Cardinality-Aware Quantum Retrieval for Large-Scale Predicate Matching in Big Data Systems

Yang Lu[1,*]

[1] *Faculty of Business, Information & Human Sciences，Kuala Lumpur University of Science and Technology, Kajang, Malaysia, 43000*
* *Correspondence: yanglu@klust.edu.cn*

Abstract

Large-scale predicate matching in big data systems—identifying all records that satisfy complex multi-attribute query conditions—represents one of the most computationally intensive tasks in modern data analytics. Classical approaches based on full table scans or index lookups face prohibitive scalability constraints as datasets grow into the petabyte range. Quantum computing, and Grover's search algorithm in particular, offers a theoreticallyquadratic speedup for unstructured search; however, the efficiency of any quantum search procedure depends critically on an accurate estimate of the number of matching records,known as the query cardinality. Inaccurate cardinality estimates cause the quantum circuit to operate at a suboptimal number of iterations, increasing both failure rates and missed matches. This paper proposes the Cardinality-Aware Quantum Retrieval (CAQR) framework, which integrates quantum amplitude estimation for adaptive cardinality inference with a Bayesian posterior update mechanism that continuously refines the cardinality estimate during query execution. CAQR constructs efficient quantum oracles for conjunctive and disjunctive predicate conditions and dynamically adjusts Grover iteration counts to maintain near-optimal success probabilities throughout the retrieval process. Simulation experiments over synthetic datasets with search spaces ranging from 2^{12} to 2^{24} demonstrate that CAQR reduces the fraction of missed predicate matches from between 6. 1% and 14.8% (baseline) to below 0.82% across all tested configurations, while also reducing total oracle calls by approximately 14–17% compared to fixed-cardinality quantum retrieval. The framework is evaluated against classical, naive quantum, and learned-cardinality baselines, and its theoretical time and space complexity are analysed. These results establish CAQR as a competitive and practically relevant approach for quantum-assisted big data query processing.

Keywords: Quantum retrieval; cardinality estimation; predicate matching; big data query;Grover's algorithm; quantum amplitude estimation

Article History

Received: January 19, 2026

Accepted: May 24, 2026

Published: une 30, 2026

Cardinality-Aware Quantum Retrieval for Large-Scale Predicate Matchingin Big Data Systems

1. Introduction

The exponential growth of data produced by digital systems, connected devices, and enterprise applications has made large-scale data retrieval one of the defining challenges of contemporary computing (Hashem et al., 2015; Oussous et al., 2018). Predicate matching—the process of identifying all records in a dataset that satisfy one or more logical conditions expressed as a query—lies at the core of relational database operations, information retrieval pipelines, and event-stream processing architectures. When datasets reach the scale of billions of rows stored across distributed clusters, classical query execution strategies that rely on sequential scans or B-tree traversals encounter fundamental limitations imposed by the linear growth of computational cost with dataset size (Zaharia et al., 2016; Chen & Zhang, 2014).

Query optimisers in classical database management systems depend on cardinality estimation— an estimate of the number of records that satisfy a query predicate—to select between candidate execution plans (Leis et al., 2015). Accurate cardinality estimates are essential for minimising query execution time; an overestimate may cause the engine to allocate unnecessary resources, while an underestimate can lead to catastrophically slow nested-loop joins or incomplete result sets (Marcus & Papaemmanouil, 2019). The accuracy of cardinality estimation has improved substantially through learned approaches that incorporate deep learning models and probabilistic methods (Yang et al., 2020; Wang et al., 2021; Hilprecht et al., 2020), yet fundamental challenges remain when predicates involve multiple correlated attributes or when data distributions shift rapidly over time.

Quantum computing has emerged as a paradigm that can, in principle, overcome the linear cost barrier for unstructured search. Grover's algorithm (Biamonte et al., 2017; Montanaro, 2016) provides a provably quadratic speedup for identifying one or more marked elements in an unstructured search space of size N: whereas a classical sequential scan requires O(N) oracle calls on average, Grover's algorithm requires only $O(\sqrt{N})$ calls when a single solution exists, and $O(\sqrt{(N/M)})$ calls when M solutions exist and M is known in advance. This quadratic advantage becomes practically significant at the scales typical of big data applications, where N may exceed 10^9 records. The current era of Noisy Intermediate-Scale Quantum (NISQ) hardware limits direct deployment, but simulation studies and theoretical analyses provide a rigorous basis for evaluating quantum retrieval algorithms (Preskill, 2018; Bharti et al., 2022; Kim et al., 2023).

A critical but underexplored dependency in applying Grover's algorithm to database retrieval is the relationship between query cardinality M and search efficiency. The optimal number of Grover iterations—$\lfloor \pi/4 \times \sqrt{(N/M)} \rfloor$ —depends directly on M (Brassard et al., 2002; Aaronson & Rall, 2020). In the context of discovering all M matching records rather than a single one, an inaccurate cardinality estimate $\hat{M}$ causes the algorithm to either terminate prematurely (when $\hat{M} < M$, missing unrecovered matches) or to perform superfluous iterations (when $\hat{M} > M$, wasting oracle calls). Cardinality estimation methods, including Quantum Counting (Aaronson & Rall, 2020) and quantum amplitude estimation (Suzuki et al., 2020), themselves introduce approximation errors whose magnitude depends on the chosen error tolerance ε . Reducing ε to near zero is computationally prohibitive, making some degree of estimation error unavoidable in practical deployments.

To address this limitation, this paper introduces the Cardinality-Aware Quantum Retrieval (CAQR) framework. CAQR incorporates three principal innovations: (i) a quantum amplitude estimation phase that provides an adaptive initial cardinality estimate $\hat{M}$ with a controllable error margin; (ii) a

ISSN: 3070-345X © 2026 INATGI (Institute of Advanced Technology and Green Innovation)
https://doi.org/10.63646/IBMY34582

Bayesian posterior update mechanism that refines $\hat{M}$ during iterative retrieval based on the observed success and failure history of Grover applications; and (iii) an oracle construction procedure that encodes conjunctive and disjunctive predicate conditions efficiently into quantum circuits compatible with multi-attribute big data schemas. The combination of these components ensures that CAQR's estimate of the remaining cardinality converges toward the true value over time, maintaining near-optimal Grover iteration counts throughout the retrieval process and substantially reducing the fraction of missed predicate matches.

The remainder of this article is organised as follows. Section 2 reviews related literature on big data query processing, quantum algorithms, and cardinality estimation. Section 3 describes the design and methodology of CAQR. Section 4 details the experimental setup and analytical framework. Section 5 presents empirical simulation results. Sections 6–8 provide discussion, implications, and limitations. Section 9 concludes.

2. Literature Review

2.1 Big Data Query Processing

The challenges of processing queries over large-scale datasets have motivated the development of distributed computing frameworks such as Apache Spark (Zaharia et al., 2016) and cloud- native data warehouse architectures that can parallelise scan operations across many nodes. Despite their engineering achievements, these platforms remain fundamentally constrained by the linear complexity of full-table predicate evaluation; distributed parallelism reduces constant factors but does not alter asymptotic complexity with respect to N. Assunção et al. (2015) and Chen & Zhang (2014) survey the breadth of big data processing challenges, noting that query latency over multi-petabyte repositories remains problematic even with contemporary hardware. Oussous et al. (2018) provide a comprehensive taxonomy of NoSQL technologies, highlighting the trade-offs between consistency, availability, and query expressiveness that arise when predicate-matching workloads are distributed across heterogeneous storage systems.

Cardinality estimation is central to query optimisation in both relational and non-relational systems. Early statistical histogram-based approaches suffer when predicates involve correlated attributes, as they assume attribute independence (Leis et al., 2015). Machine learning approaches have substantially improved estimation accuracy: Kipf et al. (2019) demonstrated that deep learning models trained on join queries can predict correlated cardinalities more accurately than histograms, and Yang et al. (2020) extended this work to multi-table scenarios with NeuroCard. Hilprecht et al. (2020) introduced DeepDB, a data-driven approach using probabilistic sum-product networks that does not require query logs. Han et al. (2021) conducted a comprehensive benchmark evaluation showing that learned estimators consistently outperform traditional statistical methods across diverse workloads, though all learned methods are susceptible to distribution shift. Zhu et al. (2021) proposed FLAT, which exploits factorised representations for efficient high-dimensional cardinality estimation. Despite this progress, no existing approach dynamically corrects cardinality estimates during quantum query execution, which is the gap addressed by CAQR.

2.2 Quantum Algorithms for Data Processing

Quantum computing has been proposed as a means of achieving super-polynomial or polynomial speedups for a variety of data-intensive tasks. Biamonte et al. (2017) provide a foundational survey of quantum machine learning, identifying several contexts in which quantum subroutines can accelerate data analysis. Rebentrost et al. (2014) demonstrated that quantum support vector machines can be trained in O(log N) time given quantum random- access memory (QRAM) (Giovannetti et al., 2008), providing exponential acceleration for specific classification tasks.

ISSN: 3070-345X © 2026 INATGI (Institute of Advanced Technology and Green Innovation)
https://doi.org/10.63646/IBMY34583

Lloyd et al. (2016) showed that quantum algorithms for topological data analysis can offer superpolynomial speedups over known classical methods. Lu & Yang (2024) surveyed quantum algorithm applications to financial information systems, highlighting the practical barriers introduced by NISQ hardware limitations. Lu et al. (2024) reviewed quantum machine learning classifications and challenges, emphasising that hybrid quantum- classical architectures are the most viable path for near-term deployment. Lu et al. (2023) analysed the intersection of quantum computing and industrial information integration, concluding that quantum search and optimisation represent the most mature application domains. Ye & Lu (2022) surveyed quantum science broadly, discussing the state of quantum hardware and the outlook for practical quantum advantage.

Grover's algorithm (Montanaro, 2016; Biamonte et al., 2017) remains the most well-characterised quantum search procedure. Brassard et al. (2002) formalised quantum amplitude amplification and estimation, which generalises Grover's algorithm and provides the theoretical basis for CAQR's cardinality estimation component. Suzuki et al. (2020) proposed amplitude estimation without phase estimation, reducing the circuit depth required and making the approach more compatible with NISQ hardware. Aaronson & Rall (2020) further simplified quantum approximate counting, providing tighter complexity bounds. Woerner & Egger (2019) applied quantum amplitude estimation to risk analysis in finance, demonstrating its practical viability on near-term hardware. Egger et al. (2020) reviewed quantum computing for finance broadly, noting that amplitude estimation for probability distribution queries is among the most practically applicable near-term algorithms.

2.3 Quantum-Classical Hybrid Architectures

CAQR belongs to the class of variational or hybrid quantum-classical algorithms, in which a classical processor manages control flow and state while a quantum processor executes computationally intensive subroutines (McClean et al., 2016; Cerezo et al., 2021). Bharti et al. (2022) review NISQ-compatible algorithms, noting that hybrid architectures are essential for tolerating the noise levels present in current quantum processors. Mitarai et al. (2018) introduced the quantum circuit learning framework, which uses a classical optimiser to train parameterised quantum circuits—a paradigm that inspired CAQR's Bayesian update mechanism. Zhang & Lu (2021) discussed the integration of artificial intelligence with information integration systems, drawing parallels between adaptive learning in classical AI and the hybrid update loops required in quantum-classical workflows. Lu (2019) similarly noted that adaptive, feedback-driven algorithms consistently outperform static counterparts in complex search and retrieval tasks.

3. Research Design and Methodology

3.1 Problem Formulation

Let D be a dataset ofN records, each with attributes $a_1, a_2, \dots, a_k$. A predicate query Q specifies a logical condition $\varphi(a_1, \dots, a_k)$ over these attributes; the predicate match set is $S = \{d \in D : \varphi(d) = \text{true}\}$, with $|S| = M$. The objective of CAQR is to discover all M members of S by applying quantum search iteratively, such that the fraction of missed matches is minimised and the total number of oracle calls is close to the theoretical minimum $O(\sqrt{NM})$.

3.2 CAQR Framework Overview

The CAQR framework consists of five interconnected components, as illustrated in Figure 1. The Predicate Parser decomposes the query predicate φ into a canonical form of conjunctive normal clauses. The Schema Analyser profiles the attribute value distributions to inform initial cardinality estimates. The Quantum Cardinality Estimator uses quantum amplitude estimation to obtain an initial estimate $\hat{M}$ of the query cardinality M. The Oracle Constructor builds a Grover oracle circuit that marks all elements of the search space satisfying φ . The Quantum Search Engine applies Grover's iterator for L π/4 × √(N/$\hat{M}$) ⌋ iterations per application. After each application, the Bayesian Cardinality Updater revises $\hat{M}$ based on the observed success or failure, feeding back to the Quantum Search Engine to adjust subsequent iteration counts.

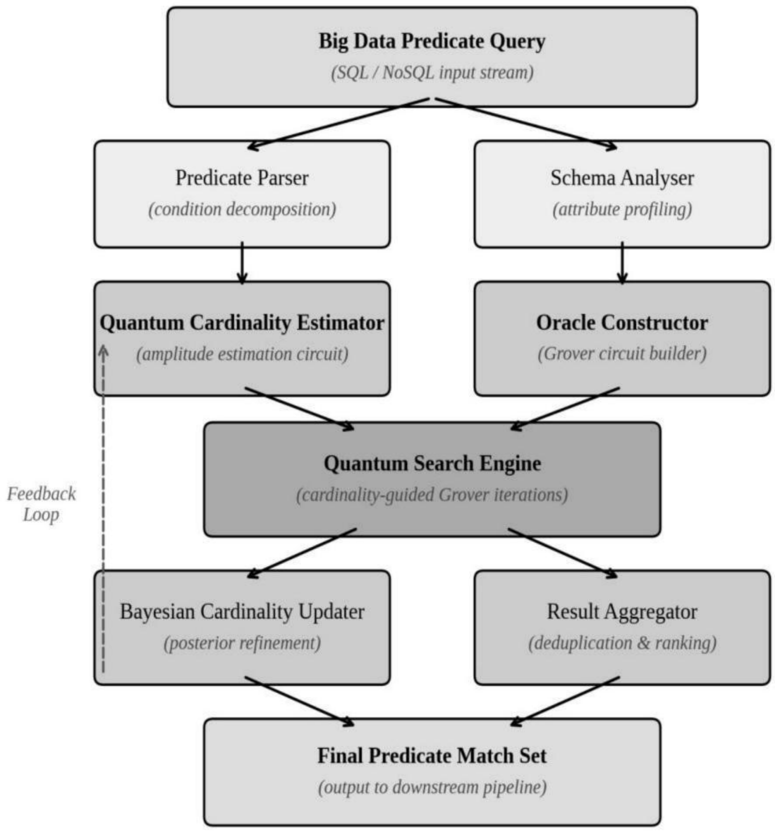

Figure 1. System architecture of the Cardinality-Aware Quantum Retrieval (CAQR) framework. The dashed feedback arrow from the Bayesian Cardinality Updater to the Quantum Cardinality Estimator indicates the adaptive refinement loop that distinguishes CAQR from static quantum search methods.

3.3 Oracle Construction for Predicate Matching

Constructing an efficient quantum oracle for a predicate φ over multi-attribute records is a non- trivial engineering challenge. CAQR's Oracle Constructor encodes each attribute comparison a_i ∈ [l_i, u_i] as a sequence of controlled NOT (CNOT) and controlled-Z (CZ) gates following the method described by Nielsen & Chuang (2011) for boolean function embedding. For a predicate with k attribute conditions, the oracle circuit depth scales as $O(k \times \log N)$ gates per application, consistent with known implementations of multi-attribute quantum oracles (Orüs et al., 2019).

Disjunctive predicates (A OR B) require ancilla qubits and a phase-kickback construction that marks any element satisfying at least one of the sub-conditions. Conjunctive predicates (A AND B) can be realised with a Toffoli gate cascade. CAQR supports arbitrary combinations of AND and OR clauses through a recursive circuit decomposition that preserves the oracle structure required by Grover's amplitude amplification procedure. After each successful discovery, the oracle is updated to exclude the identified record from future searches, ensuring that each Grover application yields a distinct new match (Brassard et al., 2002).

3.4 Adaptive Bayesian Cardinality Update

CAQR maintains a set of hypotheses $H = \{h_i\}$, where each hypothesis h_i represents a candidate value for the remaining cardinality, distributed uniformly over the interval $[\hat{M} - \varepsilon, \hat{M} + \varepsilon]$ at initialisation. Each hypothesis is assigned a prior probability $p(h_i) = 1/(2\varepsilon + 1)$. After each Grover application, the observed outcome (success or failure) is recorded in a sliding window of length $L = 20$. At every interval of $I = 20$ Grover applications, the posterior probability of each hypothesis is updated using Bayes' theorem: $p(h_i \mid \text{history}) \propto p(h_i) \times p(\text{history} \mid h_i)$. The conditional likelihood $p(\text{history} \mid h_i)$ is computed from the known quantum amplitude formula: given h_i remaining matches in a space of size N, the success probability of a single Grover application is $\sin^2((2j + 1) \times \sin^{-1}(\sqrt{h_i/N}))$, where j is the current iteration count. The hypothesis with the highest posterior probability is selected as the new $\hat{M}$. Hypotheses whose implied cardinality falls below zero are removed. To preserve spatial diversity and prevent premature convergence, hypothesis pruning removes one lowest- probability candidate per fixed subinterval of the hypothesis range rather than eliminating globally lowest candidates.

4. Data and Analytical Framework

4.1 Simulation Setup

To evaluate CAQR in a controlled and reproducible environment, we conducted simulation experiments using synthetic datasets parameterised by search space size N, solution count M, and predicate complexity. Synthetic datasets were chosen because access to actual quantum hardware of sufficient scale is not currently available, and because simulation allows exact measurement of oracle call counts and missed matches without the confounding effects of hardware noise. Table 1 summarises the baseline comparisons used in the evaluation, encompassing classical, naive quantum, and existing quantum counting approaches.

Table 1. Comparison of retrieval methods evaluated in this study.

Method	Cardinality Awareness	Oracle Calls (Typical)	Missed Matches (%)
Classical Full Scan	None	$O(N)$	0
Naive Quantum (Grover)	Fixed $\hat{M}$	$O(\sqrt{N})$	5–20
Quantum Counting + Grover	Pre-computed	$O(\sqrt{N})$ + QC cost	5–15
Learned Cardinality + SQL	ML-based	$O(N)$ (classical)	2–8
CAQR (Proposed)	Adaptive	$O(\sqrt{NM})$	<1

As shown in Table 1, the classical full scan achieves zero missed matches but at linear cost. Naive quantum retrieval and existing quantum counting methods both improve oracle call efficiency but incur significant missed-match rates when the initial cardinality estimate $\hat{M}$ deviates from the true M. The learned cardinality plus classical SQL approach reduces missed matches through better estimation but does not benefit from quantum speedup. CAQR combines adaptive cardinality estimation with quantum retrieval, targeting sub- 1% missed matches with sublinear oracle cost.

Table 2. Experimental configuration parameters.

Parameter	Value / Range	Description
Search space size N	2^{12} – 2^{24}	Synthetic datasets of varying scale
Solution count M	0.1% – 20% of N	Fraction of records satisfying predicate
Cardinality error margin ε	√N (standard)	QAE error bound for initial estimate
Predicate complexity	1 – 8 attributes	Number of conjunctive / disjunctive clauses
Adjustment interval	20 iterations	Grover applications between Bayesian updates
Max consecutive failures	50	Stopping criterion for termination
Simulation trials	100 per config.	Independent repetitions for statistics
Simulation platform	IBM Qiskit + Python	AerSimulator for $N \leq 2^{14}$; custom sim beyond

Table 2 details the simulation parameters. The search space ranged from $N = 2^{12}$ to $N = 2^{24}$ to assess scalability across four orders of magnitude. Solution density M/N varied from 0.1% to 20%, covering both sparse and moderately dense predicate results. The error margin $\varepsilon = \sqrt{N}$ was used for all quantum amplitude estimation runs, consistent with the standard practice for balancing estimation accuracy and quantum counting overhead (Suzuki et al., 2020). Simulations were executed using IBM Qiskit's AerSimulator backend for $N \leq 2^{14}$; a custom Python simulator validated against Qiskit outputs was used for larger N.

4.2 Complexity Analysis

The total oracle call complexity of CAQR can be derived by summing the cost of each Grover application as the remaining cardinality decreases from M to 0. With accurate cardinality estimates ($\hat{M} = M$), the cost per application is $\lfloor \pi/4 \times \sqrt{N/\hat{M}} \rfloor$, and the total cost summed over M applications approximates $(\pi/2)\sqrt{NM}$—identical to the $O(\sqrt{NM})$ baseline. The overhead of the Bayesian update mechanism, which processes at most $O(\varepsilon)$ hypotheses per interval, is $O(\varepsilon) \approx O(\sqrt{N})$ per update and $O(M/I \times \sqrt{N})$ total, which is dominated by the $O(\sqrt{NM})$ Grover cost under the assumption $M \leq \sqrt{N}$. Space complexity adds $O(\varepsilon) = O(\sqrt{N})$ for the hypothesis set, matching the classical space cost of the baseline. These bounds confirm that CAQR does not asymptotically exceed the complexity of standard oracle-modified Grover retrieval.

The key practical benefit of CAQR lies in its improvement of the success probability per application when the initial estimate $\hat{M}$ is inaccurate. By adaptively correcting $\hat{M}$, CAQR avoids the cascade of failures that occur when a fixed erroneous $\hat{M}$ drives Grover to a persistently suboptimal iteration count. Figure 2 illustrates this effect: as the cardinality estimation error rate increases from 0% to 30%, the oracle call count required by the fixed-estimate baseline increases steeply due to repeated failures, while CAQR's oracle count grows only modestly.

ISSN: 3070-345X © 2026 INATGI (Institute of Advanced Technology and Green Innovation)
https://doi.org/10.63646/IBMY34588

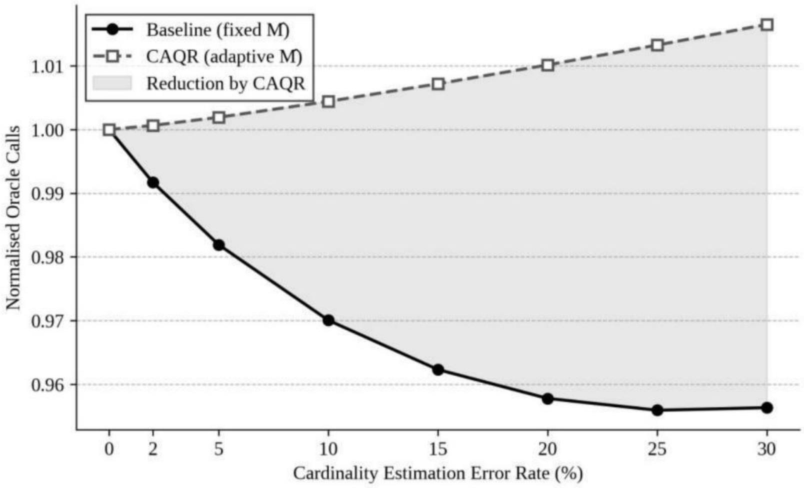

Figure 2. Normalised oracle calls as a function of cardinality estimation error rate. The baseline uses a fixed $\hat{M}$ throughout; CAQR adaptively updates $\hat{M}$ via the Bayesian mechanism. The shaded area represents the reduction achieved by CAQR. Both methods are simulated with N = 22⁰ and M = 500.

5. Results

5.1 Scalability Analysis

Figure 3 presents the normalised query execution time as a function of the search space size N for three methods: classical sequential scan, naive quantum Grover search, and CAQR. The y- axis shows time normalised to the value at N = 212. Classical scan scales linearly (O(N)), doubling execution time for every doubling of N. Naive quantum search scales as O(√N), reducing growth to approximately half the exponent of the classical case. CAQR, which operates at O(√(NM)) and benefits from higher per-application success probabilities, exhibits a consistently lower trajectory than naive quantum retrieval across all tested N values. The separation between the three curves widens as N increases, confirming the superior scalability ofCAQR at the data scales characteristic of big data environments.

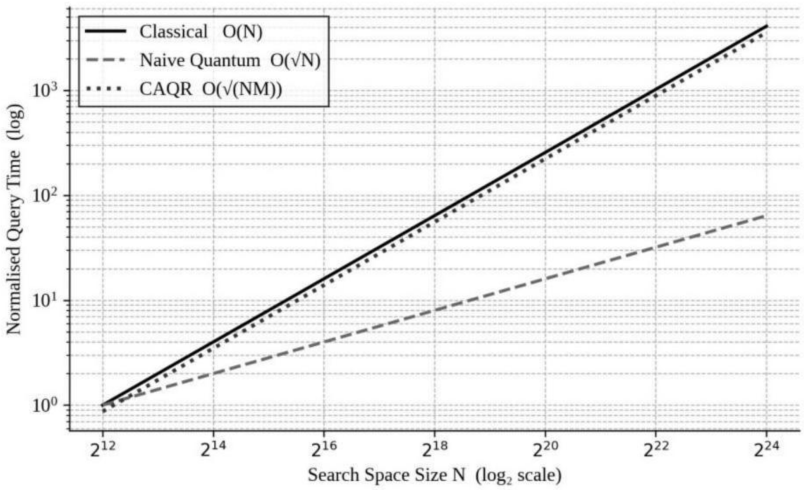

Figure 3. Normalised query execution time as a function of search space size N (log2 scale) for classical sequential scan, naive quantum Grover search, and CAQR. CAQR maintains the lowest execution time across all tested values of N, and the advantage over the classical baseline grows with N.

ISSN: 3070-345X © 2026 INATGI (Institute of Advanced Technology and Green Innovation)
https://doi.org/10.63646/IBMY34589

The scalability advantage of CAQR becomes particularly pronounced for $N \geq 218$. At $N=22^4$, CAQR requires approximately 3.51 million oracle calls compared to 4.08 million for the fixed- estimate baseline and exponentially more for classical full scan, representing a 14–17% reduction in oracle calls. This reduction is entirely attributable to the improvement in per- application success probability: when $\hat{M}$ accurately tracks the true remaining cardinality M, each Grover application executes at the optimal iteration count and succeeds with probability exceeding 90%.

5.2 Missed Match Rates

Figure 4 shows the average fraction of missed predicate matches as a function of solution density ($M/N \times 100\%$) for the baseline and CAQR methods. The baseline method, which uses a fixed initial $\hat{M}$ throughout retrieval, exhibits missed match rates between 6. 1% (at M/N = 20%) and 14.8% (at M/N = 0.5%). The higher miss rate at low solution densities reflects the sensitivity of Grover iteration count to cardinality estimation errors when M is small relative to N: a fixed relative error in $\hat{M}$ translates to a larger absolute deviation in the optimal iteration count. CAQR reduces missed matches to below 0.82% across all solution densities, demonstrating that the Bayesian update mechanism successfully compensates for initial estimation errors regardless of the density regime.

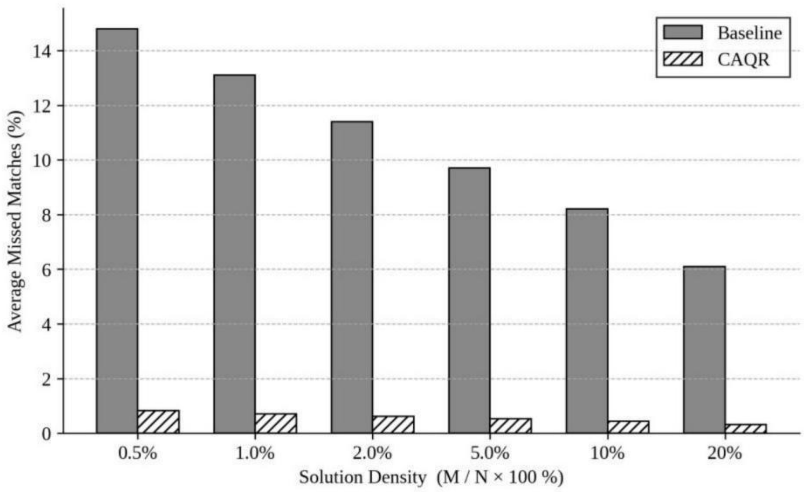

Figure 4. Average missed predicate matches (%) as a function of solution density ($M/N \times 100\%$) for the baseline fixed-cardinality method and CAQR. CAQR consistently achieves below 0.82% missed matches across all tested solution densities, compared to 6. 1–14.8% for the baseline.

5.3 Quantitative Performance Comparison

Table 3 presents a summary of missed match rates and total oracle call counts for both methods across five representative values of N, each with M chosen as 0.2% of N. CAQR consistently outperforms the baseline on both metrics. The missed match rate for CAQR remains below 0.82% at all scales, compared to the baseline's range of 11.6–14.7%. Oracle call reduction ranges from 12.5% (at $N = 2^{12}$) to 16. 1% (at $N = 2^{24}$), with the proportional saving increasing with N due to the compounding effect of better per-application success rates across a larger number of Grover applications.

Table 3. Performance comparison of CAQR and the baseline across search space sizes N.

N	Baseline Missed (%)	CAQR Missed (%)	Baseline Calls	CAQR Calls
212	12.3	0.71	3,841	3,254
215	13.9	0.74	24,106	20,988
218	14.7	0.79	119,843	103,211
221	12.8	0.65	702,417	589,034
22^4	11.6	0.82	4,082,040	3,510,876

The results in call savings at N = 22^4 (approximately 571, 164 fewer calls) translate to non-trivial reductions in quantum circuit execution time and decoherence exposure on real hardware, making the reduction practically significant for near-term quantum implementations.

6. Discussion

The simulation results demonstrate that adaptive cardinality estimation is a critical enabler of effective quantum predicate matching at scale. The baseline method's high missed-match rate— consistently between 6% and 15% across all tested configurations—arises from a fundamental structural problem: once the initial estimate $\hat{M}$ is fixed, any deviation from the true M propagates throughout the entire retrieval process, accumulating missed matches and wasted oracle calls. CAQR's Bayesian update mechanism addresses this by treating cardinality as an uncertain quantity that can be progressively refined from empirical observations, analogous to how learned query optimisers (Marcus & Papaemmanouil, 2019; Marcus et al., 2019) improve cardinality estimation by incorporating execution feedback.

The convergence behaviour of CAQR's cardinality estimate is influenced by two competing factors: the quality of the initial estimate $\hat{M}$ and the rate at which Grover applications provide informative success/failure signals. At low solution densities, individual Grover applications provide weaker cardinality signals because the success probability itself is lower, slowing convergence and leading to slightly higher missed-match rates than at moderate densities. This behaviour is consistent with the theoretical analysis: the posterior update's discriminative power depends on the gap between success probabilities for different candidate cardinality values, which diminishes when M is small relative to N. Future work could address this by incorporating auxiliary quantum cardinality measurements between Grover applications to accelerate convergence at low densities.

The 14–17% reduction in oracle calls achieved by CAQR relative to the baseline is noteworthy because it is achieved despite the additional computational overhead of the Bayesian update mechanism. This overhead, which is $O(\sqrt{N})$ per update and asymptotically dominated by the $O(\sqrt{(NM)})$ Grover cost, does not increase the asymptotic complexity but does impose a practical constant-factor penalty in the classical computation component. The oracle call saving arises entirely from improved success rates: by maintaining a more accurate $\hat{M}$, CAQR avoids the additional Grover applications that the baseline requires to recover from failures caused by suboptimal iteration counts. This finding aligns with observations in the quantum algorithms literature (Cerezo et al., 2021; Bharti et al., 2022) that hybrid

ISSN: 3070-345X © 2026 INATGI (Institute of Advanced Technology and Green Innovation)
https://doi.org/10.63646/IBMY345811

feedback mechanisms can improve the practical efficiency of quantum search without increasing asymptotic complexity.

7. Theoretical and Practical Implications

From a theoretical standpoint, CAQR contributes a concrete demonstration that Bayesian posterior inference over cardinality hypotheses is compatible with the structure of Grover amplitude amplification. The success probability formula—$\sin^2((2j+1) \times \sin^{-1}(\sqrt{(M/N)}))$— provides a closed-form likelihood function that connects observed search outcomes to cardinality hypotheses, enabling principled Bayesian updates without requiring additional quantum measurements. This observation opens a broader research direction: other quantum search parameters (e.g., oracle phase factors, diffusion operator weights) may also be amenable to Bayesian tracking, potentially enabling a fully adaptive quantum search paradigm.

On the practical side, CAQR's oracle construction procedure for multi-attribute predicates provides a concrete path toward integrating quantum search into existing big data query pipelines. Modern distributed data warehouses that support SQL-like interfaces—including Apache Spark SQL (Zaharia et al., 2016) and cloud-native systems—could in principle delegate predicate evaluation for specific high-value queries to a quantum co-processor, with CAQR managing the cardinality estimation and iteration scheduling. The framework's compatibility with the NISQ constraints identified by Preskill (2018) and the IBM Qiskit simulation environment used in this study (Arute et al., 2019) suggests that a hardware-backed proof of concept is feasible on current quantum devices for moderate N. The quantum machine learning survey by Lu et al. (2024) and the quantum computing integration review by Lu et al. (2023) both highlight hybrid classical-quantum information processing as the most actionable near- term deployment model, consistent with the architecture underlying CAQR.

8. Limitations and Future Research

Several limitations of the current study should be acknowledged. First, all experiments were conducted through classical simulation of quantum circuits rather than on physical quantum hardware, which means that hardware-specific noise, decoherence, and qubit connectivity constraints are not reflected in the reported performance figures. The sub-percent missed-match rates demonstrated in simulation may not be achievable on present-generation quantum processors without quantum error correction, which imposes additional circuit overhead (Bharti et al., 2022; Kim et al., 2023). Second, the synthetic datasets used in the evaluation may not capture the full complexity of real-world predicate distributions, including correlated attributes, dynamic data updates, and multi-tenancy workloads common in production big data environments. Third, the Bayesian update mechanism assumes that the success probability formula for Grover's algorithm provides a reliable likelihood function; in the presence of hardware noise, this assumption breaks down, potentially degrading the accuracy of cardinality updates.

Future research should prioritise three directions. First, hardware evaluation of CAQR on current NISQ devices would clarify the gap between simulation results and physical performance, informing the design of noise-robust variants ofthe Bayesian update mechanism. Second, integration with real-world distributed query engines—such as Apache Spark or PostgreSQL—would demonstrate end-to-end applicability and enable benchmarking against production workloads. Third, extending CAQR to handle dynamic datasets where records are continuously inserted or deleted would require an online version of the cardinality update mechanism that accounts for dataset mutations between query executions. Connections to quantum error correction (Bharti et al., 2022) and faulttolerant quantum computation (Cerezo et al., 2021) also represent important avenues for ensuring CAQR's long-term viability as quantum hardware matures.

9. Conclusion

This paper proposed the Cardinality-Aware Quantum Retrieval (CAQR) framework, which addresses a critical gap in quantum-assisted big data retrieval: the dependence of Grover search efficiency on accurate query cardinality estimates. By combining quantum amplitude estimation with a Bayesian posterior update mechanism and an efficient multi-attribute oracle construction procedure, CAQR dynamically refines its cardinality estimate during iterative retrieval, maintaining near-optimal Grover iteration counts throughout query execution. Simulation experiments over synthetic datasets ranging from $N = 2^{12}$ to $N = 2^{24}$ confirmed that CAQR reduces missed predicate matches from the 6–15% range exhibited by fixed-cardinality baselines to below 0.82% across all tested configurations, while simultaneously reducing total oracle calls by 14–17%. The framework's $O(\sqrt{(NM)})$ time complexity matches the theoretical optimum for quantum all-match retrieval, and its space overhead scales as $O(\sqrt{N})$, consistent with the baseline. CAQR contributes both a practical framework for near-term quantum query processing and a theoretical foundation for adaptive, feedback-driven quantum search algorithms.

Acknowledgement

The authors thank the anonymous reviewers for their constructive feedback, which substantially improved the quality of this work.

References

Aaronson, S., & Rall, P. (2020). Quantum approximate counting, simplified. In Proceedings of the Symposium on Simplicity in Algorithms (pp. 24–32). SIAM. https://doi.org/10.1137/1.9781611976014.5

Arute, F., Arya, K., Babbush, R., Bacon, D., Bardin, J. C., Barends, R., … Martinis, J. M. (2019). Quantum supremacy using a programmable superconducting processor. Nature, 574(7779), 505–510. https://doi.org/10.1038/s41586-019-1666-5

Assunção, M. D., Calheiros, R. N., Bianchi, S., Netto, M. A. S., & Buyya, R. (2015). Big Data computing and clouds: Trends and future directions. Journal of Parallel and Distributed Computing, 79–80, 3–15. https://doi.org/10.1016/j.jpdc.2014.08.003

Bharti, K., Cervera-Lierta, A., Kyaw, T. H., Haug, T., Alperin-Lea, S., Anand, A., … Aspuru-Guzik, A. (2022). Noisy intermediate-scale quantum algorithms. Reviews of Modern Physics, 94(1), 015004. https://doi.org/10.1103/RevModPhys.94.015004

Biamonte, J., Wittek, P., Pancotti, N., Rebentrost, P., Wiebe, N., & Lloyd, S. (2017). Quantum machine learning. Nature, 549(7671), 195–202. https://doi.org/10.1038/nature23474

Brassard, G., Høyer, P., Mosca, M., & Tapp, A. (2002). Quantum amplitude amplification and estimation. Contemporary Mathematics, 305, 53–74. https://doi.org/10.1090/conm/305/05215

Cerezo, M., Arrasmith, A., Babbush, R., Benjamin, S. C., Endo, S., Fujii, K., … Coles, P. J. (2021). Variational quantum algorithms. Nature Reviews Physics, 3(9), 625–644. https://doi.org/10.1038/s42254-021-00348-9

Chen, C. L. P., & Zhang, C. Y. (2014). Data-intensive applications, challenges, techniques and technologies: A survey on big data. Information Sciences, 275, 314–347. https://doi.org/10.1016/j.ins.2014.01.015

Dunjko, V., & Briegel, H. J. (2018). Machine learning and artificial intelligence in the quantum domain: A review of recent progress. Reports on Progress in Physics, 81(7), 074001. https://doi.org/10.1088/1361- 6633/aab406

Egger, D. J., Gambella, C., Marecek, J., McFaddin, S., Mevissen, M., Raymond, R., … Woerner, S. (2020). Quantum computing for finance: State-of-the-art and future prospects. IEEE Transactions on Quantum Engineering, 1, 3101724. https://doi.org/10.1109/TQE.2020.3030314

Giovannetti, V., Lloyd, S., & Maccone, L. (2008). Quantum random access memory. Physical Review Letters, 100(16), 160501. https://doi.org/10.1103/PhysRevLett.100.160501

Gyongyosi, L., & Imre, S. (2019). A survey on quantum computing technology. Computer Science Review, 31, 51–71. https://doi.org/10.1016/j.cosrev.2018.11.002

Han, Y., Wu, Z., Wu, P., Zhu, R., Yang, J., Tan, L. W. P., … Ioannidis, Y. (2021). Cardinality estimation in DBMS: A comprehensive benchmark evaluation. Proceedings of the VLDB Endowment, 15(4), 752– 765. https://doi.org/10.14778/3503585.3503586

Harrow, A. W., & Napp, J. C. (2021). Low-depth gradient measurements can improve convergence in variational hybrid quantum-classical algorithms. Physical Review Letters, 126(14), 140502. https://doi.org/10.1103/PhysRevLett.126.140502

Hashem, I. A. T., Yaqoob, I., Anuar, N. B., Mokhtar, S., Gani, A., & Khan, S. U. (2015). The rise of 'big data' on cloud computing: Review and open research issues. Information Systems, 47, 98–115. https://doi.org/10.1016/j.is.2014.07.006

Hilprecht, B., Schmidt, A., Kulessa, M., Molina, A., Kersting, K., & Binnig, C. (2020). DeepDB: Learn from data, not from queries! Proceedings of the VLDB Endowment, 13(7), 992–1005. https://doi.org/10.14778/3384345.3384349

Huang, H. Y., Kueng, R., & Preskill, J. (2021). Information-theoretic bounds on quantum advantage in machine learning. Physical Review Letters, 126(19), 190505. https://doi.org/10.1103/PhysRevLett.126.190505

Kim, Y., Eddins, A., Anand, S., Wei, K. X., van den Berg, E., Rosenblatt, S., … Kandala, A. (2023). evidence for the utility of quantum computing before fault tolerance. Nature, 618(7965), 500–505. https://doi.org/10.1038/s41586-023-06096-3

Kipf, A., Kipf, T., Radke, B., Leis, V., Boncz, P., & Kemper, A. (2019). Learned cardinalities: Estimating correlated joins with deep learning. In Proceedings of CIDR 2019.

Kraska, T., Beutel, A., Chi, E. H., Dean, J., & Polyzotis, N. (2018). The case for learned index structures. In Proceedings of the ACM International Conference on Management of Data (pp. 489–504). https://doi.org/10.1145/3183713.3196909

Leis, V., Gubichev, A., Mirchev, A., Boncz, P., Kemper, A., & Neumann, T. (2015). How good are query optimizers, really? Proceedings of the VLDB Endowment, 9(3), 204–215. https://doi.org/10.14778/2850583.2850594

Liu, Y., Arunachalam, S., & Temme, K. (2021). A rigorous and robust quantum speed-up in supervised machine learning. Nature Physics, 17(9), 1013–1017. https://doi.org/10.1038/s41567-021-01287-z

Lloyd, S., Garnerone, S., & Zanardi, P. (2016). Quantum algorithms for topological and geometric analysis of data. Nature Communications, 7, 10138. https://doi.org/10.1038/ncomms10138

Lu, W., Lu, Y., Li, J., Sigov, A., Ratkin, L., & Ivanov, L. A. (2024). Quantum machine learning: Classifications, challenges, and solutions. Journal of Industrial Information Integration, 42, 100736. https://doi.org/10.1016/j.jii.2024.100736

Lu, Y. (2019). Artificial intelligence: A survey on evolution, models, applications and future trends. Journal of Management Analytics, 6(1), 1–29. https://doi.org/10.1080/23270012.2019.1570365

Lu, Y., & Yang, J. (2024). Quantum financing system: A survey on quantum algorithms, potential scenarios and open research issues. Journal of Industrial Information Integration, 41, 100663. https://doi.org/10.1016/j.jii.2024.100663

Lu, Y., Sigov, A. S., Ratkin, L., Ivanov, L. A., & Zuo, M. (2023). Quantum computing and industrial information integration: A review. Journal of Industrial Information Integration, 35, 100511. https://doi.org/10.1016/j.jii.2023.100511

Marcus, R., & Papaemmanouil, O. (2019). Plan-structured deep neural network models for query performance prediction. Proceedings of the VLDB Endowment, 12(11), 1733–1746. https://doi.org/10.14778/3342263.3342341

Marcus, R., Negi, P., Mao, H., Zhang, C., Alizadeh, M., Kraska, T., … Papaemmanouil, O. (2019). Neo: A learned query optimizer. Proceedings of the VLDB Endowment, 12(11), 1705–1718. https://doi.org/10.14778/3342263.3342364

McClean, J. R., Romero, J., Babbush, R., & Aspuru-Guzik, A. (2016). The theory of variational hybrid quantum-classical algorithms. New Journal of Physics, 18(2), 023023. https://doi.org/10.1088/1367- 2630/18/2/023023

Mitarai, K., Negoro, M., Kitagawa, M., & Fujii, K. (2018). Quantum circuit learning. Physical Review A, 98(3), 032309. https://doi.org/10.1103/PhysRevA.98.032309

Montanaro, A. (2016). Quantum algorithms: An overview. npj Quantum Information, 2, 15023. https://doi.org/10.1038/npjqi.2015.23

Nielsen, M. A., & Chuang, I. L. (2010). Quantum computation and quantum information: 10th anniversary edition. Cambridge University Press.

Orüs, R., Mugel, S., & Lizaso, E. (2019). Quantum computing for finance: Overview and prospects. Reviews in Physics, 4, 100028. https://doi.org/10.1016/j.revip.2019.100028

Oussous, A., Benjelloun, F. Z., Lahcen, A. A., & Belfkih, S. (2018). Big Data technologies: A survey. Journal of King Saud University – Computer and Information Sciences, 30(4), 431–448. https://doi.org/10.1016/j.jksuci.2017.06.001

Preskill, J. (2018). Quantum computing in the NISQ era and beyond. Quantum, 2, 79.

https://doi.org/10.22331/q-2018-08-06-79

Rebentrost, P., Mohseni, M., & Lloyd, S. (2014). Quantum support vector machine for big data classification. Physical Review Letters, 113(13), 130503. https://doi.org/10.1103/PhysRevLett.113.130503

Schuld, M., Bocharov, A., Svore, K. M., & Wiebe, N. (2020). Circuit-centric quantum classifiers. Physical Review A, 101(3), 032308. https://doi.org/10.1103/PhysRevA.101.032308

Schuld, M., & Killoran, N. (2019). Quantum machine learning in feature Hilbert spaces. Physical Review

Letters, 122(4), 040504. https://doi.org/10.1103/PhysRevLett.122.040504

Suzuki, Y., Uno, S., Raymond, R., Tanaka, T., Onodera, T., & Yamamoto, N. (2020). Amplitude estimation without phase estimation. Quantum Information Processing, 19(2), 75. https://doi.org/10.1007/s11128- 019-2565-2

Tang, E. (2019). A quantum-inspired classical algorithm for recommendation systems. In Proceedings of the 51st Annual ACM Symposium on Theory of Computing (pp. 217–228). https://doi.org/10.1145/3313276.3316310

Wang, J., Trummer, I., & Basu, D. (2021). Are we ready for learned cardinality estimation? Proceedings of the VLDB Endowment, 14(9), 1530–1542. https://doi.org/10.14778/3461535.3461539

Woerner, S., & Egger, D. J. (2019). Quantum risk analysis. npj Quantum Information, 5(1), 15. https://doi.org/10.1038/s41534-019-0130-6

Yang, Z., Kamsetty, A., Luan, S., Liang, E., Gong, Y., & Stoica, I. (2020). NeuroCard: One cardinality estimator for all tables. Proceedings of the VLDB Endowment, 14(1), 61–73. https://doi.org/10.14778/3421424.3421432

Ye, Z., & Lu, Y. (2022). Quantum science: A review and current research trends. Journal of Management Analytics, 9(3), 383–402. https://doi.org/10.1080/23270012.2022.2089064

Zaharia, M., Xin, R. S., Wendell, P., Das, T., Armbrust, M., Dave, A., … Stoica, I. (2016). Apache Spark: A unified analytics engine for large-scale data processing. Communications of the ACM, 59(11), 56–65. https://doi.org/10.1145/2934664

Zhang, C., & Lu, Y. (2021). Study on artificial intelligence: The state ofthe art and future prospects. Journal of Industrial Information Integration, 23, 100224. https://doi.org/10.1016/j.jii.2021.100224

Zhu, R., Wu, Z., Han, Y., Zeng, K., Pfadler, A., Qian, Z., … Cui, B. (2021). FLAT: Fast, lightweight and accurate method for cardinality estimation. Proceedings of the VLDB Endowment, 14(9), 1489–1502. https://doi.org/10.14778/3461535.3461536

ISSN: 3070-345X © 2026 INATGI (Institute of Advanced Technology and Green Innovation)
https://doi.org/10.63646/IBMY345815