t\right)}{\Delta_t^2}K_t \\
&= \frac{z^2_{\alpha_t/2}}{4K_t}\frac{1}{\varepsilon_t^2} - \mu_{0,t}(1-\mu_{0,t})\cdot\frac{K_t}{\Delta_t^2}.
\end{aligned} \tag{39}$$

[5]Note the inequality in the first step holds regardless of the relationship between $\bar{X}_t$ and $\mu_{0,t}$, as no matter whether $\bar{X}_t \ge \mu_{0,t}$ (*i.e.*, $\mu_{\mathrm{post},t}$ lies between $\mu_{0,t}$ and $\bar{X}_t$) or $\bar{X}_t < \mu_{0,t}$ (*i.e.*, $\mu_{\mathrm{post},t}$ lies between $\bar{X}_t$ and $\mu_{0,t}$), the second term attains a maximum value $(\bar{X}_t-\mu_{0,t})\cdot(1-\bar{X}_t-\mu_{0,t})$ at $\mu_{\mathrm{post},t}=\mu_{0,t}$.

Inequality (39) demonstrates that the complexity of Normal-BIQAE is lower than that of the standard-scheduled IQAE by a positive term, which is in agreement with the hypothesis that injection of a Bayesian approach achieves a complexity reduction that outweighs the extra cost due to retargeting.

To further quantify the magnitude of this improvement, we consider the prior distribution preparation process as outlined in Algorithm 3. Using the mean and variance of the posterior distribution of stage $t-1$ according to Appendix Theorem 6, we show

$$\Delta_t^2 = \left(\frac{K_t}{K_{t-1}}\right)^2 \frac{\mu_{0,t}(1-\mu_{0,t})}{\mu_{post,t-1}(1-\mu_{post,t-1})} \cdot \sigma^2_{\text{post},t-1}.$$

As in the derivation of $\sigma_{\text{post},t}$ in Eq. (36), we have

$$\frac{\sigma^2_{\text{post},t-1}}{\mu_{\text{post},t-1}(1-\mu_{\text{post},t-1})} = \frac{4K_{t-1}^2\varepsilon_{t-1}^2}{z^2_{\alpha_{t-1}/2}}.$$

Thus, we conclude

$$\Delta_t^2 = \left(\frac{K_t}{K_{t-1}}\right)^2 \mu_{0,t}(1-\mu_{0,t}) \cdot \frac{4K_{t-1}^2\varepsilon_{t-1}^2}{z^2_{\alpha_{t-1}/2}}.$$

Choosing $\alpha_t = \alpha_{t-1} = \frac{\alpha}{T}$ and applying the union bound according to the convention of IQAE [93], we derive the prior variance

$$\Delta_t^2 = \mu_{0,t}(1-\mu_{0,t}) \cdot \frac{4K_t^2\varepsilon_{t-1}^2}{z^2_{\alpha_t/2}}.$$

Finally, via substitution of the expression for Δ_t^2 into the quantum sample complexity inequality in Eq. (39), we establish the desired upper bound result

$$N_{\text{oracle},t} \leq \frac{z^2_{\alpha_t/2}}{4K_t}\left(\frac{1}{\varepsilon_t^2} - \frac{1}{\varepsilon_{t-1}^2}\right).$$

C.7 Preparation of the Prior in Beta-BIQAE

Proof of Appendix Theorem 8. We prove the consistency of the estimator obtained in PreparePriorBeta Algorithm 6 by Theorem 2.2 of ref. [122], which we reproduce below in the context of Beta-BIQAE for ease of reference:

Theorem 2.2. *Given the following Assumptions A1–A3b, the MLE estimator derived in Algorithm 6, $(\hat{\alpha}_R, \hat{\beta}_R)$, almost surely converges to (α^*, β^*) as $R \to \infty$.*

Assumption A1. *The independent random variables $\tilde{Y}_i$, $i = 1, \ldots, R$, have a common distribution with a measurable density function g.*

Assumption A2. *The considered beta family has density functions $q_{\alpha,\beta}(\tilde{y})$ that are measurable in $\tilde{y}$ for every (α, β) in Θ, a compact subset of $\mathbb{R}^2$, and continuous in (α, β) for every $\tilde{y}$.*

Assumption A3a. $\mathbb{E}\left[\log g\left(\tilde{Y}_i\right)\right]$ *exists and* $|\log q_{\alpha,\beta}(\tilde{y})| \leq m(\tilde{y})$ *for all* (α, β) *in* Θ, *where* $m(\tilde{Y}_i)$ *is integrable.*

Assumption A3b. $D_{\text{KL}}(g \| q_{\alpha,\beta})$ *has a unique minimum at* (α^*, β^*) *in* Θ.

Thus, to employ Theorem 2.2 of ref. [122] we first verify Assumptions A1–A3 hold for Beta-BIQAE. According to Algorithm 6, in Beta-BIQAE, the sample drawn from the posterior distribution obtained at the end of Stage t is

$$Y_1, Y_2, \cdots, Y_R \overset{\text{i.i.d.}}{\sim} \text{Beta}\left(a_{\text{post},t}, b_{\text{post},t}\right),$$

and the sample $\tilde{Y}_1, \tilde{Y}_2, \cdots, \tilde{Y}_R$ obtained from the mapping in the algorithm is related to the original sample $Y_1, \cdots, Y_R$ as

$$Y_i = \sin^2\left(K_t \cdot f_{t+1}(\tilde{Y}_i)\right), \quad \forall i \in \{1, \ldots, R\}. \tag{40}$$

Assumption A1 that the transformed samples are i.i.d. and possess a density function is therefore satisfied since the original samples (Y_i's) are i.i.d. and follow a beta distribution, the transformation Eq. (40) is differentiable, and the density function of the transformed sample follows as Lemma 2:

Lemma 2. *The density function of $\tilde{Y}_i$ is*

$$\begin{aligned} g(\tilde{y}) =& \frac{\left[\sin^2\left(K_t \cdot f_{t+1}(\tilde{y})\right)\right]^{a_{post,t}-1} \cdot \left[\cos^2\left(K_t \cdot f_{t+1}(\tilde{y})\right)\right]^{b_{post,t}-1}}{B\left(a_{post,t}, b_{post,t}\right)} \cdot \frac{K_t}{K_{t+1}} \cdot \frac{1}{\sqrt{\tilde{y}(1-\tilde{y})}} \\ & \cdot \left|\sin\left(K_t \cdot f_{t+1}(\tilde{y})\right) \cos\left(K_t \cdot f_{t+1}(\tilde{y})\right)\right|. \end{aligned}$$

Proof of Lemma 2. Denote the mapping from Y_i to $\tilde{Y}_i$ by φ [given in Eq. (34)] such that the density function is expressed as

$$g(\tilde{y}) = p_Y\left(\varphi^{-1}(\tilde{y})\right) \cdot \left|\frac{d}{d\tilde{y}}\varphi^{-1}(\tilde{y})\right|,$$

where $p_Y(\cdot)$ is the density function of Y_i, and, according to Eq. (40), φ^{-1} is the inverse mapping

$$\varphi^{-1}(\tilde{y}) = \sin^2\left(K_t \cdot f_{t+1}(\tilde{y})\right).$$

Direct calculations then lead to

$$\frac{d}{d\tilde{y}}\left[\varphi^{-1}(\tilde{y})\right] = 2\sin\left(K_t \cdot f_{t+1}(\tilde{y})\right) \cdot \cos\left(K_t \cdot f_{t+1}(\tilde{y})\right) \cdot K_t \cdot f'_{t+1}(\tilde{y}),$$

which expressed in terms of the derivative of f_k in Eq. (32) is

$$\begin{aligned} & f'_{t+1}(\tilde{y}) = f'_{k_{t+1}}(\tilde{y}) = (-1)^{l_{t+1}} \cdot \frac{1}{2K_{t+1}} \cdot \frac{1}{\sqrt{\tilde{y}(1-\tilde{y})}} \\ \Rightarrow & \left|\frac{d}{d\tilde{y}}\left[\varphi^{-1}(\tilde{y})\right]\right| = \frac{K_t}{K_{t+1}} \cdot \frac{1}{\sqrt{\tilde{y}(1-\tilde{y})}} \cdot \left|\sin\left(K_t \cdot f_{t+1}(\tilde{y})\right) \cdot \cos\left(K_t \cdot f_{t+1}(\tilde{y})\right)\right|. \end{aligned}$$

Moreover, since Y_i follows the beta distribution $\mathrm{Beta}\left(a_{\mathrm{post},t}, b_{\mathrm{post},t}\right)$,

$$p_Y\left(\varphi^{-1}\left(\tilde{y}\right)\right) = \frac{\left[\varphi^{-1}(\tilde{y})\right]^{a_{\mathrm{post},t}-1} \cdot \left[1-\varphi^{-1}(\tilde{y})\right]^{b_{\mathrm{post},t}-1}}{B\left(a_{\mathrm{post},t}, b_{\mathrm{post},t}\right)}.$$

Combination of the expressions for the gradient of the inverse mapping $\varphi^{-1}(\tilde{y})$ and $p_Y\left(\varphi^{-1}(\tilde{y})\right)$ thus yields the expression for the density function of $\tilde{Y}_i$

$$\begin{aligned} g(\tilde{y}) =& \frac{\left[\sin^2\left(K_t \cdot f_{t+1}(\tilde{y})\right)\right]^{a_{\mathrm{post},t}-1} \cdot \left[\cos^2\left(K_t \cdot f_{t+1}(\tilde{y})\right)\right]^{b_{\mathrm{post},t}-1}}{B\left(a_{\mathrm{post},t}, b_{\mathrm{post},t}\right)} \cdot \frac{K_t}{K_{t+1}} \cdot \frac{1}{\sqrt{\tilde{y}(1-\tilde{y})}} \\ & \cdot \left|\sin\left(K_t \cdot f_{t+1}(\tilde{y})\right) \cos\left(K_t \cdot f_{t+1}(\tilde{y})\right)\right|. \end{aligned}$$

Assumption A2 is likewise satisfied due to the chosen beta family and parameter space: Specifically, the parameter space Θ in Theorem 8 is a finite rectangle and thus a compact set, and the density function of the beta distribution

$$q_{\alpha,\beta}(\tilde{y}) = \frac{1}{B(\alpha,\beta)}\, \tilde{y}^{\alpha-1}(1-\tilde{y})^{\beta-1}$$

is continuous in $(\alpha, \beta) \in (0, \infty)^2$ for every $\tilde{y}$.

To confirm Assumption A3a is satisfied, we first verify the existence of $\mathbb{E}[\log(g(\tilde{Y}_i))]$. According to Lemma 2, $g(\tilde{Y}_i)$ can be expressed as

$$\begin{aligned} g(\tilde{Y}_i) &= \frac{Y_i^{a_{\text{post},t}}(1-Y_i)^{b_{\text{post},t}}}{B\left(a_{\text{post},t}, b_{\text{post},t}\right)} \cdot \sqrt{\frac{Y_i(1-Y_i)}{\tilde{Y}_i(1-\tilde{Y}_i)}} \cdot \frac{K_t}{K_{t+1}} \\ \Rightarrow \log g(\tilde{Y}_i) &= \left(a_{\text{post},t} + \frac{1}{2}\right)\log Y_i + \left(b_{\text{post},t} + \frac{1}{2}\right)\log(1-Y_i) \\ &\quad - \frac{1}{2}\log \tilde{Y}_i - \frac{1}{2}\log(1-\tilde{Y}_i) \\ &\quad - \log B\left(a_{\text{post},t}, b_{\text{post},t}\right) + \log\left(\frac{K_t}{K_{t+1}}\right). \end{aligned}$$

Direct calculation then shows

$$\mathbb{E}[\log Y_i] = \psi\left(a_{\text{post},t}\right) - \psi\left(a_{\text{post},t} + b_{\text{post},t}\right),$$

where ψ is the digamma function; and since $1 - Y_i \sim \text{Beta}\left(b_{\text{post},t}, a_{\text{post},t}\right)$,

$$\mathbb{E}[\log\left(1-Y_i\right)] = \psi\left(b_{\text{post},t}\right) - \psi\left(a_{\text{post},t} + b_{\text{post},t}\right).$$

It is then sufficient to show that $\mathbb{E}[\log \tilde{Y}_i]$ and $\mathbb{E}[\log(1-\tilde{Y}_i)]$ are finite, which is proven in Lemma 2:

Lemma 3. *$\mathbb{E}[\log \tilde{Y}_i]$ and $\mathbb{E}[\log(1-\tilde{Y}_i)]$ are finite.*

Proof of Lemma 3. By Lemma 2,

$$\begin{aligned} \mathbb{E}[|\log \tilde{Y}_i|] &= \int_0^1 |\log \tilde{y}| \cdot g(\tilde{y}) d\tilde{y} \\ &\leq \frac{-1}{B\left(a_{\text{post},t}, b_{\text{post},t}\right)} \int_0^1 \frac{\log \tilde{y}}{\sqrt{\tilde{y}(1-\tilde{y})}} d\tilde{y} \\ &= \frac{2\pi \log 2}{B\left(a_{\text{post},t}, b_{\text{post},t}\right)}, \end{aligned}$$

and

$$\mathbb{E}[|\log(1-\tilde{Y}_i)|] \leqslant \frac{-1}{B\left(a_{\text{post},t}, b_{\text{post},t}\right)} \cdot \int_0^1 \frac{\log(1-\tilde{y})}{\sqrt{\tilde{y}(1-\tilde{y})}} d\tilde{y} = \frac{2\pi \log 2}{B\left(a_{\text{post},t}, b_{\text{post},t}\right)}.$$

We then confirm the remainder of Assumption A3a holds: Given that

$$|\log q_{\alpha,\beta}(\tilde{y})| \leqslant |\alpha - 1| \cdot |\log \tilde{y}| + |\beta - 1| \cdot |\log(1-\tilde{y})| + |\log B(\alpha, \beta)|,$$

we define the control function

$$\begin{aligned} m(\tilde{y}) \coloneqq &\max\left\{|\alpha_{\min} - 1|, |\alpha_{\max} - 1|\right\} \cdot |\log \tilde{y}| + \\ &\max\left\{|\beta_{\min} - 1|, |\beta_{\max} - 1|\right\} \cdot |\log(1-\tilde{y})| + \\ &\max_{(\alpha,\beta)\in\Theta} |\log B(\alpha, \beta)|. \end{aligned}$$

Note that $m(\tilde{y})$ is well-defined because $B(\alpha, \beta)$ is continuous on $(0, +\infty)^2$ and Θ is a rectangle in $(0, +\infty)^2$. By Lemma 3, $m(\tilde{Y}_i)$ is integrable, and thus $m(\tilde{y})$ is a valid control function.

Finally, we verify Assumption A3(b). The Kullback–Leibler divergence of $q_{\alpha,\beta}$ from g can be expressed as

$$D_{\text{KL}}(g \| q_{\alpha,\beta}) = \int g(\tilde{y}) \log g(\tilde{y})\, d\tilde{y} - \int g(\tilde{y}) \log q_{\alpha,\beta}(\tilde{y})\, d\tilde{y},$$

where g does not depend on (α, β). Thus, Assumption A3(b) is equivalent to the assumption that

$$h(\alpha, \beta) \coloneqq -\mathbb{E}[\log q_{\alpha,\beta}(\tilde{Y}_i)]$$

has a unique minimum in Θ, which we prove as follows:

Since

$$h(\alpha, \beta) = \log B(\alpha, \beta) - \mathbb{E}[\log \tilde{Y}_i] \cdot (\alpha - 1) - \mathbb{E}[\log(1 - \tilde{Y}_i)](\beta - 1),$$

the finiteness of $h(\alpha, \beta)$ is guaranteed by Lemma 3; according to Lemma 4 below, $\log B(\alpha, \beta)$ is strictly convex on $(0, +\infty)^2$; and since the remaining terms are affine, the function $h(\alpha, \beta)$ is strictly convex on $(0, +\infty)^2$. The unique minimizer of $h(\alpha, \beta)$ over Θ is thus implied by the fact that Θ is a closed compact set.

Lemma 4. *$B(\alpha, \beta)$ is strictly log-convex on $(0, +\infty)^2$.*

Proof. Let $\alpha_1, \beta_1, \alpha_2, \beta_2 \in (0, +\infty), u, v \geqslant 0, u + v = 1$. We have

$$\begin{aligned}
B\left(u\left(\alpha_1, \beta_1\right) + v\left(\alpha_2, \beta_2\right)\right) &= B\left(u\alpha_1 + v\alpha_2, u\beta_1 + v\beta_2\right) \\
&= \int_0^1 t^{u\alpha_1 + v\alpha_2 - 1}(1-t)^{u\beta_1 + v\beta_2 - 1} dt \\
&= \int_0^1 t^{u(\alpha_1 - 1) + v(\alpha_2 - 1)}(1-t)^{u(\beta_1 - 1) + v(\beta_2 - 1)} dt \\
&= \int_0^1 \underbrace{\left[t^{\alpha_1 - 1}(1-t)^{\beta_1 - 1}\right]^u}_{f_1(t)} \cdot \underbrace{\left[t^{\alpha_2 - 1}(1-t)^{\beta_2 - 1}\right]^v}_{f_2(t)} dt \\
&= \int_0^1 f_1(t) f_2(t) dt \\
&\leq \left[\int_0^1 f_1^{\frac{1}{u}}(t) dt\right]^u \cdot \left[\int_0^1 f_2^{\frac{1}{v}}(t) dt\right]^v \\
&= \left[\int_0^1 t^{\alpha_1 - 1}(1-t)^{\beta_1 - 1} dt\right]^u \cdot \left[\int_0^1 t^{\alpha_2 - 1}(1-t)^{\beta_2 - 1} dt\right]^v \\
&= B^u\left(\alpha_1, \beta_1\right) \cdot B^v\left(\alpha_2, \beta_2\right),
\end{aligned}$$

where the sixth step follows from Hölder's inequality. Therefore, $B(\alpha, \beta)$ is log-convex on $(0, +\infty)^2$.

To show that $B(\alpha, \beta)$ is strictly log-convex, we further note that the equality in the sixth step holds if and only if $\lambda |f_1|^{\frac{1}{u}} = \mu |f_2|^{\frac{1}{v}}$ almost everywhere for some constants $\lambda, \mu \geq 0$ not simultaneously zero. With substitution of the expressions for f_1 and f_2, this condition becomes, for some constants $\lambda, \mu \geq 0$, not both zero, almost everywhere

$$\begin{aligned}
& \lambda\, t^{\alpha_1 - 1}(1-t)^{\beta_1 - 1} = \mu\, t^{\alpha_2 - 1}(1-t)^{\beta_2 - 1} \\
\Leftrightarrow \quad & t^{\alpha_1 - \alpha_2}(1-t)^{\beta_1 - \beta_2} = \text{const} \\
\Leftrightarrow \quad & \alpha_1 = \alpha_2, \quad \beta_1 = \beta_2.
\end{aligned}$$

That is, the equality holds if and only if $(\alpha_1, \beta_1) = (\alpha_2, \beta_2)$, which implies that the log-convexity of $B(\alpha, \beta)$ is strict.

□