Title	Trade-off between Fidelity and Latency for Entanglement Routing Design in Quantum Networks
Author(s)	NGUYEN, Thu Trang
Citation	
Issue Date	2025-09
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/20031
Rights	
Description	Supervisor: リム 勇仁, 先端科学技術研究科, 修士 (情報科学)

Abstract

Quantum networks intersect a paradigmatic shift in communication technologies, offering unprecedented capabilities for secure communication and distributed computation through the exploitation of quantum mechanical phenomena. These networks facilitate revolutionary and unlimited of possibility applications such as quantum key distribution (QKD), distributed quantum computing (DQC), quantum teleportation (QT), quantum clock synchronization (QCS), etc. —all fundamentally predicated on quantum entanglement as their operational cornerstone. Moreover, on the contrary with classical computer communication protocol, quantum networks have the fundamental no-cloning theorem in the quantum mechanic prohibits conventional signal amplification techniques, necessitating the development of quantum-native routing strategies that operate within these unique constraints.

However, the practical realization of quantum networks until now still faces a crucial challenge: The reliability and efficient distribution of high-fidelity entangled quantum states across long distances. Due to this distribution, the process faces intrinsic limitations arising from quantum decoherence, photonic loss in transmission channels, and the inherently probabilistic nature of entanglement generation protocols. In quantum communication, fidelity serves as a fundamental metric, representing the degree of similarity or correlation between the actual distributed entangled state and an ideal maximally entangled Bell pair. Fidelity is a crucial measure of the quality of quantum states in a network since it directly affects the reliability and effectiveness of quantum communication protocols. High fidelity ensures that the transmitted entangled states maintain their quantum properties, closely resembling the ideal Bell pair, and are thus suitable for performing secure and efficient quantum operations.

Latency, on the other hand, refers to the total time delay involved in the entire process of entanglement establishment, from the initial creation of the entangled pairs to the final verification and confirmation of the entanglement quality. This includes several distinct components: the physical propagation delay, which accounts for the time it takes for quantum information to travel through the communication channel (such as optical fibers); the entanglement generation time, which involves the probabilistic creation of entangled states between quantum nodes; and the quantum memory storage time, which reflects the duration for which quantum states

are held in memory while awaiting further operations such as entanglement swapping or purification. Although significant research has yielded solutions aimed at achieving high-fidelity entanglement despite these obstacles, the equally critical factor of time delay, or latency, in quantum routing links remains largely unaddressed in current investigations. This oversight is notable, given that latency is a crucial performance bottleneck for real-world quantum applications. Together, fidelity and latency define the efficiency and effectiveness of quantum networks. Achieving a balance between these two factors is a fundamental challenge in quantum routing, wherein high fidelity typically necessitates increased complexity of the routing processes, and conversely, high latency often compromises the quality of the quantum state, so it needs many purification steps.

This thesis addresses this critical challenge by introducing Q_{FiLa} (Quantum Fidelity-Latency), a novel quantum link metric specifically engineered to optimize the fundamental trade-off between quantum state fidelity and communication latency in entanglement routing protocols. Q_{FiLa} tackle this dichotomy by providing a unified optimization framework that enables adaptive, application-aware routing decisions essential for next-generation quantum communication infrastructures. Our research systematically integrates the Q_{FiLa} metric within the established Q-LEAP routing framework, enabling dynamic path selection through parameterized weighted combinations of fidelity and latency objectives. This integration facilitates adaptive routing strategies that can be tailored to diverse network conditions and heterogeneous application requirements. Specifically, Q_{FiLa} empowers routing protocols to intelligently prioritize either minimal latency (critical for time-sensitive applications such as real-time QKD and quantum sensing) or maximal fidelity (essential for high-precision applications including quantum teleportation and fault-tolerant distributed quantum computing), thereby aligning network resource allocation with application-specific requirements.

Before proceeding with the analysis of the Q_{FiLa} metric, we first examine the latency within the Nested Purification Protocol (NPP). This investigation allows us to understand the overall structure of latency in a specific purification protocol. Subsequently, we will extend our analysis to explore latency in the context of individual link routing, providing a detailed examination of latency at a more granular level within the quantum network architecture. To evaluate Q_{FiLa}'s performance characteristics, we conduct comprehensive simulation studies in both realistic and synthetic quantum network topologies. The Japan Photonic Network Model (JPNM) serves as our primary real-world benchmark, accurately reflecting the complexities and constraints of contemporary telecommunications infrastructure. Additionally, we add the simulation to another level by employing systematically generated random topologies with varying scales and connection densities to assess the metric's generalizability across diverse network configurations. Our experimental evaluation demonstrates that Q_{FiLa}-enhanced routing algorithms achieve substantial performance improvements over baseline Q-LEAP implementations, exhibiting significant reductions in end-to-end latency and computational overhead while maintaining equivalent throughput characteristics. These efficiency gains are particularly pronounced in high-fidelity operational regimes, attributable to Q_{FiLa}'s capacity to prioritize near-optimal communication links and effectively prune the routing search space through intelligent weighting mechanisms—an advantage that becomes increasingly critical in large-scale networks where routing computational overhead represents a primary performance bottleneck.

Furthermore, Q_{FiLa} demonstrates remarkable adaptability across diverse network topologies and scales. Comprehensive sensitivity analyses conducted on random network topologies with varying scale factors consistently reveal robust performance characteristics, even in sparse and topologically irregular configurations. This versatility underscores Q_{FiLa}'s practical viability for deployment across both established telecommunications-inspired quantum networks and emerging decentralized quantum internet architectures.

A principal contribution of this work lies in demonstrating that joint optimization of fidelity and latency parameters yields substantial improvements in overall quantum network performance metrics. Q_{FiLa} provides a computationally efficient, algorithmically flexible, and architecturally scalable solution for addressing this fundamental optimization challenge, thereby establishing a foundation for advanced quality-aware quantum routing protocols. Beyond these immediate contributions, this thesis identifies promising research directions for incorporating additional performance metrics—including hop count, entanglement generation success probability, quantum memory coherence time, and purification protocol overhead—into future routing optimization frameworks, advancing toward more comprehensive and resource-efficient quantum network protocols.

Despite these significant advances, several limitations remain that warrant future investigation. The current study operates under assumptions of static network topologies and homogeneous quantum link characteristics, without explicit modeling of dynamic network conditions, hardware heterogeneity, or entanglement purification failure mechanisms. Future research should incorporate adaptive link characterization based on real-time network feedback, explore multi-flow concurrent routing scenarios, and investigate hybrid classical-quantum routing architectures. Additionally, the development of machine learning-based adaptive parameter tuning for Q_{FiLa} weight optimization represents a particularly promising avenue for further performance enhancement.

In conclusion, this thesis introduces Q_{FiLa} as a novel and demonstrably effective quantum link metric that substantially advances the state-of-the-art in quantum network routing through joint optimization of fidelity and latency parameters. The proposed approach provides a robust theoretical and practical foundation for scalable, quality-aware quantum communication systems, contributing significantly to the realization of practical quantum internet infrastructure.