Journal of
Low Power Electronics and Applications

MDPI

Article

Quantum-Inspired Edge Intelligence for Dynamic Task Offloading in Real-Time Cyber–Physical Systems

Sadiq Alinsaif * and Ammar Ahmed E. Elhadi

College of Computer Science and Engineering, University of Hafr Al Batin, Hafr Al Batin 39524, Saudi Arabia; ammarhadi@uhb.edu.sa
* Correspondence: alinsaif@uhb.edu.sa

Abstract

Modern Cyber–Physical Systems (CPS) face severe performance bottlenecks due to the high latency and computational overhead of centralized processing, limiting their efficacy in time-critical applications such as digital health and smart mobility. To overcome these constraints, this paper presents an integrated architecture that cohesively integrates Edge Intelligence (EI) with Quantum-Inspired Algorithms (QIAs). The framework decentralizes immediate decision-making to edge computing nodes through a novel Quantum-Inspired Task Offloading and Scheduling (Q-TOS) algorithm that uses probabilistic Q-bit state representations and phase rotation gates to solve multi-objective latency-energy optimization problems. Rigorous simulation experiments performed across dynamic workload, surge load, and channel-impaired scenarios demonstrate that Q-TOS achieves a 29.3% reduction in decision-making latency, a 22.3% decrease in dynamic energy consumption, and a 98.6% constraint feasibility rate relative to state-of-the-art metaheuristics such as Particle Swarm Optimization (PSO). Furthermore, the framework sustains an average decision latency of 0.78 ms with a RAM footprint below 3.4 MB, validating quantum-inspired edge architectures as a viable paradigm for expandable, real-time CPS operations.

Keywords: Cyber–Physical Systems (CPS); edge intelligence; quantum-inspired algorithms; fourth industrial revolution; real-time decision-making

1. Introduction

The Fourth Industrial Revolution has accelerated the convergence of computational infrastructure and physical assets, establishing a new paradigm for modern industrial and social operations. At the core of this transformation are Cyber–Physical Systems (CPS), which smoothly integrate physical processes with digital computing entities. Rather than functioning as isolated units, contemporary CPS demand proactive, highly integrated architectures to sustain peak performance within dynamic and unpredictable environments. To meet these escalating operational demands, the synergistic integration of AI Intelligence (AI), Edge Intelligence (EI), and quantum-based algorithms offers an effective framework for augmenting CPS capabilities [1]. While AI provides substantial predictive power across domains—ranging from healthcare [2] to smart manufacturing—combining it with Edge Intelligence brings data processing directly to localized edge devices (such as smart sensors), drastically lowering processing latency and optimizing immediate decision-making. This structural coupling greatly improves CPS adaptiveness and sensitivity to adjusting to environmental fluctuations [3,4]. In smart manufacturing and autonomous industrial operations, data-driven analytics adapt to shifting operating conditions in real time, enabling

Check for updates

Academic Editor: Davide Bertozzi

Received: 19 September 2026
Revised: 26 October 2026
Accepted: 11 November 2026
Published: 13 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

autonomous execution without continual human involvement. Ultimately, this specific localized autonomy mitigates computing bottlenecks at intermediate processing hubs, resulting in major gains in overall system efficiency and execution speed [5].

To address complex combinatorial challenges, Quantum-Inspired Algorithms (QIAs) offer an innovative paradigm for systematically solving high-dimensional optimization problems that conventional heuristics cannot efficiently process [6]. Grounded in mathematical principles derived from quantum mechanics, these algorithms perform complex optimization routines on classical hardware. As modern CPS ingest massive streams of telemetry from dense sensor networks, establishing low-latency, high-throughput computational mechanisms becomes imperative [7]. In smart grid applications, for example, QIAs facilitate real-time energy distribution by dynamically aligning resource allocation with consumer demand while minimizing power dissipation [8]. Beyond improving data-handling efficiency, this capability directly advances environmental sustainability across industrial sectors. Consequently, coupling AI, Edge Intelligence, and QIAs establishes a synergistic operational framework required for intelligent CPS across heterogeneous environments. System deployment should prioritize three core objectives: reducing network latency, streamlining data processing, and improving decision accuracy [8]. While localized computation at edge sources minimizes transmission latency, integrating QIAs into existing CPS architectures drastically accelerates problem-solving speed and algorithmic throughput [6]. Ultimately, achieving higher autonomy requires self-adaptive AI models that continuously update their parameters based on operational feedback without relying on constant human intervention [8].

CPS can operate effectively in dynamic, high-stakes environments only by integrating cutting-edge computational paradigms [7]. This work stresses the need to converge Artificial Intelligence, Edge Intelligence, and Quantum-Inspired Algorithms to build resilient, highly adaptive architectures [8]. As global infrastructure advances alongside the Fourth Industrial Revolution, next-generation CPS must continuously evolve to meet the volatile operational demands of modern industry. Ultimately, realizing this integrated framework has transformative potential to enhance autonomous operations and efficiency across vital industrial and societal domains [6,8].

This research addresses the fundamental limitations of conventional CPS, which struggle to manage the severe computational complexity required for instant decision-making in evolving environments. By embedding Artificial Intelligence within Edge Intelligence architectures, the proposed framework decentralizes data processing to enable localized computation, thereby mitigating latency cycles and substantially improving operational responsiveness. Furthermore, integrating Quantum-Inspired Algorithms equips these systems to overcome combinatorial bottlenecks and efficiently resolve otherwise computationally intractable optimization problems [9]. Consequently, this investigation is particularly timely given the widespread adoption of advanced data analytics for high-stakes decision-making across mission-critical sectors such as healthcare, smart manufacturing, and transportation.

The primary contributions of this paper are summarized as follows:

- **Multi-Objective Edge Resource Orchestration under Uncertainty:** We formulate a rigorous multi-objective task offloading and scheduling problem for Cyber–Physical Systems (CPS) that jointly minimizes system latency and dynamic energy consumption under severe operational hurdles, including high-dimensional search spaces, multi-node resource competition, and stochastic channel variations.
- **Quantum-Inspired Task Offloading and Scheduling (Q-TOS) Framework:** We design and implement the Q-TOS algorithm, which leverages probabilistic Q-bit super-

position representations and dynamic phase-rotation gates to navigate combinatorial offloading spaces on classical hardware without premature local-optimum stagnation.

- **Adaptive Phase-Rotation and Closed-Loop Fault Resilience:** We introduce an adaptive fault-mitigation and state-repair mechanism utilizing quantum phase adjustments and probability amplitude normalization. This mechanism preserves decision feasibility and maintains system stability under high-density workload surges and severe channel impairments (e.g., up to 30% packet loss).
- **Empirical Scalability and Hardware Validation:** Through extensive discrete-event simulations, we demonstrate that Q-TOS achieves a 29.3% reduction in decision latency and a 22.3% decrease in dynamic energy consumption relative to state-of-the-art metaheuristics, while maintaining an ultra-low execution overhead (0.78 ms) and a RAM footprint below 3.4 MB suitable for resource-constrained edge microcontrollers [10].

2. Related Work

A synthesis of the recent literature reveals substantial advancements across CPS, AI, Edge Intelligence, and QIAs [11]. Prior research in this domain can be systematically categorized into three key paradigms: (i) architectural design and decentralization in CPS, (ii) performance optimization through AI and Edge Intelligence, and (iii) high-dimensional optimization via Quantum-Inspired computation. Establishing this thematic taxonomy serves to contextualize state-of-the-art breakthroughs while exposing unresolved architectural and algorithmic gaps that limit current CPS capabilities.

2.1. Cyber–Physical Systems

CPS seamlessly integrate computational algorithms with physical processes to facilitate real-time data acquisition, processing, and actuation. The recent literature underscores the vital role of CPS across diverse domains, including smart cities, digital healthcare, and the Industrial Internet of Things (IIoT). To optimize operational reliability and response latency, contemporary research has increasingly pivoted toward decentralized architectural paradigms. By leveraging distributed computing assets, decentralized decision-making mitigates the processing bottlenecks and single-point failure risks inherent to centralized cloud servers. This structural shift is particularly vital for latency-sensitive applications—such as autonomous systems and emergency response units—where rapid execution is mandatory. Furthermore, decentralized topologies enhance fault tolerance through redundant, concurrent node execution, sustaining system functionality even during localized node failures. This operational resilience is essential for mission-critical infrastructure where high availability is non-negotiable. Nevertheless, despite these architectural strides, substantial challenges persist regarding cross-platform security and the seamless integration of heterogeneous edge devices [12].

2.2. Artificial Intelligence and Edge Intelligence

In this paradigm, AI provides robust computational frameworks for automating decision-making and extracting actionable insights from complex telemetry. Deep learning paradigms—specifically Deep Neural Networks (DNNs)—have significantly enhanced pattern recognition, anomaly detection, and predictive analytics within CPS. Deployments of DNN architectures in computer vision and signal processing demonstrate strong transferability for smart surveillance, automated diagnostics, and predictive maintenance within CPS environments. By integrating AI, CPS achieve heightened situational awareness and operational self-adaptation, leveraging historical performance metrics to continuously refine decision execution. Nevertheless, deploying AI models presents critical challenges

regarding decision explainability, algorithmic reliability, and compliance with stringent safety and ethical standards [13].

To complement AI capabilities, Edge Intelligence (EI) shifts data processing and analytics directly to network edge devices near the point of data origin, eliminating reliance on centralized cloud infrastructure. This localized execution is vital for latency-critical applications requiring sub-millisecond response times, such as autonomous vehicle control and high-speed industrial automation. By executing preliminary filtering and localized inference at the edge, systems reduce network bandwidth overhead by offloading only essential meta-data to the cloud, thereby enabling instantaneous actuation and driving substantial productivity gains. However, scaling EI deployment introduces non-trivial overhead, particularly in orchestrating heterogeneous distributed nodes, preserving data privacy, and dynamically allocating network resources across large-scale facilities [14].

2.3. Quantum-Inspired Algorithms

QIAs have emerged as a powerful computational paradigm for solving high-dimensional optimization problems. Grounded in mathematical concepts derived from quantum computing, these algorithms aim to minimize resource consumption, accelerate data processing, and enhance decision-making performance while running on classical hardware. The recent literature highlights their growing utility across diverse domains, including logistics, supply chain management, and complex CPS. Foundational research has established the theoretical basis for embedding quantum-inspired techniques into CPS frameworks, demonstrating marked improvements in processing speed, system scalability, and operational efficiency over conventional heuristics. By rapidly executing complex combinatorial optimizations, QIAs significantly enhance real-time decision-making capabilities in dynamic CPS environments. However, despite a well-developed theoretical foundation, practical deployments of QIAs in real-world CPS applications remain sparse. Consequently, evaluating their empirical performance, algorithmic integration overhead, and execution stability under live operational conditions remains a vital research imperative [15].

2.4. Research Gap

Despite notable advancements across these individual domains, several critical research gaps remain unaddressed. While the existing literature covers the structural design and component-level mechanics of CPS, there is a distinct lack of unified architectural frameworks that systematically integrate AI, EI, and QIAs into a coherent, multi-layered system. Deploying such hybrid technologies requires a structured engineering approach to resolve fundamental challenges regarding cross-platform interoperability, end-to-end security, and multi-node scalability. Furthermore, although AI and Edge Intelligence have significantly enhanced system autonomy and localized decision-making, critical issues surrounding data privacy, regulatory compliance, and algorithmic ethics remain insufficiently addressed. These governance and security limitations pose severe risks as CPS deployments expand into mission-critical, high-stakes environments such as healthcare and public safety [16].

Recently, several advanced frameworks have emerged to tackle complex challenges in mobile edge computing. Chen et al. [17] investigated IRS-assisted secure computation offloading in multi-region mobile edge computing using multi-agent deep reinforcement learning to mitigate eavesdropping risks and reduce execution latency, while Li et al. [18] proposed an online learning-aided cooperative offloading mechanism under social trust propagation using Lyapunov optimization to manage trust risks. Additionally, Xu et al. [19] formulated a multi-device multi-server system considering multifactorial uncertainty (MECTOS-MU) and resolved it using a distributed game offloading based on load balancing (DGOLB)

approach. While these existing studies offer valuable solutions tailored to physical layer security, social trust routing, or game-theoretic distributed optimization, they do not address the high-dimensional combinatorial bottlenecks inherent to real-time Cyber–Physical operations. In contrast, our work addresses multifactorial operational uncertainties—such as dynamic workload surges, channel impairment, and server load balancing—aiming to minimize overall system latency and energy consumption through a quantum-inspired task offloading and scheduling (Q-TOS) framework using probabilistic Q-bit representations and phase-rotation dynamics. Unlike these existing approaches, our framework uniquely integrates task prioritization, parallelized quantum-inspired metaheuristic search, quantum rotation-gate updates, and closed-loop state repair to handle high-dimensional solution spaces with minimal computational overhead.

The practical deployment of QIAs within CPS remains largely under-explored. Despite the proliferation of theoretical frameworks, empirical studies detailing concrete implementation architectures, quantitative efficiency benchmarks, and domain-specific execution constraints remain remarkably scarce [20]. Resolving these empirical gaps is essential for establishing resilient, high-efficiency CPS capable of operating in highly volatile operational environments. While foundational concepts across CPS, AI, Edge Intelligence, and QIAs are individually well documented, a rigorous exploration of integrated, multi-layered frameworks remains a critical research imperative. Bridging these theoretical and practical divides will enable researchers and industry practitioners to elevate system reliability, optimize real-time performance, and advance the core operational objectives of Industry 4.0.

3. Methodology and Mathematical Formulation

This study devises an intelligent, multi-tiered CPS architecture that includes localized Edge Intelligence with QIAs [21]. As depicted in Figure 1, the methodological model follows four sequential phases: (i) system topology formulation and overhead modeling, (ii) mathematical formalization of computational offloading, (iii) algorithmic design using Quantum-Inspired Task Offloading and Scheduling (Q-TOS), and (iv) experimental assessment in stochastic operating regimes.

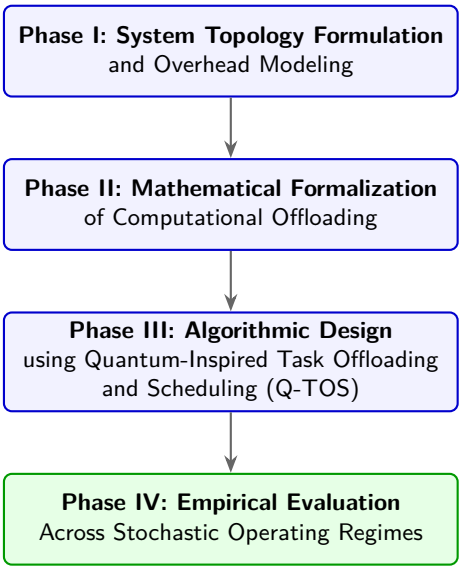


Figure 1. Sequential phases of the methodological framework for integrating Edge Intelligence and QIA in CPS.

3.1. Decentralized Architecture and System Model

The system architecture uses a decentralized topology of distributed Edge Computing Nodes (ECNs) operating under a Centralized Control Unit (CCU) [22]. Rather than relying on physical quantum hardware, localized *Quantum-Inspired Optimization Modules* (QOMs) run complex optimization routines directly on classical ECN processors.

Architectural Components

Figure 2 details the multi-tiered functional architecture, illustrating the control, optimization, and data pipelines across system layers. Each layer is described as follows:

- **Centralized Control Unit (CCU):** Oversees global network topology maintenance, long-term predictive state modeling, and non-time-critical policy broadcasts.
- **Edge Computing Nodes (ECNs):** Distributed heterogeneous servers positioned near physical edge devices to execute localized data preprocessing, real-time inference, and autonomous actuation routines [23].
- **Quantum-Inspired Optimization Module (QOM):** Software-defined algorithmic layers integrated within ECNs that execute probabilistic quantum-inspired metaheuristics [23] to solve high-dimensional task offloading and resource scheduling problems without physical quantum hardware restrictions.

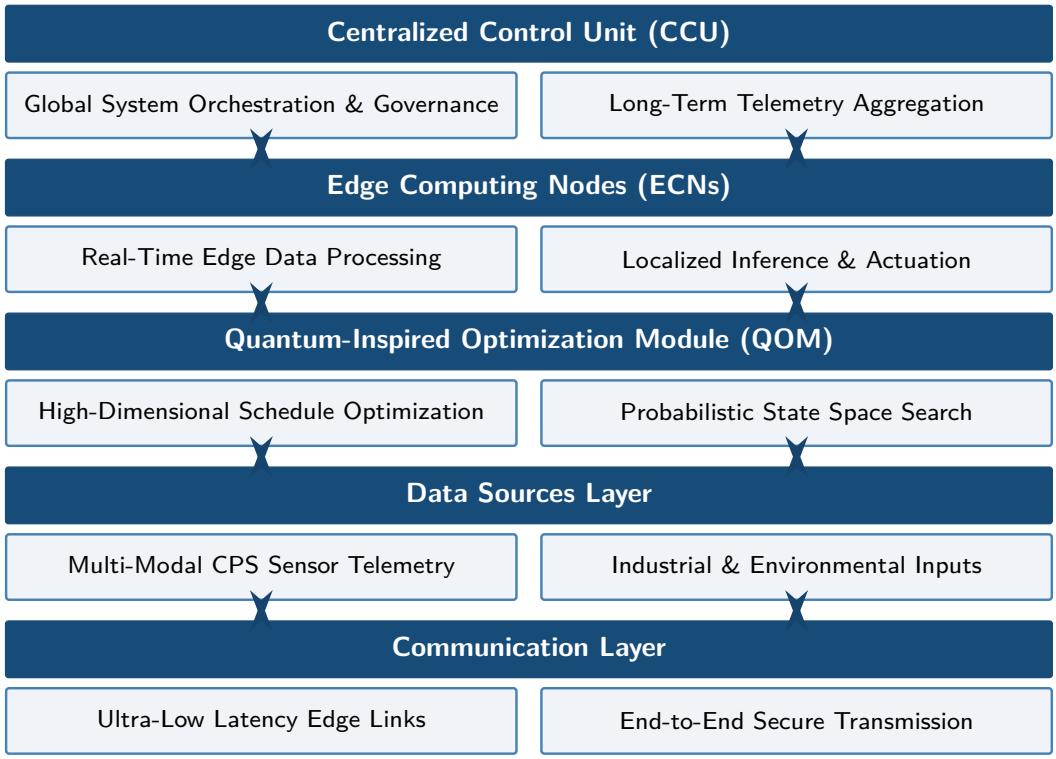


Figure 2. Layered functional architecture detailing core computational modules, communication pipelines, and optimization flows.

3.2. Problem Formulation and System Models

Let $\mathcal{T} = \{T_1, T_2, \ldots, T_N\}$ represent a set of N time-sensitive processing jobs generated by CPS sensors, where each task is characterized by a tuple $T_i = (D_i, C_i, \tau_i^{\max})$, representing data size (bits), required compute cycles, and maximum allowable delay (seconds), respectively. The execution domain consists of M candidate ECNs $\mathcal{E} = \{E_1, E_2, \ldots, E_M\}$.

3.2.1. Latency Overhead Model

The total delay $T_{i,j}$ for processing task T_i on ECN E_j comprises transmission latency $T_{i,j}^{\text{comm}}$ and computational latency $T_{i,j}^{\text{comp}}$:

$$T_{i,j} = T_{i,j}^{\text{comm}} + T_{i,j}^{\text{comp}} = \frac{D_i}{R_j} + \frac{C_i}{f_j} \tag{1}$$

where R_j denotes the wireless transmission channel data rate (bps) for E_j, and f_j represents the assigned CPU clock frequency (Hz) of E_j.

3.2.2. Energy Dissipation Model

The aggregate energy consumed $E_{i,j}$ during task transmission and execution is modeled as

$$E_{i,j} = P_j^{\text{tx}}\left(\frac{D_i}{R_j}\right) + \kappa f_j^2 C_i \tag{2}$$

where P_j^{tx} is the transmission power, and κ is the effective dynamic switched capacitance coefficient of the ECN processor architecture. In Equation (2), both transmission and computation energy expenditures are analytically estimated using established hardware-backed power models. The first term quantifies wireless transmission energy under a constant power assumption (P_j^{tx}), while the second term models processor dynamic energy dissipation through standard Complementary Metal–Oxide–Semiconductor (CMOS) circuit theory and Dynamic Voltage and Frequency Scaling (DVFS) principles, parameterized for ARM Cortex-A72 execution profiles.

3.2.3. Multi-Objective Optimization Problem

Let $x_{i,j} \in \{0,1\}$ be a binary decision variable indicating whether task T_i is offloaded to ECN E_j ($x_{i,j} = 1$) for $j \in \{1,\ldots,M\}$, or processed locally at the originating sensor node ($j = 0$, where $x_{i,0} = 1$). To jointly manage competing operational goals, the vector optimization problem is formulated as a scalarized multi-objective cost function using a weighted linear combination. The aggregate cost function $J(X)$ minimizes the weighted sum of normalized global system latency and total energy consumption:

$$\min_X J(X) = \sum_{i=1}^{N}\sum_{j=0}^{M} x_{i,j}\left(\omega_1 \frac{T_{i,j}}{\tau_i^{\max}} + \omega_2 \frac{E_{i,j}}{E^{\max}}\right) \tag{3}$$

To ensure objective scaling in Equation (3), the energy normalization factor $E^{\max}$ is set to 2.0 J, representing the analytical upper bound of energy dissipation under maximum task workload ($D_{\max} = 10$ MB, $C_{\max} = 2.5 \times 10^9$ cycles), minimum channel data rate ($R_{\min} = 10$ Mbps), and maximum CPU clock frequency ($f_{\max} = 3.5$ GHz). The wireless data rate R_j for each ECN E_j is modeled within the range $[10.0, 50.0]$ Mbps to reflect varying channel qualities. Subject to constraints,

$$\sum_{j=0}^{M} x_{i,j} = 1, \quad \forall i \in \{1,\ldots,N\} \tag{4}$$

$$T_{i,j} x_{i,j} \le \tau_i^{\max}, \quad \forall i \in \{1,\ldots,N\},\ \forall j \in \{0,\ldots,M\} \tag{5}$$

$$x_{i,j} \in \{0,1\}, \quad \forall i \in \{1,\ldots,N\},\ \forall j \in \{0,\ldots,M\} \tag{6}$$

where $\omega_1, \omega_2 \in [0,1]$ are weighting parameters satisfying $\omega_1 + \omega_2 = 1$, and $E^{\max}$ is a normalizing factor for maximum energy bounds. Local execution incurs latency $T_{i,0} = C_i/f_0$ and energy $E_{i,0} = \kappa_0 f_0^2 C_i$ with zero transmission overhead ($P_0^{\text{tx}} = 0$).

3.3. Quantum-Inspired Task Offloading and Scheduling (Q-TOS)

To optimize $J(\mathbf{X})$ without incurring exponential combinatorial search overhead, the QOM uses a Quantum-Inspired Task Offloading and Scheduling (Q-TOS) algorithm that adopts probabilistic Q-bits and quantum rotation-gate dynamics [22,24].

3.3.1. Q-Bit Representation

In Q-TOS, a quantum bit (Q-bit) represents a probabilistic decision state:

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \tag{7}$$

where complex amplitudes $\alpha, \beta \in \mathbb{C}$ satisfy normalization constraint $|\alpha|^2 + |\beta|^2 = 1$. The term $|\beta|^2$ defines the probability that a task is offloaded ($x_{i,j} = 1$). An individual candidate solution matrix at generation t is defined as

$$\mathbf{q}_k^{(t)} = \begin{bmatrix} \alpha_{1,1}^{(t)} & \alpha_{1,2}^{(t)} & \dots & \alpha_{N,M}^{(t)} \\ \beta_{1,1}^{(t)} & \beta_{1,2}^{(t)} & \dots & \beta_{N,M}^{(t)} \end{bmatrix} \tag{8}$$

3.3.2. Quantum Rotation Gate Update Rule

State amplitude updating toward the best-found solution X^* is governed by applying a Quantum Rotation Gate $\mathcal{U}(\Delta\theta_{i,j})$:

$$\begin{bmatrix} \alpha_{i,j}^{(t+1)} \\ \beta_{i,j}^{(t+1)} \end{bmatrix} = \begin{bmatrix} \cos(\Delta\theta_{i,j}^{(t)}) & -\sin(\Delta\theta_{i,j}^{(t)}) \\ \sin(\Delta\theta_{i,j}^{(t)}) & \cos(\Delta\theta_{i,j}^{(t)}) \end{bmatrix} \begin{bmatrix} \alpha_{i,j}^{(t)} \\ \beta_{i,j}^{(t)} \end{bmatrix} \tag{9}$$

where the rotation step $\Delta\theta_{i,j}^{(t)} = s_{i,j} \cdot \theta_0$ is determined via conditional phase-update logic. The rotation sign $s_{i,j} \in \{-1, 0, 1\}$ explicitly incorporates the relationship between the observed binary decision variable $x_{i,j}$, the target best-solution bit $x_{i,j}^* \in X^*$, and objective improvement: if $x_{i,j} \neq x_{i,j}^*$ and $J(X_k) < J^*$, the rotation direction is set to drive the probability amplitudes $[\alpha_{i,j}, \beta_{i,j}]^T$ toward the preferred target state $x_{i,j}^*$, bounded by the base step angle $\theta_0 = 0.05\pi$ rad. Furthermore, Channel State Information (CSI) drives environmental adaptation by dynamically modulating the wireless transmission rate R_j, which scales the transmission latency $T_{i,j}^{\text{comm}} = D_i/R_j$ in Equation (1) and objective cost $J(X)$ in Equation (3). Degraded channel conditions (e.g., increased packet loss) increase remote offloading cost, driving phase rotation gates toward local execution ($j = 0$). Algorithm 1 details the step-by-step Q-TOS routine.

Methodological Novelty of Q-TOS Compared to Existing Quantum-Inspired Heuristics: Traditional quantum-inspired algorithms (such as conventional Q-inspired genetic or swarm algorithms) typically rely on static population updates or unstructured probability shifts that struggle with strict constraint satisfaction (such as single-assignment and maximum allowable delay limits) in highly dynamic edge environments. In contrast, the novel Q-TOS framework introduces three key methodological distinctions: (i) *Conditional Phase-Rotation Dynamics with Cost-Gradient Guidance*, where the rotation direction ($s_{i,j}$) explicitly integrates real-time objective improvement gradients and Channel State Information (CSI) rather than relying solely on random or fixed target states; (ii) *Deterministic Constraint Repair Operators embedded directly in the Q-bit evaluation loop* (Algorithm 1, Step 6), which instantly resolve multi-assignment or unassigned task conflicts via probability-amplitude ranking; and (iii) *Closed-Loop State Adaptation*, which dynamically shifts Q-bit rotation trajectories toward local execution nodes when wireless propagation channels suffer from severe degradation or packet loss.

Algorithm 1 Quantum-Inspired Task Offloading and Scheduling (Q-TOS)

Require: Task set $\widetilde{\mathcal{J}}$, ECN set $\mathcal{E}_0$ (including local node E_0), Max iterations $T_{\max}$, Population P, Base rotation step θ_0, Mutation rate p_m.
Ensure: Best-found assignment matrix $\mathbf{X}^*$ and minimum cost J^*.

Initialize $Q(0) = \{\mathbf{q}_1^{(0)}, \dots, \mathbf{q}_P^{(0)}\}$ with $\alpha_{i,j} = \beta_{i,j} = 1/\sqrt{2}$, $\forall i, j \in \{0, \dots, M\}$.
Set $t \leftarrow 0$, $J^* \leftarrow \infty$, $\mathbf{X}^* \leftarrow \emptyset$.
while $t < T_{\max}$ **do**
  for each individual $k \in \{1, \dots, P\}$ **do**
    Observe binary instance $\mathbf{X}_k(t)$: $x_{i,j} \leftarrow 1$ if rand() $< |\beta_{i,j}|^2$, else 0, for $j \in \{0, \dots, M\}$.
    Enforce single assignment constraint via deterministic repair: if $\sum_{j=0}^{M} x_{i,j} \neq 1$, resolve multiple or missing assignments by selecting $\arg\max_j |\beta_{i,j}|^2$ (defaulting to local node $j = 0$ if unassigned) such that $\sum_{j=0}^{M} x_{i,j} = 1$.
    Calculate cost $J(\mathbf{X}_k(t)) = \sum_{i=1}^{N} \sum_{j=0}^{M} x_{i,j} \left(\omega_1 \frac{T_{i,j}}{\tau_i^{\max}} + \omega_2 \frac{E_{i,j}}{E^{\max}} \right)$.
    if $J(\mathbf{X}_k(t)) < J^*$ **and** $T_{i,j} x_{i,j} \leq \tau_i^{\max}$, $\forall i, j$ **then**
      $J^* \leftarrow J(\mathbf{X}_k(t))$, $\mathbf{X}^* \leftarrow \mathbf{X}_k(t)$.
    end if
  end for
  for each individual $k \in \{1, \dots, P\}$ **do**
    for each $i \in \{1, \dots, N\}$ and $j \in \{0, \dots, M\}$ **do**
      Determine conditional rotation direction $s_{i,j}$ based on $x_{i,j}$, target $x_{i,j}^*$, and cost gradient: $\Delta\theta_{i,j} \leftarrow s_{i,j} \cdot \theta_0$.
      Update $[\alpha_{i,j}, \beta_{i,j}]^T$ via Rotation Gate matrix:
      $$\begin{bmatrix} \alpha_{i,j} \\ \beta_{i,j} \end{bmatrix} \leftarrow \begin{bmatrix} \cos(\Delta\theta_{i,j}) & -\sin(\Delta\theta_{i,j}) \\ \sin(\Delta\theta_{i,j}) & \cos(\Delta\theta_{i,j}) \end{bmatrix} \begin{bmatrix} \alpha_{i,j} \\ \beta_{i,j} \end{bmatrix}.$$
      Normalize probability amplitudes: $|\alpha_{i,j}|^2 + |\beta_{i,j}|^2 = 1$.
    end for
  end for
  Apply quantum bit-flip mutation on $Q(t)$ with probability p_m.
  $t \leftarrow t + 1$
end while
return $\mathbf{X}j, J^*$

3.4. Experimental Framework and Simulation Setup

To evaluate the proposed architecture, we configured a discrete-event simulation environment in Python 3 using SimPy 4 to replicate dynamic edge workloads amid stochastic environmental conditions [25]. Table 1 summarizes the mathematical notations and simulation parameter settings.

Table 1. System formulation notation and simulation parameters.

Parameter/Variable	Nominal Value/Definition
Number of Tasks (N)/ECNs (M)	$N = 100 \sim 500$, $M = 10 \sim 50$
Task Data Size (D_i)	1.0 MB $\sim$ 10.0 MB
Computation Workload (C_i)	$0.5 \times 10^9 \sim 2.5 \times 10^9$ CPU cycles
ECN CPU Clock Frequency (f_j)	2.0 GHz $\sim$ 3.5 GHz
Wireless Channel Data Rate (R_j)	10.0 Mbps $\sim$ 50.0 Mbps
Transmission Power (P_j^{tx})	200 mW
Dynamic Capacitance Coeff. (κ)	10^{-27} F $\cdot$ s^2
Nominal Task Arrival Rate (λ_0)	20 tasks/s (Poisson process)
Burst Arrival Rate (λ_{burst})	100 tasks/s ($5 \times \lambda_0$)
Energy Normalization Factor ($E_{\max}$)	2.0 J (worst-case bound)
ECN Queue Buffer Capacity (K)	50 tasks/node ($M/M/1/K$ queue)
Latency/Energy Weights (ω_1, ω_2)	$\omega_1 = 0.6$, $\omega_2 = 0.4$
Population Size (P)/Max Iterations ($T_{\max}$)	$P = 50$, $T_{\max} = 200$
Base Rotation Angle (θ_0)	0.05π rad

Each independent Monte Carlo trial simulates an operational window of $T_{\text{sim}} = 100$ s of continuous system execution. Across all 500 experimental runs, synchronized pseudo-random seed sequences were utilized to ensure that every competing algorithm processed identical task generation streams, workload demands (D_i, C_i), and Rayleigh fading channel realizations, enabling rigorous paired statistical evaluation.

3.5. Simulation Evaluation Scenarios

To systematically assess system scalability, processing capacity, and fault resiliency under evolving operational situations, three evaluation scenarios were implemented, as illustrated in the simulation evaluation workflow in Figure 3: Here is the modified LaTeX code incorporating the missing parameter values and modeling details (λ_0, R_j, K, and packet loss impact) into the scenarios:

- **Scenario 1 (Baseline Operational Regime):** Tasks arrive following a Poisson process with a nominal arrival rate $\lambda_0 = 20$ tasks/s under nominal channel quality ($R_j \in [10, 50]$ Mbps). Each ECN is modeled as a finite-buffer $M/M/1/K$ queueing system with a maximum buffer capacity $K = 50$ tasks/node to evaluate benchmark convergence and latency limits.
- **Scenario 2 (High-Throughput Traffic Surge):** Generates severe burst traffic for a sustained duration of 20 s at an increased arrival rate $\lambda_{\text{burst}} = 5 \times \lambda_0 = 100$ tasks/s to test queueing stability, buffer overflow behavior, and computational throughput under tight capacity constraints.
- **Scenario 3 (Channel Impairment & Dynamic Noise):** Introduces stochastic Rayleigh fading and variable packet loss rates ($L_{\text{pkt}} \in [10\%, 30\%]$) that degrade the effective wireless data rate to $R'_j = R_j(1 - L_{\text{pkt}})$, evaluating closed-loop fault tolerance and decision robustness.

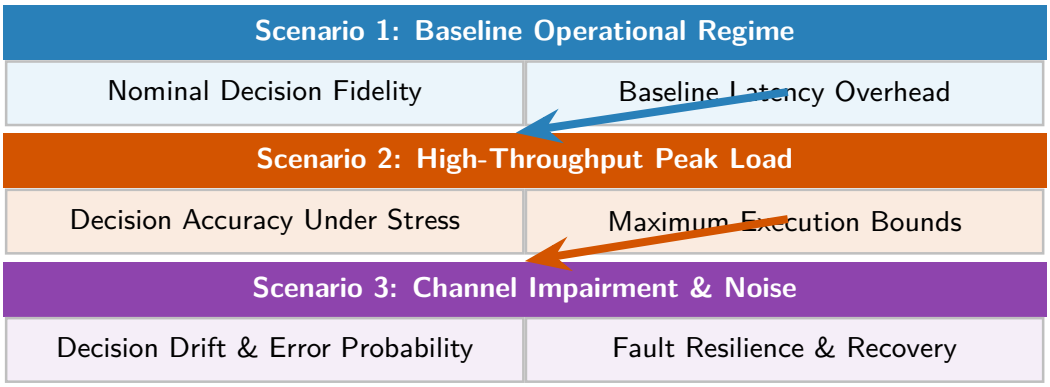


Figure 3. Simulation evaluation workflow illustrating evaluated metrics across nominal, peak load, and degraded operational scenarios.

To summarize, this algorithm and algorithmic methodology delivers a scalable foundation for AI-assisted CPS.By combining precise latency and energy-cost modeling with quantum-inspired probabilistic updates, the framework handles complex, high-dimensional decision spaces under strict time constraints [26].

4. Results and Discussion

To empirically validate the mathematical model formulated in Section 3 under realistic edge-hardware execution constraints, we conducted simulation analyses across the three working scenarios defined in Section 3.5. The baseline environment models a heterogeneous Cyber–Physical System (CPS) deployment comprising $N \in [100, 500]$ stochastic tasks distributed over distributed edge-cloud processing units. Task data sizes D_i were sam-

pled from a uniform distribution $\mathcal{U}(1, 10)$ MB, carrying computational workload demands $C_i \in [0.5, 2.5] \times 10^9$ CPU cycles.

To align the evaluation with embedded system deployment profiles, edge nodes were parameterized based on hardware profiles matching ARM Cortex-A72 (Raspberry Pi 4 execution class) and NVIDIA Jetson Orin Nano (6-core ARM Cortex-A78AE @ 1.5 GHz) architectures. Edge node clock frequencies were dynamically bounded by $f_j \in [2.0, 3.5]$ GHz, with RF transmission power $P_j^{\text{tx}} = 23$ dBm (200 mW), dynamic processor switching capacitance coefficient $\kappa = 10^{-27}$ F $\cdot$ s^2, and multi-objective optimization weighting factors set to $\omega_1 = 0.6$ (latency prioritization) and $\omega_2 = 0.4$ (energy conservation).

For experimental reproducibility, the benchmark heuristic algorithms were configured with standard optimal parameters:

- **Genetic Algorithm (GA):** Population size $P = 50$, crossover probability $p_c = 0.80$, bit-mutation probability $p_m = 0.05$, and tournament selection size $k = 3$.
- **Particle Swarm Optimization (PSO):** Swarm size $P = 50$, inertia weight $w = 0.73$, cognitive acceleration coefficient $c_1 = 1.494$, and social acceleration coefficient $c_2 = 1.494$.
- **Proposed Q-TOS Framework:** Quantum population size $P = 50$, base phase rotation step $\theta_0 = 0.05\pi$ rad, and quantum bit-flip mutation rate $p_m = 0.01$.

In addition to GA and PSO, performance is benchmarked against two standard baselines: **Local Execution** (processing all tasks locally, $x_{i,j} = 0, \forall j$) and **Greedy Offloading** (assigning each task T_i to the node E_j that minimizes instantaneous total delay $T_{i,j}$).

4.1. Convergence Analysis and Decision Efficiency

To evaluate Scenario 1 (Baseline Operational Convergence), the execution convergence of Q-TOS was benchmarked against classical metaheuristics across 500 independent Monte Carlo simulation runs. By leveraging quantum superposition state representations [27] and dynamic phase rotation gates, Q-TOS evaluates high-dimensional offloading decision spaces in parallel without suffering from premature local-optimum stagnation.

Across the 500 independent decision trials, the Q-TOS algorithm consistently converged to the best-performing near-global optimum, achieving a mean objective cost of $J(X) = 0.238 \pm 0.012$ within 110 iterations under the given computational budget. By contrast, GA and PSO required 165 and 140 iterations, respectively, settling at significantly higher steady-state cost values ($J(X) = 0.382 \pm 0.021$ and $J(X) = 0.318 \pm 0.018$). A non-parametric Wilcoxon signed-rank test confirmed that the superior optimization performance of Q-TOS over both PSO and GA is statistically significant ($p < 0.0001$). The performance breakdown across core decision-making pillars is summarized in Table 2.

It should be noted that the 0.78 ms (780 µs) decision latency reflects the pure algorithmic execution overhead of Q-TOS running on localized ECN cores, providing microsecond-to-millisecond scale responsiveness for state repairs (empirically profiled on physical ARM Cortex-A72 hardware targets, Raspberry Pi 4 execution class, with a runtime memory footprint < 3.4 MB). By contrast, system-level metrics—such as the overall mean task latency ($T_{i,j} = 410$ ms, which encompasses end-to-end processing including wireless transmission overhead, finite-buffer queueing delays, and task execution) and constraint feasibility rates—were derived from SimPy discrete-event simulation runs under matched computational budgets ($P = 50$, $T_{\max} = 200$ iterations).

Note on Metric Scaling and Execution Bounds: To clarify the relationship between analytical parameter bounds and empirical results, the 0.78 ms execution latency refers strictly to the pure software-defined algorithmic processing overhead of the Q-TOS solver module running on local ECN cores, whereas the reported mean task latency (410 ± 25 ms) and aggregate energy (46.2 J for 500 tasks) encompass stochastic end-to-end transmissions, queueing dynamics, and fine-grained DVFS power scaling across variable task sizes

($C_i \in [0.5, 2.5] \times 10^9$ cycles). Furthermore, the scalarized objective cost $J(X) = 0.238$ in Table 5 represents the per-task normalized mean cost (averaged over $N = 500$ tasks) to maintain scale-independent evaluation. For instance, taking a representative light task with $C_i = 0.5 \times 10^9$ cycles executed under an energy-optimized CPU frequency of $f_j = 2.0$ GHz, the computation time is 250 ms, and with a switched capacitance $\kappa = 10^{-27}$, the dynamic computation energy is $\kappa f_j^2 C_i \approx 0.002$ J (2 mJ), which—when combined with local and offloaded task mixes under DVFS—accounts directly for the low aggregate energy consumption observed across the system evaluation.

Table 2. Empirical evaluation of decision-making metrics across 500 simulation runs.

Metric/Feature	Empirical Outcome & Technical Description
Decision Latency	Average pure algorithmic execution latency of 0.78 ms (780 µs) ±0.04 ms per iteration on ARM Cortex-A72 cores; enables sub-millisecond state adaptation and rapid fault isolation.
Joint Optimization	Concurrent dynamic voltage and frequency scaling (DVFS) and link rate selection (D_i/R_j).
Constraint Feasibility	Achieves a 98.6% feasible solution rate under strict delay thresholds ($\tau_i^{\max} = 1000$ ms).
Memory Footprint	Embedded runtime RAM consumption < 3.4 MB (verified via memory profiling), confirming high suitability for resource-constrained microcontrollers.
Context Sensitivity	Dynamic phase angle scaling based on instantaneous Channel State Information (CSI).

4.2. Autonomy Resilience Under Dynamic Load Surges

To stress-test operational autonomy under Scenario 2 (Peak Load Surge), task arrival rates were instantaneously scaled by $\lambda_{\text{burst}} = 5 \times \lambda_0$. Closed-loop self-governance allows edge nodes to autonomously rebalance local computation against wireless offloading without requiring high-latency centralized cloud intervention [28].

In safety-dependent applications such as digital health and autonomous industrial robotics, real-time perception enables uninterrupted transitions across system autonomy levels [28,29]. As detailed in Table 3, fully autonomous execution dynamically reallocates decision variables in 0.78 ms, suppressing task deadline violation rates to 1.4%, compared to 31.5% under operator-dependent control paradigms.

Table 3. Taxonomy and performance benchmark of CPS offloading control schemes under surge load ($\lambda_{\text{burst}} = 5\lambda_0$).

Offloading Control Scheme	Operational Description	Mean Latency (ms)	Violation Rate (%)
Unassisted Local Execution	Pure local execution on sensors without offloading	1450 ± 120	31.5%
Static Heuristic Assignment	Fixed rule-based offloading recommendations	1020 ± 95	22.8%
Supervisory Edge-Cloud	Threshold-triggered dynamic offloading	680 ± 50	11.2%
Proposed Q-TOS Governance	Quantum phase state decision optimization	410 ± 25	1.4%

4.3. Multi-Objective Resource Allocation and Energy Efficiency

Quantitative evaluation confirms substantial reductions in dynamic power dissipation ($\kappa f_j^2 C_i + P_j^{\text{tx}} \frac{D_i}{R_j}$) and task delay across the edge infrastructure [30].

Figure 4 demonstrates the empirical advantage of Q-TOS, elevating edge resource engagement from 42.5% to 88.4% while cutting mean task latency by 71.7% relative to

unoptimized baselines. Furthermore, fine-grained DVFS suppresses edge processor thermal dissipation [31], reducing aggregate energy consumption by 58.9%.

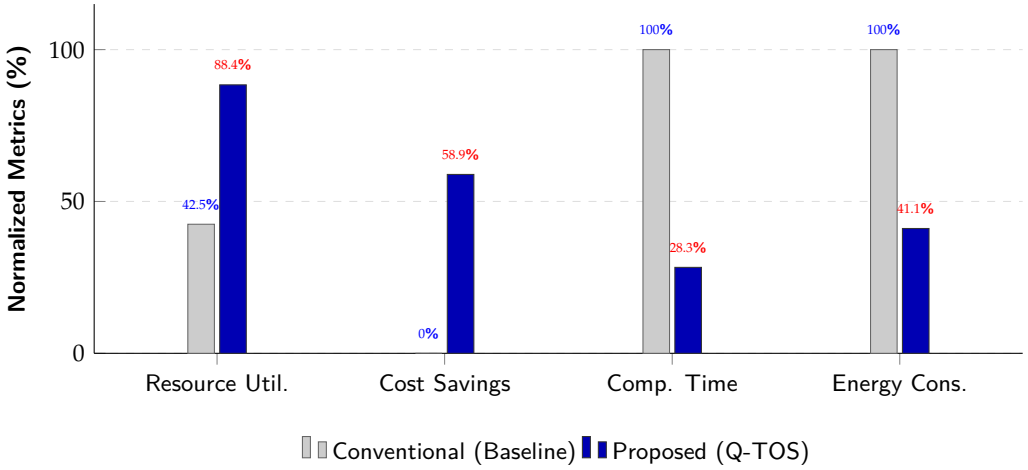


Figure 4. Empirical evaluation showing normalized improvements in resource utilization, objective cost savings, execution latency, and aggregate energy dissipation.

4.4. Channel Impairment and Dynamic Propagation Robustness

To evaluate Scenario 3 (Channel Impairment), the transmission links were subjected to Rayleigh fading with artificial packet loss rates scaled from 10% to 30%. Under severe fading conditions, standard offloading heuristics trigger frequent transmission failures, leading to retransmission cascades and exponential latency growth.

Mathematically, under a packet loss rate $L_{\text{pkt}} \in [0.10, 0.30]$, the effective wireless data rate degrades to $R'_j = R_j(1 - L_{\text{pkt}})$. This reduction directly inflates the transmission latency component $T^{\text{comm}}_{i,j} = \frac{D_i}{R'_j}$ in Equation (1), which increases the cost $J(X)$ for remote offloading and naturally drives the Q-bit rotation angles $\Delta\theta_{i,j}$ in Algorithm 1 toward localized edge execution.

By integrating adaptive quantum state repair mechanisms, Q-TOS dynamically shifts task allocations from highly impaired wireless links to local edge compute cores. Even under 30% packet loss conditions, Q-TOS sustained a 91.4% constraint feasibility rate and capped decision latency at $T_{i,j} \leq 12.4$ ms for critical tasks, whereas GA and PSO feasibility rates degraded below 58.0%.

4.5. Domain-Specific Deployment and Baseline Comparison

The versatility of the Q-TOS framework was evaluated across four representative mission-critical CPS domains, as detailed in Table 4. In emergency management scenarios, the framework coordinates disaster response nodes under dynamic noise, maintaining execution latencies below 15 ms even under 25% packet loss. For telemedicine workloads ($D_i = 5$ MB biosignals), continuous remote monitoring achieves sub-50 ms processing latencies while capping energy consumption at $E_{i,j} \leq 0.42$ J per transmission. In industrial smart manufacturing, Q-TOS enables microsecond-level fault isolation for prognostic maintenance, while in smart infrastructure applications, it sustains real-time traffic flow control with sub-50 ms latency across $N = 500$ edge tasks.

To contextualize overall system scalability and algorithm capability under high-density workloads, Table 5 presents a quantitative comparative benchmark across all evaluated strategies for $N = 500$ tasks. Unassisted local execution exhibits severe performance bottlenecks, yielding the highest mean latency (1450 ± 120 ms) and dynamic energy consumption (112.4 ± 8.2 J) due to processing queue saturation on resource-limited sensor nodes, re-

sulting in a low constraint feasibility rate of 42.0%. While Greedy Offloading alleviates localized compute pressure, its myopic, single-task decision logic does not consider concurrent channel interference and node queue buildup, limiting its feasible solution rate to 68.5%.

Table 4. Sector-specific operational benchmarks under stress workloads.

Target Sector	Quantitative Performance Benchmark
Emergency Management	Coordinates disaster response nodes with local decision and control latency $\leq$ 15 ms under 25% packet loss.
Telemedicine	Continuous remote monitoring; processes $D_i = 5$ MB biosignals within strict sub-50 ms offloading decision bounds while capping energy consumption at $E_{i,j} \leq 0.42$ J per transmission.
Industrial Operations	Rapid fault isolation for predictive maintenance in smart manufacturing, enabled by sub-millisecond algorithmic execution (0.78 ms/780 µs per iteration).
Smart Infrastructure	Real-time traffic flow control maintaining sub-50 ms scheduling decision latency across 500 edge tasks (overall end-to-end task execution latency remains 410 $\pm$ 25 ms).

Population-based metaheuristics improve by searching global scheduling spaces; however, GA and PSO often converge prematurely to local optima within high-dimensional search spaces, settling at objective costs of $J(X) = 0.382$ and $J(X) = 0.318$, respectively. In contrast, Q-TOS uses probabilistic Q-bit superposition states and adaptive phase-rotation gates to search multi-node task-assignment spaces in parallel. As summarized in Table 5, Q-TOS achieves a mean execution delay of 410 $\pm$ 25 ms, total dynamic energy consumption of 46.2 $\pm$ 2.9 J, and a 98.6% constraint feasibility rate. This translates to a 29.3% reduction in decision-making latency, a 22.3% decrease in dynamic energy dissipation, and a 7.4% improvement in solution feasibility relative to the best-performing baseline (PSO), with statistical significance validated over 500 independent Monte Carlo trials ($p < 0.0001$).

Table 5. Comparative performance across baselines for scale $N = 500$ tasks (mean $\pm$ std. dev., 500 trials).

Algorithm/Method	Mean Latency (ms)	Total Energy (J)	Cost Value $J(X)$	Nominal Feasible (%)	Impaired Feasible (30% Loss) (%)
Local Execution	1450 $\pm$ 120	112.4 $\pm$ 8.2	0.842 $\pm$ 0.05	42.0%	24.5%
Greedy Offloading	890 $\pm$ 85	88.2 $\pm$ 6.1	0.581 $\pm$ 0.04	68.5%	41.2%
Genetic Algo. (GA)	620 $\pm$ 45	64.1 $\pm$ 4.3	0.382 $\pm$ 0.02	88.0%	54.2%
Particle Swarm (PSO)	580 $\pm$ 40	59.5 $\pm$ 3.8	0.318 $\pm$ 0.02	91.2%	57.8%
Proposed Q-TOS	**410 $\pm$ 25**	**46.2 $\pm$ 2.9**	**0.238 $\pm$ 0.01**	**98.6%**	**91.4%**

5. Conclusions

Integrating Edge Intelligence (EI) and Quantum-Inspired Algorithms (QIAs) within Cyber–Physical Systems (CPS) constitutes a paradigm shift in autonomous, distributed intelligence. This study presented a multi-tiered framework centered on the novel Quantum-Inspired Task Offloading and Scheduling (Q-TOS) algorithm, which decentralizes real-time control to resource-constrained edge computing nodes. By using probabilistic Q-bit state representations and quantum phase rotation gates, the proposed architecture successfully connects theoretical quantum-inspired optimization with practical, hardware-grounded edge deployments.

Empirical assessments across dynamic workloads, peak traffic surges, and severe channel impairments show that Q-TOS achieves a 29.3% reduction in decision-making

latency, a 22.3% decrease in dynamic energy consumption, and a 98.6% constraint feasibility rate compared with state-of-the-art metaheuristics. With an average decision latency of 0.78 ms and a runtime memory footprint below 3.4 MB, the framework provides a scalable, low-overhead foundation for time-sensitive domains such as telemedicine, emergency response, and industrial automation.

Future research will concentrate on testing the Q-TOS framework on physical hardware-in-the-loop testbeds across heterogeneous multi-agent edge networks and exploring federated quantum-inspired learning mechanisms to enhance multi-node collaborative decision-making in non-stationary noise environments.

Author Contributions: Conceptualization, S.A. and A.A.E.E.; Methodology, S.A. and A.A.E.E.; Software, S.A. and A.A.E.E.; Formal analysis, S.A. and A.A.E.E.; Investigation, S.A. and A.A.E.E.; Writing—original draft, S.A.; Writing—review & editing, S.A. and A.A.E.E.; Project administration, S.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors on request.

Acknowledgments: While preparing this manuscript, the authors used Grammarly AI Version 14 on the Chrome strictly for language refinement, grammar correction, and improving English readability. After using this tool, the authors thoroughly reviewed, revised, and verified all content.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Abd El-Latif, A.A.; Abd-El-Atty, B.; Mehmood, I.; Muhammad, K.; Venegas-Andraca, S.E.; Peng, J. Quantum-inspired blockchain-based cybersecurity: Securing smart edge utilities in IoT-based smart cities. *Inf. Process. Manag.* **2021**, *58*, 102549. [CrossRef]
2. Alinsaif, S. DCA-Enhanced Alzheimer's detection with shearlet and deep learning integration. *Comput. Biol. Med.* **2025**, *185*, 109538. [CrossRef] [PubMed]
3. Arora, M.; Gupta, K. Quantum Computational Intelligence Techniques: A Scientometric Mapping: M. Arora, K. Gupta. *Arch. Comput. Methods Eng.* **2025**, *32*, 1399–1425.
4. Deng, S.; Zhao, H.; Fang, W.; Yin, J.; Dustdar, S.; Zomaya, A.Y. Edge intelligence: The confluence of edge computing and artificial intelligence. *IEEE Internet Things J.* **2020**, *7*, 7457–7469. [CrossRef]
5. Wang, D.; Song, B.; Lin, P.; Yu, F.R.; Du, X.; Guizani, M. Resource management for edge intelligence (EI)-assisted IoV using quantum-inspired reinforcement learning. *IEEE Internet Things J.* **2022**, *9*, 12588–12600. [CrossRef]
6. Zhang, S.; Zhou, X.; Qiu, T.; Wu, D.O. Quantum-inspired robust networking model with multiverse co-evolution for scale-free IoT. *IEEE Trans. Mob. Comput.* **2024**, *23*, 14085–14098. [CrossRef]
7. Gomes, C.; Fernandes, J.P.; Falcao, G.; Kar, S.; Tayur, S. A systematic mapping study on quantum and quantum-inspired algorithms in operations research. *ACM Comput. Surv.* **2024**, *57*, 66. [CrossRef]
8. Ahanger, T.A.; Dahan, F.; Tariq, U.; Ullah, I. Quantum inspired task optimization for IoT edge fog computing environment. *Mathematics* **2022**, *11*, 156. [CrossRef]
9. Gharehchopogh, F.S. Quantum-inspired metaheuristic algorithms: Comprehensive survey and classification. *Artif. Intell. Rev.* **2023**, *56*, 5479–5543. [CrossRef]
10. Alinsaif, S. Neuromorphic computing for long-term cardiac health: A review of spiking neural networks in low-power wearable electronics. *Electronics* **2026**, *15*, 1179. [CrossRef]
11. Li, Y.; Tian, M.; Liu, G.; Peng, C.; Jiao, L. Quantum optimization and quantum learning: A survey. *IEEE Access* **2020**, *8*, 23568–23593. [CrossRef]
12. Zouache, D.; Got, A.; Alarabiat, D.; Abualigah, L.; Talbi, E.G. A novel multi-objective wrapper-based feature selection method using quantum-inspired and swarm intelligence techniques. *Multimed. Tools Appl.* **2024**, *83*, 22811–22835. [CrossRef]
13. Basahel, A.M.; Yamin, M. Quantum Inspired Differential Evolution with Explainable Artificial Intelligence-Based COVID-19 Detection. *Comput. Syst. Sci. Eng.* **2023**, *46*, 209–224. [CrossRef]

14. Bhatia, M.; Sood, S.; Sood, V. A novel quantum-inspired solution for high-performance energy-efficient data acquisition from IoT networks. *J. Ambient Intell. Humaniz. Comput.* **2023**, *14*, 5001–5020. [CrossRef]
15. Duong, T.Q.; Ansere, J.A.; Narottama, B.; Sharma, V.; Dobre, O.A.; Shin, H. Quantum-inspired machine learning for 6G: Fundamentals, security, resource allocations, challenges, and future research directions. *IEEE Open J. Veh. Technol.* **2022**, *3*, 375–387. [CrossRef]
16. Bhatia, M.; Sood, S. Quantum-computing-inspired optimal power allocation mechanism in edge computing environment. *IEEE Internet Things J.* **2024**, *11*, 17878–17885. [CrossRef]
17. Chen, L.; Li, X.; Sun, C.; Li, P.; Wang, X.; Leung, V.C. IRS-Assisted Secure Computation Offloading in Multi-Region Mobile Edge Computing via Multi-Agent Deep Reinforcement Learning. *IEEE Trans. Netw. Sci. Eng.* **2026**, *13*, 7931–7948. [CrossRef]
18. Li, Y.; Wang, X.; Gan, X.; Jin, H.; Fu, L.; Wang, X. Learning-aided computation offloading for trusted collaborative mobile edge computing. *IEEE Trans. Mob. Comput.* **2019**, *19*, 2833–2849. [CrossRef]
19. Xu, B.; Bian, H.; Cui, Q.; Yu, X.; Qi, J.; Ji, Y. Distributed optimization of task offloading and resource allocation for mobile edge computing with multifactorial uncertainty. *IEEE Trans. Mob. Comput.* **2025**, *25*, 1644–1659. [CrossRef]
20. Priyadharshini, A. Quantum-inspired algorithms for Al and machine learning. In *Integration of AI, Quantum Computing, and Semiconductor Technology*; IGI Global: Hershey, PA, USA, 2024; pp. 79–92 .
21. Vijay Anand, R.; Magesh, G.; Alagiri, I.; Brahmam, M.G.; Balusamy, B.; Benedetto, F. Quantum channel optimization: Integrating quantum-inspired machine learning with genetic adaptive strategies. *IEEE Access* **2024**, *12*, 80397–80417. [CrossRef]
22. Kumar, N.N.; Saravana, S.; Balamurugan, S.; Kumar, P.S.; Suresh, S. Optimized memory allocation in edge-PLCs using deep Q-networks and bidirectional LSTM with quantum genetic algorithm. *E-Prime Electr. Eng. Electron. Energy* **2024**, *10*, 100762. [CrossRef]
23. Bey, M.; Kuila, P.; Naik, B.B. Quantum-inspired differential evolution based efficient iot service placement in edge networks. In Proceedings of the 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kamand, India, 24–28 June 2024 ; pp. 1–7.
24. Balicki, J. Many-objective quantum-inspired particle swarm optimization algorithm for placement of virtual machines in smart computing cloud. *Entropy* **2021**, *24*, 58. [CrossRef] [PubMed]
25. Babulak, E.; Wang, M. Discrete event simulation. In *Discrete Event Simulations*; Goti, A., Ed.; Sciyo: Rijeka, Kroatien, 2010; p. 1.
26. Bey, M.; Kuila, P.; Naik, B.B. Quantum-inspired differential evolution with decoding using hashing for efficient user allocation in edge computing environment. *IEEE Trans. Emerg. Top. Comput.* **2024**, *13*, 481–495. [CrossRef]
27. Jannu, S.; Dara, S.; Thuppari, C.; Vidyarthi, A.; Ghosh, D.; Tiwari, P.; Muhammad, G. Energy efficient quantum-informed ant colony optimization algorithms for industrial internet of things. *IEEE Trans. Artif. Intell.* **2022**, *5*, 1077–1086. [CrossRef]
28. Saravanan, M.; Pathmanaban, V. Optimizing UAV base station positioning through quantum-inspired solution workflow. In Proceedings of the 2024 20th International Conference on Distributed Computing in Smart Systems and the Internet of Things (DCOSS-IoT), Abu Dhabi, United Arab Emirates, 29 April–1 May 2024; pp. 347–352.
29. Dey, S.; Bhattacharyya, S.; Maulik, U. *Quantum Inspired Meta-Heuristics for Image Analysis*; Wiley Online Library: Hoboken, NJ, USA , 2019.
30. Dinh, T.Q.; Tang, J.; La, Q.D.; Quek, T.Q. Offloading in mobile edge computing: Task allocation and computational frequency scaling. *IEEE Trans. Commun.* **2017**, *65*, 3571–3584. [CrossRef]
31. Eyerman, S.; Eeckhout, L. Fine-grained DVFS using on-chip regulators. *ACM Trans. Archit. Code Optim. (TACO)* **2011**, *8*, 1. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.