Quantum Nested Sampling

S M A Allahyar, Theodorous Yianni

I. Introduction

This paper introduces a groundbreaking approach to nested sampling by incorporating quantum computing, which offers significant speedups in three key areas of computational time complexity. First, we enhance the replacement of dead points with live points using a quantum search algorithm, reducing the complexity from $O(N)$ to $O(\sqrt{N})$, where N is the number of points. Secondly, a quantum summation algorithm is employed for rapid loglikelihood evaluation, potentially reducing the time complexity from linear to sub-linear scaling. Lastly, Quantum Superposed Multi-Run Nested Sampling is proposed, utilizing quantum parallelism to achieve an error reduction that scales with $1/N$, surpassing the classical $1/\sqrt{N}$ scaling.

II. Nested Sampling and Time Complexity

Nested sampling, a computational approach for Bayesian analysis, operates by initializing a set of points (live points) from the prior distribution. The algorithm iteratively replaces the lowest likelihood point with a new point from a higher likelihood region. This process estimates the density of states and computes the posterior and evidence, with associated errors.

The computational time complexity of nested sampling, denoted as T, is directly proportional to the number of live points (n_{live}), the time complexity of loglikelihood evaluation ($T_{L(\theta)}$), an implementation-specific factor ($T_{\text{Impl.}}$), and the Kullback-Leibler divergence between the prior and the posterior ($D_\pi\{P\}$). Mathematically, it can be represented as:

$$T \propto n_{\text{live}} \times T_{L(\theta)} \times T_{\text{Impl.}} \times D_\pi\{P\}$$

The time complexity of the likelihood evaluation and the implementation-specific factor are key contributors, particularly in multi-dimensional parameter spaces. Strategies to reduce the runtime involve optimizing likelihood evaluation, improving replacement efficiency, and enhancing prior to posterior compression. In this paper we introduce three quantum algorithms to reduce the contributing factors of nested sampling runtime:

1) $T_{\textbf{Impl.}}$: the time complexity of the replacement of a dead point (reducing $T_{\text{Impl.}}$).
2) $T_{L(\theta)}$: We will reduce the time complexity of evaluating the loglikelihood function by utilising a quantum summation algorithm that significantly speeds up summation of large amounts of terms.
3) **Quantum Superposed Multi-Run Nested Sampling**

Nested sampling's accuracy is linked to the accumulated Poisson noise, influencing the error in the evidence logarithm ($\ln Z$). This relationship showcases that faster compression not only speeds up the process but also enhances accuracy and precision.

III. The Nested Sampling Algorithm

Nested sampling, a robust tool for Bayesian inference, excels in computing evidence Z and sampling from the posterior $\mathcal{P}$, especially in high-dimensional spaces. Its efficiency stems from a probabilistic relationship between the likelihood function L and the prior volume X. The prior volume of the nested sampling algorithm at its ith iteration is defined by:

$$X_i = \int_{L(\theta)>L_i} \pi(\theta) d\theta, \tag{1}$$

which is an integral of the prior over the region contained within an iso-likelihood contour, $L(\theta) = L_i$. The algorithm proceeds as follows:

1) **Initialization**: Sample N 'live points' from the prior distribution $\pi(\theta)$. Set the initial prior volume X_0 to 1.
2) **Iteration**: At each iteration i, remove the point with the lowest likelihood L_i (termed 'dead point').
3) **Replacement**: Replace the dead point with a new point, sampled uniformly from the prior with the constraint $L(\theta_{\text{new}}) > L_i$. Classically, the method to find a new point subject to this constraint is to uniformly randomly generate a set of parameters, θ, which is a vector. Then we evaluate the likelihood at this point in the parameter space and if it satisfies the constraint, $L(\theta_{\text{new}}) > L_i$, then we select it as a replacement for the 'dead point'. Otherwise we keep uniformly randomly generating a new set of parameters, θ, and check the constraint until it is satisfied. These, θ, are generated and checked for the constraint one by one. This process involves the probabilistic reduction of the prior volume, as described by:
$$X_{i+1} = X_i t_{i+1},$$
where t_i is drawn from a power law distribution.
4) **Approximation of Prior Volume**: The prior volume X_i and its approximation through a set of simulated ts, using the power law distribution, are given by:
$$\log X_i \approx -\frac{i \pm \sqrt{i}}{N}.$$
5) **Evidence Computation**: Convert the multidimensional integral into a single-variable integral:
$$Z = \int_0^1 L(X) dX,$$
and evaluate it using a weighted sum of dead points.

6) **Termination**: The algorithm concludes when the estimated remaining posterior mass is below a predefined threshold.
7) **Posterior Sampling**: Generate posterior samples using the weights assigned to each dead point.

This orthodox nested sampling, as proposed by Skilling, serves as a foundation for more advanced variants, such as gradient nested sampling, which are further discussed in the subsequent chapters.

IV. Quantum Enhancement in Nested Sampling

Our novel approach integrates quantum computing into nested sampling, significantly reducing the implementation time complexity, $T_{\text{Impl.}}$. By employing a variant of Grover's algorithm, a quantum search algorithm, we achieve a quadratic speedup in searching the parameter space.

Grover's algorithm, a quantum search technique, locates specific items within an unsorted database with a quadratic speedup compared to classical algorithms. In our context, the algorithm optimizes the search for new live points within higher likelihood regions in nested sampling.

As shown in [1], there is a quantum algorithm based on Grover's algorithm, which can find an element that satisfies some particular search criterion, out of a list of elements where the proportion of elements that satisfies this search criterion is unknown. This can be applied to nested sampling as a way of finding a point that exceeds a likelihood threshold in order to replace the dead point in step 3 of the algorithm.

This works by initialising a large random set of points and then applying the quantum search algorithm, which randomly picks out of the points which exceeds the likelihood threshold.This process effectively reduces the $T_{\text{Impl.}}$ component of the time complexity. The quantum oracle encodes the likelihood function, identifying states exceeding a defined likelihood threshold. The diffusion operator then amplifies the probability amplitudes of these states, enhancing the efficiency of finding new live points, as will be explained in the next section.

The integration of quantum computing into nested sampling not only accelerates the computational process but also maintains the accuracy and robustness of the algorithm. This quantum enhancement opens new possibilities for tackling high-dimensional and complex problems in Bayesian analysis.

V. Introduction to Grover's Algorithm

Grover's algorithm stands as a cornerstone in quantum computing, offering a paradigm shift in the efficiency of searching unsorted databases. By leveraging the principles of quantum superposition and entanglement, Grover's algorithm achieves a quadratic speedup over classical counterparts. Specifically, it can locate a unique item within an unsorted database of N items in $O(\sqrt{N})$ operations, compared to $O(N)$ in a classical scenario.

Grover's algorithm uses amplitude amplification, which is the principle of how a general quantum search algorithm works. First we initialise the state $|s\rangle$, which is an equal superposition of states $|x_i\rangle$, where each $|x_i\rangle$ represents an element x_i in the search database. Let the operator U represent the quantum oracle, which transforms $|x_i\rangle$ as $U|x_i\rangle = -|x_i\rangle$ if x_i satisfies the search criterion, and $U|x_i\rangle = |x_i\rangle$ otherwise. Let the operator U_s be the operator which changes the sign of the amplitude of only the $|s\rangle$ state, this operator is the diffusion operator.

Then, the operator $Q = -U_s U$ increases the amplitude of states which satisfy the search criterion. If the proportion of elements in the search database which satisfy the criterion is a, then the number of times Q would need to be applied would be of order $\frac{1}{\sqrt{a}}$ [1]. A classical unstructured search algorithm would require a number of searches of order $\frac{1}{a}$.

VI. Quantum-Assisted Point Selection in Nested Sampling

For point selection in nested sampling, the oracle U would have to be able to calculate the likelihood function of a point and determine if it's above the likelihood threshold. This means that if a is the proportion of points above the likelihood threshold, the likelihood function would need to be calculated roughly $\frac{1}{\sqrt{a}}$ times. The search database would be a set of N randomly chosen points θ, and the number of points would depend on a. Even if we don't know a exactly, we an find a rough estimate of a (explain how), and set N to be a number significantly larger than $\frac{1}{a}$. This is because if the database has no element which satisfies the criterion, it would be a waste of resources to search it, and the number of iterations it takes to find such an element is asymptotically constant for large N. The only costs of making N large are the time taken to initialise the database on the quantum computer, which is $O(log(N))$, and the complexity of each application of the diffusion operator, which is $O(log(N))$. To be more exact, if $|s\rangle$ is of the form $|s_1\rangle \otimes |s_2\rangle \otimes ...|s_n\rangle$, where each $|s_i\rangle$ is made up of m qubits and is a superposition of two states, then $N = 2^n$ and there are $\binom{2^m}{2}^n$ possible initialisations of N states. The complexity of initialising $|s\rangle$ would be $O(mn) = O(log(N)m)$. This means we wouldn't want to make N too large, but the optimal value of N would also depend on the complexity of evaluating the likelihood, specifically, a more complicated likelihood would mean N should be larger.

In practice, this means that the quantum-enhanced nested sampling algorithm, or Quantum Nested Sampling (QNS), can identify new live points with greater efficiency than classical methods. QNS promises to mitigate the curse of dimensionality in high-dimensional problems, maintaining the robustness of Bayesian inference while drastically reducing computational overhead.

Through Quantum Nested Sampling, the integration of quantum algorithms like Grover's into traditional statistical mechanics methods heralds a new era of computational capability, redefining the limits of data analysis and inference.

VII. Quantum Log-Likelihood Evaluation in Cosmology

The practical implementation of quantum algorithms for cosmological data analysis necessitates the efficient evaluation of the log-likelihood function. For cosmological models, the log-likelihood often entails the summation of a substantial number of terms, where each term is a function of model parameters (θ), observed data points (d_i), and associated uncertainties (σ_i). Specifically, the Gaussian log-likelihood function in cosmology can be expressed as:

$$\ln L(\theta) = -\frac{1}{2} \sum_i \left[\ln(2\pi\sigma_i^2) + \frac{(d_i - m_i(\theta))^2}{\sigma_i^2} \right]$$

Here, $m_i(\theta)$ represents the model predictions for the observational data points. The quantum acceleration of this computation leverages the quantum summation algorithm, allowing for a potential sub-linear scaling with the number of terms. This capability not only hastens the log-likelihood evaluation but also supports the analysis of more extensive datasets, propelling cosmological inquiry to previously unattainable levels of precision and scale. The quantum summation algorithm, therefore, is a pivotal advancement for the field, presenting a quantum computational approach to a classically intensive problem.

As in [2], there is a quantum algorithm to estimate $\sum_i f(x_i)$, and the mean squared error of this estimate scales as $\frac{1}{q}$, where q is the number of iterations. Firstly, we initialise the state $|\psi\rangle\otimes|0\rangle = (\frac{1}{\sqrt{N}} \sum_{i=1}^{N} |x_i\rangle)\otimes|0\rangle = (P\otimes I)|\mathbf{0}\rangle$, where I is the identity operator and P initialises the superposition of $|x_i\rangle$ states. We assume that $f(x_i) \in [0,1]\ \forall x_i$, if not, then we can transform f so that this assumption holds, then at the end we perform the reverse transformation to our estimate. Then let R be the operator which transforms the state $|x_i\rangle \otimes |0\rangle$ as $R|x_i\rangle \otimes |0\rangle = |x_i\rangle \otimes (\sqrt{1-f(x_i)}|0\rangle + \sqrt{f(x_i)}|1\rangle$. Let the operator $A = R(P \otimes I)$ and $A|\mathbf{0}\rangle = \frac{1}{\sqrt{N}} \sum_i^N (|x_i\rangle \otimes (\sqrt{1-f(x_i)}|0\rangle + \sqrt{f(x_i)}|1\rangle) = cos\phi|\Phi_0\rangle|0\rangle + sin\phi|\Phi_1\rangle|1\rangle$, where ϕ depends on $\sum_i f(x_i)$. Now, $\sum_i f(x_i)$ can be estimated by estimating ϕ. Let U be the operator which flips the sign of amplitude only if the last qubit is $|1\rangle$. Let U_0 be the operator which flips the sign of the amplitude only if the last qubit is $|0\rangle$. Then $Q = -AU_0A^{-1}U$. As shown in [1], repeated applications of Q scale up ϕ as

$$Q^j R(|\psi\rangle \otimes |0\rangle) = cos((2j+1)\phi)|\Phi_0\rangle|0\rangle \quad (2)$$

$$+sin((2j+1)\phi)|\Phi_1\rangle|1\rangle \quad (3)$$

. [1] shows how a quantum Fourier transform can be used to estimate ϕ, and thus $\sum_i f(x_i)$.

VIII. Quantum Superposed Multi-Run Nested sampling

In classical nested sampling, one can perform multiple runs and take the average as a way of reducing error. Each run can be thought of as a function of a set of random seed variables. The average of N runs can be written as $\frac{1}{N} \sum_{i=1}^{N} f(x_i)$, where each x_i represents a random set of seed variables, and the function f calculated the result of running the nested sampling algorithm with a set of seed variables as the input.

Now we can see that by initialising a superposition of states of the form $|x_i\rangle$ which each represent a set of seed variables, and by building a unitary operator R which can compute the output of a nested sampling run with x_i as the input, we can use the same type of quantum summation algorithm as in the last section to find the average of multiple nested sampling runs more efficiently. Specifically, using the quantum method, the error scales as $\frac{1}{N}$, whereas classically it scales as $\frac{1}{\sqrt{N}}$.

.

IX. Conclusion

The integration of quantum algorithms into nested sampling marks a significant advance, reducing computational time complexities across several fronts. The utilization of Grover's search algorithm offers a quadratic speedup in identifying new live points. The quantum summation algorithm provides an efficient method for evaluating the loglikelihood function over large datasets. Furthermore, Quantum Superposed Multi-Run Nested Sampling employs quantum parallelism to enhance the precision of Bayesian evidence estimation. These contributions collectively forge a path for more efficient and accurate high-dimensional data analysis in cosmological and other complex scientific inquiries.

References

[1] G. Brassard, P. Høyer, M. Mosca, and A. Tapp, "Quantum amplitude amplification and estimation," *Quantum Computation and Information*, p. 53–74, 2002. [Online]. Available: http://dx.doi.org/10.1090/conm/305/05215

[2] S. Herbert, "Quantum monte carlo integration: The full advantage in minimal circuit depth," *Quantum*, vol. 6, p. 823, Sep. 2022. [Online]. Available: http://dx.doi.org/10.22331/q-2022-09-29-823