Doctoral Dissertation Academic Year 2025

Resource Management in Heterogeneous Quantum Repeater Networks

Graduate School of Media and Governance
Keio University

Naphan Benchasattabuse

Naphan Benchasattabuse
ORCID: 0000-0003-4475-3015

ABSTRACT

In this thesis, I explore whether it is possible to build a unified Quantum Internet architecture that supports different types of quantum repeaters—especially the two most distinct and seemingly incompatible ones: memory-based quantum repeaters and all-photonic, memoryless repeaters. These technologies have traditionally been developed with the aim of becoming the single dominant solution, but I ask: Can they work together in the same network? What kind of architecture would support both? And how can simulation help us understand what is needed to manage such a network at scale? To address these questions, I propose an architecture based on an existing recursive network design and programmable RuleSet-based protocols that can coordinate diverse hardware components. I introduce a new emitter-photon building block to bridge memory-based and all-photonic segments, and show how classical networking abstractions can be extended to manage quantum operations. While I have developed a simulation tool grounded in these architectural principles and validated it against existing simulators and analytical models, a full-scale investigation of the resource trade-offs and performance implications remains future work. Nevertheless, the results so far suggest that a unified, heterogeneous quantum network is not only possible but increasingly practical with current technologies—though ongoing experimental progress will be essential to fully realize this vision.

Keywords: Quantum Networking, Quantum Repeater, Network Protocols, Quantum Algorithms, Optimization Algorithms

ACKNOWLEDGEMENTS

I am very fortunate to have Rodney Van Meter and Michal Hajdušek as my advisors. Rod always reminded me not just to focus on the cool and fun ideas, but to always think of the impact and solve the real problems that would advance the field. He taught me, and everyone else in our group, to think beyond the short term, to look ahead and identify the challenges that truly matter. Rod has changed how I view research, how I look at problems, how I think, and how I write. Michal was always helpful in answering my sometimes dumb, basic questions about quantum information; he helped me refine my thinking, make my ideas concrete, and polish rough concepts from scratch, and was invaluable in preparing my papers and figures.

I would like to thank my thesis committee, Rodney Van Meter, Michal Hajdušek, Takahiko Satoh, Shota Nagayama, Hiroyuki Kusumoto, and Hideyuki Kawashima, for their valuable advice and comments in shaping this thesis.

I would also like to express my gratitude to Andreas Bärtschi and Stephan Eidenbenz for being good company, wonderful hosts, and great mentors during my internship at Los Alamos National Laboratory. Andreas and Stephan helped me learn to think with the rigor and formality of a computer scientist, a perspective I greatly needed. I would also like to thank Luis Pedro García-Pintos, Nathan Lemons, John Golden, and Yiğit Subaşi for the fruitful discussions that led to a paper included in this thesis.

My PhD would not have started without the IBM Quantum Challenge 2019, and for that, I thank the organizers for an event that made me realize I wanted to pursue a PhD and a research career in the field of quantum computation. I would also like to thank the MEXT Super Global scholarship program for supporting the second and third year of my graduate school journey, which enabled me to study without financial burden.

I would also like to express my appreciation to Ryosuke Satoh, Yasuhiro Okura, Sitong Liu, Takaaki Matsuo, Kentaro Teramoto, Shigetora Miyashita, Kento Samuel Soon, Poramet Pathumsoot, Monet Tokuyama, and all the members of AQUA for being good friends. Thank you for the good food, for exploring the city, and for all the enjoyable times in Japan during my PhD. I also thank my high school friends in the same journey of PhD here in Japan, Sopak Supakul and Panchanit Ubolsaard, for our occasional meetups, fun times, and mutual encouragement.

Finally, I am forever indebted and grateful to my parents, Somrak and Narongchai Benchasattabuse, and to my girlfriend, Thamolwan Pathomkasikul, for their unwavering emotional support. I could not have navigated all the obstacles and struggles to arrive at where I am today without you.

PUBLISHED CONTENT AND CONTRIBUTIONS

N. Benchasattabuse, M. Hajdušek, and R. Van Meter, "Bridging All-photonic and Memory-based Quantum Repeaters," submitted for publication, 2025.
N.B. conceived the project, developed key ideas of the architecture and the protocol, conducted simulation, and led the writing of the manuscript.

N. Benchasattabuse, A. Bärtschi, L. P. García-Pintos, J. Golden, N. Lemons, and S. Eidenbenz, "Lower bounds on the number of rounds of the quantum approximate optimization algorithm required for guaranteed approximation ratios," *Physical Review A*, vol. 111, no. 6, p. 062 411, Jun. 9, 2025, DOI: 10.1103/PhysRevA.111.062411.
N.B. developed key ideas of the proofs, participated in the conceiving of the project, completing the proofs, and led the writing of the manuscript.

N. Benchasattabuse, M. Hajdušek, and R. Van Meter, "Integrating Entanglement Purification into All-Photonic Quantum Repeaters," in *2025 IEEE International Conference on Quantum Computing and Engineering (QCE)*, to be published, Apr. 25, 2025, DOI: 10.48550/arXiv.2504.18121.
N.B. conceived the project, developed key ideas of the framework, conducted simulation, and led the writing of the manuscript.

J. Chung et al., "Cross-Validating Quantum Network Simulators," in *IEEE INFOCOM 2025 - IEEE Conference on Computer Communications*, *Workshop on Quantum Networked Applications and Protocols*, to be published, Apr. 2, 2025, DOI: 10.48550/arXiv.2504.01290.
N.B. participated in conceiving the project, developing key ideas of validation scenarios, conducting simulations, and writing the manuscript.

N. Benchasattabuse, M. Hajdušek, and R. Van Meter, "Engineering Challenges in All-Photonic Quantum Repeaters," *IEEE Network*, vol. 39, no. 1, pp. 132–139, Jan. 2025, DOI: 10.1109/MNET.2024.3411802.
N.B. conceived the project, developed key ideas, conducted numerical results, and led the writing of the manuscript.

N. Benchasattabuse, M. Hajdušek, and R. Van Meter, "Architecture and Protocols for All-Photonic Quantum Repeaters," in *2024 IEEE International Conference on Quantum Computing and Engineering (QCE)*, vol. 01, Sep. 2024, pp. 1879–1889, DOI: 10.1109/QCE60285.2024.00217.
N.B. conceived the project, developed key ideas of the architecture and the protocol, conducted simulation, and led the writing of the manuscript.

N. Benchasattabuse, T. Satoh, M. Hajdušek, and R. Van Meter, "Amplitude Amplification for Optimization via Subdivided Phase Oracle," in *2022 IEEE International Conference on Quantum Computing and Engineering (QCE)*, Sep. 2022, pp. 22–30, DOI: 10.1109/QCE53715.2022.00020.
N.B. conceived the project, developed key ideas of the analysis, conducted numerical simulation, and led the writing of the manuscript.

R. Satoh et al., "QuISP: A Quantum Internet Simulation Package," in *2022 IEEE International Conference on Quantum Computing and Engineering (QCE)*, Sep. 2022,

pp. 353–364, DOI: 10.1109/QCE53715.2022.00056.
N.B. participated in developing the simulator, conducting simulations, and writing the manuscript.

R. Van Meter et al., "A Quantum Internet Architecture," in *2022 IEEE International Conference on Quantum Computing and Engineering (QCE)*, Broomfield, CO, USA: IEEE, Sep. 2022, pp. 341–352, DOI: 10.1109/QCE53715.2022.00055.
N.B. participated in developing semantics and models of the protocol, providing concrete implementations of the protocol, and writing the manuscript.

TABLE OF CONTENTS

Abstract . . . iii
Acknowledgements . . . iv
Published Content and Contributions . . . v
Table of Contents . . . vi
List of Illustrations . . . ix
List of Tables . . . xi
Nomenclature . . . xii
Chapter I: Introduction . . . 1
1.1 Motivations . . . 1
1.2 Overview . . . 2
Chapter II: Preliminaries on stabilizer formalism and graph states . . . 9
2.1 Useful mathematical and quantum information concepts . . . 9
2.2 Stabilizer formalism . . . 12
2.3 Graph states . . . 15
Chapter III: Quantum networks . . . 21
3.1 Introduction . . . 21
3.2 Tasks of quantum networks . . . 22
3.3 Quantum link architectures . . . 27
3.4 Quantum repeaters . . . 29
3.5 Error management . . . 32
3.6 Challenges in realizing quantum networks . . . 35
Chapter IV: Architecture and protocols for memory-based quantum repeaters . . . 39
4.1 Design decisions . . . 40
4.2 Defining quantum network services . . . 42
4.3 Expressing connection semantics via RuleSets . . . 44
4.4 Defining quantum network protocols with RuleSets . . . 48
4.5 Networking . . . 57
4.6 Internetworking and recursive architecture . . . 65
4.7 Quantum routing software architecture (QRSA) . . . 69
4.8 Simulation as evidence for architectural design . . . 73
4.9 Discussion . . . 83
Chapter V: Integrating all-photonic repeaters into the architecture . . . 89
5.1 Preliminaries to understanding all-photonic quantum repeaters . . . 90
5.2 Original all-photonic quantum repeaters . . . 93
5.3 Half-RGS building block . . . 98
5.4 Generation of half-RGS . . . 103
5.5 The full RGS protocol . . . 105
5.6 Fidelity derivation of the end-to-end Bell pairs . . . 112
5.7 Purification-enhanced RGS scheme . . . 118
5.8 Performance evaluation . . . 121
5.9 Numerical results . . . 123
5.10 Discussion and future outlook . . . 126

Chapter VI: An application of distributed quantum computation: an optimization algorithm perspective . . . 131
6.1 Introduction: why study a heuristic optimization algorithm? . . . 132
6.2 Background on quantum optimization algorithms . . . 133
6.3 Amplitude amplification with subdivided phase oracle for optimization problems . . . 135
6.4 Performance analysis for common objective distributions . . . 138
6.5 Lower bounds on QAOA runtimes . . . 142
6.6 Lower bounds on QAOA with a Grover-style mixer . . . 144
6.7 Lower bounds applied to search problems . . . 148
6.8 Implications for network architecture and resource management . . . 152
6.9 Conclusion . . . 154
Chapter VII: Discussion and conclusion . . . 155
7.1 Summary of contributions . . . 155
7.2 Revisiting the research questions . . . 156
7.3 Future work and open questions . . . 157
7.4 Perspective . . . 159
Bibliography . . . 161
Appendix A: List of Papers and Presentations . . . 195
A.1 Peer-Reviewed Journals . . . 195
A.2 Conference Papers (Peer-Reviewed) . . . 195
A.3 Posters with Proceedings (Peer-Reviewed) . . . 197
A.4 Posters without Proceedings (Peer-Reviewed) . . . 197
A.5 Manuscripts under Review and In Preparation . . . 197
A.6 Talks . . . 198

LIST OF ILLUSTRATIONS

Number *Page*

2.1 Examples of graph states . . . 16
3.1 Illustrative topology of a quantum network . . . 24
3.2 Heralded link-level entanglement generation . . . 25
3.3 Comparison of primary quantum link architectures . . . 27
3.4 Illustrations of entanglement generation at the link level and inside quantum repeater between memories . . . 30
4.1 Workflow of the RuleSet execution cycle . . . 46
4.2 Resource flow as an effect of a Rule . . . 48
4.3 Two-pass connection setup within a network . . . 49
4.4 A conceptual depiction of RuleSets installed in nodes for a four-hop connection . . . 55
4.5 A conceptual depiction of purification pumping Rules . . . 56
4.6 A schematic of a full quantum router . . . 62
4.7 Recursive architecture illustration of QRNA . . . 66
4.8 Two-pass connection setup with rewrite in a recursive internetwork . . . 67
4.9 A depiction of quantum routing software architecture . . . 70
4.10 SeQUeNCe software framework . . . 75
4.11 Message diagram for link-level entanglement generation in QuISP and SeQUeNCe . . . 78
4.12 Plots showing the fidelities of swapped Bell pairs between output of QuISP and its analytical model varying noise types and strength . . . 80
4.13 Plots showing the fidelities of swapped Bell pairs between output of SeQUeNCe and its analytical model varying noise types and strength . . . 80
4.14 Plots showing time to establish 1000 Bell pairs between QuISP and SeQUeNCe . . . 81
4.15 Plots showing fidelity of two-hop Bell pairs with varying error types and rate between QuISP and SeQUeNCe . . . 82
5.1 Examples of graph states . . . 91
5.2 An illustration of a connection path segment in the RGS scheme network 94
5.3 Overview of the RGS scheme . . . 95
5.4 An illustration of a half-RGS and its transformation to a biclique RGS . . . 99
5.5 A comparison between two end node architectures for RGS schemes . . . 100
5.6 Virtual link abstraction for RGS scheme segments . . . 102

5.7 An illustration of half-RGS generation via deterministic quantum emitters and side effects . . . 104
5.8 A depiction of Pauli frame correction procedure for generation agnostic RGS protocol . . . 107
5.9 An illustration of side effect propagation for a single arm of half-RGS generation . . . 109
5.10 Classical communication bits comparison One-Stage and Two-Stage Pauli frame correction of RGS protocol . . . 111
5.11 An illustration of equivalent side effects on two-qubit graph states . . . 113
5.12 Schematic of two-way entanglement purification and its optimistic variant 119
5.13 A depiction of entanglement purification on half-RGSs and purification circuits utilized in purification of bipartite graph states . . . 120
5.14 An example of flexible purification scheduling enabled by purification-enhanced RGS . . . 124
5.15 Fidelity comparison between different RGS purification strategies given photon depolarizing error probability . . . 124
5.16 End-to-end Bell pair generation rate ratio between RGS purification strategies . . . 125
6.1 2D complex plot of amplitude evolution with an SPO . . . 138
6.2 Amplitude amplification performance for various objective distributions . 139
6.3 Performance sensitivity of the SPO algorithm to the parameter k . . . 141

LIST OF TABLES

Number *Page*

3.1 Generations of quantum repeaters and their error management strategies . 30
4.1 Inter-node messages in the RuleSet-based protocol framework 52
4.2 Action clauses in the RuleSet-based protocol framework 52
4.3 Event triggers in the RuleSet-based protocol framework 53
5.1 The stabilizer of graph states before XX measurement 92

NOMENCLATURE

Approximation ratio. A performance measure used to evaluate an approximation algorithm's average-case behavior. For a given optimization problem, it is expressed as the multiplicative factor comparing the average output of the algorithm against the true optimal solution. A value close to one indicates that the algorithm consistently finds solutions that are, on average, very close to the optimal result, signifying a high-quality approximation.

Entanglement purification. An error management technique that improves the fidelity of noisy entangled pairs by consuming multiple lower-fidelity copies to produce fewer, higher-fidelity ones.

Graph states. Quantum states represented as mathematical graphs where vertices symbolize qubits, and edges represent entanglement relations between qubits. Graph states are a subset of stabilizer states.

Link-level entanglement. An entangled quantum state shared between neighbor nodes in the network through physical quantum link without involving intermediate quantum repeater nodes.

Livelock. A type of resource starvation where changes in local state indicate that progress is being made with respect to local events, while the overall global task is not progressing.

Quantum emitter. A stationary qubit that can be controlled to sequentially emit multiple photons, with each emission coherently extending the entanglement to form a single multipartite state between the emitter and the entire sequence of photons.

Quantum memory. A physical system capable of storing quantum information for a period of time that preserves the coherence of quantum state and any entanglement the state shared with other system.

Quantum network. A collection of interconnected quantum devices that can establish entangled states between distant parties, enabling tasks that are impossible to achieve with classical communication alone.

Quantum network interface card (QNIC). A quantum analogue to a classical Network Interface Card (NIC), the QNIC manages a node's quantum communication links and the communication qubits involved. Its fundamental distinction is its capacity for local quantum processing on the states it handles, effectively transforming into a specialized, limited-capability quantum computer optimized for network protocols. The QNIC is also typically equipped to interact and coordinate operations with additional computation or algorithmic qubits and processing units housed within the larger network node.

Quantum repeater. A device essential for extending the reach of entanglement distribution beyond direct transmission limits, serving as an intermediate node in the network topology and implementing error management strategies.

Repeater graph state (RGS). A specific multipartite photonic graph state that serves as the central resource in a time-reversed memoryless quantum repeater architecture, characterized by a set of inner qubits connected as a complete graph, with each inner qubit also entangled with a corresponding outer qubit for near-deterministic link-level entanglement creation.

Stabilizer formalism. A compact way to describe and efficiently simulate a class of quantum states, known as stabilizer states, which are simultaneously +1 eigenstates of a set of commuting Pauli operators.

Chapter 1

INTRODUCTION

1.1 Motivations

Quantum communication is an indispensable element of quantum information science. It enables the exchange of quantum states and, more importantly, the creation of entangled states shared across distant parties. Such entanglement is essential for scaling up quantum computers, enhancing quantum sensing technologies, strengthening privacy in communications and securing remote or cloud resources, and may ultimately reshape our relationship with the Internet. While significant progress has been made in the physics of entanglement distribution, building quantum networks in practice also requires tight coordination between quantum and classical components, as well as mechanisms to accommodate diverse and evolving technologies. In particular, how the classical side of the network should coordinate, control, and manage resources while remaining flexible enough to support future hardware presents significant engineering challenges.

In this thesis, I will describe our progress in developing architectures, communication protocols, and resource management strategies for entanglement-based quantum networks. By approaching these systems from an engineering perspective, we begin to see which aspects need to be revisited and worked out to move from theory to practice. Drawing on classical network design, emphasizing the role of classical communication exchanges, and considering the realities of engineering practical systems, I have identified and made progress in addressing the following questions.

How do we design a unified network architecture that supports heterogeneous quantum repeaters?

Quantum repeaters are the backbone of long-distance quantum communication [1], yet they come in diverse forms, ranging from memory-based repeaters with quantum memories to emerging all-photonic designs that rely entirely on flying qubits. These differences pose a challenge: how can we design a unified, end-to-end network architecture that supports such fundamentally distinct technologies? Unlike classical networks, where early flexibility gave way to the rigidity of the TCP/IP stack—especially the narrow waist of the IP layer, which has become nearly impossible to change without breaking the Internet—we have the opportunity to avoid locking ourselves into a premature standard. To accommodate heterogeneity and future evolution, we need abstractions and interfaces that support interoperability, modularity, and adaptability across the stack. What architectural principles can enable this? What abstractions should we define to

bridge fundamentally different hardware assumptions? How can we design a quantum network stack that remains flexible as the technology landscape continues to evolve?

How can simulation guide the design and resource management of scalable quantum networks?

Simulation has long been a powerful tool for evaluating complex systems, especially when analytical models fall short or when experimentation is costly or infeasible. In quantum networking, where physical testbeds are still limited in scale and flexibility, simulation becomes a key enabler for exploring network behavior, validating architectural choices, and refining protocol designs. Yet quantum systems introduce unique challenges—such as non-determinism, error correlations, and the need to track entangled states—which make classical simulation nontrivial. A good simulation platform must strike a balance between physical fidelity and computational efficiency, enabling meaningful predictions while remaining usable at large scale. What models and abstractions are necessary to capture the key behaviors of quantum networks? How do we validate simulations against real experiments? What kind of metrics, outputs, or insights should simulations deliver to inform the next generation of quantum network design?

How can all-photonic quantum repeaters become practical and interoperable with memory-based designs?

All-photonic quantum repeaters [2] offer a compelling vision: high-rate entanglement distribution without relying on fragile or expensive quantum memories. While traditionally considered futuristic, advances in deterministic photon sources and photonic entanglement generation have brought them closer to experimental reality. However, engineering challenges remain—particularly when moving beyond idealized settings to realistic regimes that include loss, decoherence, and imperfect hardware. Moreover, most existing designs treat all-photonic repeaters as self-contained chains, often disconnected from broader network architectures based on memory-equipped nodes. To harness their advantages in practice, we need strategies to improve their robustness and integrate them with existing memory-based infrastructures. Can all-photonic repeaters operate effectively in high-error environments? How do we connect them to end nodes with memory-based capabilities, and what protocols or abstractions enable seamless interoperability? What new hybrid architectures become possible when both paradigms work together?

1.2 Overview

This thesis begins by reviewing the mathematical foundations and concepts of quantum state representation and error modeling used in quantum networking (chapter 2),

followed by an overview of quantum networks and the engineering challenges in their realization (chapter 3). The work then presents a unified network architecture for memory-based quantum repeaters, building on the principles of quantum recursive network architecture and demonstrating its effectiveness through simulation (chapter 4). This framework is subsequently extended to support all-photonic quantum repeaters—previously thought to be incompatible with memory-based systems—by detailing their integration into a single heterogeneous network. This extension provides new protocol designs, addresses engineering challenges lacking in prior conceptual models, and examines resource and error management to move all-photonic designs toward practical implementation (chapter 5). With these architectural foundations established, the thesis shifts its focus to applications, using quantum optimization algorithms as a case study to examine how their runtime behavior translates into demands on network resources (chapter 6). The thesis concludes with a discussion of open problems and future directions (chapter 7).

In the following, I provide a chapter-by-chapter summary that details the background, methodology, and main results of this thesis. Each chapter's introduction will provide more in-depth motivation and context for the topics discussed.

Chapter 2: Preliminaries on stabilizer formalism and graph states

This chapter provides essential background knowledge. It will briefly review key concepts from quantum information theory, entanglement and Bell pairs, quantum channels, Clifford groups, and fidelity of quantum states. Furthermore, it will introduce fundamental building blocks relevant to quantum networks, including an overview of stabilizer formalism, graph states, and the basics of quantum error correction codes that underpin advanced repeater designs.

Chapter 3: Quantum networks

This chapter delves into the foundational principles of quantum repeater networks, establishing the state of the art and the core motivations for this work. The promise of a global Quantum Internet is driven by revolutionary applications—from quantum key distribution (QKD) [3, 4] and distributed sensing [5, 6] to private delegated computation [7–9] and large-scale quantum computing—all of which rely on distributing entanglement as their key enabling task. However, the fundamental obstacle of photon loss limits the distance of direct communication, making quantum repeaters indispensable. While repeaters solve the distance problem using core operations like entanglement swapping, this process in turn introduces accumulated operational errors that must be managed. To provide context for the solutions developed in this thesis, this chapter surveys these challenges and provides a classification of different quantum repeater generations (1G, 2G, 3G) [10], outlining their distinct approaches to error management and setting the

stage for the architectural work that follows.

Chapter 4: Architecture and protocols for memory-based quantum repeaters

While experimental progress in quantum networking has been advancing rapidly [11–26], the architectural work required to build scalable, interoperable networks has lagged behind, with most frameworks focusing on single-network scenarios rather than a complete, end-to-end design [27–35]. To address this research gap, this chapter details a comprehensive architecture for memory-based networks. The framework builds upon the foundational principles of the Quantum Recursive Network Architecture (QRNA) by Van Meter, Touch, and Horsman [36], which adapts the recursive design of the classical Internet [37–39] to the quantum domain. The core idea of this recursion is to use layers of abstraction to manage complexity; for instance, an entire network at a lower layer can be represented as a single virtual node at a higher layer. This allows network details to be hidden when crossing administrative boundaries and greatly simplifies protocol reasoning. The chapter details the collaborative work that established these principles in a cohesive framework, presented in two complementary papers at the *2022 IEEE International Conference on Quantum Computing and Engineering (QCE)*.

The first of these papers, led by Rodney Van Meter, introduced the overarching *A Quantum Internet Architecture* [40]. The second, led by Ryosuke Satoh, detailed the software realization of this architecture—the Quantum Routing Software Architecture (QRSA)—and its implementation in the Quantum Internet Simulation Package (QuISP) [41]. Built on the OMNeT++ discrete-event simulator, QuISP implements the principles of QRNA and QRSA using modular, task-oriented software components that differ from a classical OSI-like layered model often utilized in other models of quantum network protocols [27]. My specific contribution within this effort was to concretize the framework by formalizing the RuleSet protocol's inter-node messages and execution logic, as well as improving the quantum state and error models for the QuISP simulator.

To validate this simulation-based approach, the final part of the chapter presents a rigorous cross-validation study of QuISP against another leading simulator, SeQUeNCe [42], developed by the University of Chicago and Argonne National Laboratory. In this collaborative work, led by Joaquin Chung and Michal Hajdušek and published at the *IEEE INFOCOM 2025 - IEEE Conference on Computer Communications, Workshop on Quantum Networked Applications and Protocols*, we show that despite different protocol implementations, both simulators agree on network behavior where analytical models are available [43]. This result establishes the credibility of our architectural model and the simulation tools used to evaluate it.

Chapter 5: Integrating all-photonic repeaters into the architecture

While most repeater designs rely on quantum memories, the distinct paradigm of mem-

oryless, all-photonic quantum repeaters [2, 44] offers a compelling alternative, particularly for applications like QKD where end nodes can process incoming photons without storage. However, a key research issue has been that these schemes have been studied in isolation, with significant engineering challenges and a lack of a clear path for integration with memory-based networks. This chapter presents my novel contributions, detailed in a series of lead-author publications [45–48] with my collaborators Michal Hajdušek and Rodney Van Meter, that aim to make all-photonic repeaters practical and interoperable, thereby realizing the vision of a truly heterogeneous repeater network that is a central theme of this thesis.

The investigation begins with our work published in *IEEE Network* [46], which provides a tutorial on all-photonic repeaters and a systematic analysis of the engineering challenges that are often overlooked. This work identified key issues in timing synchronization, integration with memory-equipped end nodes, and the classical communication overhead required for each Bell pair. A key result was showing that the classical post-processing could be made dramatically more efficient, reducing communication costs by orders of magnitude.

Building on this analysis, the chapter introduces a novel building block, the half-Repeater Graph State (half-RGS), which is constructed from deterministic quantum emitter interfaces [49, 50]. This modular design, published at the *2024 IEEE International Conference on Quantum Computing and Engineering (QCE)* [45], solves nearly all of the challenges raised in our *IEEE Network* paper. It simplifies photonic state generation, resolves the issue of tight global timing, and provides a natural interface to memory-based systems, allowing an all-photonic chain to be abstracted as a single virtual link within the QRNA model. Finally, the chapter details how this building block can be used to solve a decade-long open problem raised in the original all-photonic repeater paper by Azuma, Tamaki, and Lo [2]: how to incorporate entanglement purification. In a paper to be published at the *2025 IEEE International Conference on Quantum Computing and Engineering (QCE)* [48], we provide a proof-of-concept analysis showing that by using optimistic purification [51–53], this primitive creates a purification-enhanced RGS scheme that improves fidelity without sacrificing the high generation rates that make all-photonic repeaters attractive.

Chapter 6: An application of distributed quantum computation: an optimization algorithm perspective

With the architectural framework established, a critical research issue remains: understanding the workloads that a quantum network must support. It is widely believed that distributed quantum computation will be necessary to scale quantum computing beyond a single device [54, 55], but there is a lack of concrete traffic models for such scenar-

ios. Focusing on optimization as a key application area where quantum advantage is anticipated with a few hundred logical qubits [56], this chapter bridges the gap between network architecture and application requirements.

The chapter first details a novel optimization algorithm based on a subdivided phase oracle, a work with my collaborators Takahiko Satoh, Michal Hajdušek, and Rodney Van Meter, published at the *2022 IEEE International Conference on Quantum Computing and Engineering (QCE)* [57]. This algorithm utilizes knowledge of the problem's objective value distribution to modify the standard Grover search, and can be understood as a specific instance of the broader QAOA framework.

The main contribution, however, is a rigorous analysis of QAOA itself. In a work with my collaborators Andreas Bärtschi, Luis Pedro García-Pintos, John Golden, Nathan Lemons, and Stephan Eidenbenz, published in *Physical Review A* [58], we utilize speed limit results from quantum annealing [59] to derive some of the first analytical lower bounds on QAOA's runtime. The key result of this work is showing that the required runtime is not monolithic; it can range from constant-time for easy problems to polynomial or even square-root scaling for unstructured search. Using these findings, I argue about the workload implications for a distributed datacenter model. This chapter discusses how factors like problem partitioning, the choice of fault-tolerant computational model, and trust assumptions fundamentally alter the nature of the network traffic and the type of entangled resources required, providing crucial insights that inform the design of realistic traffic models.

Chapter 7: Discussion and conclusion

Finally, I summarize the primary contributions of this thesis, revisits the main research questions in light of the presented results, and discusses the broader implications of this work for the development of the Quantum Internet. It will also outline promising directions for future research and address remaining open challenges. This structured approach is intended to guide the reader from fundamental principles to the specific contributions of this thesis and their significance for the field.

How to Read This Thesis

This thesis is written with the assumption that the reader is already familiar with the basics of quantum information and quantum computing, including concepts such as manipulating quantum states, quantum circuit notation, mixed states, and state fidelities. Prior extensive experience with quantum networking, classical networking theory, or specific repeater architectures is not a strict requirement, as this work aims to briefly cover the necessary foundational concepts for these areas.

Each chapter, particularly those detailing our contributions, endeavors to begin with

a high-level intuition and motivation before delving into the technical specifics. This approach is intended to allow readers interested primarily in the overarching ideas and outcomes to access them readily, while still providing the necessary depth for those wishing to understand the detailed mechanisms. Readers with significant prior experience in quantum networks, particularly those familiar with repeater generations and basic protocols, may find it productive to start with chapter 4 where our primary architectural contributions begin, after reviewing the preliminaries in chapter 2 and chapter 3 as needed.

Chapter 2

PRELIMINARIES ON STABILIZER FORMALISM AND GRAPH STATES

This thesis is written with the goal of being accessible to a broad audience in STEM, especially those with backgrounds in engineering, computer science, physics, or related fields at the senior undergraduate level. I aim to assume as little prior knowledge as possible, while still moving efficiently through the material. Readers are expected to be comfortable with basic quantum information concepts—such as representing quantum states with state vectors and density matrices, interpreting quantum circuits, and understanding the action of common quantum gates. In terms of mathematics, I assume familiarity with foundational linear algebra concepts like linear independence, orthogonality, bases, eigenvalues, and eigenvectors, as well as a working knowledge of probability theory and set notation. Some basic understanding of graph theory will also be helpful—knowing what a graph is, how adjacency matrices or lists work, and concepts like paths and spanning trees. For any essential ideas that are easy to forget or less commonly used outside of specialized work, I will provide a brief reminder in the relevant sections, so that readers do not need to pause to look them up.

2.1 Useful mathematical and quantum information concepts

In this section, I will give a brief review of some relevant concepts and definitions we will use to derive my results in this thesis.

Entanglement and Bell pairs

Assuming a basic familiarity with writing quantum states in bra-ket notation [60], let us briefly review the concept of entanglement. Entanglement is a property of two or more qubits in which measurements in certain bases will always give correlated results. That is, knowing the results of a measurement on a subset of the qubits can determine the measurement outcomes for the remaining unmeasured qubits. The smallest unit of entanglement, and one that will appear repeatedly in this thesis, is the *Bell pair*—an entangled quantum state of two qubits. And the most common form of Bell pair is the $|\Phi^+\rangle$ which is given by

$$|\Phi^+\rangle = \frac{1}{2}\left(|00\rangle + |11\rangle\right). \tag{2.1}$$

As we can see, measuring any of the two qubits of $|\Phi^+\rangle$ in the Z basis will determine the results of the other qubit.

A two-qubit system has four basis states. While quantum states are typically written

using the computational basis, sometimes it is more useful to use a different, fully entangled basis known as the *Bell basis*. The four basis states of the Bell basis are given by

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}\left(|00\rangle + |11\rangle\right), \tag{2.2}$$

$$|\Phi^-\rangle = \frac{1}{\sqrt{2}}\left(|00\rangle - |11\rangle\right), \tag{2.3}$$

$$|\Psi^+\rangle = \frac{1}{\sqrt{2}}\left(|01\rangle + |10\rangle\right), \tag{2.4}$$

$$|\Psi^-\rangle = \frac{1}{\sqrt{2}}\left(|01\rangle - |10\rangle\right). \tag{2.5}$$

This basis is useful because the four states can be transformed into one another (ignoring the global phase) by applying a single-qubit Pauli gate to just one of the two qubits. For example, the relationships between the states include

$$(Z \otimes I)\,|\Phi^+\rangle = |\Phi^-\rangle, \quad (X \otimes I)\,|\Phi^+\rangle = |\Psi^+\rangle, \quad \text{and} \quad (Y \otimes I)\,|\Phi^+\rangle = i\,|\Psi^-\rangle. \tag{2.6}$$

Quantum channels

When dealing with real-world quantum systems, errors from noise and environmental interactions will inevitably alter the quantum states. A *quantum channel* is the formalism used to characterize and describe these transformations. Because a channel can change a pure state into a statistical mixture of possible outcomes, it is necessary to use the density matrix representation, ρ, to describe the state of the system.

Formally, a quantum channel is a completely positive and trace-preserving (CPTP) map, $\mathcal{E}$. The most common way to represent this map is via the operator-sum, or Kraus, representation, given by

$$\mathcal{E}(\rho) = \sum_k A_k \rho A_k^\dagger, \tag{2.7}$$

where the $\{A_k\}$ are the set Kraus operators for the channel. To ensure that the total probability is conserved, these operators must satisfy the trace-preserving condition $\sum_k A_k^\dagger A_k = I$.

Conceptually, it is often simpler to think of each Kraus operator A_k as representing a possible error that can occur, with the probability of that specific outcome being $\mathrm{Tr}\left(A_k \rho A_k^\dagger\right)$. It is important to note that the Kraus representation for a given channel is not unique, thus the probability of any single event k is not a fixed property of the channel, but is dependent on both the chosen set of operators and the input state ρ.

In this thesis, we will limit our discussion primarily to *Pauli channels*. For this type of channel, the errors can be described as combinations of single-qubit Pauli gates acting

on the state with certain probabilities. An example of a single-qubit Pauli channel acting on a state ρ is

$$\mathcal{E}(\rho) = (1 - p_x - p_y - p_z)\rho + p_x X\rho X + p_y Y\rho Y + p_z Z\rho Z, \tag{2.8}$$

where p_x, p_y, and p_z denote the probability that an X, Y, or Z error occurs. This model can be extended to multi-qubit states; for a general two-qubit state, there are 15 possible Pauli error terms.

Fidelity of quantum states

To quantify how close an experimentally prepared quantum state is to an ideal target state, we use a metric called *fidelity*. It provides a single number, between 0 and 1, that captures how well a state has been prepared or preserved, with a fidelity of 1 indicating a perfect match.

For the comparison between a potentially mixed quantum state ρ and a pure target state $|\psi\rangle$, the fidelity is formally defined as the overlap between the two states. This can be expressed in two equivalent ways:

$$F(\rho, |\psi\rangle) = \langle\psi|\,\rho\,|\psi\rangle = \mathrm{Tr}(\rho\,|\psi\rangle\,\langle\psi|). \tag{2.9}$$

Note that other definitions of fidelity exist in the literature, such as the square root of this value or a more general form for comparing two mixed states [60]. However, for the purposes of this thesis, we will limit ourselves to the definition above.

As the focus of this thesis is the distribution of Bell pairs, it is useful to consider this definition in that context. Given that the four Bell states in eqs. (2.2) to (2.5) form a complete and orthogonal basis, any single-qubit Pauli error acting on a Bell pair simply transforms it into another Bell state. Consequently, the fidelity of a noisy Bell pair with respect to a target state (e.g., $|\Phi^+\rangle$) is equivalent to the probability of measuring it in the Bell basis and obtaining the expected outcome.

Clifford groups and Clifford gates

Assuming that we know the basic quantum gates used in simple quantum programs. Many of the familiar gates, Hadamard (H), not (X), phase (S), and controlled-not (CX/CNOT) gates, are classified as Clifford gates. So what are Clifford gates and what is a Clifford group?

A *group* (in the Clifford group) is a set G with an operation that takes two elements of G and produce a new element inside G. The group operation is a function $G \times G \to G$ that needs to also satisfy certain properties; (1) associativity, (2) identity, and (3) inverse. Some examples of groups with their accompanied operation are integers with

addition, non-zero real numbers with multiplication, and a set of bijective function (or permutation) over n items.

For the purpose of this thesis, we only need to understand the key consequence of Clifford operators forming a group—that is, circuits composed entirely of Clifford gates can often be simplified or "compressed." Specifically, a sequence of Clifford operations applied to the initial state $|0\rangle^{\otimes n}$ can only produce quantum states that belong to a small, finite subset of the vast n-qubit Hilbert space. This special property gives rise to two powerful descriptive tools that are used extensively in this thesis: the stabilizer formalism and the equivalent graph state formalism, which are covered next.

2.2 Stabilizer formalism

A very useful class of quantum states that we will see repeatedly in this thesis is the class of *stabilizer states* [61]. These states possess structure that allows for efficient classical description and simulation under certain operations [62]. Stabilizer states are states that can be prepared by applying a quantum circuit composed entirely of Clifford gates to the all-zero initial state $|0\rangle^{\otimes n}$. The circuits that prepare stabilizer states are called stabilizer circuits. Unlike arbitrary quantum states—which generally require classical resources scaling exponentially with the number of qubits n to store their amplitudes—n-qubit stabilizer states admit an efficient classical representation. Specifically, an n-qubit stabilizer state $|\psi\rangle$ is uniquely defined as the simultaneous +1 eigenstate of a set of n independent, commuting Pauli operators $\{g_1, g_2, \ldots, g_n\}$, known as the generating set or the *stabilizer generators* (or simply the generators).

Formally, the set of all Pauli operators that *stabilize* a given state—that is, operators that map the state back to itself—form what is called the stabilizer group $\mathcal{S}$ of $|\psi\rangle$. This group, $\mathcal{S}$, is an abelian subgroup of the n-qubit Pauli group $\mathcal{P}_n$. For a pure n-qubit stabilizer state, $\mathcal{S}$ contains exactly 2^n elements. But rather than listing all of them, we can describe the group efficiently using its generating set (the generators) given by

$$\mathcal{S} = \langle g_1, g_2, \ldots, g_n \rangle . \tag{2.10}$$

When we want to be explicit about which state the stabilizer group stabilizes, we will use a subscript to denote the state for which the stabilizers stabilize, e.g., $\mathcal{S}_{|\psi\rangle}$.

Note that, in many contexts where strict mathematical precision is not required, people may refer to the generating set simply as the stabilizer set, or simply the stabilizers. Stabilizer formalism plays a crucial part in various areas of quantum information science, including quantum error correction [61], efficient simulation of Clifford circuits [62, 63], and quantum communication and networking—the primary focus of this thesis.

In a nutshell, the stabilizer formalism is a compact way to describe stabilizer states by writing down their stabilizer generators.

Stabilizer states and circuits are utilized (sometimes in conjunction with oracle) in elementary topics of quantum information science such as quantum teleportation [64], the Bernstein-Vazirani [65], Simon [66] and Deutsch-Jozsa [67] algorithms or quantum key distribution [3, 4]. For example, the Bell state and 5-qubit GHZ state,

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), \tag{2.11}$$

$$|\mathrm{GHZ}\rangle_5 = \frac{1}{\sqrt{2}}(|00000\rangle + |11111\rangle), \tag{2.12}$$

can be described with the stabilizer groups defined by

$$\mathcal{S}_{|\Phi^+\rangle} = \langle XX, ZZ \rangle, \tag{2.13}$$

$$\mathcal{S}_{|\mathrm{GHZ}\rangle_5} = \left\langle \begin{matrix} XXXXX, \\ ZZIII, \\ IZZII, \\ IIZZI, \\ IIIZZ \end{matrix} \right\rangle \tag{2.14}$$

respectively. (Note that I omitted the tensor product symbols.) The two stabilizer groups can also be written with a more compact way, dropping the I terms and denoting the index of the qubit that the operator acts on, as

$$\mathcal{S}_{|\Phi^+\rangle} = \langle X_0X_1, Z_0Z_1 \rangle, \tag{2.15}$$

$$\mathcal{S}_{|\mathrm{GHZ}\rangle_5} = \langle X_0X_1X_2X_3X_4, Z_0Z_1, Z_1Z_2, Z_2Z_3, Z_3Z_4 \rangle, \tag{2.16}$$

respectively. This short notation will be mainly used throughout this thesis.

Representing subspaces with the stabilizer formalism

Beyond describing unique quantum states, the true power of the stabilizer formalism lies in its ability to describe *subspaces* of the total Hilbert space. This is the foundation of its use in quantum error correction [61]. If an n-qubit system is described by a stabilizer group with only $n - k$ independent generators, this group does not stabilize a single vector but rather a 2^k-dimensional subspace. This subspace is large enough to encode k logical qubits, providing a framework for storing and protecting quantum information.

The set of all operators that commute with the stabilizer group $\mathcal{S}$ is called its *normalizer*. Operators that are in this normalizer but not in $\mathcal{S}$ itself are known as *logical operators*. For each logical qubit, there is a pair of such logical Pauli operators, typically denoted (X_L, Z_L), that act on the encoded information without disturbing the stabilizing conditions.

A simple but illustrative example is the three-qubit bit-flip code, which uses $n = 3$ physical qubits to encode $k = 1$ logical qubit. The logical basis states are defined as

$|0\rangle_L = |000\rangle$ and $|1\rangle_L = |111\rangle$. The stabilizer group for this code's subspace must stabilize both of these states. Two independent generators for this group are $g_1 = Z_1Z_2$ and $g_2 = Z_2Z_3$. The stabilizer group is $\mathcal{S} = \langle Z_1Z_2, Z_2Z_3\rangle$. The logical operators are operators that transform the logical states; for instance, the logical X operator, $X_L = X_1X_2X_3$, flips $|0\rangle_L \leftrightarrow |1\rangle_L$. Notice that X_L commutes with both g_1 and g_2. This relationship is often summarized using the $[[n, k, d]]$ *notation*, where n is the number of physical qubits, k is the number of logical qubits, and d is the code distance (the weight of the smallest logical operator). The three-qubit bit-flip code is therefore an $[[3, 1, 1]]$ code. This is because while the code can detect any single-qubit X (bit-flip) error, a single-qubit Z (phase-flip) error, such as Z_1, acts as a logical Z operator and is therefore undetectable by the code's stabilizers.

Measurements on stabilizer states

A powerful feature of the stabilizer formalism is that the effect of single-qubit Pauli measurements can be tracked efficiently. The state resulting from such a measurement is also a stabilizer state (or a state within a stabilizer-defined subspace), and we can determine its new set of generators by applying a simple update rule. There are two main cases to consider, depending on whether the measurement operator P commutes or anti-commutes with every generator in the stabilizer group $\mathcal{S}$.

Case 1: P commutes with all generators of $\mathcal{S}$. When the measurement operator commutes with the entire stabilizer group, it belongs to the normalizer of $\mathcal{S}$. This leads to two distinct sub-cases.

Sub-case 1a: Measuring a stabilizer. If P is an element of the stabilizer group itself (up to a phase), the measurement outcome is deterministic. Applying P to the state $|\psi\rangle$ will always yield the eigenvalue associated with that stabilizer (typically +1 by convention). The state after measurement is unchanged, as are its stabilizer generators. For example, measuring the operator X_1X_2 on the Bell state $|\Phi^+\rangle$, which is stabilized by $\langle X_1X_2, Z_1Z_2\rangle$, will always yield the outcome +1, leaving the state as $|\Phi^+\rangle$.

Sub-case 1b: Measuring a logical operator. If the state resides in a code space (defined by $n-k$ generators for $k > 0$) and P is a logical operator (i.e., it commutes with $\mathcal{S}$ but is not an element of $\mathcal{S}$), the measurement corresponds to reading out a logical qubit. In this scenario, the measurement outcome is dependent on the actual state. The measurement projects the logical qubit into one of its basis states and reduces the dimension of the code space by half. The new stabilizer group is formed by adding the measured operator, $\pm P$, to the generating set, increasing the number of generators by one. For example, consider the [[3, 1, 1]] code space stabilized by $\langle Z_1Z_2, Z_2Z_3\rangle$. Measuring the logical Z operator, $P = Z_1$, on the logical state $|+\rangle_L = (|000\rangle + |111\rangle)/\sqrt{2}$ will yield +1 or -1 with equal probability. If the outcome is +1, the state collapses to $|000\rangle$, and the new

stabilizer set becomes $\langle Z_1Z_2, Z_2Z_3, +Z_1\rangle$, which is equivalent to $\langle Z_1, Z_2, Z_3\rangle$.

Case 2: P anti-commutes with some generators of $\mathcal{S}$. In this case, the measurement outcome is always random, with a 50% probability of being +1 and 50% for -1. The measurement projects the state onto an eigenstate of P, so the post-measurement state is different, and its stabilizer group must be updated. The update rule is as follows:

1. Find any generator g_i in the generating set that anti-commutes with the measurement operator P.
2. Replace g_i in the generating set with $\pm P$, where the sign is determined by the classical measurement outcome (+1 or -1).
3. For every other generator g_j ($j \neq i$) that also anti-commutes with P, replace it with the product $g_j \cdot g_i$. This new generator will now commute with P.
4. All generators that already commuted with P are left unchanged.

As an example, let us measure the first qubit of a Bell state $|\Phi^+\rangle$ in the Z-basis, so $P = Z_1$. The initial generators are $g_1 = X_1X_2$ and $g_2 = Z_1Z_2$. The measurement operator Z_1 anti-commutes with g_1 but commutes with g_2. Following the rule, we replace g_1 with $\pm Z_1$ (depending on the outcome). If the outcome is +1, the new generating set is $\langle +Z_1, Z_1Z_2\rangle$. This set can be simplified by multiplying the second generator by the first, giving the equivalent set $\langle Z_1, Z_2\rangle$, which stabilizes the state $|00\rangle$. This ability to efficiently calculate the post-measurement state is what makes the classical simulation of Clifford circuits possible.

2.3 Graph states

Another useful representation of quantum states is the graph state formalism [68, 69]. They are quantum states represented as mathematical graphs $G(V, E)$, where V is the set of vertices and E is the set of edges. In the graph state formalism, a vertex represents a qubit initialized in $|+\rangle$ while an edge represents an entanglement relation between two qubits given by the controlled-phase gate (CZ). Formally, a graph state $|G(V, E)\rangle$ represented by a graph G is given by

$$|G(V,E)\rangle = \prod_{\{u,v\}\in E} CZ(u,v)\,|+\rangle^{\otimes|V|}. \tag{2.17}$$

Interestingly, graph states are a subset of stabilizer states. They can be described via the stabilizer formalism with stabilizer generators given by

$$g_j = X_j \bigotimes_{v\in N_j} Z_v \tag{2.18}$$

associated with each vertex $j \in V$ in the graph G, where N_j denotes the set of neighbors of vertex j. Furthermore, any stabilizer state is equivalent to (possibly non-unique) graph states up to local Clifford gates—that is, applications of single-qubit Clifford gates take the graph state to the stabilizer state. In other words, we can extend the graph state definition to include a single layer of local Clifford operators applied on qubits of the graph state $|G;\underline{C}\rangle$ with

$$|\psi\rangle = |G;\underline{C}\rangle = |G;C_1,C_2,\ldots,C_n\rangle = \bigotimes_{i=1}^{n} C_i\,|G\rangle\,, \tag{2.19}$$

where $C_i \in C_1$ is a single qubit Clifford gate acting on qubit i, with C_1 being the local Clifford group (single qubit Clifford gates). We call the Clifford operator acting on each vertex of the graph the "Clifford side effect" or just "side effect" when the context is clear. These side effects also known as vertex operators in some literature [70]. An example of a stabilizer state with two graph representations is shown in Fig. 2.1.

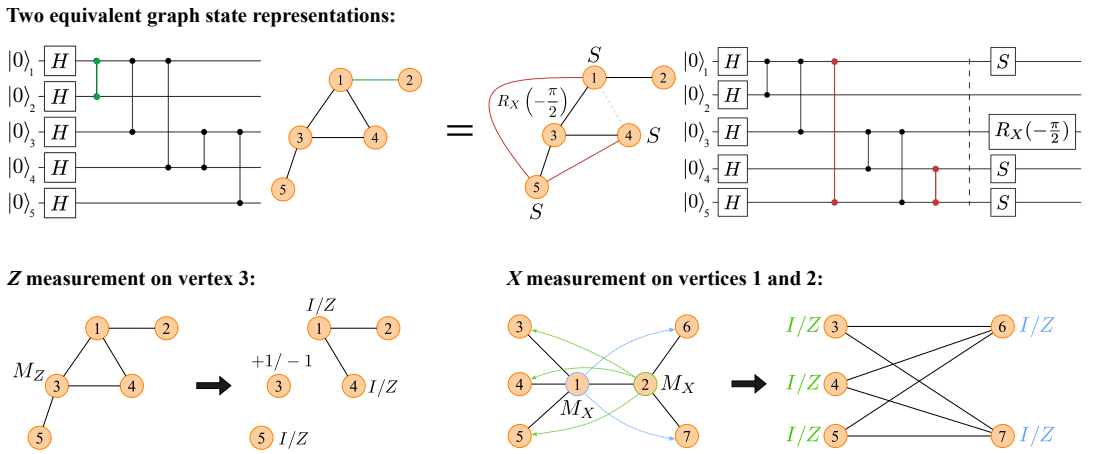


Figure 2.1: Two graph state representations of the same quantum state are depicted at the top. Here, vertices represent qubits in the quantum circuit, and the application of a controlled-phase gate between two qubits corresponds to an edge between the corresponding vertices. An example of such an edge is highlighted in green in the top left. The graph state on the top right is obtained by altering the description through local complementation on vertex 3. This process deletes any existing edges between neighboring vertices of vertex 3, such as the edge $(1,4)$ in this case, and introduces new edges that were previously absent, highlighted in red as $(1,5)$ and $(4,5)$. The application of the depicted Clifford operations ensures that despite the differing graph representations, the quantum states remain identical in both cases. Visualization of a Z measurement with side effects is presented in the bottom left, while the impact of two X measurements (XX measurement) on a different graph is illustrated in the lower right. The Clifford side effects (I/Z) of qubits 3, 4, and 5 depend on the measurement outcome of qubit 2, while those of qubits 6 and 7 hinge on the outcome of qubit 1, as indicated by the blue and green arrows. (From [46].)

In many works that use graph states as a tool for analysis, the focus is on the graph's structure itself, as this is often sufficient for proving key properties of the state. In this

context, it is common to work under the principle of local Clifford (LC) equivalence, where two graph states are considered equivalent if they can be transformed into one another by applying only single-qubit Clifford gates. This effectively means we can disregard the Clifford side effects and work with the simpler state definition in eq. (2.17), treating the side effects as details that can be corrected later. By extension, the related concept of local unitary (LU) equivalence is also used in certain scenarios.

In this thesis, when I talk about graph states, I will be mainly using the extended definition in eq. (2.19) since working on tracking an exact state will be easier to reason about. A few times when we only want to focus on the structure of the graph—when concerning only whether or not the state are still entangled or certain paths between pair of vertices in the graph—I will drop the consideration of side effects and focus on the equivalent properties of the graph state.

Local complementations

As we have seen briefly earlier, a stabilizer state may have multiple graph state representations. One useful operation that can be performed on a graph is the *local complementation* [71]. By definition, local complementation operation $\tau_a(G)$ is an operation on a graph G specified by a vertex a which transform the graph by complementing a subgraph induced by the neighborhood N_a of a—that is, the edge between each pair of neighbors of a is toggled. In simple words, for each pair of neighbors (u, v) of a, if there exists an edge $\{u, v\} \in E$, we remove it, while if u and v are not previously connect by an edge, we add edge $\{u, v\}$ to the edge set. An example of local complementation is shown in Fig. 2.1.

Performing local complementation on graph states also changes the Clifford side effects on qubits not just the graph connectivity, as given by (recall gate definitions in section 2.1)

$$|\tau_a(G)\rangle = \sqrt{-iX} \bigotimes_{v \in N_a} \sqrt{iZ} \, |G\rangle \, . \tag{2.20}$$

And for its action on the graph states with side effects, we get

$$|\tau_a(G); \underline{C}\rangle = \left(\bigotimes_{j \in V} C_j \right) \left(\sqrt{-iX} \bigotimes_{v \in N_a} \sqrt{iZ} \right) |G\rangle \, . \tag{2.21}$$

Notice the order of operations in the above equation, that the Clifford side effects are applied *after* the operations introduced via local complementation. Although we use the term "apply" or "application" with local complementation, in this context (and in most context of quantum information science), it is a means to change the classical state description and no actual physical operations are performed on the state. In this view, local complementation is a tool that we utilize to change our mental description of the state (or the tracked state), and only when we want to use the state after, i.e., then we need to account for these side effects.

Local complementation is a very powerful tool in changing the graph representation of the state and is crucial in optimizing graph state preparations, which has wide range of applications in quantum information science, including quantum computations [72, 73], compilations of quantum programs [74, 75], and quantum networking applications [2, 44, 76, 77].

Measurements on graph states

Performing Pauli basis measurements on graph states preserve the graph states representation. This seems like a trivial assertion, since performing Pauli basis measurements on a stabilizer state gives us another stabilizer state and we have already established that any stabilizer state can be represented with graph states. But conveniently, the action of Pauli measurements on graph states—when considering at the level of local Clifford equivalent—can be explained simply with operations on graphs (vertex deletion and local complementation).

First, let us look at the action of Z basis measurement on graph states without any side effects. The resulting graph after the measurement is the same graph but with the measured vertex removed. (Or if we consider non-destructive projective measurement, we delete all the edges incident to the measured vertex.) To see this, we can rewrite the graph state as

$$|G\rangle = \prod_{\{u,v\}\in E} CZ(u,v)\,|+\rangle^{\otimes|V|} \tag{2.22}$$

$$= \left(\prod_{v\in N_a} CZ(a,v)\right)\left(\prod_{\substack{\{u,v\}\in E\\ u,v\neq a}} CZ(u,v)\right)|+\rangle^{\otimes|V|} \tag{2.23}$$

$$= \left(\prod_{v\in N_a} CZ(a,v)\right)|G-a\rangle \tag{2.24}$$

$$= |0\rangle_a\,|G-a\rangle + |1\rangle_a \prod_{v\in N_a} Z_v\,|G-a\rangle\,. \tag{2.25}$$

We can now see that the effect of measuring a in the Z basis has the same effect as deleting vertex a from the graph while the side effects, depending on the measurement outcome, are possible Z on the neighbor vertices.

We can also derive the resulting state after Z measurement $M_{a,Z}$ on qubit a with stabilizer formalism. To see this, recall the generator set defined from each vertex (eq. (2.18)). Notice that every generator g_i, where $i \neq a$, commutes with the measurement operator $P^a_{Z,\pm}$. Only g_a anticommutes with the measurement, thus, we can replace g_a in the generator set with $\pm Z_a$ where the sign depends on the measurement outcome. For each g_v with $v \in N_a$ in the original graph, we can then update them to remove the Z_a term

by multiplying it with the newly added $\pm Z_a$ term. Notice that the updated generators $g \in \{ \pm Z_a g_v \mid v \in N_a \}$ will have a minus sign in front of them if we had obtain -1 as the measurement result. This is the same as if we define the generator set of the graph $G - a$ with eq. (2.18) but with Z_v acting on g_v for $v \in N_a$ of the original graph.

Now that we have covered Z measurement, let us start with the intuition for Y and X measurements. Recall that we can use local complementation to change the representation of the graph, and more importantly, change the side effects of vertices. The idea is, say we want to perform X measurement on qubit a, to transform the graph state with series of local complementations such that the side effect C_a of qubit a satisfies

$$P^a_{X,\pm} C_a = C'_a P^a_{Z,\pm}, \tag{2.26}$$

where C'_a is a Clifford operator to fix the resulting projection from Z axis back to X axis. A straightforward example of C_a would be H. The Clifford fix C'_a can usually be omitted when we are only interested in the outcome or when we use destructive measurements.

Using the above intuition, we can define the actions of projective measurements in the Z, Y, and X basis as

$$P^a_{Z,\pm} |G\rangle = |Z, \pm\rangle \otimes U^a_{Z,\pm} |G - a\rangle , \tag{2.27}$$
$$P^a_{Y,\pm} |G\rangle = |Y, \pm\rangle \otimes U^a_{Y,\pm} |\tau_a(G) - a\rangle , \tag{2.28}$$
$$P^a_{X,\pm} |G\rangle = |X, \pm\rangle \otimes U^a_{X,\pm} |\tau_b(\tau_a \circ \tau_b(G) - a)\rangle , \tag{2.29}$$

where $P^a_{B,\pm}$ denotes measurement on qubit a with basis B and with eigenvalue result ± 1, b denotes a vertex chosen from neighborhood of a assuming a is not an isolated vertex and $U^a_{B,\pm}$ denotes added side effects introduced by the measurements. These side effects are given by [68, 69]

$$U^a_{Z,+} = I, \qquad U^a_{Z,-} = \bigotimes_{v \in N_a} Z_v, \tag{2.30}$$
$$U^a_{Y,+} = \bigotimes_{v \in N_a} \sqrt{-iZ}_v, \qquad U^a_{Y,-} = \bigotimes_{v \in N_a} \sqrt{+iZ}_v, \tag{2.31}$$
$$U^a_{X,+} = \sqrt{+iY}_b \bigotimes_{v \in N_a \setminus (N_b \cup b)} Z_v, \qquad U^a_{X,-} = \sqrt{-iY}_b \bigotimes_{v \in N_b \setminus (N_a \cup a)} Z_v. \tag{2.32}$$

The choices of side effects to change the basis of intended measurement into Z are chosen by noticing that

$$P^a_{Y,\pm} \sqrt{iX} = \sqrt{-iX} P^a_{Z,\pm}, \tag{2.33}$$
$$P^a_{X,\pm} \sqrt{iX} \sqrt{iZ} = \sqrt{-iZ} \sqrt{-iX} P^a_{Z,\pm}. \tag{2.34}$$

Note that although the action of local complementation in eq. (2.20) suggests that we need to perform local complementation three times to realize $\sqrt{iX}$, we need to perform

it only once, since the resulting states differ only by a global phase. Also, notice that to perform an X measurement, we actually only need two local complementation sequences, but the way we defined it above requires us to perform it three times. The reason for this is most likely due to [68, 69] wanting to constrains the side effects to only rotations in the Y and Z axes.

Chapter 3

QUANTUM NETWORKS

Quantum networking holds the promise to revolutionize the way information is processed and communicated across distributed systems. By harnessing quantum mechanical phenomena such as superposition and entanglement, these networks are envisioned to enable new modes of communication, sensing, and computation that are fundamentally beyond the reach of classical systems. The possibility of interconnecting quantum devices in a similar manner to today's classical Internet has sparked intense interest, with scientists envisioning the transformative capabilities that a future Quantum Internet could unlock. In this chapter, I will give an overview of the motivations behind quantum networking and briefly summarize key foundational concepts in quantum networks.

To achieve this, the chapter is structured as follows. We first define what a quantum network is and the key tasks it must perform. We then delve into the essential hardware and protocol building blocks, including different link architectures and the role of quantum repeaters. Finally, we will survey the critical challenge of error management and conclude with an overview of the major open problems that define the frontier of this exciting field.

3.1 Introduction

Let us begin with the pressing question, "What is a quantum network?" For the purposes of this thesis, a *quantum network* [78–81] refers to a way of connecting quantum devices such that they can accomplish tasks that would be impossible if they were limited to communicating only through classical information.

At its core, a quantum network—or more broadly, the field of *quantum communication*—enables today's communication networks to be supplemented with quantum mechanical phenomena. This includes not only the transmission and reception of quantum information, typically encoded in qubits, but also the ability to establish entangled states between multiple parties across the network.

The most well-known application of quantum communication is quantum key distribution (QKD) [3, 4, 82, 83], which enables two parties in a network to create a shared string of secret classical bits that can later be used as a key for encryption. This task of establishing a shared secret key cannot be achieved through classical communication alone. Today, there are deployed networks that support QKD [84–88] (often referred to as QKD networks), along with commercially available quantum sources and detectors that enable qubit transmission over optical fibers or free-space links. It is important

to note, however, that many QKD implementations operate without requiring entanglement between the communicating parties as certain QKD protocols can be realized by transmitting unentangled quantum states over the channel.

Beyond QKD, however, a rich landscape of applications emerges that fundamentally relies on the distribution of *entanglement*. Each application imposes unique demands on the underlying network architecture. For instance, quantum sensing and metrology promise enhancements such as improved telescope resolution [89, 90], ultra-precise clock synchronization [91–93], vibration detection [94], and physical location verification [95]. These capabilities require distributed entangled states with high fidelity, low timing jitter, and robustness to loss [5, 6, 96–98].

In cryptography and secure communication, entanglement unlocks new functionalities such as quantum secret sharing [99, 100], certified data deletion [101], quantum commitments [102, 103], and multi-party primitives like conference key agreement, leader election, and secure voting [104–109]. These protocols not only require entanglement but often rely on multipartite or high-dimensional entangled states, placing further demands on network capabilities and trust models.

Finally, distributed quantum computing represents a natural culmination of these trends. By connecting quantum processors over a network, one can realize composite systems that surpass the capabilities of any single device [55, 110–112]. Entanglement also enables blind quantum computation [7–9], homomorphic encryption [113, 114], and quantum interactive proofs [115], where a client can securely delegate quantum computations while preserving the privacy of their data and algorithm. These applications demand not only robust entanglement distribution but also reliable error correction, secure authentication, and sophisticated coordination across the network stack.

The diversity of these applications makes it clear that a quantum network must be more than just a channel for sending qubits. It must be a sophisticated system for creating, distributing, and managing entanglement. The following section breaks down the fundamental tasks required to build such a system.

3.2 Tasks of quantum networks

To fulfill the promise of the applications described above, a quantum network must perform a series of fundamental tasks. These tasks form a logical hierarchy, starting from the definition of the network itself, to generating entanglement over a single link, extending that entanglement across the network, and managing the inevitable errors that arise. This section will outline each of these core operational requirements.

Unlike in classical networks, where communication involves sending and receiving messages, directly transmitting quantum information from one node to another is often

not the best approach. Quantum states are inherently fragile, and due to the no-cloning theorem [116–118], if a quantum message is lost or corrupted during transmission, it cannot be recovered. This makes quantum communication particularly susceptible to noise and loss, especially for states that are difficult or costly to prepare.

For this reason, the most fundamental task of a quantum network is to establish generic shared entangled states between distant parties. These shared entangled states serve as fundamental resources for various quantum communication tasks—most notably, quantum teleportation [64], which enables the transmission of quantum data (quantum channel) without the need to physically send the qubits themselves. The smallest such entangled resource that enables quantum communication is the *Bell pair*, which can also serve as a building block for creating more complex multipartite entangled states.

Definition of a network

When I refer to a quantum network in this thesis, I mean a collection of interconnected quantum devices—referred to as network nodes—that are capable of generating, processing, or storing quantum states. At an abstract level, a quantum network can be modeled as a graph where vertices represent network nodes and edges denote physical quantum links between them. Depending on the level of abstraction, some nodes (e.g., those without quantum memories or only used for routing photons through) may be omitted (as depicted in Fig. 3.1). In such cases, the topology of the graph reflects only the essential components for the protocol under discussion.

Throughout this thesis, the network graphs I present generally correspond to a level of abstraction in which only memory-equipped nodes are included with exceptions when we need emphasis on link types and in chapter 5 where I discuss all-photonic quantum repeaters. An edge between two nodes implies that they are capable of generating Bell pairs—i.e., there exists a physical quantum channel between them—and such an edge corresponds to what I will refer to as an elementary link [119–121] that creates *link-level entanglement*.

Having defined the network in terms of nodes and links, we now turn to the first operational task: generating entanglement across one of these elementary links.

Link-level entanglement generation

As is common in introductory treatments of quantum communication [60], foundational protocols such as quantum teleportation typically assume that two parties, Alice and Bob, begin with a pre-shared Bell pair. In a practical network, this entanglement is not pre-existing but must be actively generated. This can be achieved in several ways, for instance, by having one party create a Bell pair between a stationary quantum memory and a flying qubit, which is then sent to the other party through a direct physical channel.

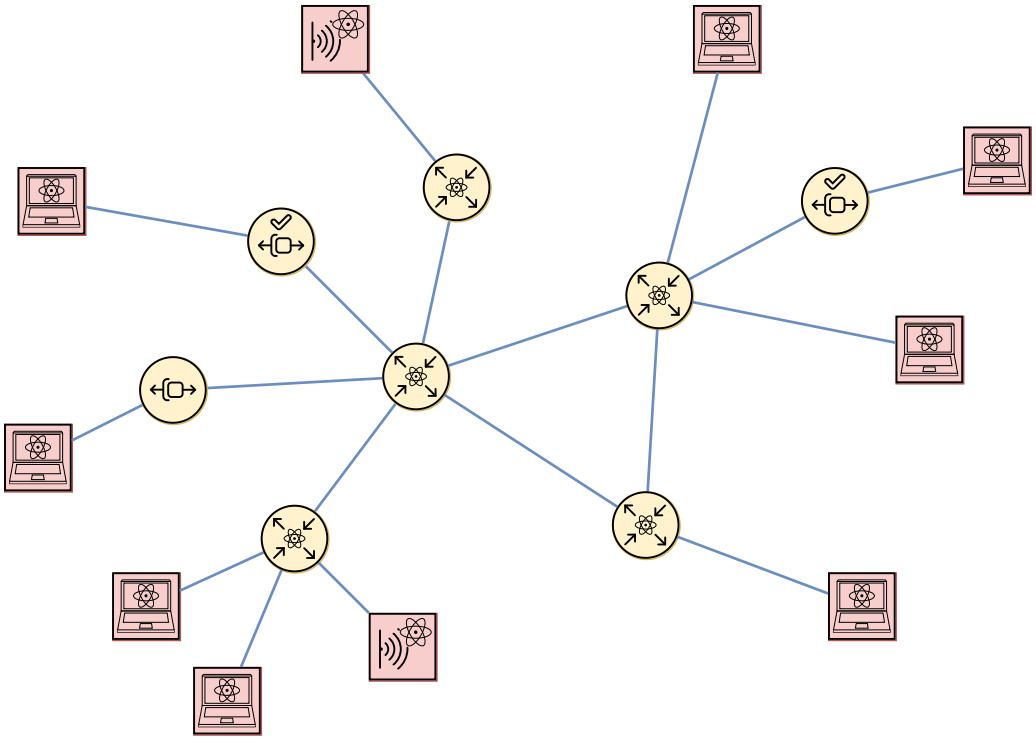


Figure 3.1: An illustrative topology of a quantum network. The network consists of end nodes, depicted as red squares (including both computational and sensor types), and intermediate nodes, shown as yellow circles (representing quantum repeaters and routers). A complete legend and detailed descriptions of each node's function are discussed in chapter 4.

The entanglement established between adjacent nodes through such a process is referred to as *link-level entanglement*.

Formally, link-level entanglement is an entangled state created between nodes that are neighbors in the network graph, meaning they are connected by a direct physical quantum channel such as an optical fiber [119, 120]. One of the most common methods for this is heralded entanglement generation, which uses a central measurement station to create entanglement between two distant memories without the direct transmission of a memory-entangled qubit between them, as illustrated in Fig. 3.2.

While generating entanglement over a single link is a foundational capability, its reach is severely limited by physical constraints.

Entanglement extension

Although photons are excellent carriers of quantum data for transmission due to their weak interaction with the environment, which helps maintain quantum coherence, there is a fundamental limit to how far entanglement can be distributed over optical fiber because of photon loss. Unlike in free-space communication, where photons can travel

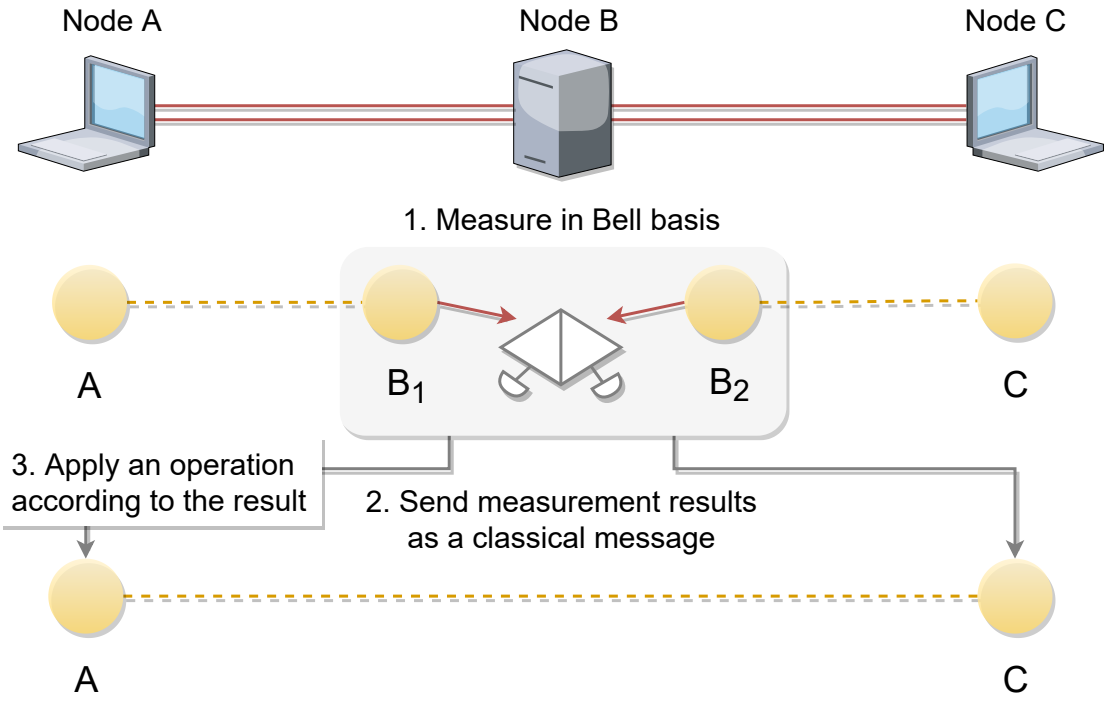


Figure 3.2: A heralded scheme for generating link-level entanglement, characteristic of a Memory-Interference-Memory (MIM) link architecture. Two adjacent nodes, Alice (A) and Charlie (C), each entangle a local quantum memory with a photon. These photons are sent to a Bell-state analyzer (BSA) node (B). A successful measurement at the BSA projects the two distant memories into an entangled state, which is announced, or heralded, by a classical signal. (From [41].)

astronomical distances, the transmission probability in fiber-optic channels decays exponentially with distance. For instance, a typical low-loss optical fiber has an attenuation of about 0.2 dB per kilometer. This translates to a photon arrival probability of roughly 39% at 20 km, 10% at 50 km, and only 1% at 100 km. These figures, which consider only the intrinsic loss of the fiber, already severely limit the distance over which Bell pairs can be reliably distributed. In practice, the total loss is even greater due to system inefficiencies, such as imperfect coupling between fibers and other components.

These limitations are formalized by bounds on point-to-point quantum communication—most notably the PLOB [122] and TGW [123] bounds—which are often referred to as the *repeaterless bound*. They characterize the maximum rate of extractable secret key or entanglement distribution achievable without intermediate quantum nodes, and make clear that direct transmission alone cannot scale to enable a global quantum Internet.

To overcome this exponential attenuation and extend the reach of entanglement distribution, *quantum repeaters* are essential [1, 124, 125]. These devices break up long communication paths into shorter segments, distribute entanglement over each, and then utilize *entanglement swapping* [126] to effectively "stitch together" these segments—allowing entanglement to be extended over much longer distances.

Quantum repeaters are cornerstones of quantum networks and have remained a central focus of research in the field. We will examine them in more detail later in section 3.4.

Successfully extending entanglement across multiple hops, however, introduces a new critical challenge: the accumulation of errors.

Error management

Besides the entanglement distribution rate, an even more critical metric is the fidelity of the distributed states. High-fidelity entangled state is essential for many quantum networking applications that I have previously mentioned, which require strict fidelity thresholds. If the fidelity of link-level entanglement is not high enough, performing entanglement swapping will only further degrade the final state's fidelity due to error accumulation. To address this, we can apply either *entanglement purification* [127–130]—which consumes multiple lower fidelity copies to produce fewer, higher-fidelity ones—or perform *quantum error-correction* [131–133], which encode quantum state into a logical state to protect it from future errors, provided the base fidelity is above a certain threshold.

We will discuss both techniques further in sections 3.4 and 3.5, as error management is one of the core responsibilities of quantum repeaters.

While the quantum tasks of generation, extension, and error management are paramount, they must be orchestrated by a robust classical control layer responsible for overall network operations.

Network operations

On the classical side, the network also needs to remain operational. It must collectively handle application requests from end nodes (users) and make progress toward completing them based on certain objectives—such as minimizing the waiting time of end nodes or prioritizing requests by assigning them ranks. Each node must also monitor its links to ensure they are still operational, while participating in routing, multiplexing, and maintaining operational security.

Most of these tasks are carried out by quantum repeaters. As I will outline later (in section 3.4), the responsibilities of quantum repeaters go beyond simply extending the range of distributed entangled states. In certain network models [134–137], tasks like request scheduling, routing, and multiplexing are delegated to a central controller. We will discuss various network models and architectures later in chapter 4.

In summary, the operation of a quantum network relies on a stack of fundamental tasks, from defining the network topology to managing classical control operations. Having outlined these essential functions, we now turn our attention to the specific physical and

protocol-level implementations that realize them, beginning with a survey of different quantum link architectures.

3.3 Quantum link architectures

Establishing link-level entanglement between adjacent quantum nodes is the first crucial step in any repeater-based quantum network. Various physical link implementations, often termed *quantum link architectures*, have been proposed [119, 120, 138–142]. This section will review a few of these key architectures. They differ in hardware requirements, specific operational procedures, and detailed performance characteristics, such as entanglement generation rate and achievable fidelity. Understanding these differences is crucial for guiding the selection of an appropriate architecture based on the components available at hand.

Memory-interference-memory link

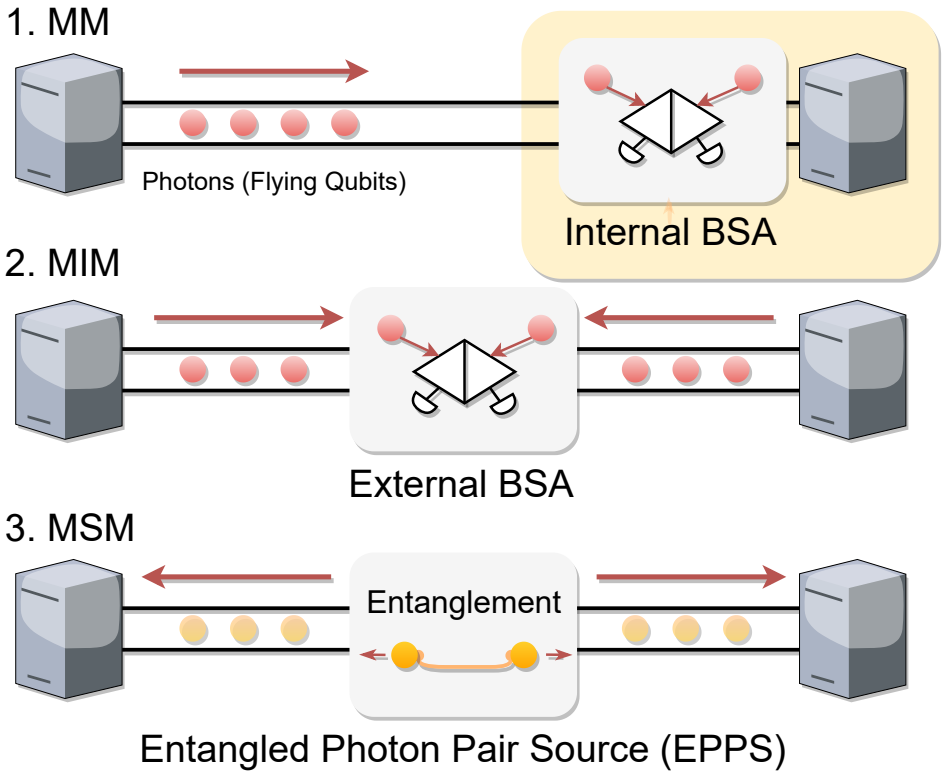


Figure 3.3: A comparison of three primary quantum link architectures for heralded entanglement generation. **(1) Memory-Memory (MM):** An asymmetric architecture where the Bell-state analyzer (BSA) is co-located at the receiver node, which can simplify network topology. **(2) Memory-Interference-Memory (MIM):** A symmetric architecture where two nodes each entangle a local memory with a photon and send the photons to a central BSA for a joint measurement. **(3) Memory-Source-Memory (MSM):** A central entangled photon pair source (EPPS) distributes entangled photons to each node, where they are subsequently stored in quantum memories. (From [41].)

The memory-interference-memory (MIM) link is also referred to as the midpoint inter-

ference link [138–140]. In this scheme, two nodes first generate entanglement between their respective quantum memory and a photon. These photons are then sent through an optical fiber to a central Bell-state analyzer (BSA). The photons meet at the BSA and undergo a Bell-state measurement (BSM). Upon a successful BSM, the entanglement is effectively transferred, creating an entangled Bell pair between the two remote quantum memories. An example of MIM link is shown in Fig. 3.3 A notable drawback is that BSM outcomes are inherently probabilistic. When implemented with linear optics, the BSM can only distinguish two of the four Bell states, which fundamentally limits the theoretical success probability to 50%, even with ideal hardware. Methods to surpass this 50% limit exist, but they require more complex protocols involving more demanding quantum states, specialized hardware, or the use of non-linear interactions, which will be discussed in section 5.10.

Memory-memory link

The memory-memory (MM) link [143] is sometimes referred to as the sender-receiver model. It operates similarly to the MIM link, but with a key architectural change. The BSA is not a standalone station but is instead located inside one of the nodes. This modification simplifies both network communication and its topology. This arrangement allows the receiver node to potentially operate with fewer memories because they can be recycled more quickly for subsequent entanglement generation rounds due to near-instantaneous feedback. This architecture is shown in Fig. 3.3.

Memory-source-memory link

The memory-source-memory (MSM) link, or midpoint source architecture (shown in Fig. 3.3), takes a different approach by starting with a source that directly creates entangled photon pairs [119, 141]. An entangled photon pair source (EPPS) is placed between the two end nodes. The photons generated by the EPPS are sent to the two nodes at the ends of the link. At each node, which must house a BSA, a measurement is performed to entangle the incoming photon with a local quantum memory. This can be achieved by having memories emit their own photons to interfere with the incoming ones or by using absorptive memories to directly store the photons [144]. While requiring two BSAs reduces the theoretical maximum success rate to 25% (compared to MIM's 50%), its key advantage is that memories that fail to entangle can be quickly recycled, significantly reducing their idle time. This can boost the overall generation rate, especially when the number of available memories is limited. However, the protocol requires careful design, as poorly tuned parameters for the EPPS generation rate and memory recycling time can severely degrade or even halt the generation process [145].

Sneakernet link

A drastically different approach is the sneakernet link, which is analogous to how shipping physical hard drives can be faster for transferring massive datasets than sending them over the internet [120]. This architecture proposes establishing a quantum link by physically transporting quantum memories. Entanglement is first created between two memories locally, and then one memory is shipped to a distant location. This is considered a viable option for distributing large volumes of high-fidelity entanglement, particularly given the potentially low entanglement generation rates in early-stage quantum networks.

Vacuum beamguide link

A novel approach for long-distance communication is the vacuum beamguide (VBG) link [142]. Unlike optical fibers, a VBG guides light using an array of precisely aligned lenses housed within an evacuated tube. This design overcomes the fundamental limit of material absorption in fiber by having photons travel primarily through vacuum. The VBG promises ultrahigh transparency and an exceptionally low attenuation rate, potentially on the order of 10^{-4} dB/km. Such low loss could enable ground-based, repeater-less quantum communication over continental scales ($>$ 1000 km). It offers a high-bandwidth, robust channel that is isolated from environmental perturbations. While constructing a VBG requires significant investment and engineering to maintain vacuum and lens alignment—similar in scale to the LIGO project [146]—it represents a feasible pathway to transform and realize quantum links and quantum networks using current technology.

The choice of link architecture is thus a foundational design decision, balancing hardware complexity against performance. We have surveyed protocols that strategically place sources and measurement stations (MIM, MM, MSM) and alternative concepts that circumvent channel loss entirely (Sneakernet, VBG). Each of these methods establishes a single quantum link with its own characteristic rate and fidelity. Having surveyed how these individual links can be established, we now zoom out to consider the pivotal devices that connect these links together: the quantum repeaters.

3.4 Quantum repeaters

As previously discussed, point-to-point entanglement distribution via optical fiber cannot scale to a global quantum Internet. Even within a metropolitan area network, establishing a direct point-to-point connection between every pair of end nodes is impractical. Quantum repeaters [1, 124, 129] help alleviate both of these problems. Their main task is to store link-level entanglement and perform entanglement swapping to create end-to-end entangled states, while also serving as the intermediate nodes that form the network topology. Fig. 3.4 illustrates these two fundamental operations. Part (a)

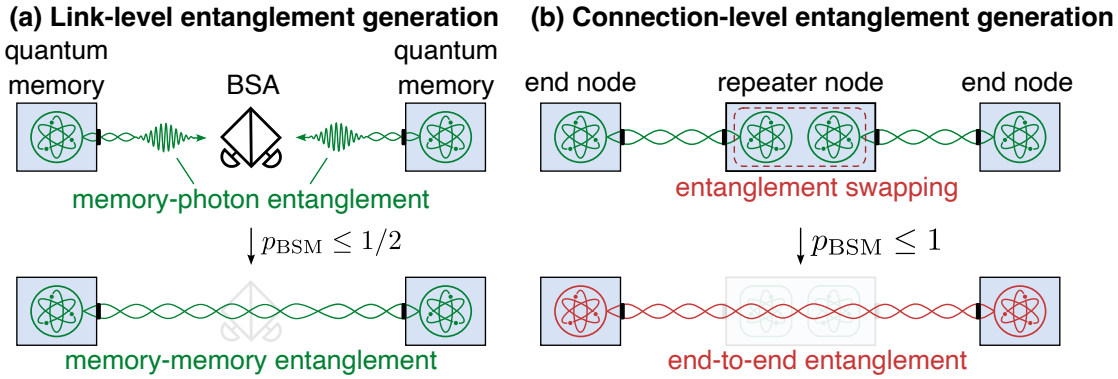


Figure 3.4: (a) Entanglement generation using a photonic *Bell-state measurement* (BSM) at the *Bell-state analyzer* (BSA) in the *memory-interference-memory* (MIM) link architecture. Success probability is denoted by p_{BSM}. (b) After successful generation of two neighboring links, the repeater performs entanglement swapping on its quantum memories to create an end-to-end entangled connection. (From [43].)

shows the first step, where entanglement is generated over an elementary link using a scheme like the memory-interference-memory (MIM) architecture. Once two adjacent links have succeeded, the repeater performs the second task, entanglement swapping, which connects the two links into one longer end-to-end entanglement, as depicted in part (b).

In this thesis, I will make a clear distinction in terminology to precisely define the roles of network nodes. I define a quantum repeater as a node that connects to exactly two neighboring nodes, its primary function being to extend entanglement along a line. I will term a node with more than two connections a *quantum router*, as it is responsible for directing entanglement across different paths in a network. Finally, while other literature sometimes uses the term "quantum switch," I will reserve the term switch for a purely optical device (without quantum memories) that physically routes photons from one link to another without performing measurements.

Classifications of repeaters

Muralidharan et al. [10] categorize quantum repeaters into three generations based on how they manage photon loss and operational errors, as summarized in Table 3.1.

Table 3.1: Generations of quantum repeaters and their error management strategies.

Gen	Photon Loss Management	Error Management
1G	Heralded Entanglement Gen.	Heralded Entanglement Purification
2G	Heralded Entanglement Gen.	Quantum Error Correction (QEC)
3G	Quantum Error Correction (QEC)	Quantum Error Correction (QEC)

First-generation (1G) repeaters manage both error types using heralded, probabilistic mechanisms [124, 129]. Link-level Bell pairs are heralded by successful Bell-state

measurements (BSMs), and operational errors are typically addressed via heralded entanglement purification protocols.

Second-generation (2G) repeaters still rely on heralded generation to manage photon loss but use quantum error-correcting codes (QEC) to handle operational errors [143, 147–153]. This involves encoding physical link-level Bell pairs into logical Bell pairs. 2G networks place stringent demands on the fidelity of local gates and memory coherence times, as QEC codes only suppress errors if the physical error rates are below a certain threshold.

Third-generation (3G) repeaters use QEC to handle both photon loss and operational errors [44, 154–159]. In contrast to 1G and 2G networks, 3G repeaters enable a store-and-forward transmission model for quantum data, akin to packet switching [44, 155, 159]. This requires decoding and re-encoding photonic quantum states at each repeater station, demanding extremely high-fidelity components and potentially denser repeater placement.

Note that this generational framework is not the only classification in the literature. Repeaters can also be classified by their communication model (e.g., entanglement swapping vs. packet forwarding) or by their underlying physical platform [160].

Choosing the right generation

The classification of repeaters into generations was introduced to reflect a progression in technological capability [10, 161]. Higher generations correspond to more demanding hardware requirements but, in return, enable different error handling strategies that can unlock significantly higher entanglement distribution rates.

However, a higher generation number does not inherently imply better performance in all contexts. The choice of generation is a design decision aimed at matching the error management strategy to the characteristics of the available hardware. Each generation operates optimally within a specific regime. For example, if a network has high-fidelity gates and long-lived memories but suffers from high fiber loss, 2G repeaters may outperform 3G designs. Similarly, if the initial link-level entanglement is very noisy, 1G repeaters using purification are likely to achieve the best possible end-to-end fidelity. Thus, the framework is a guide for optimizing network performance with a given technology, not a strict hierarchy of quality.

In recent years, the boundaries between these generations have begun to blur. Advancements in quantum error correction demonstration [162–164] and the realization that real-world systems must handle multiple error types simultaneously suggest that future networks will likely employ hybrid strategies, combining techniques from all three generations.

All-photonic quantum repeaters

Distinct from memory-based repeater paradigms, all-photonic quantum repeaters represent another major approach to extending quantum communication distances. These schemes, most notably the repeater graph state (RGS) based approach [2], aim to circumvent the need for long-lived quantum memories at intermediate nodes altogether. Instead, they rely on the creation and manipulation of large, highly entangled multipartite photonic states that span multiple repeater segments. Entanglement is then established via a sequence of adaptive measurements performed on these photonic states. While offering potential advantages such as high speed and avoidance of memory decoherence, all-photonic repeaters present their own unique set of challenges, particularly concerning the generation of large photonic resource states and the management of photon loss. A detailed exploration of all-photonic repeaters and how our architectural framework can be extended to integrate them is the subject of chapter 5.

Having categorized the different generations of repeaters based on their approach to errors, we now delve deeper into the specific mechanisms they use for this critical function.

3.5 Error management

In any real-world quantum system, imperfections give rise to errors that degrade the fidelity of quantum resources. In the context of quantum networking, we face two primary sources of errors. The first is the transmission channel error, dominated by photon loss, which limits the distance of communication as discussed in section 3.4. The second, and the focus of this section, is operational error. This includes decoherence in quantum memories, imperfect gate operations, and noise from the generation hardware itself. The result is that the initial *base fidelity* of a freshly generated Bell pair may be insufficient for a given application, and as these states are connected via entanglement swapping, errors from each link accumulate, further degrading the fidelity of the final end-to-end state.

To mitigate operational errors, quantum networks employ two complementary techniques: quantum error correction (QEC) and entanglement purification. QEC proactively protects quantum states from future degradation, while purification reactively distills high-fidelity pairs from ensembles of noisy ones. This section will review both of these principal strategies, beginning with the foundational concepts of QEC.

Quantum error correction

Quantum error correction (QEC) is a technique that enables a system to resist a bounded number of errors by encoding the information of a single logical qubit into a redundant state of multiple physical qubits. Errors can then be identified and corrected non-

destructively via syndrome measurements (recall section 2.2).

However, QEC is not a panacea; its effectiveness hinges on the physical error rate being below a certain *error threshold*. If this condition is met, logical errors can be suppressed. Above the threshold, the correction procedure introduces more noise than it removes [133]. This is why QEC is typically associated with higher-generation repeaters, which assume higher-fidelity components.

In the context of quantum repeaters, QEC serves two distinct roles that define the different repeater generations. Second-generation (2G) repeaters use QEC primarily to protect against local operational errors while the link-level entanglement generation is still heralded. The process involves performing remote operations to entangle two logical qubits that reside in separate repeater nodes, assuming the initial link-level Bell pairs used are already of high quality [147]. Third-generation (3G) repeaters, in contrast, extend the use of QEC to combat channel loss directly. The first 3G strategy performs remote operations using codes that tolerate some photon loss, effectively treating a failed link-generation attempt as a correctable erasure error while maintaining a high threshold for operational errors [154, 165, 166]. The second 3G strategy uses loss-tolerant codes to directly encode and transmit flying logical qubits, making the initial link generation itself robust to photon loss [2, 44, 167].

Probabilistic entanglement purification via error detection

When the base link-level fidelity does not satisfy the QEC threshold, a different strategy is required. In the context of first-generation repeaters, where loss is high and base fidelity is low, networks can resort to probabilistic entanglement purification via error detection [127, 128, 130]. This process boosts the fidelity of noisy Bell pairs by consuming multiple copies to produce a smaller number of higher-fidelity ones. This "detect and discard" approach serves as an error detection mechanism, where some pairs are sacrificed to certify the quality of others. Because it is often more efficient to generate a new purified state than to repair a faulty one, this method is well-suited for preparing known target states like Bell pairs.

This method of purification is probabilistic and often referred to as *two-way heralded entanglement purification*. The "two-way" nature refers to the need for classical messages to be exchanged between parties to herald the success or failure of a purification attempt. The main strategy is based on joint stabilizer measurements across multiple copies of Bell pairs [127, 128]. This approach is powerful because its theoretical fidelity threshold is 50%, assuming noiseless quantum operations. Other two-way protocols have also been studied, such as entanglement breeding, which requires ideal Bell pairs as a catalyst [127], or error filtration, which relies on non-destructive measurements [168].

Recent work has extended this stabilizer-based approach by more explicitly using the

structure of quantum error-correcting codes. In these protocols, syndrome extraction is used to detect errors, but instead of applying a correction, a detected error still triggers the probabilistic "discard" step [169, 170]. This approach is notable because it can achieve a non-vanishing distillation rate in the asymptotic limit—a property known as a constant-rate yield—offering potential performance advantages over traditional methods.

Deterministic entanglement purification with QEC codes

In contrast to the probabilistic, two-way protocols, another class of purification is deterministic and requires only one-way classical communication. In these schemes, often called *one-way purification*, there is no probabilistic discarding of resources. The key distinction lies in the classical communication flow: while both parties perform local measurements on their share of the entangled states, only one party sends their measurement results to the other. The receiving party then uses this one-way information to determine and apply the final corrective operations to distill the high-fidelity states.

The foundational protocol for this approach is *entanglement hashing*, which was first proposed for the asymptotic limit of an infinite ensemble of Bell pairs [171]. Hashing can, in principle, distill perfect entanglement with a high rate. However, its reliance on random quantum codes makes the classical decoding process complex, and its intolerance to any operational noise makes the original protocol impractical for real-world hardware with a finite number of resources [172].

Recent approaches improve on this by employing structured quantum error-correcting codes to deterministically map many noisy Bell pairs into fewer, higher-fidelity logical ones. In particular, quantum low-density parity-check (qLDPC) codes [173] have emerged as a promising foundation. These codes support efficient decoding and maintain a constant distillation rate, even with finite resources [174].

Despite these advantages, one-way purification has a key limitation: it demands a higher initial fidelity of the Bell pairs. Unlike two-way protocols, which tolerate base fidelities down to 50%, qLDPC-based one-way purification schemes require infidelities below 5% to yield a net fidelity gain, assuming gate and measurement errors around 0.1%. This positions one-way purification closer to the operating regime of full QEC, making it powerful but currently more demanding in practice.

System-level error management strategies

In practice, while a single technique like two-way purification is sufficient to achieve high fidelity from low-fidelity base resources, this approach is often not optimal across the entire fidelity range. Limited by communication latency, two-way purification protocols tend to be less efficient in terms of throughput at higher fidelities compared to other

approaches. In these regimes, deterministic one-way purification or direct QEC-based schemes often become more resource-efficient and yield a higher rate. This reality motivates a hybrid, multi-stage strategy, creating what can be termed a purification pipeline. In this model, two-way protocols would first be used to elevate low-fidelity pairs to an intermediate fidelity, after which more efficient one-way or QEC protocols could take over to achieve the final target fidelity. The specific architecture of such a pipeline, including the crossover points between different protocols, is highly dependent on the system's noise model and underlying hardware technology.

A more difficult challenge lies in coordinating these techniques across the entire connection path to produce an end-to-end Bell pair. This requires solving a joint optimization problem involving both purification and entanglement swapping schedules, while managing trade-offs among fidelity, rate, and memory decoherence [78, 175–181]. Ultimately, system-level error management must draw on a complementary toolkit. Two-way purification plays an essential role in 1G repeaters, where it distills usable states from highly noisy physical resources. In 2G and 3G repeaters, QEC becomes the central tool for protecting already high-fidelity states from additional errors and enabling fault-tolerant operations. As quantum networks become more practical, hybrid strategies will likely dominate—blurring the boundaries between repeater generations and combining multiple error management techniques to adapt dynamically to the system's needs.

3.6 Challenges in realizing quantum networks

While the basic principles of quantum communication are well established, translating these principles into scalable, reliable, and efficient quantum networks presents a wide array of practical challenges. These obstacles, spanning from physical hardware limitations to complex system-level control, represent the primary bottlenecks to deploying quantum technologies in real-world environments. This section provides an overview of the most pressing engineering and architectural challenges, which collectively motivate the architectural work presented in the subsequent chapters of this thesis.

Physical layer and link-level challenges

At the most fundamental level, a quantum network is built upon the ability to generate high-quality entanglement over physical links. However, this process is probabilistic and fraught with difficulty. For memory-based systems, generating link-level entanglement is often hampered by photon loss in optical fibers and imperfections in sources and detectors. For all-photonic repeaters, the challenge lies in creating large, multipartite graph states with flying qubits. While deterministic generation via quantum emitters is theoretically efficient and has been demonstrated in small scale, current hardware often relies on probabilistic photonic fusion, which can incur a large resource overhead.

Beyond generation, a network's performance depends on having a clear and continuous understanding of its physical components. This requires protocols for *link characterization*—constantly monitoring parameters like fidelity, latency, and generation rate to inform dynamic decision-making. Furthermore, the lack of *standardized network elements* is a key barrier to building modular and interoperable multi-vendor networks. Without agreed-upon standards for physical interfaces, photon wavelengths, or detector characteristics, scaling beyond bespoke laboratory testbeds remains difficult.

The difficulty of overcoming these physical-layer hurdles in experimental hardware, where iteration is slow and costly, underscores the critical need for high-fidelity simulation. Exploring the network-wide impact of link-level noise, loss, and component imperfections is a task perfectly suited for a robust simulation platform, directly motivating the research question on **how simulation can guide the design and validation of scalable quantum networks**.

Network control and resource management

Beyond the physical layer, a quantum network's efficiency is determined by how it manages its scarce and fragile resources, which requires a sophisticated classical control plane. This introduces several interdependent optimization problems. The core challenge is the co-optimization of *scheduling, routing, and multiplexing*. Nodes must efficiently schedule entanglement attempts, select optimal routes that balance fidelity and rate, and fairly multiplex limited resources like quantum memories among competing users.

These decisions depend on having *effective cost metrics for routing* that go beyond classical measures like hop-count to include quantum parameters such as fidelity, noise channel characteristics, and coherence time. Furthermore, because link performance can fluctuate, the network must support *real-time monitoring and adaptation*, allowing protocols to adjust dynamically to changing conditions. The interplay between these tasks is complex; for instance, routing choices for an all-photonic path might prioritize different metrics than for a memory-based path, and the "circuit-switched" nature of the all-photonic approach presents unique multiplexing challenges that are not yet fully understood.

Finally, with multiple users sharing the network, *ensuring fair resource allocation* is a critical socio-technical problem. Mechanisms must be developed to arbitrate access to scarce resources, enforce priorities, and provide equitable service, all while accounting for the probabilistic nature of quantum operations. Developing and testing these sophisticated control strategies in the face of such complexity is almost intractable without powerful modeling tools, highlighting how **simulation becomes indispensable for designing and validating effective resource management protocols**.

System-level and architectural hurdles

At the highest level, significant architectural hurdles must be overcome to ensure the network functions as a coherent, scalable system. A primary obstacle is *managing heterogeneity and ensuring interoperability*. Real-world quantum networks will inevitably be built from a diverse mix of hardware, repeater technologies, and protocol generations. Creating a unifying framework that can abstract away these differences while still leveraging their unique capabilities is a central architectural problem.

This challenge is amplified by the need for *tight timing and synchronization*, as many quantum protocols require sub-nanosecond precision between distant nodes. Furthermore, designing *robust end-to-end protocols* is complicated by the enormous classical communication overhead some schemes introduce. All-photonic repeaters, for instance, can require processing a massive volume of classical data at the end nodes, creating a potential bottleneck on classical processing power. Finally, the question of *end-node participation* reveals a deep architectural disconnect, as it remains unclear how memory-equipped end nodes can best interface with a memory-less all-photonic repeater, or how two networks built from different technologies should be connected.

These disconnections at the system level between different technologies and operational paradigms crystallize the core questions that motivate this thesis. They force us to ask: **How do we design a unified architecture to manage this heterogeneity?** And, as a key part of that, **how can all-photonic quantum repeaters become practical and interoperable with memory-based designs?**

The need for a holistic architecture

These challenges are deeply interconnected and do not exist in isolation. Addressing them requires a holistic architectural vision that bridges hardware diversity, supports dynamic protocol adaptation, and enables performance reasoning at multiple levels of abstraction. Such an architecture must be modular, programmable, and support scalable control, without being tightly coupled to a single hardware platform or repeater generation. In the following chapters, we propose such a solution: a unifying framework built on the Quantum Recursive Network Architecture (QRNA) and programmable RuleSet-based protocols, designed to manage heterogeneous technologies and pave the way for a flexible and scalable Quantum Internet.

Chapter 4

ARCHITECTURE AND PROTOCOLS FOR MEMORY-BASED QUANTUM REPEATERS

Quantum communication is advancing rapidly, with experimental demonstrations ranging from laboratory-scale to metropolitan-scale testbeds [11–26]. While many efforts have explored the design of protocols at various abstraction layers [27–35, 182–184], few have integrated these components into a coherent, end-to-end architectural framework [185–188]. If the history of the classical Internet has taught us anything, it is that hardware and architecture must evolve hand-in-hand; otherwise, the network risks becoming a patchwork of incompatible solutions that hinder scalability. At present, architectural development is lagging behind hardware progress. It is therefore essential to begin establishing architectural foundations now—foundations that can both guide hardware development and adapt to future applications. While multiple network architectures will likely coexist, a true Quantum Internet will ultimately require them to interoperate under a unified framework.

In this chapter, I will outline our attempt at developing a comprehensive architecture for networks utilizing memory-based quantum repeaters. Our framework builds upon and concretizes two powerful, pre-existing conceptual ideas. The first is the architectural principle of the *quantum recursive network architecture (QRNA)*, an adaptation by Van Meter, Touch, and Horsman [36] of the classical Recursive Network Architecture (RNA) developed by Touch [37–39]. QRNA is motivated by the principle that managing small, independent networks is easier than managing a single monolithic one; it uses layers of abstraction to hide internal complexity and simplify protocol reasoning. The second foundational idea is the RuleSet-based protocol, initially proposed by Matsuo et al. [189, 190] to manage errors for link-level entanglement. Our collaborative work, detailed in this chapter, was to extend and concretize these concepts into a complete, end-to-end framework.

Portions of this chapter have been adapted from the following publications:

- R. Van Meter et al., "A Quantum Internet Architecture," in *2022 IEEE International Conference on Quantum Computing and Engineering (QCE)*, Broomfield, CO, USA: IEEE, Sep. 2022, pp. 341–352, ISBN: 978-1-6654-9113-6, DOI: 10.1109/QCE53715.2022.00055.
- R. Satoh et al., "QuISP: A Quantum Internet Simulation Package," in *2022 IEEE International Conference on Quantum Computing and Engineering (QCE)*, Sep. 2022, pp. 353–364, DOI: 10.1109/QCE53715.2022.00056.
- J. Chung et al., "Cross-Validating Quantum Network Simulators," in *IEEE INFOCOM 2025 - IEEE Conference on Computer Communications*, *Workshop on Quantum Networked Applications and Protocols*, to be published, Apr. 2, 2025, DOI: 10.48550/arXiv.2504.01290. (Accepted by The First Workshop on Quantum Networked Applications and Protocols QuNAP 2025)

These concepts are brought to operational life by the Quantum Routing Software Architecture (QRSA), a modular framework designed to manage connection establishment, execute RuleSet-driven operations, handle routing decisions, and control the underlying quantum hardware. A key part of this effort was the development of the Quantum Internet Simulation Package (QuISP), which implements these architectural principles. My specific contribution within this collaborative effort was to mature the simulator by completing its error models and improving its performance, and to concretize the architecture by formalizing the RuleSet protocol for end-to-end connections.

The chapter is structured as follows. It begins with an overview of the fundamental design decisions inherent in building any quantum network (section 4.1). Subsequent sections (sections 4.2, 4.3, 4.5 and 4.6) then systematically present the concrete protocols and mechanisms our framework provides. The specifics of the Quantum Routing Software Architecture (QRSA) are detailed in section 4.7. To validate our simulation-based approach, the chapter then presents a rigorous cross-validation study of QuISP against SeQUeNCe [42], another leading simulator, in section 4.8. This result establishes the credibility of our architectural model and the simulation tools used to evaluate it. Finally, the architecture is considered as a whole, and we discuss how other proposals in the literature can be viewed within the same framework (section 4.9).

4.1 Design decisions

Building a coherent and scalable architecture for the Quantum Internet requires addressing several foundational design questions, many of which share similarities to the challenges faced during the development of the classical Internet. The choices made at this stage are critical, as they profoundly influence the network's core functionality, its capacity to support diverse applications, the efficiency of resource utilization, the complexity of network management, and ultimately, its overall robustness and scalability. Establishing these architectural underpinnings is essential not only for guiding near-term hardware development but also for ensuring the network can adapt to future technological advancements and operational insights.

Navigating the design space involves making deliberate choices across several key areas. These decisions, which we will explore in detail throughout this chapter within the context of our proposed architecture, are briefly introduced below.

The fundamental network service: The fundamental services provided by the network determine how useful the network is to higher-layer applications. What is the primary quantum resource provided to users? Should the network focus solely on distributing Bell pairs, delegating the task of building more complex states to applications? Or should it offer richer services like the direct generation of multipartite states (e.g., graph states or W-states) or even qubit teleportation? This choice impacts protocol complexity

and the division of complex tasks between the network and user space. (Section 4.2)

The nature of connections: How should end-to-end entanglement be established and maintained? Should paths be predetermined and resources reserved (akin to circuit-switching), which better suits repeater schemes requiring intricate scheduling of swapping and purification (like 1G/2G repeaters or certain all-photonic designs)? Or should entanglement be extended hop-by-hop (like packet-switching), offering flexibility but potentially sacrificing guarantees, fidelity, and efficiency? Furthermore, we must define the stateful nature of these connections, including the signaling mechanisms (handshakes) for setup and teardown, and whether control is centralized or distributed. (Section 4.3)

Application interface: How do end-user applications interact with the network? Defining a clear Application Programming Interface (API)—a quantum equivalent of the classical socket—is crucial for abstracting the underlying network complexity and enabling application development by separating appliation logic from the complexity of network resource management and the distributed quantum operations required to generate link-level Bell pairs, perform purifications, or entanglement swapping. (Section 4.2)

Conveying user requests: How do applications communicate their needs (e.g., target fidelity, rate, state type, number of pairs) to the network? This involves designing request protocols and establishing clear naming and tracking conventions for quantum resources, especially when multiplexing resources among multiple concurrent applications. (Section 4.3)

Network node roles: What distinct functional roles should different network components play? Defining clear roles for end nodes, various generations and classes of quantum repeaters and quantum routers (memory-based and all-photonic), and other devices is essential for modularity and interoperability, separating technological capabilities from architectural function. (Section 4.5)

Routing strategy: How are paths selected through the network to connect end nodes? While seemingly analogous to classical routing, quantum routing algorithms must also account for unique challenges like qubit decoherence, probabilistic operations, biased quantum error channels, and the tight coordination between quantum operations and classical signaling. (Section 4.5)

Resource multiplexing: How are shared quantum resources—memories, channels, entangled states—allocated among competing user requests? Various strategies exist, from reserving resources (circuit-switching) to multiplexing with dynamic allocation based on time, buffer space, or statistical priority, each with implications for efficiency, fairness, and Quality of Service (QoS). Mechanisms for authentication, authorization, and accounting (AAA) are also intrinsically linked to resource allocation. (Section 4.5)

Security considerations: Beyond the inherent security features like QKD, quantum networks introduce novel vulnerabilities [191–193]. We must consider threats ranging from physical side-channel attacks on devices to architectural vulnerabilities like denial-of-service on quantum resources or impersonation via classical control channels, particularly in the context of distributed quantum protocols. (Section 4.5)

Internetworking: How can independently operated quantum networks, potentially based on different technologies or protocols, connect and interoperate to form a larger, global Quantum Internet? This requires standardized protocols, cross-domain addressing schemes, and mechanisms for establishing trust and translating services between networks. (Section 4.6)

While this overview of design decisions is not exhaustive, addressing these fundamental questions provides a framework for constructing a robust and adaptable quantum network architecture. We will delve into these areas in the following sections, presenting our approach based on the RuleSet-based protocols and the Quantum Recursive Network Architecture (QRNA) to provide concrete solutions and mechanisms addressing these design decisions.

4.2 Defining quantum network services

Entanglement is the resource that fuels quantum applications such as QKD, teleportation, quantum sensing, delegated quantum computation, or distributed quantum computation. Continuous, reliable and efficient replenishment of this resource is one of the primary tasks of a quantum network. However, entangled states come in many shapes and sizes [194–199].

This section delves into the first fundamental design decision: defining the services a quantum network offers to its users. We first explore the challenges and trade-offs inherent in this choice and then describe the approach taken within our proposed architectural framework.

Fundamental network service and semantics

The design of a quantum network must begin with a clear definition of its fundamental services—what quantum states or capabilities the network is expected to provide to end users. These decisions determine the complexity of the protocols at the network layer and the applications that run above it. A minimalist design may treat *Bell pairs* as the primary network-level service. Bell pairs serve as the smallest unit of entanglement and the foundation for nearly all quantum communication protocols. Restricting the service to Bell pair distribution simplifies the network's responsibilities. However, this approach shifts complexity to the applications, which must construct multipartite or fault-tolerant states themselves and manage the coordination overhead that entails.

At the other end of the spectrum, networks may offer richer services such as multipartite entangled states (e.g., GHZ, W-states [200], graph states [68, 69]), or fault-tolerant state teleportation. While applications can, in theory, synthesize these states from Bell pairs, direct network-level support may offer efficiency gains and reduce sensitivity to noise by internalizing complex procedures like direct graph state generations or supporting the delivery of error-correcting code encoded logical qubits.

In our architecture, we adopt Bell pair distribution as the core network service, as it allows for a well-scoped, foundational architectural framework. As we will see later in section 4.9, our architecture is also extensible to multipartite entanglement services. These extensions build naturally on the base protocols discussed here and suggest a path for future network capabilities.

Importantly, the semantics of Bell pair distribution are not merely those of passive delivery. Even in this basic model, distributed quantum computation occurs along the path via entanglement swapping, possibly combined with purification at intermediate repeater nodes. A proper service definition must account for this processing, as it directly affects fidelity, latency, and trust assumptions in the network.

In addition to quantum state delivery, timing information is often a critical part of the service. Applications in distributed quantum sensing [5, 6, 89, 96–98] and clock synchronization [201] require precise knowledge of when entanglement was established or when measurement events occurred. Hence, high-precision timestamps may need to be bundled into the service interface offered by the network.

In summary, while this work focuses on Bell pair distribution as the baseline network service, it should be emphasized that its semantics involve nontrivial quantum processing along the route and may include auxiliary data like timestamps. This careful service definition sets the stage for scalable and extensible quantum network architectures.

Quantum sockets

With Bell pair distribution defined as the fundamental service of the quantum network, the next question is how applications at the end nodes interact with this service. This is the role of the Quantum Socket API [202]—an abstraction layer that mediates between quantum applications and the underlying network infrastructure, much like how classical socket APIs decouple application logic from network protocols in the Internet.

The Quantum Socket provides a uniform interface for creating entanglement-based connections between nodes. Applications can use it to request Bell pairs without needing to manage the probabilistic behavior of physical-layer protocols or the complexity of entanglement swapping and purification. If the network supports richer services beyond Bell pair delivery, such as teleporting quantum states or accessing shared randomness,

these could also be provided via the socket API.

To accommodate a range of use cases, the API is designed to be node-type agnostic, supporting a variety of application endpoints. We can categorize applications into three types [203]: (i) extracting correlated classical data (for QKD and sensor applications), (ii) distributed Clifford operations where correction operations can be performed posthoc, and (iii) distributed non-Clifford operation where states need full confirmation of success and correction before being consumed. For extracting correlated classical results, interaction with the network is largely synchronous from the application's point of view, as classical results are obtained directly after measurement. These applications are tolerant of the stochastic timing of Bell pair delivery and can selectively discard measurements performed on unentangled states without affecting their correctness. In contrast, computational applications (the second and third types) typically require asynchronous interfaces to coordinate with their local quantum processes. This coordination involves complex decision-making, such as whether to wait for a resource or proceed, and how to handle errors—either by propagating corrections forward in time or by reversing the computation to a state prior to the consumption of the faulty resource. Developing robust programming models that handle this variability, including callbacks or deferred execution, remains an open research problem and is beyond the scope of this thesis.

By exposing a consistent API to applications regardless of application and node types, the Quantum Socket interface promotes modularity and simplifies development. It defines the boundary between application logic and the complex distributed operations that the network must perform behind the scenes. This thesis focuses on the design and operation of the quantum network *up to this boundary*—the delivery of Bell pairs or provision of core entanglement-based services—with the understanding that applications themselves are responsible for the subsequent, higher-level utilization of these resources.

4.3 Expressing connection semantics via RuleSets

With the fundamental network service defined, we now turn to the question of how end-to-end entangled connections are established and managed. This process involves a series of coordinated quantum operations across multiple, physically separated nodes—a form of *distributed computation*. Each node has access only to its local quantum resources and classical control, yet must participate in a global task to generate and maintain shared entangled states.

To address the complex problems of this distributed computation, this section details our adoption of a connection-oriented model and introduces the RuleSet protocol, built from the work in [189, 190], as the core mechanism for orchestrating the distributed operations required.

A RuleSet-driven connection-oriented model

To support flexible scheduling of complex operations and enable non-local coordination, such as entanglement purification extending beyond directly connected neighbors, our architecture employs a *connection-oriented* model. In this paradigm, an end-to-end path is established first, with all intermediate nodes along this path explicitly agreeing to participate in the protocols for distributing Bell pairs for that specific connection. This model is chosen primarily for the enhanced reliability and more predictable state fidelity it offers for the resulting end-to-end Bell pairs, particularly when contrasted with connectionless approaches which typically restrict purification to adjacent nodes and may lack end-to-end fidelity guarantees. (A more detailed comparison between connection-oriented and connectionless models is provided in section 4.9.)

To realize this connection-oriented model while maximizing node autonomy and minimizing the latency associated with synchronous control or excessive classical communication round-trips, we adopt an event-driven operational scheme orchestrated by *RuleSets* [189, 190]. Upon the establishment of a connection, each participating node is provisioned with a specific RuleSet. This RuleSet dictates the node's behavior in response to local events—such as the successful generation of link-level entanglement, the arrival of a classical message, or a timeout. This RuleSet-driven approach is central to our architecture as it enables largely decentralized control and asynchronous operation; nodes react dynamically based on their local context and pre-defined rules rather than waiting for explicit step-by-step instructions or relying on global lockstep coordination.

RuleSet protocol

A RuleSet consists of a collection of Rules where each Rule contains a *condition clause* and an *action clause*. When a Rule's condition is satisfied by the local state or event context, its corresponding action is executed. A collection of such RuleSets, one at each participating node, collectively constitutes the full distributed program. A key distinction from classical analogies like Software Defined Networking (SDN) OpenFlow [204] or P4 [205] is that our RuleSet design explicitly incorporates the management of complex, probabilistic local state, including quantum states, which is indispensable for quantum repeater functionality. This management of local state can be cast as the protocol's "side effects". A conceptual depiction of a RuleSet is shown in Fig. 4.1.

Once a quantum resource, such as a link-level Bell pair, is assigned to a specific RuleSet for a connection, that assignment is typically exclusive and persists until the resource is consumed or released by that RuleSet. The initial assignment of resources to RuleSets falls under the purview of a separate multiplexing scheme (will be discussed in section 4.5). Consequently, RuleSets, along with any qubits at a node that are currently assigned to a particular connection, constitute vital *connection state* that must be main-

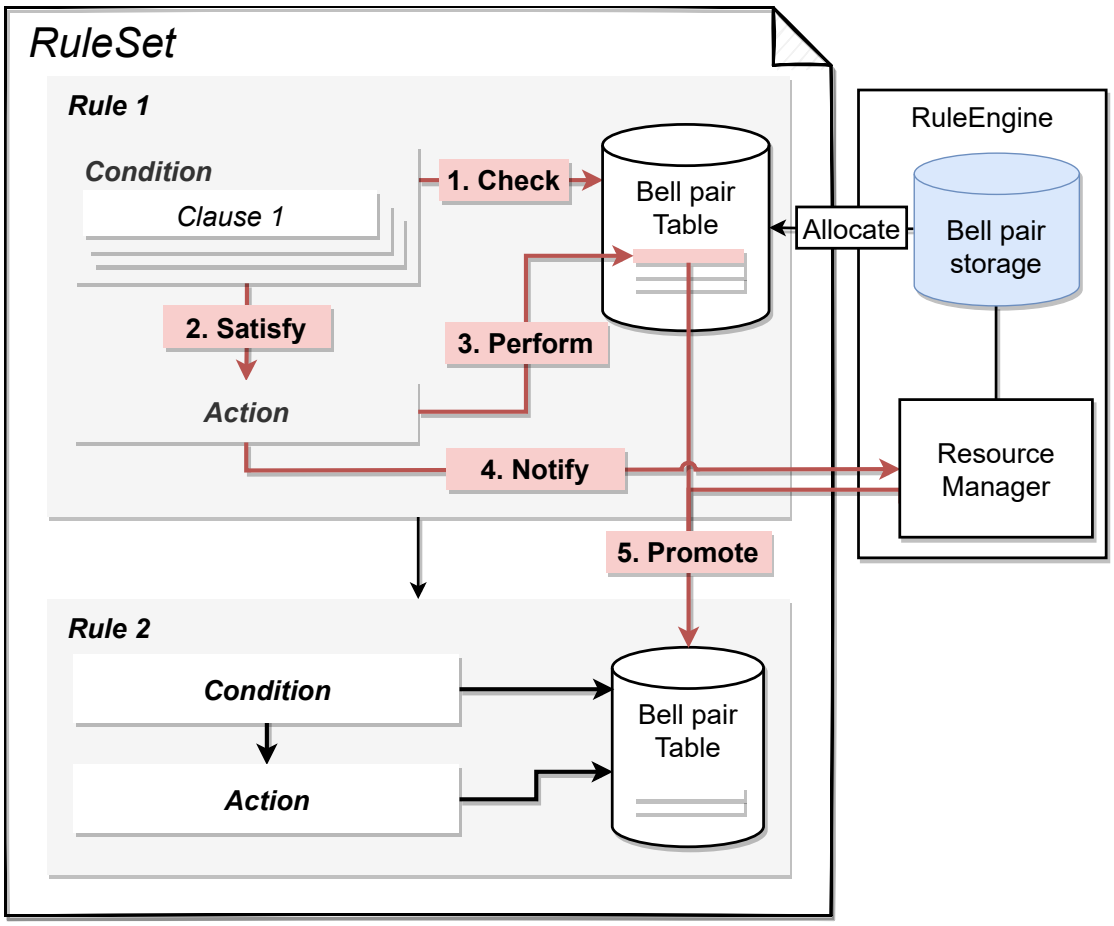


Figure 4.1: Workflow of the RuleSet execution cycle. The process unfolds in a deterministic sequence: (1) The Rule Engine begins an evaluation phase, checking the Condition clauses of its Rules sequentially. (2) If a Rule's conditions are fully satisfied, the process advances to the next step; otherwise, the engine waits for a new event (e.g., resource allocation) to trigger a new evaluation cycle. (3) The selected Rule executes its Action clause. (4) Upon completion, the Action notifies the Rule Engine of the outcome. (5) Finally, the Rule Engine performs a state update, promoting the resource to its next stage in the protocol. (From [41].)

tained at each repeater or router along the path. While the scalability implications of such statefulness require careful assessment (see section 4.5 for security considerations), maintaining some form of state at intermediate nodes appears unavoidable for most approaches to robust quantum networking.

A fundamental principle of this model is that there is one RuleSet dedicated to each connection traversing a node. Once a quantum resource, such as a link-level Bell pair, is assigned to a RuleSet, it is also tagged with a label (called a *tag*) and given a *sequence number* to track its role in the protocol. Tags are used to define resource pools, and the movement of resources from one tag to another encodes the logical progression of the protocol. The tag ordering is unidirectional and intentional: it serves as a scaffold for reasoning about RuleSet behavior and ensures termination of the

protocol. The sequence number, meanwhile, helps maintain resource ordering within a tag and supports coordination across nodes without additional messaging. This prevents mismatches, such as two nodes selecting different Bell pairs for the same operation, a problem known as leapfrogging [78].

Thus, a RuleSet acts as a self-contained program governing the lifecycle of entangled states for a connection. Resources flow from one tag (pool) to another as operations proceed. When a Rule consumes or updates a resource, it either moves it to a new tag (signaling advancement in the protocol), or marks it as defunct (releasing physical qubits for link-level generation reuse). This unidirectional tag flow ensures that every quantum state associated with a RuleSet either progresses toward successful end-to-end delivery or is discarded without ambiguity. Note that this model does not eliminate the possibility of livelocking, where Bell pairs are repeatedly allocated and appear to progress through tag transitions, but are eventually all discarded without reaching the final tag or being delivered to the application. An example of resource flow within a RuleSet is depicted in Fig. 4.2. Because each connection is managed by its own RuleSet and associated quantum resources, nodes must maintain a small but critical amount of connection state, including active RuleSets and their associated qubits. While this statefulness has implications for scalability (discussed in section 4.6), it appears unavoidable for robust and flexible quantum networking, particularly when enabling asynchronous and autonomous repeater operations.

Two-pass connection setup

To establish a connection in our architecture, we employ a two-pass protocol for establishing connections [206]. This process separates resource discovery and path selection from the distribution and activation of the operational rules. The first pass, or outbound leg, is initiated by the node requesting the connection (the *Initiator*). As the connection request message traverses the network towards the intended *Responder*, guided by underlying routing algorithms, it gathers information about the links along paths, such as available quantum resources, node capabilities, and estimated link fidelities.

Upon reaching the Responder, all the collected path and resource information is processed. The Responder then takes on the crucial role of devising the specific RuleSets required for every intermediate node (and the end nodes) along the chosen path to collectively establish (distributed computation along the path) the end-to-end entangled state. This centralized generation of RuleSets by the Responder allows for sophisticated optimization and coherent planning of the entire distributed protocol. Following this, the second pass, or return leg, involves the distribution of these tailored RuleSets from the Responder back to each node along the selected path. Once all nodes have received and installed their respective RuleSets, the distributed operations for the connection can

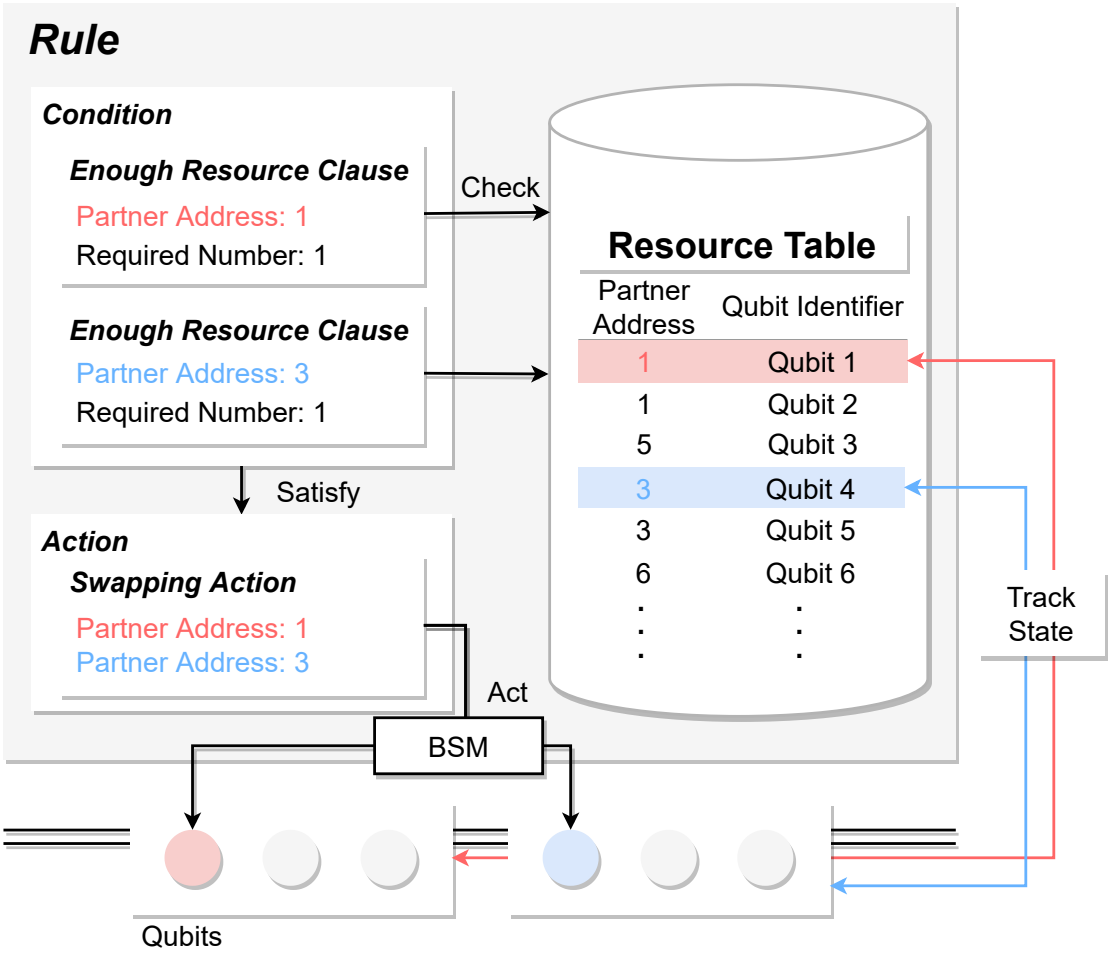


Figure 4.2: An illustration of resource flow as a result of a Rule's execution. The diagram shows how a resource is first evaluated against a Rule's Condition clause. Upon a successful match, the corresponding Action clause is executed, causing the resource to transition to a new state or tag.

commence, driven by local events and the logic outlined in the rules.

This two-pass connection setup is illustrated for a single network in Fig. 4.3. As in the classical Internet, we expect that most connections will be initiated by a client node reaching out to a server. Placing the responsibility for RuleSet creation with the Responder (often the server) provides a natural point for service providers to innovate. By developing more efficient or robust RuleSets, providers can offer superior connection performance (e.g., faster setup, higher fidelity, better resource utilization), creating a basis for competitive differentiation.

4.4 Defining quantum network protocols with RuleSets

Now that we have a conceptual understanding of what a RuleSet is, this section delves into a detailed specification of the RuleSet protocol itself. We will dissect the core components that constitute individual Rules, examine the mechanisms for quantum state representation and classical inter-node communication that RuleSets rely upon, and

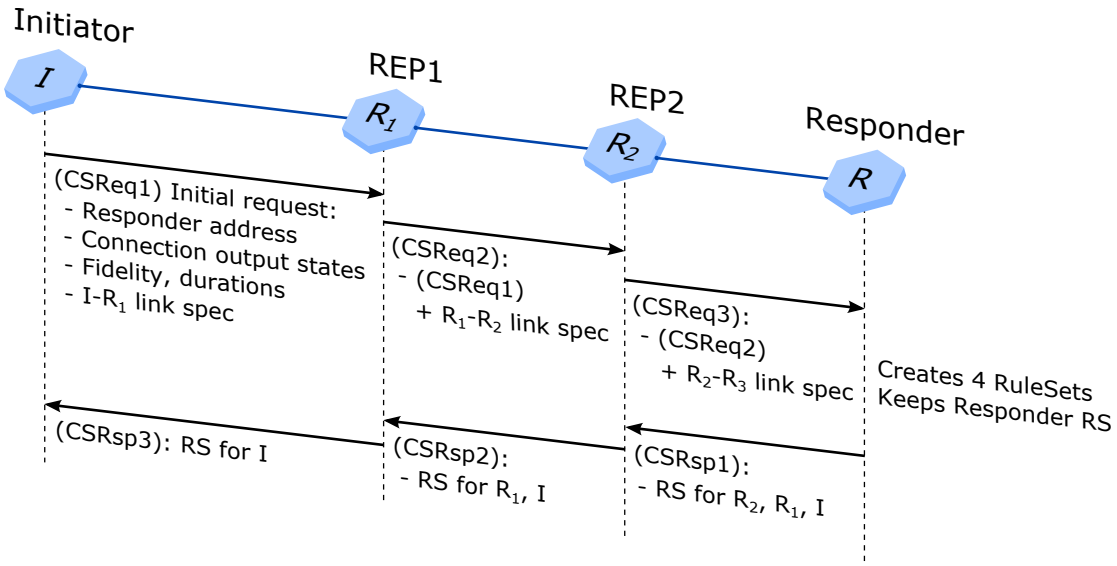


Figure 4.3: Two-pass connection setup within a network along a path. (From [40].)

illustrate their operational semantics through practical examples. This will provide a comprehensive understanding of how RuleSets function as a programmable framework for distributed quantum operations.

The focus of this specification is on the *logical structure and operational semantics* of the RuleSet protocol. This includes detailing the essential information content and parameters that messages must convey between nodes for the protocol to function correctly. However, we will not prescribe specific low-level implementation choices such as particular message serialization formats (e.g., JSON, protocol buffers, or custom binary representations), precise bit-level encodings, or the underlying classical network transport mechanisms (e.g., TCP/IP or other datagram services) used to carry these messages. For the purposes of defining the RuleSet protocol here, we assume the existence of reliable classical communication channels capable of delivering these logically defined messages between participating quantum network nodes.

Fundamental constructs: Rules and RuleSets

At the core of the RuleSet framework are individual Rules, which define quantum operations, specify how resources are used, and state the conditions under which those operations are triggered. A RuleSet is simply an ordered list of such Rules, constructed programmatically and associated with a unique RuleSet ID (or a connection ID). Each Rule within the RuleSet also receives a Rule ID, allowing it to be referenced for inter-node communication.

The RuleSet life cycle starts and ends with the connection, thus a single connection will have a single RuleSet.

State naming and management in rulesets

Managing and agreeing upon the specific qubits for joint operations like entanglement purification or swapping among collaborating nodes are critical responsibilities handled by RuleSets. While we could draw inspiration from the IP architecture by employing network-wide unique naming for qubits, RuleSets have local execution contexts, making global naming for individual qubits unnecessary.

Instead, RuleSets use local logical names for qubits. At the lowest level, physical qubits within a QNIC can be addressed by the local node using a tuple like `<QNICAddress,QubitIndex>`. However, this level of detail is not, and should not be, visible to the RuleSet logic.

Within a RuleSet instance on a given node, the RuleSet mechanism identifies a quantum resource primarily by its *tag*—a label that categorizes the resource into a specific pool corresponding to its current role or stage in the protocol (as discussed in section 4.3). To ensure an unambiguous and absolute ordering of multiple resources that may share the same tag (e.g., several Bell pairs awaiting purification), each resource, upon being allocated to a tag, is automatically assigned a *sequence number* by the local RuleSet engine. The combination `<tag,seq no.>` thus serves as a distinct and locally unique key within that RuleSet instance for referencing a resource and associating it with relevant metadata, such as its tracked fidelity. Actions specified within Rules typically reference these resources based on their tag and their relative order within that tag, for instance, by operating on the resource with the smallest sequence number (i.e., the oldest available in that pool). This structured internal naming is vital for deterministic local operations and for preventing local operational mismatches that could lead to problems like the leapfrogging issue [78] if not handled carefully at a higher protocol design level.

The RuleSet execution mechanism itself, therefore, only requires and manages this local `<tag,seq no.>` system for resource identification within a single node's RuleSet instance; it does not impose any inherent restrictions or requirements on how shared entangled states are identified between nodes. However, we *recommend* a strategy wherein the RuleSet creator (e.g., the Responder during connection setup) devises RuleSets such that the `<tag,seq no.>` is kept consistent across the RuleSets of all nodes involved in a particular joint operation on that resource. The responsibility thus lies with the RuleSet author to ensure that RuleSets are designed—potentially using this consistent naming strategy—to correctly manage shared resources and achieve the desired end-to-end protocol behavior. We believe this minimal local naming scheme provided by the RuleSet engine, combined with judicious design of the RuleSets themselves, is sufficient for implementing a wide range of complex quantum network protocols while allowing creativity with minimal restrictions on RuleSet creator parts, as will be illustrated with examples in section 4.4.

Inter-node communication

Distributed protocols, including RuleSets, rely on classical messages for coordination. In the RuleSet framework, messages are sent as part of a Rule's Action clause, and messages are received through Condition clauses that can trigger subsequent rules. These messages are crucial for coordinating their executions and making progress toward the successful distribution of end-to-end shared Bell pairs.

Each RuleSet coordination message includes the following fields.

- A **recipient address** indicating the destination node.
- A **message context ID** to help the receiving Rule Engine match the message to the correct logical context (this is distinct from the resource tag used to identify qubits).
- A **structured content** (key-value pairs) payload, which may contain measurement results, Pauli frame updates, or other operational data.

While the underlying message framework may support a generic dictionary-like structure for its content, for clarity in protocol design and ease of RuleSet construction, we define a set of conventional message types or commands (as listed in table 4.1). These common patterns provide convenience for frequently occurring inter-node communications, making the RuleSets more human-readable and their intended logic easier to distinguish and manage. The utility of these messages is evident in standard repeater operations such as entanglement purification and swapping, for heralding success and information about Pauli frame.

Condition clauses as triggers

Condition clauses within Rules define the specific circumstances or triggers that must be met for a Rule's associated Action clause to be executed. They can be conceptualized as predicates that enable transitions or operations within the distributed protocol governed by the RuleSet. Table 4.3 enumerates the types of Condition clauses supported. A Condition clause might depend on various factors, including the arrival of specific messages, the successful generation or availability of one or more local entangled states meeting certain criteria (e.g., having a particular number of resource with specific tag label), or internal events like timeouts. For instance, a Rule designed to deliver an entangled pair to an application might have a Condition clause that matches a single Bell pair labeled with a terminal tag indicating readiness. More commonly, Condition clauses need to match two entangled states simultaneously: for example, two Bell pairs shared with the same neighbor node for purification or two Bell pairs connecting to different remote nodes for entanglement swapping, and these can be realized by looking

Table 4.1: A summary of the primary message types used for inter-node communication within the RuleSet protocol framework. These messages are sent as part of an Action clause and are used to trigger Condition clauses on remote nodes.

Name	Descriptive Name	Arguments	Comments
Remote Events (Message Transmission)			
FREE	Release Remote Qubit	Partner addr., resource IDs	Instructs a partner to release their half of a shared state.
UPDATE	Pauli Frame Update	Partner addr., resource IDs, Pauli frame correction	Notifies a partner about a Pauli frame correction.
MEAS	Measurement Outcome	Partner addr., resource IDs, result	Exchanges classical measurement outcomes between partners, primarily for purification.
TRANSFER	Entanglement Transfer	Partner addr., resource IDs	Notifies a partner about the result of an entanglement swap.

Table 4.2: A summary of the Action clauses available within the RuleSet framework. Actions are grouped into local classical software operations and local quantum hardware operations.

Name	Descriptive Name	Arguments	Comments
Local Software Actions (Classical)			
SETTIMER	Set timer	Timer ID	Use with caution, as distributed race conditions can occur.
PROMOTE	Promote qubits	Qubit IDs, Rule ID	Transfers control/ownership of qubits to a different rule or stage.
FREE	Free qubits	Qubit IDs	Releases qubits back to the pool of unallocated resources.
SET	Set variable	Variable ID, Value	Changes a RuleSet variable; useful for tracking statistics like measurement counts for tomography.
Local Hardware Actions (Quantum)			
MEAS	Measure qubits	Qubit IDs, Basis	Measures one or more qubits in a specified or randomly chosen basis.
QCIRC	Apply quantum circuit	Qubit IDs, Circuit	Applies a general unitary operation. Used for BSM, purification, swapping, and encoding logical qubits.

Table 4.3: A summary of the primary Condition clauses that act as triggers within the RuleSet framework. These conditions are evaluated by the Rule Engine to initiate actions.

Name	Descriptive Name	Arguments	Comments
Internal State Conditions			
CMP	Compare RuleSet Variable	Variable ID, operator (e.g., ==, >), value	Evaluated when the specified internal variable is updated. The rule triggers if the new value satisfies the comparison. Used to manage protocol logic, such as counting purification rounds or messages received.
TIMER	Timer Expiration	Timer ID	Triggers a rule after a set duration has passed. Must be used with caution to handle timeouts and avoid race conditions in distributed protocols.
Quantum Resource Conditions			
RES	Resource Availability	Number of pairs, Partner addr. (or wildcard), min. Fidelity	The primary trigger for quantum operations. Matches when a specified number of Bell pairs meeting a minimum fidelity constraint are available from a given partner. Used for purification, swapping, and final delivery to an application.

at the tags. The precise logic of the condition ensures that actions are only taken when the necessary quantum and classical state prerequisites are fulfilled.

Action clauses for operations

Action clauses define the operations a node performs when the conditions of a Rule are satisfied. At this point, the Rule is guaranteed exclusive access to any quantum resources it is meant to use, as all resources must be reserved through the prerequisite Condition clause. Actions may include local quantum operations (e.g., applying quantum circuits), the generation of classical messages (to other nodes or to self), and updates to local classical state.

To ensure correctness and enable automated protocol validation, Action clauses must manage all resources they consume. This means each allocated quantum resource must, by the end of the Action, be either: (1) returned to the free resource pool, (2) assigned a new tag to indicate its updated role in the protocol, or (3) retained under its existing tag with updated metadata. Likewise, messages that triggered the Rule are considered consumed and must be either discarded or incorporated into outgoing messages or state

changes.

Although the logic inside an Action clause can be expressive, its operations are constrained to a well-defined set of primitives supported by the Rule Engine (see table 4.2) and are deliberately not Turing-complete. These constraints, which include operations like applying a local quantum circuit (QCIRC) or measuring qubits (MEAS), are a key architectural choice. By restricting the complexity of Actions, it becomes easier to analyze the protocol's properties, such as correctness, termination, and deadlock, and to prevent common bugs. Because many of these primitive quantum operations yield classical information, generating classical messages is an essential and frequent outcome of Action clauses.

Ruleset execution logic and resource management

The operational correctness of RuleSet-based protocols depends on two core properties: deterministic execution and precise resource accounting. To ensure predictable behavior, our model enforces that only one Rule fires at a time, even if multiple Rule Condition clauses are simultaneously satisfied. Rule selection proceeds by evaluating Condition clauses in a strict, predefined order—typically determined by a unique Rule ID. Once a Rule is selected, its Action clause executes atomically and to completion.

Although we assume that executing an Action clause incurs negligible latency compared to the delays from quantum operations or message transmission, this sequential execution model introduces potential performance limitations. In high-throughput regimes—where quantum resources are produced or messages arrive faster than actions can be executed—this could become a bottleneck. Addressing such scaling challenges while preserving deterministic behavior remains an open avenue for future work.

Each RuleSet instance maintains an internal resource table (or equivalent state tracker) to monitor the current state of all relevant local and remote resources (messages). This enables Rule Conditions to express complex dependencies—for example, waiting for a classical message from another node while simultaneously checking the availability of a specific class of Bell pairs. The table is updated dynamically as resources are generated, consumed, tagged, or modified by executed actions, and as messages are received.

To avoid conflicts and ensure independent progress of concurrent connections, we enforce resource encapsulation; once a quantum resource is assigned to a RuleSet instance (i.e., a connection context), it is not shared with other RuleSets. This strict partitioning prevents interference between protocols running in parallel on the same node. As will be discussed in the context of multiplexing strategies (section 4.5), such isolation allows each RuleSet to operate autonomously on its subset of resources and advance according to its local protocol logic.

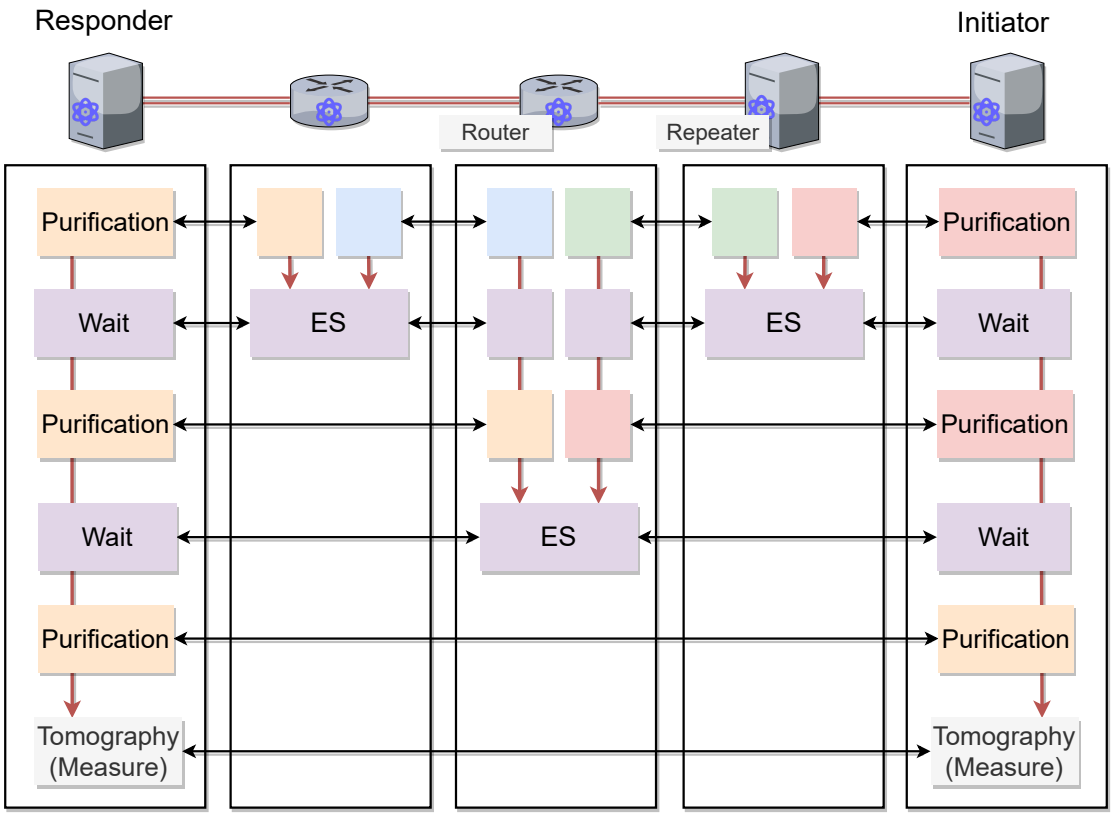


Figure 4.4: A conceptual illustration of how RuleSets could be installed for four-hop connection to create end-to-end Bell pairs. The RuleSets show flexibility by indicating purifications to be performed at three different stages of the protocol: at the link-level; before the second-swap; and end-to-end. (From [41].)

Illustrative examples of RuleSets

To make the operational principles of RuleSets more concrete, I provide simplified examples of how Rules can be defined for common quantum network tasks below. These examples illustrate the event-driven nature of RuleSet execution and show how quantum resources transition through the stages of a protocol. A conceptual illustration of a simple, sequential protocol for a four-hop connection is shown in Fig. 4.4. In this example, purifications are performed on the link-level Bell pairs, again after the first swap, and a final time on the end-to-end pair before it is consumed by a tomography application. Note that the final tomography block is merely illustrative and could be replaced by a rule that delivers the resource to any quantum application.

While this example is simplistic, it demonstrates a general principle: once the desired flow of a quantum resource is defined—from its link-level creation, through intermediate processing steps like purification and swapping, to its final consumption—a corresponding sequence of RuleSet blocks can be constructed to implement it. More complex protocols, such as those that transform Bell pairs into graph states [207], could be realized with a similar modular approach. For instance, to accommodate multiple

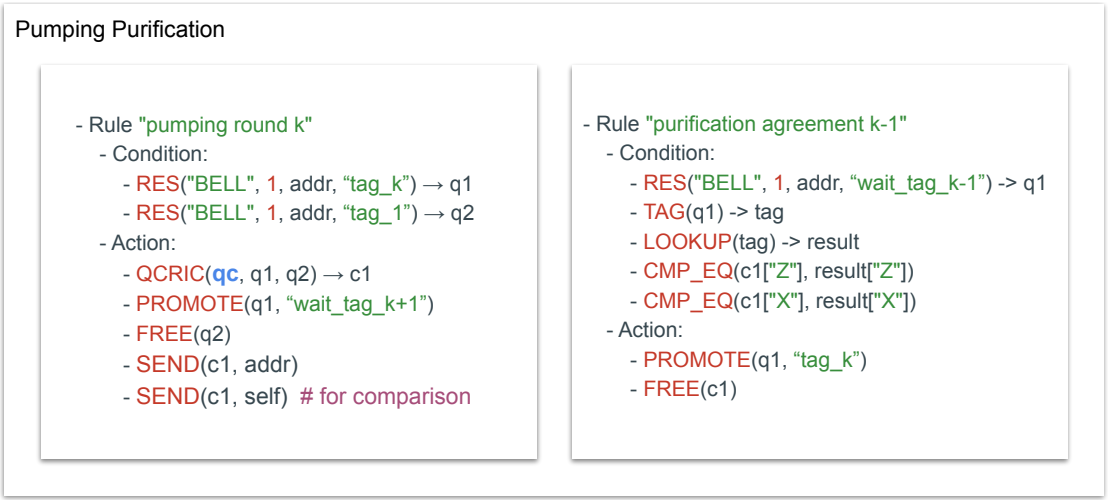


Figure 4.5: A conceptual depiction of a RuleSet for entanglement purification pumping. The diagram illustrates two key features of the protocol: the competition for base-level resources between different purification rounds (left), and the two-step process of a local action followed by a rule that waits for a heralded message to confirm the outcome (right).

rounds of purification at the link level, several purification blocks could be chained sequentially in the RuleSet.

However, more advanced purification strategies like pumping or banding require more careful construction of Rules. An illustration of the conditions and actions for entanglement purification pumping is shown in Fig. 4.5. In this model, the resource flow is not strictly sequential, as rules for different purification rounds must compete for the same pool of lower-level resources. To ensure correctness, the ordering of Rule evaluation becomes critical; for example, a Rule for a round-k purification must be evaluated before a Rule for a round-$(k-1)$ purification to ensure it gets priority access to the necessary resources.

A key implementation detail highlighted in this example is that any purification attempt is a two-step distributed process. First, a Rule performs the local quantum operation and moves the resource to a waiting state (e.g., "wait_tag_k" in Fig. 4.5). A second Rule is then triggered upon receiving the partner's heralded message. If the message content agrees with the local measurement results, this second rule confirms a successful outcome and promotes the resource to the next stage of purification.

Summary and implications of RuleSet-based semantics

In summary, the RuleSet-based mechanism provides a structured and expressive framework for defining and managing the complex connection semantics required in quantum repeater networks. By decomposing distributed protocols into sets of local, event-driven condition-action rules, this approach facilitates significant node autonomy and

asynchronous operation, which are crucial for practical network efficiency and responsiveness. The explicit inclusion of local state management, including quantum states and Stage-specific variables, allows RuleSets to capture the nuanced requirements of quantum protocols like entanglement purification and swapping.

The clear separation of concerns—with RuleSets defining the "program" for a connection, distinct elements for state naming, inter-node messaging, condition evaluation, and prescribed actions, all orchestrated by a well-defined two-pass setup protocol—promotes modularity and aids in reasoning about protocol correctness and performance. Furthermore, the ability to centrally devise RuleSets (e.g., at the Responder) offers opportunities for optimization and adaptation of protocols based on network conditions or application requirements.

While this framework offers considerable flexibility, it also highlights the inherent statefulness of quantum repeater nodes, a factor with significant implications for resource management and scalability. Assessing the resource footprint (both quantum memory and classical processing/storage for RuleSets) under various load conditions remains an important area for ongoing research. Nevertheless, the RuleSet paradigm provides a powerful abstraction for taming the complexity of distributed quantum computation inherent in establishing end-to-end entanglement, paving the way for more robust and capable quantum networks.

4.5 Networking

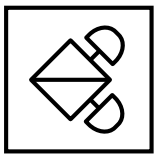
Building upon the RuleSet mechanisms for establishing and managing individual connections, this section broadens the scope to address network-level operational considerations. We will explore how to construct and operate robust quantum networks in complex topologies, considering diverse traffic patterns and the interplay of various network components. This involves defining the roles of different node types, strategies for routing and resource allocation across the network, and fundamental mechanisms for security and resource accounting.

Node types

The architecture of a quantum network is defined by its constituent nodes and their specialized functions. For clarity in describing our system, we group these nodes according to their primary contributions to network operation. In our architecture, we classify nodes into three main types: end nodes, for application interaction; repeater and router nodes, for extending entanglement and path management; and support nodes, for auxiliary operational tasks.

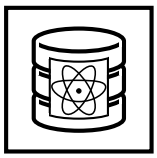
End nodes represent hosts that wish to execute a quantum application such as quantum key distribution, secret sharing and blind quantum computation [7, 9]. The technological

maturity required of an end node heavily depends on the desired application. There are four major kinds of end nodes:

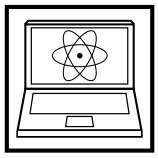


A measurement (MEAS) node is the most basic type of quantum end node, designed primarily for protocols that do not require quantum state storage. Its core capability is to receive individual photons and perform measurements on them in at least two different bases. Lacking quantum memory, MEAS nodes are well-suited for applications like quantum key distribution (QKD) or as simple terminals in certain forms of secure delegated computation protocols [8, 9], typically interacting with the network in a synchronous manner where measurement results directly yield classical data.

A sensor (SNSR) node is a specialized end node designed to utilize entangled states, often shared with distant parties, for high-precision measurements of physical quantities, for clock synchronization tasks, or for distributed sensing tasks [5, 6, 96–98]. These nodes typically feature limited quantum memory to hold working qubits (e.g., one half of an entangled pair) and possess specific quantum processing capabilities tailored for sensing protocols. Such capabilities include performing joint measurements, like Bell State Measurements (BSMs), between their stored qubits and photons that have interacted with the environment [89, 208]. While an SNSR node's internal processing is specialized, certain sensing applications may also necessitate high-rate entanglement generation from the network to achieve desired performance. For SNSR nodes, precise timing information is almost invariably a critical component of the service they provide or require, and their operation typically culminates in outputting classical data that corresponds to the sensed phenomenon.

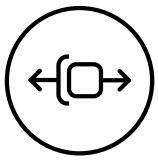


A store (STOR) node is a specialized end node whose primary function is to serve as a high-fidelity quantum data repository. Its core capabilities are the long-term storage of quantum states—often prepared and teleported from other locations—and the ability to teleport these states out on demand. While a STOR node does not require a universal gate set, it must support certain gates (e.g., Clifford gates) for active quantum error correction to preserve the stored quantum data. This includes using quantum error-correcting codes to protect against decoherence, along with mitigation strategies for correlated errors from events such as cosmic ray strikes [209–214]. In a network context, STOR nodes may function as data servers, enabling asynchronous applications where valuable states are prepared and stored for later retrieval.

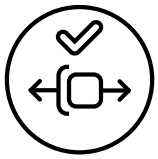


A computational (COMP) node represents a full-fledged quantum processing endpoint within the network. Equipped with quantum memory and additional algorithmic qubits, it can store, manipulate, and perform complex computations on quantum states received from the network or generated locally. COMP nodes support a wide range of advanced quantum network applications [215–218], including distributed quantum algorithms, more general forms of blind quantum computation [7, 9, 113, 114], and potentially fault-tolerant quantum computing [219–222], often requiring asynchronous interfaces to coordinate their local quantum workloads with network operations.

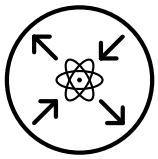
Next, we consider quantum repeater and router nodes, which are essential for overcoming distance limitations and managing network paths. We classify these into three primary types, briefly summarized here (for a more detailed foundational discussion, please refer to section 3.4).



A first-generation repeater (REP1) is a network node with two quantum interfaces, whose main task is to extend entanglement. Its primary operations include generating link-level Bell pairs with its neighbors, performing entanglement swapping to connect these segments, and managing errors through heralded entanglement purification on physical qubits along the connection path.



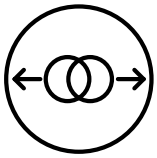
A second-generation repeater (REP2) also focuses on entanglement swapping to bridge distances but generally requires a larger quantum memory capacity and higher fidelity local quantum operations than a REP1. It utilizes quantum error correction (QEC) to manage operational errors in conjunction with heralded link-level entanglement generation. Instead of, or in addition to, purifying link-level physical Bell pairs, a REP2 node operates on logical qubits, where quantum information is encoded across multiple physical qubits to protect it from errors. This approach inherently demands more sophisticated hardware and advanced computational capabilities for encoding, decoding, and error correction routines during swapping.



A quantum router (RTR) is a more complex and versatile node, possessing all capabilities of a quantum repeater and typically featuring three or more quantum network interfaces, enabling it to make sophisticated path selection decisions in complex topologies. Architecturally, an RTR may consist of multiple line cards and a quantum backplane, allowing it to run a full suite of network operation protocols. Beyond basic repeating functions like entanglement swapping, an RTR can govern network borders potentially interfacing between different repeater generations or technologies, participate in generating multi-

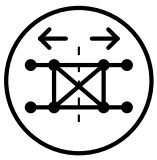
partite entangled states if the network provides such a service [207, 223–225], and may act as a Responder in connection setups by rewriting or generating new RuleSets for different network domains (as will be discussed further in section 4.6).

The last type, supporting nodes, provide essential functions that enable or enhance the operations of end nodes and repeater nodes in the entanglement distribution process. Key types of support nodes we list are as follows.

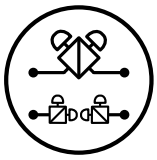


An entangled photon pair source (EPPS) is a device dedicated to generating pairs of entangled photons, commonly through processes like Spontaneous Parametric Down-Conversion (SPDC). These entangled photons are then typically distributed over quantum channels to be captured or measured at link endpoints, forming the initial resource for entanglement-based protocols. EPPS nodes can be deployed in various scenarios, including terrestrial fiber links or free-space satellite-to-ground communication [226–229].

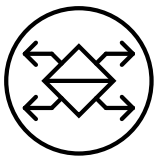
A Bell state analyzer (BSA) is a crucial component for performing projective measurements on two incoming photons, ideally projecting their combined state into one of the four Bell states. BSAs are fundamental for realizing photonic entanglement swapping [126], a primary process that creates link-level entanglement, used particularly to convert memory-photon entanglement into memory-memory entanglement between distant quantum memories. The efficiency and complexity of a BSA depend on the optical implementation and the specific photonic qubit encoding used.



A Repeater Graph State Source (RGSS) is a specialized source that generates multipartite entangled photonic states, specifically tailored for all-photonic (memory-less) quantum repeater architectures [2, 49, 50, 167]. An RGSS typically distributes segments of the generated repeater graph state to adjacent network nodes, where subsequent measurements on these photonic qubits are performed to establish long-distance entanglement without relying on quantum memories.



An advanced Bell state analyzer (ABSA) represents a more sophisticated version of a BSA, particularly required in advanced all-photonic repeater protocols based on repeater graph states [2, 49, 50, 167]. Unlike basic BSAs, an ABSA must be capable of performing measurements on single or multiple photons in dynamically selectable bases. The choice of measurement basis often depends on the outcomes of prior measurements within the network and the specific structure of the repeater graph state being utilized, implying more complex hardware and real-time classical control logic.



An optical switch (OSW) is a device that can passively route photons from input optical fibers or paths to different output paths without performing measurements on them [230]. OSWs, which can be based on technologies like nanomechanical systems or nanophotonic circuits [230], can be integrated as components within other node types (e.g., routers or complex end nodes) or can function as standalone elements in the network to dynamically reconfigure optical pathways.

This list is by no means exhaustive but covers the main components of a quantum network. The division into end, repeater and support nodes is not mutually exclusive, as there may be some overlap in functionality. For example, the RGSS may be viewed as a type of repeater node as well, as it realizes the task of entanglement swapping and link-level generation for all-photonic quantum repeater (discuss in section 3.4 and chapter 5). The RGSS and ABSA may participate in the connection planning process; the simpler BSA, on the other hand, is tasked only with notifying two nodes about the success of entanglement creation, and need not be visible to nodes farther away in the path.

Routing

Routing in quantum networks is the process of determining the communication path between a given set of end nodes. This typically involves two distinct routes: a quantum path consisting of quantum nodes for entanglement distribution, and a separate classical path connecting the control mechanisms of these quantum devices [81]. It is generally assumed that nodes with quantum channel connections also possess a corresponding classical channel. As discussed previously, a quantum router's primary task includes participating in finding these routes. Fig. 4.6 depicts such a router, capable of interfacing with multiple link types or different repeater generations and potentially possessing workspace memory for advanced processing like multipartite entanglement generation or complex error management. Selecting a route, or connection path, from one end node to another is often termed the entanglement routing problem [231, 232].

A common approach to quantum network routing is to adapt classical shortest-path finding algorithms, which shifts the primary challenge to defining an appropriate cost metric for use with algorithms like Dijkstra's. This method, sometimes referred to as qDijkstra [233, 234], often employs a cost metric combining the fidelity and generation rate of link-level Bell pairs. While the rate is relatively easy to characterize, fidelity is a more difficult metric to obtain accurately in practice. It requires constant link monitoring, as noise sources in quantum links can drift and change over time. Quantum state tomography [235–237] provides an accurate but resource-intensive way to measure the fidelity of generated Bell pairs at the link level. Consequently, developing lower-cost means of characterizing quantum states is an active research area [238–241], and

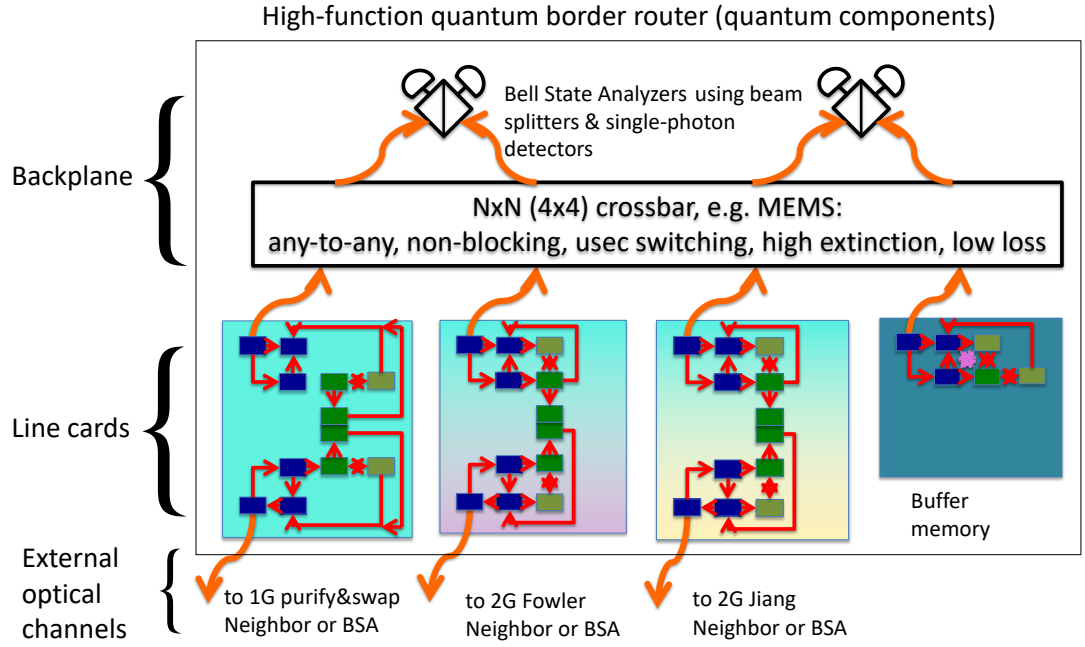


Figure 4.6: A full quantum router with hardware architecture similar to today's commercial Internet routers will have QNICs (line cards) coupled via a backplane consisting of optical ports, an optical switch, and Bell State Analyzer measurement devices. Using the BSAs, the qubits in the backplane buffers at the top of the line cards are entangled while the transceiver qubits in the lower portion attempt to create entanglement with neighboring nodes. Once both backplane and neighbor entangled states are made, entanglement swapping is used within each line card to splice the long-distance entanglement. A number of steps in hardware complexity (and cost) will exist between the minimal configuration of and this one. (From [40].)

the broader field of quantum state and process learning remains dynamic with many open questions [242]. Despite these measurement challenges, including fidelity in link metrics allows route calculations to automatically account for the trade-off between high-rate, low-fidelity links and low-rate, high-fidelity ones. The qDijkstra approach, using such metrics, has demonstrated good agreement between calculated path costs and simulated throughput across various paths with heterogeneous links [234].

Although qDijkstra shows good results for certain link characteristics, it is not always optimal, and defining a universally effective and easily determinable cost metric remains an open problem. Using only link fidelity and rate as costs is often insufficient, as the specific shape and characteristics of noise also play a vital role in the final end-to-end entanglement fidelity [233]. Recognizing these limitations, recent research has begun to move beyond simple qDijkstra-like approaches, attempting to integrate multiple factors—such as the probabilistic nature of entanglement swapping, memory decoherence during waiting periods, and purification scheduling—into more comprehensive global cost metrics [243]. These newer strategies draw from a diverse toolkit of opti-

mization and algorithmic techniques. For example, some utilize linear programming and mixed-integer constrained linear programming [244–246], while others employ greedy algorithms [134, 247, 248]. Furthermore, AI-based approaches [249, 250] and various multi-path routing methodologies [134, 137, 251–256] are also being actively investigated. Thus, defining appropriate and readily determinable cost metrics continues to be an important area of exploration for efficient quantum network routing [257].

Multiplexing and resource allocation

The term multiplexing can carry different meanings depending on context. At the physical layer of a quantum network, it might refer to techniques such as wavelength-division multiplexing (WDM) or frequency-division multiplexing (FDM) in optical fibers or spatial multiplexing via multiple cores or modes [25, 136, 258–262]. However, in this thesis, multiplexing refers specifically to *network-level resource multiplexing*: the dynamic allocation of shared quantum resources—such as quantum memories at repeater nodes, link-level Bell pairs, or intermediate entangled states—to multiple, concurrent end-to-end user connections.

Several multiplexing strategies have been proposed to manage these quantum network resources effectively [263–265]. In a circuit-switching approach, resources along the entire path are reserved exclusively for a given request for the full duration of its lifetime. Time-division multiplexing (TDM) alternates resource allocation across requests in a round-robin per Bell pair or time-sliced manner. Buffer-space multiplexing takes a more dynamic approach, assigning available quantum memories to requests as they arrive, with per-connection buffers resembling the notion of network slicing [266]. Finally, statistical multiplexing—inspired by best-effort forwarding in classical networks—allocates resources probabilistically, using heuristics based on request type, urgency, or priority. This can lead to higher overall throughput under load by letting separate parts of the network operate independently while sharing bottlenecks [267].

Each of these schemes comes with trade-offs. Circuit switching, while simple and reliable, underutilizes resources in high-load regimes. Buffer-space and statistical multiplexing have shown higher aggregate throughput in simulation studies, especially in scenarios with heterogeneous traffic and competing demands [264, 265]. Statistical multiplexing in particular has been shown to outperform circuit switching by exploiting temporal variations in demand. Xiao et al. [268] compared buffer-space and TDM approaches, although their study focused on throughput with sequential entanglement swapping and did not account for fidelity. These investigations remain mostly limited to small-scale networks. More recently, mixed-integer linear programming (MILP) techniques have been applied to jointly optimize routing and multiplexing [269, 270], suggesting a path toward more integrated control.

However, many open questions remain. Current models often neglect the complexities introduced by traffic variability, fidelity constraints, or the interaction between routing and resource allocation. There is concern that uncoordinated multiplexing could lead to phenomena similar to congestion collapse, where short-range connections monopolize limited entanglement resources, starving long-range paths. Multiplexing decisions must therefore be made in conjunction with routing logic and security mechanisms—including authentication, authorization, and accounting (the AAA model, discussed next). Non-blocking multiplexing protocols are preferable where possible, but any protocol that allows for extended resource occupancy must be governed by policies that consider user identity, priority, and possibly even payments or credit-based quotas.

Authentication, authorization, and accounting

As previously discussed, the gap between available quantum resources and anticipated demand is likely to necessitate fixed or semi-fixed resource allocation in early-stage quantum networks. Under such constraints, mechanisms for authentication, authorization, and accounting (AAA) will be essential components of the network architecture [271]. These systems will not only manage access control and usage tracking but also help enforce policies that determine who can reserve scarce resources and under what conditions.

In the absence of a carefully designed AAA framework, economic factors alone may dictate access, potentially privileging users or institutions with the ability to pay the most. To ensure equitable participation and avoid early monopolization, it will be important to explore AAA architectures that incorporate fairness criteria, priority based on mission-critical applications, or other policy-driven metrics beyond purely financial bids.

Security

Quantum mechanics offers the promise of unprecedented confidentiality between communicating parties—an advantage that has made quantum key distribution (QKD) a major focus in both theoretical physics and computer science. However, the intense focus on QKD has shaped a somewhat narrow view of quantum network security, one that overlooks broader architectural and protocol-level vulnerabilities. While it is true that quantum mechanics enables new capabilities—such as the detection of eavesdropping and certain forms of tampering—it also introduces entirely new attack surfaces [192, 193, 272, 273].

Some of these threats are rooted in the physical nature of quantum systems. For example, although QKD is in principle information-theoretically secure, real-world implementations are susceptible to side-channel attacks [273, 274], Trojan-horse attacks [275,

276], and vulnerabilities such as photon number splitting [277], which arise due to imperfections in sources and detectors [278]. These are not weaknesses of the protocol itself but of the hardware, software, and deployment environment—a fact that holds true for classical cryptographic systems as well. Other classes of vulnerabilities stem from network architecture and protocol design choices, such as denial-of-service (DoS) attacks on key quantum channels or repeaters, and impersonation of nodes via classical control messages. These threats highlight the need to consider both quantum-specific and classical attack vectors in any serious security model.

Moreover, distributed quantum applications—such as secure multi-party computation, quantum consensus protocols [218], and leader election [104, 108]—rely heavily on assumptions about trust, timing, and node availability. Any deviation from these assumptions can be exploited by adversaries, potentially undermining the correctness of the protocol. Thus, even in quantum networks, security is not just about encryption; it also involves resilience against disruption, impersonation, and strategic manipulation of the protocol environment.

All of this reinforces the critical importance of robust authentication and tamper-resistance mechanisms in early quantum network deployments. Whether classical privacy is necessary in all contexts is still an open question, but given the history of security oversights in classical networks, it is clear that a comprehensive security framework should be in place early—ideally well before operational quantum networks scale. These security concerns are closely intertwined with multiplexing strategies and AAA (authentication, authorization, and accounting) systems. In particular, as discussed previously, any form of persistent or exclusive resource allocation (such as circuit switching) will inevitably raise issues of identity, access rights, and possibly even economic fairness—each of which must be secured against misuse and abuse.

4.6 Internetworking and recursive architecture

The Internet is often idealized as a two-level architecture: external gateway protocols like BGP manage routing between networks, while internal gateway protocols such as OSPF and IS-IS handle routing within them. In reality, the structure is far more intricate. Tunneling protocols, virtual networks, switched Ethernet segments (which rely on Spanning Tree Protocol despite being “link-layer”), and modern virtualization approaches like 5G network slicing [266] have shaped the Internet into a complex, multi-tiered system with ad hoc inter-layer interactions.

Given the opportunity to redesign the system from first principles [33]—and with the benefit of hindsight—we would likely prefer a unified framework that explicitly supports clean abstraction boundaries and systematic composition across layers.

One such framework is the Recursive Network Architecture (RNA), proposed by Touch

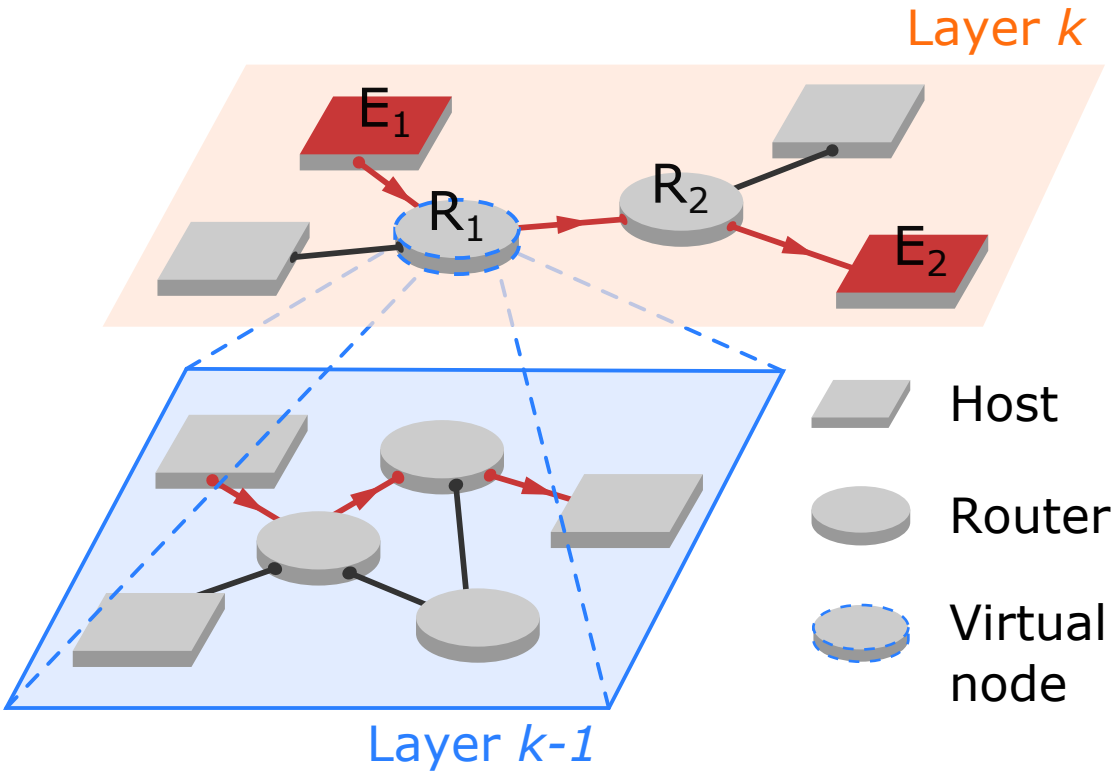


Figure 4.7: QRNA adopts a fully recursive architecture that enables an entire network to be virtualized as a node. QRNA can extend down to the link layer, and interoperates across internally heterogeneous networks as long as they conform to the protocol at their borders [279]. (From [40].)

et al. [37–39]. Inspired by RNA, Van Meter, Touch, and Horsman proposed the Quantum Recursive Network Architecture (QRNA) [36] as a foundational model for the Quantum Internet which we are building upon. QRNA offers a scalable blueprint for connecting logically and physically heterogeneous networks while preserving autonomy, privacy, and security through recursion.

Recursion in QRNA permeates both naming (section 4.4) and routing (section 4.5), producing a hierarchy of address spaces modeled as a directed acyclic graph. This hierarchical structure supports scalable routing, localized policy enforcement, and layered isolation—key to privacy and trust.

In traditional internetworking, connections are typically either boundary-to-boundary (for transit) or boundary-to-end-node (for termination). QRNA unifies these under a recursive abstraction; both are viewed as end-to-end connections at different levels of the stack. As shown in Fig. 4.7, a host node E_1 initiating a connection to E_2 at layer k perceives a direct path via routers R_1 and R_2. When the request reaches the border router, it is translated and embedded into layer $k-1$, where the path is resolved internally before being passed back up to layer k. This nesting can continue recursively across multiple layers, allowing composition of arbitrarily complex internetworks.

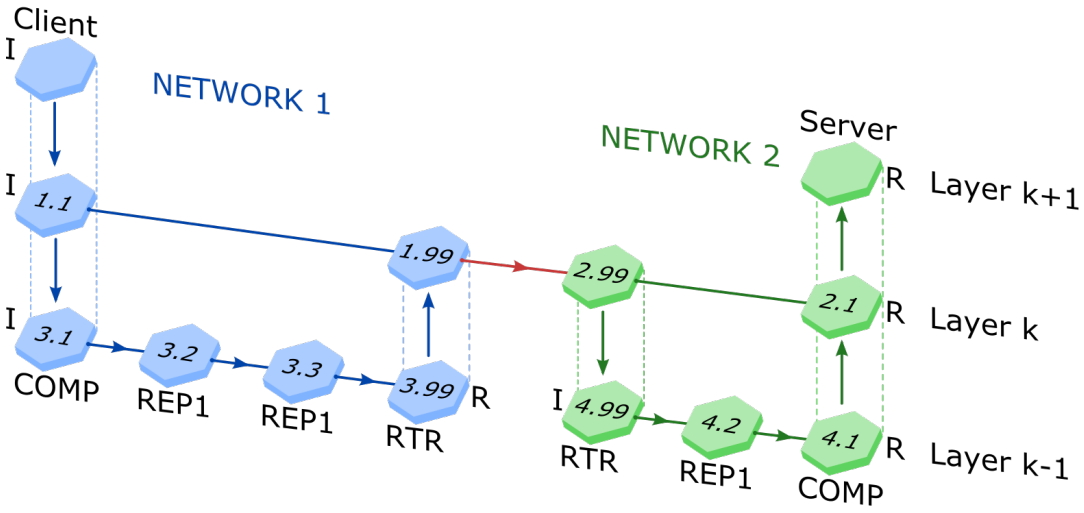


Figure 4.8: Two-pass connection setup in a recursive internetwork. Arrows indicate how the connection request moves across and between layers. Each recursion layer encapsulates its own Initiator (I) and Responder (R), resulting in multiple I-R pairs at layer $k-1$. (From [40].)

Connection setup with rewrite

The recursive model must integrate with the two-pass connection setup described in section 4.3. We achieve this via *RuleSet rewriting* at recursion boundaries—specifically, border routers translate detailed link information from internal layers into abstracted virtual links for higher layers.

This process is illustrated in Fig. 4.8. To preserve network autonomy and scalability, each border router rewrites its segment of the path into a single logical hop. This abstraction mimics how BGP hides internal topology while advertising route metrics. The border router acts as a Responder to the upstream Initiator, reporting a performance metric (e.g., Bell pair generation rate at a given fidelity) that summarizes its internal network's behavior.

In the Quantum Internet, especially in first- and second-generation repeaters, connections are stateful distributed computations. All intermediate nodes participate in operations such as purification and entanglement swapping, requiring connections to be pre-established and maintained via state along the entire path.

During setup, each node receives a Responder (destination) address and selects the next hop based on local policy. If the received destination address is a virtual one (i.e., not a direct link-level neighbor), the node translates the connection into a request for the layer below. At each recursion level, new Initiator-Responder pairs are instantiated. When a request reaches a Layer $k-1$ Responder (e.g., nodes 3.99 or 4.1 in Fig. 4.8), a provisional RuleSet is constructed. This RuleSet governs local behavior and is distributed backward along the setup path. Upon approval by the destination, the final RuleSet is confirmed

and installed.

This recursive encapsulation provides several benefits: **(i)** It bounds the scope of routing and state information, enhancing scalability and autonomy. **(ii)** It allows each administrative domain to innovate locally within the RuleSet framework. **(iii)** It enables systematic reasoning about complex paths via layered abstraction. **(iv)** It supports interoperability between different repeater generations, enabling seamless deployment of new technologies [280]. **(v)** It facilitates interworking between fundamentally different network architectures [279].

Routing, multiplexing, and AAA

Routing in QRNA generalizes classical inter-domain models by requiring a local routing protocol at each recursion level. While classical Internet routing relies on BGP and internal protocols like OSPF, QRNA uses full recursion, with each layer managing its own topology, policy, and performance metrics.

At higher recursion levels, path selection is typically guided by virtual link costs, often using the qDijkstra metric. This abstraction enables scalable inter-network routing by treating entire networks as single links at higher layers. When a network is virtualized in this way, it exposes performance metrics—such as entanglement generation rate or fidelity—that reflect internal policies. For instance, a network may choose to perform additional purification internally to present a higher-fidelity virtual link to upper layers. Prior work on purification scheduling and swap optimization [177, 281–286] has shown that performing purification closer to the physical link layer can significantly improve end-to-end throughput and fidelity.

At the lowest levels, within individual networks, routing strategies may vary and are not limited to distributed models like qDijkstra. Centralized or policy-driven approaches can often yield better performance and may be preferable depending on local infrastructure and goals (see section 4.5).

An intriguing direction for future research at higher recursion levels is the adaptation of segment routing techniques to quantum networks. In classical networks, segment routing [287–289] enables a source to encode a sequence of forwarding instructions—called segments—directly into the packet, allowing for flexible, programmable routing without requiring per-flow state at intermediate nodes. Applied to QRNA, this could involve embedding segment identifiers corresponding to recursive virtual links, streamlining route computation and enabling richer policy expression, such as fidelity constraints or preferred purification strategies.

Multiplexing is more challenging. As with classical networks, many concurrent connections must share infrastructure, and mechanisms for fairness, scheduling, and resource

allocation are essential. Unfortunately, the classical Internet provides limited guidance: while inter-domain QoS mechanisms have been discussed for decades, they remain sparsely deployed [290, 291]. As such, scalable, secure multiplexing remains a critical open research problem in quantum networking.

4.7 Quantum routing software architecture (QRSA)

The operational capabilities of quantum repeater nodes—particularly the execution of complex, stateful protocols like RuleSets within the broader principles of the Quantum Recursive Network Architecture (QRNA)—are realized through a well-defined software framework. Classical control software plays a critical role in designing and deploying efficient repeater-based quantum networks.

To this end, we propose the Quantum Routing Software Architecture (QRSA), a modular and extensible framework that provides the essential functionality for managing quantum resources, executing distributed protocols, and interfacing with physical hardware. QRSA serves as the node-level realization of the QRNA principles discussed in earlier sections, with software components designed to reflect the abstractions introduced in RuleSets and the distributed coordination requirements of quantum networking.

As illustrated in Fig. 4.9, QRSA consists of five primary components: the Connection Manager (CM), Routing Daemon (RD), Hardware Monitor (HM), Rule Engine (RE), and Realtime Controller (RC). Together, these modules form the minimal functional configuration needed for a quantum repeater or end node to participate in a scalable quantum network. The following subsections describe each component in detail.

Connection Manager

The Connection Manager (CM) is responsible for orchestrating connection establishment between an Initiator and Responder, spanning all intermediate nodes along the selected path. Its duties include handling incoming connection requests, assessing feasibility based on current resource availability and node policies, and coordinating the two-pass connection setup protocol detailed in sections 4.3 and 4.6.

During the outbound pass of the connection setup, the CM at each intermediate node appends local link information—such as available QNICs and link characteristics—to the traversing request. Once the annotated request reaches the Responder node and is accepted, the CM's most critical role begins; generating the RuleSets for each node along the path [292].

The CM at the Responder uses the aggregated link information to determine whether the network can meet the requested connection parameters (e.g., target fidelity, number of Bell pairs). If feasible, it constructs RuleSets that encode the protocol logic and resource coordination for each participating node and distributes them during the return pass. If

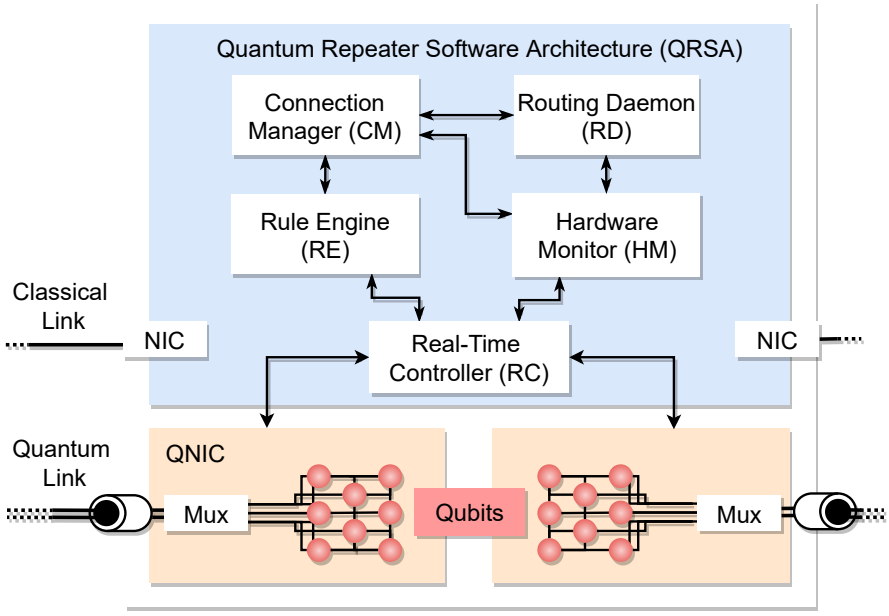


Figure 4.9: The Quantum Routing Software Architecture (QRSA) comprises five interacting software components: Connection Manager (CM), Routing Daemon (RD), Hardware Monitor (HM), Rule Engine (RE), and Realtime Controller (RC), managing both classical and quantum communication interfaces (QNICs). (From [41].)

the connection cannot be supported, the CM returns a rejection to the Initiator.

Routing Daemon

The Routing Daemon (RD) maintains path information to other nodes in the network, focusing on quantum-capable links. It builds and updates routing tables by exchanging reachability or link-state information with peer daemons at neighboring nodes via a suitable routing protocol (see section 4.5). The RD integrates local link metrics provided by the Hardware Monitor, such as entanglement rate or fidelity, and reflects network topology constraints or policies.

The RD's information is vital during connection setup. When the CM receives a request to establish a connection, it interfaces with the RD to determine the next hop toward the Responder. This interaction can be a direct query, or it can be achieved by the CM reading from a shared data structure, such as a routing table, that the RD maintains and continuously updates. The level of detail in this data can vary depending on the routing algorithm, ranging from minimal reachability information to full path-level metrics.

Hardware Monitor

The Hardware Monitor (HM) is responsible for tracking the performance and operational characteristics of local quantum hardware and link interfaces. Since link quality directly impacts the feasibility and efficiency of end-to-end quantum protocols, the HM provides timely updates on the status of each QNIC, including entanglement generation success

probabilities, link-level fidelities, noise models, and memory decoherence parameters.

This information is consumed by both the Routing Daemon (RD) and the Connection Manager (CM). The RD uses it to update link cost metrics and inform routing decisions, while the CM relies on it to tailor RuleSets to prevailing link conditions. For example, if a link exhibits high loss or reduced fidelity, the CM may increase the number of purification rounds or opt for alternative purification circuits. In the case of second- and third-generation repeaters, the HM may also provide logical qubit fidelity data under different error-correcting code distances.

Continuously characterizing links, however, can be resource intensive—especially in dynamic environments where link characteristics fluctuate due to external factors. Moreover, some diagnostic procedures (such as quantum state tomography) may temporarily render a link unavailable during the measurement process, reducing overall network throughput.

Recent work suggests alternatives that can reduce this overhead. For instance, there are methods proposed for inferring link properties through passive observation of entanglement purification outcomes [239–241]. This avoids the need to suspend normal operation while still gaining useful information. Similar ideas have been applied to quantum error correction, where techniques have been developed to estimate noise properties from syndrome data [293, 294], offering relevance for second- and third-generation repeaters.

Developing efficient, non-intrusive methods for detecting and adapting to link-level changes remains an important open problem. Advances in this area would improve real-time responsiveness to hardware fluctuations, ultimately enhancing the robustness and performance of quantum networks.

Rule Engine

The Rule Engine (RE) is the component responsible for executing RuleSets at each node. Once a RuleSet is delivered by the CM, the RE instantiates and manages its lifecycle, monitoring available resources and triggering actions when their associated conditions are met. The RE tracks the state of local qubits—via tags and sequence numbers as described in section 4.4—as well as classical messages, which may originate from the network or other software components.

When a Rule's condition clause is satisfied, the RE executes the associated action clause. This may involve invoking quantum operations via the Realtime Controller, updating resource metadata (e.g., re-tagging or releasing qubits), or generating new classical messages. In this way, the RE drives the autonomous, event-driven logic of the repeater node in accordance with the RuleSet's semantic specification.

A key function of the RE is managing resource multiplexing, ensuring that qubits are allocated and released according to active RuleSets. It also participates in connection teardown—when an end node determines that a RuleSet has completed, a teardown message is propagated through the path to instruct intermediate nodes to clean up associated state. Designing efficient teardown and resource agreement protocols is non-trivial but beyond the scope of this thesis, we refer curious readers to check some preliminary work in [292].

Realtime Controller

The Realtime Controller (RC) is the lowest layer in QRSA, responsible for translating logical commands from the RE into precise control signals for the quantum hardware. It executes quantum operations such as gate applications, Bell measurements, qubit initialization, and photon emissions. Because many quantum operations are time-sensitive, the RC must meet strict real-time performance constraints.

The RC generates the appropriate physical signals (e.g., laser pulses, microwave pulses, or hardware triggers) and synchronizes timing with neighboring nodes during link-level entanglement generation. It also handles qubit re-initialization, measurement readout, and low-level coordination of hardware components. In this way, it serves as the operational interface between abstract protocol logic and the concrete capabilities of the quantum device.

Component interactions and overall operation

The five components of QRSA work in concert to support distributed quantum networking. The process typically begins with the Connection Manager receiving a connection request. It queries the Routing Daemon for the next hop, consults the Hardware Monitor for link quality, and determines if the path can support the desired protocol. If the Responder approves the connection, the CM generates and distributes RuleSets to all participating nodes.

At each node, the Rule Engine loads and executes the RuleSet, monitoring resource availability and triggering quantum operations through the Realtime Controller. The RC, in turn, interfaces directly with the hardware to perform gate operations, entanglement generation, measurements, and qubit resets. Classical messages generated during protocol execution are routed via the system's classical NICs and interpreted by the REs at destination nodes.

This modular software architecture is designed to be extensible, interoperable, and suitable for various generations of quantum repeater protocols. Each component has clearly defined responsibilities and interfaces, enabling clean abstraction layers and enabling research into protocol-level behavior, resource allocation, and network scheduling. QRSA

also serves as the implementation target for the QuISP simulator (section 4.8), which models these components in a realistic time-driven simulation environment. We believe QRSA represents a minimal, yet powerful, foundation for the implementation of scalable quantum repeater nodes.

4.8 Simulation as evidence for architectural design

The architectural framework presented in this thesis, comprising the Quantum Recursive Network Architecture (QRNA), RuleSets, and the Quantum Routing Software Architecture (QRSA), aims to provide a scalable and modular approach for designing with the ability to analyze quantum networks. Throughout this chapter, we have described the architectural principles and protocol semantics in detail. In this section, we turn to simulation as a means of evaluating the behavioral fidelity and practicality of these ideas.

Simulation plays two complementary roles: it allows us to reason about the correctness of protocol behavior under realistic conditions, and it enables empirical validation of our models through comparison with analytical predictions and alternative implementations. In what follows, we first describe how RuleSets enable local reasoning about distributed behavior, then introduce the QuISP simulator [41, 189], explain its timing and error models, and present the results of a comparative validation study [43] with another quantum network simulator, the SeQUeNCe simulator [42, 295, 296].

Reasoning with RuleSets and the role of simulation

In an ideal world, the long-term validation of a network architecture would come through widespread real-world adoption. However, many well-designed systems—whether processors, operating systems, or communication protocols—have failed to gain traction due to external factors unrelated to technical merit. More importantly, retrospective success offers little help when evaluating architectures prospectively. For quantum networks, where deployments are fledging and experimental feedback is currently limited, it is crucial to develop abstractions and tools that support structured reasoning and prospective validation.

A core motivation for RuleSets, introduced in section 4.4, is to provide a programmable and analyzable abstraction for specifying node behavior in a distributed quantum network. RuleSets define node logic through event-driven condition-action rules, enabling modular and locally inspectable implementations of complex protocols such as entanglement swapping and purification. Their structure supports reasoning about protocol semantics, resource state transitions, timing constraints, and inter-node coordination.

Beyond supporting the implementation of correct behavior, RuleSets also help identify failure scenarios. Timing-sensitive issues such as leapfrogging or premature resource

discard (see section 4.4) can be addressed by carefully designing condition clauses and discard timers relative to classical messaging delays. This makes RuleSets a practical framework for protocol design that balances abstraction with operational precision.

Nevertheless, this kind of design-time reasoning does not guarantee correctness in deployed systems. The real behavior of a network depends on the execution semantics provided by the underlying QRSA software stack, including the Rule Engine, classical messaging, and physical resource control. Bugs or mismatches may arise from several sources:

- Misinterpretation of Rule Engine behavior by the RuleSet author;
- Implementation flaws in QRSA components (e.g., Connection Manager or Rule Engine);
- Incorrect assumptions about timing, failure handling, or hardware error rates;
- Unanticipated interactions or edge-case behaviors in concurrent execution.

To mitigate these risks, simulation plays a critical role. It allows us to evaluate protocol behavior under noisy, asynchronous, and resource-constrained conditions—revealing issues that are difficult to anticipate analytically. In this way, simulation complements human reasoning and provides an essential bridge between protocol design and practical implementation, supporting the iterative refinement of both logic and architecture.

QuISP and the QRSA execution model

To support simulation-based study of RuleSets and QRNA, we developed *Quantum Internet Simulation Package (QuISP)*, an open-source quantum network simulator implemented on top of the OMNeT++ discrete event simulation engine [297, 298]. QuISP is explicitly designed to reflect the QRSA model described in section 4.7, with a modular structure comprising the five core software components: the Connection Manager (CM), Routing Daemon (RD), Hardware Monitor (HM), Rule Engine (RE), and Realtime Controller (RC).

Each quantum node inside QuISP operates autonomously using RuleSets to guide its logic. When a condition is satisfied (e.g., a message arrives, a resource becomes available), an associated action is triggered; this may involve performing a local quantum operation, sending a message, or tagging or discarding a qubit. QuISP tracks both quantum and classical state over time, simulating the behavior of nodes in full accordance with QRSA semantics.

Importantly, while QuISP supports the execution of full RuleSet-based logic, our current use of the simulator has focused on small-scale networks, primarily linear repeater

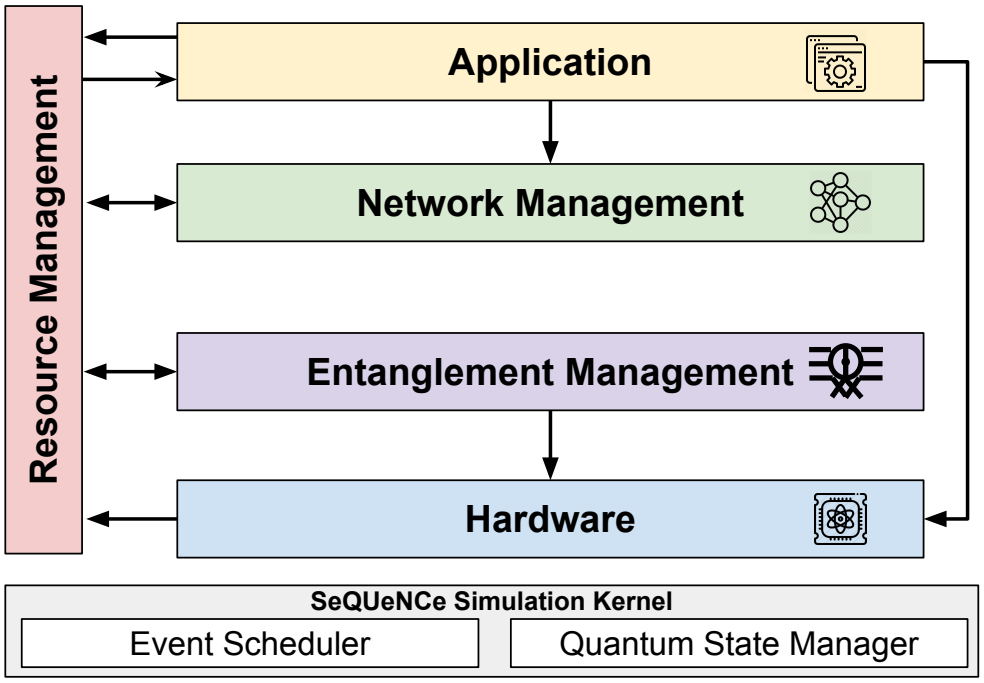


Figure 4.10: SeQUeNCe software framework comprises a simulation kernel and five functional modules. (From [43].)

chains. The primary goal has been to confirm that the RuleSet abstraction and QRSA model result in the expected low-level behavior in the presence of timing delays, resource constraints, and noise. The simulator provides a platform for testing both protocol logic and architectural behavior under these controlled settings. It is of interest to extend the use cases of QuISP to the study of emergent behaviors of large-scale quantum network in the future.

SeQUeNCe and its execution model

To assess the validity of QuISP's physical and timing models, we conducted a comparative study with *SeQUeNCe: Simulator of QUantum Network Communication*, an independently developed open-source quantum network simulator by Argonne National Laboratory and the University of Chicago [42, 295, 296]. Although SeQUeNCe does not implement the QRSA model or use RuleSet-based protocols, it operates with a similar rule-based resource manager that governs quantum memory allocation and coordination between its internal modules (see Fig. 4.10).

SeQUeNCe is structured around a modular architecture that enables flexible customization and reuse of simulation components. It consists of five primary modules—hardware, entanglement management, resource management, network management, and application—built atop a discrete event simulation kernel. This kernel handles event scheduling and includes a quantum state manager that maintains the evolving quantum state of all components. Quantum states in SeQUeNCe can be represented in

either bra-ket or density matrix formalism, supporting detailed modeling of quantum operations, noise, and decoherence.

The hardware module models the physical characteristics of quantum components such as QNICs and quantum memories, including their errors and limitations. The entanglement management module implements protocols for generating and managing entangled pairs. Resource and network management modules coordinate local and inter-node resource usage using classical control messages. SeQUeNCe's resource manager follows a rule-based logic to update memory state, functionally similar to the memory handling mechanism of QuISP's Rule Engine.

The application module models end-user quantum programs and their service demands. Additionally, SeQUeNCe integrates analytical noise models for efficient simulation of noisy bipartite entanglement, enabling simulations to maintain both physical realism and scalability. Together, these components allow SeQUeNCe to simulate a wide range of quantum network scenarios with high configurability and physical fidelity.

Timing models of QuISP and SeQUeNCe

Both QuISP and SeQUeNCe employ discrete-event simulation frameworks to capture the timing behavior of quantum and classical network operations, but they differ in execution strategy and modeling granularity.

QuISP models quantum operations, photon emission, and classical message propagation as discrete events scheduled with configurable system parameters. Link-level delays are determined by physical distance and the speed of light in fiber, while additional classical processing latencies can be specified for components such as the Rule Engine. Qubits experience time-dependent decoherence using a sliced exponential noise model: the qubit's lifetime is divided into short slices δt, and probabilistic error channels are applied slice by slice, closely approximating continuous error accumulation.

Link-level entanglement generation in QuISP begins with a connection setup, during which the desired number of Bell pairs N_{Bell} is specified. After synchronization, nodes emit photon trains toward a Bell-state analyzer (BSA), with photons separated by a time interval t_{sep}. Each round consists of photon emission, BSM processing, and classical feedback. The duration of one round is

$$T_{\text{round}} = 2T_0 + 10t_{\text{sep}} + (N_{\text{mem}} - 1)t_{\text{sep}}, \tag{4.1}$$

where $T_0 = \max(L_A, L_B)/c$ is the one-way signal delay from BSA to the farthest node and N_{mem} denoting the number of memories the node holds. The expected number of rounds needed to generate N_{Bell} pairs is

$$k = \left\lceil \frac{N_{\text{Bell}}}{N_{\text{mem}} p_{\text{succ}}} \right\rceil, \tag{4.2}$$

where p_{succ} is the success probability of a BSM. Thus, the total expected time is

$$T_{\text{exp}}^{\text{QuISP}}(N_{\text{Bell}}) = T_{\text{setup}} + k \cdot T_{\text{round}}, \tag{4.3}$$

with $T_{\text{setup}} = 2L/c + T_0$ representing the initial setup delay.

SeQUeNCe, by contrast, initiates a fresh entanglement generation protocol for each Bell pair, using a three-way handshake between nodes and the BSA. Timing events such as photon emission and classical messaging are scheduled based on physical path lengths and configured device latencies. The handshake introduces setup latency for every round, making the per-attempt model more sensitive to communication overhead.

Each entanglement generation round in SeQUeNCe includes handshake and feedback delays totaling approximately as

$$T_{\text{round}}^{\text{SeQUeNCe}} = 4L/c. \tag{4.4}$$

The total expected time to generate N_{Bell} Bell pairs is then

$$T_{\text{exp}}^{\text{SeQUeNCe}}(N_{\text{Bell}}) = k \cdot T_{\text{round}}^{\text{SeQUeNCe}}, \tag{4.5}$$

where k is again the expected number of successful attempts. This round-by-round negotiation model simplifies protocol logic but adds latency, especially for long-distance links.

Message diagrams illustrating both timing models are shown in Fig. 4.11, helping to visualize the key differences in protocol structure and scheduling assumptions.

These differing models foreshadow the timing behaviors observed in simulation results discussed in later subsection.

Error models of QuISP and SeQUeNCe

While both simulators implement quantum operational errors, they differ significantly in modeling philosophy and granularity.

QuISP adopts a probabilistic, per-qubit error model on top of graph state representation. Each qubit is associated with an error probability vector $\vec{\pi}(t)$ that represents the likelihood of being in different error states, including Pauli errors (X, Y, Z), relaxation and excitation (R, E), and photon loss (L). The evolution of $\vec{\pi}(t)$ over time follows a time-sliced Markov process as

$$\vec{\pi}(t) = \vec{\pi}(0)Q^n, \quad \text{for } n = \left\lceil \frac{t}{\delta t} \right\rceil, \tag{4.6}$$

where Q is the single-qubit error transition matrix and δt is the time slice interval.

Memory decoherence, single-qubit gates, and photon loss are modeled using this per-qubit transition matrix Q, with the evolving error vector $\vec{\pi}(t)$ tracking accumulated

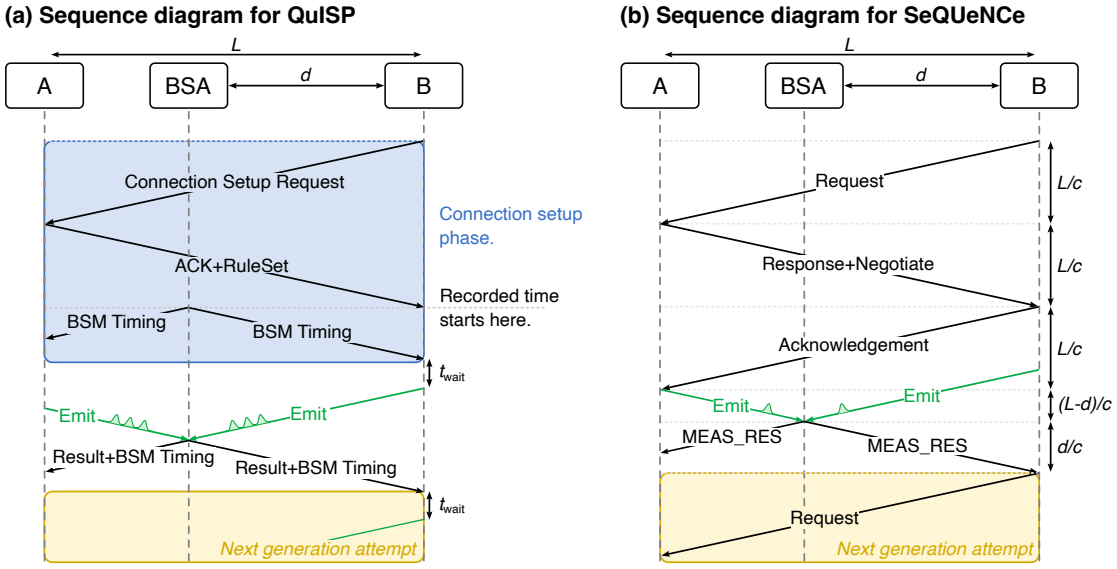


Figure 4.11: Message diagrams for link-level entanglement generation in QuISP and SeQUeNCe. (a) In QuISP's timing model, a single connection setup phase (blue) is performed once to request a batch of Bell pairs. This initiates a continuous stream of photon emission and BSA attempts (white), a process that repeats (yellow) until the requested number of Bell pairs are successfully generated. (b) In SeQUeNCe's timing model, a three-way handshake protocol is used to generate each Bell pair individually. The entire handshake process (yellow) is repeated for every subsequent pair required by the connection. (From [43].)

noise. Two-qubit gate errors, which introduce correlated noise and thus violate the independence assumption of single-qubit channels, are treated separately; with probability p_g, a two-qubit Pauli error is sampled and applied directly to the affected qubits. Measurement errors are implemented by flipping the reported classical outcome with probability p_m.

The fidelity of a two-hop Bell pair with noisy gates and measurements (ignoring decoherence) is

$$F_{\text{swap}} = (1-p_g)(1-p_m)^2 + \frac{3}{15}p_g(1-p_m)^2 + \frac{8}{15}p_g(1-p_m)p_m + \frac{4}{15}p_g p_m^2. \quad (4.7)$$

With memory decoherence over times t_1, t_2, and total classical communication delay T, the final fidelity becomes

$$F_{\text{swap, decoherence}} = F_{\text{swap}} \cdot \left(Q^{2t_1+2t_2+2T}\right)_{00}. \quad (4.8)$$

For depolarizing errors, we have a closed form expression for any integer power of the transition matrix given by

$$\left(Q^t_{\text{depol}}\right)_{00} = \left[1 + 3^{1-t}(3-4p)^t\right]/4, \quad (4.9)$$

where p is chosen such that $\left(Q^{t_{\text{coherence}}}_{\text{depol}}\right)_{00} = 1/e$.

SeQUeNCe uses analytical approximations to track fidelity. Memory decoherence is modeled using a continuous-time depolarizing channel applied to Werner states as

$$F(t) = F_{\text{in}} e^{-2t/\tau} + \frac{1 - e^{-2t/\tau}}{4}, \tag{4.10}$$

where τ is the memory coherence time, with the initial fidelity F_{in}. Gate noise is applied with probability p_{g} as a complete depolarization, while measurement errors flip the output with probability p_{m}.

For swapping two Werner states of fidelity F_1 and F_2, the post-swap fidelity is

$$\begin{aligned} F_{\text{swap,W}} = \frac{p_{\text{g}}}{4} + (1 - p_{\text{g}}) \big[& (1 - p_{\text{m}})^2 (F_1 F_2 + 3 e_1 e_2) \\ & + p_{\text{m}}(p_{\text{m}} - 2)(F_1 e_2 + e_1 F_2 + 2 e_1 e_2) \big], \end{aligned} \tag{4.11}$$

with $e_i = (1 - F_i)/3$.

These error models reflect differing design goals. QuISP prioritizes scalability by using per-qubit stochastic noise models that approximate quantum state evolution without tracking full quantum states. In contrast, SeQUeNCe uses more detailed noise modeling based on Kraus operators and density matrices, enabling higher precision and allowing for analytical tractability in some small-scale cases. Despite these differences, both simulators have been shown to produce consistent fidelity results under comparable conditions, as discussed in the next subsection.

Figs. 4.12 and 4.13 show the end-to-end fidelity results for both QuISP and SeQUeNCe under varying gate and measurement error rates, with the other error type held fixed. In both simulators, we observe excellent agreement between the simulated outputs and the respective theoretical models; eqs. (4.7) and (4.8) for QuISP, and eq. (4.11) for SeQUeNCe. Additionally, when memory decoherence is introduced via depolarizing noise, the simulated fidelity continues to closely track analytical predictions. These results confirm the effective consistency of both simulators' error models, despite their differing approaches to model errors.

Cross-validation of QuISP and SeQUeNCe: fidelity, timing, and error models

Cross-validating independent simulation platforms is essential for establishing trust in their results. Agreement between QuISP and SeQUeNCe, which were developed separately and based on different modeling philosophies, strengthens confidence in both simulators' ability to accurately capture quantum network behavior. This joint effort, presented in [43], aimed to reproduce three canonical quantum networking experiments on both platforms and compare their outputs to analytical predictions.

The three experiment scenarios studied are as follows.

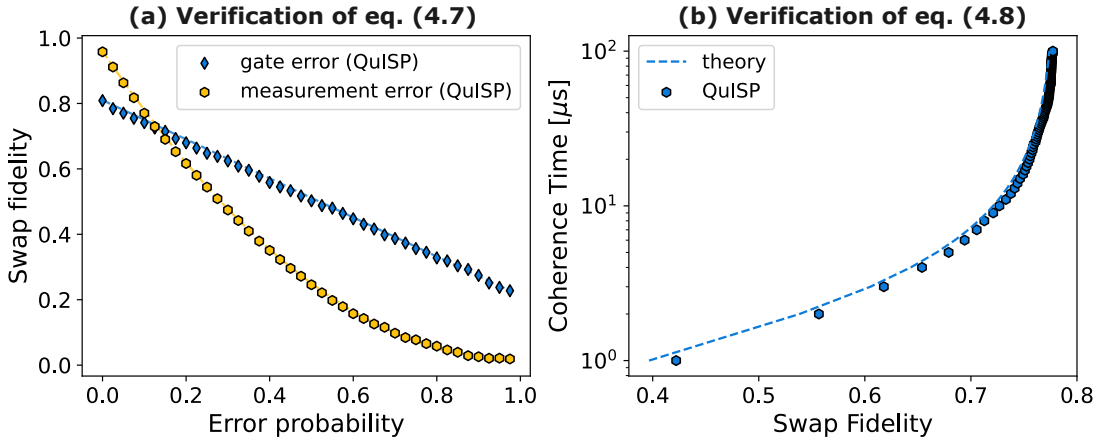


Figure 4.12: (a) Verification of end-to-end fidelity eq. (4.7) for QuISP. Blue diamonds for varying p_g and $p_m = 0.1$ and yellow hexagons for varying p_m with $p_g = 0.05$. Dashed lines represent our theoretical predictions. (b) Verification of eq. (4.8) for $p_g = 0.05$ and $p_m = 0.1$, varying degree of depolarizing noise. (From [43].)

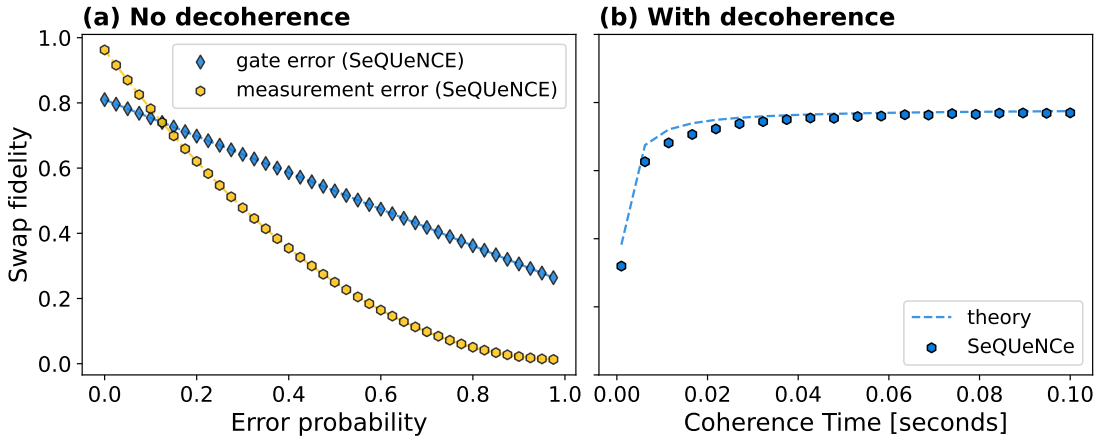


Figure 4.13: Verification of SeQUeNCe error models. (a) Without decoherence. Blue diamonds for $p_m = 0.1$ and yellow hexagons for $p_g = 0.05$. (b) Including decoherence for $p_g = 0.05$ and $p_m = 0.1$. (From [43].)

1. **Symmetric entanglement generation:** Entanglement generation over a memory–interference–memory (MIM) link, where the number of quantum memories at both end nodes is varied to evaluate scaling performance under symmetric conditions.

2. **Asymmetric entanglement generation:** End nodes each have a single memory, while the position of the Bell-state analyzer (BSA) is varied along the link to introduce asymmetry. This experiment examines how BSA placement affects timing and throughput under different classical control models.

3. **Two-hop entanglement swapping:** Establishing an end-to-end Bell pair across two hops, with each end node having a single memory and the repeater node having two (one per link). The experiment focuses on end-to-end fidelity under

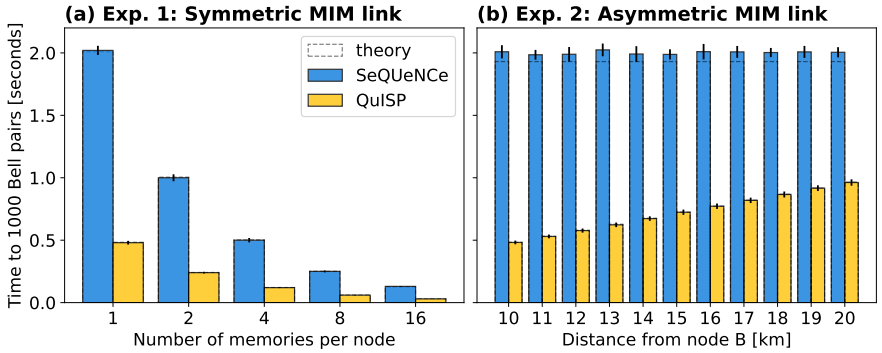


Figure 4.14: Time to generate 1000 Bell pairs for symmetric MIM link in (a), and asymmetric MIM link in (b). QuISP simulation used $t_{\text{sep}} = 1\text{ns}$. (From [43].)

varied operational errors, including gate infidelity, memory decoherence, and measurement noise.

In the first two experiments, which involved only photon loss and no operational noise, both simulators produced generation rates and completion times (for requests of 1000 Bell pairs) that matched analytical expectations derived from their respective timing models. As shown in Fig. 4.14, both simulators demonstrate the same qualitative scaling behavior in the first experiment, where increasing the number of available memories leads to improved performance.

In the second experiment, which introduces asymmetry by shifting the position of the Bell-state analyzer (BSA), the two simulators diverge in their trend lines due to differences in classical control assumptions. In SeQUeNCe, each Bell pair generation attempt is treated as an independent connection, requiring a new three-way handshake as shown in the timing diagram. This repeated setup dominates the overall timing, making the position of the BSA effectively irrelevant. In contrast, QuISP performs connection setup once at the beginning for the entire request of 1000 Bell pairs, and subsequent entanglement attempts proceed continuously. As a result, the effect of BSA placement becomes visible in QuISP: a centrally placed BSA minimizes total communication delay, while asymmetrical placement leads to longer delays per attempt and a more noticeable impact from photon loss on the longer fiber path. Despite these differences, both simulators match their own analytical models, confirming the internal consistency of their respective timing assumptions.

The third experiment, two-hop entanglement swapping, offered a more comprehensive test of distributed behavior and error modeling. This scenario introduces memory decoherence (finite T_1), gate infidelity, and measurement error. As shown in Fig. 4.15, QuISP and SeQUeNCe produced nearly identical end-to-end fidelity curves across a wide range of error parameters, once the parameters were translated to account for differences in how each simulator models noise.

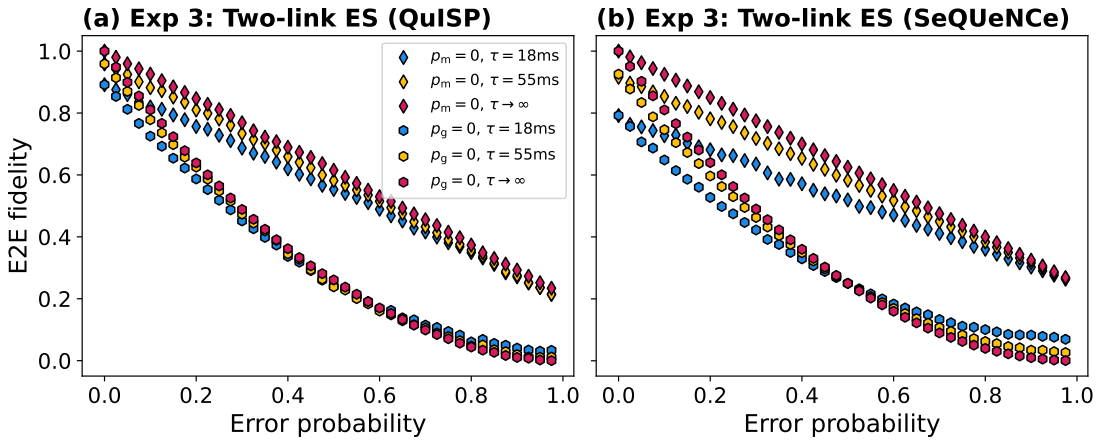


Figure 4.15: Experiment 3 (Two-hop swapping): End-to-end fidelity under varying gate and memory noise. Agreement between simulators supports the validity of each platform's error model and protocol behavior. (From [43].)

This agreement is particularly noteworthy given the distinct modeling philosophies of the two simulators. SeQUeNCe employs the density matrix formalism and analytical error models, while QuISP adopts a discretized and computationally lighter approach, using single-qubit noise channels and stochastic sampling to model correlated errors such as those arising from two-qubit gates. In the third experiment, the primary source of discrepancy lies in how gate error probabilities are defined. SeQUeNCe models a gate error of probability p as a complete depolarization of the two-qubit state with probability p, whereas QuISP applies one of 15 non-identity Pauli error terms with total probability p, excluding the identity operator. When this difference of definition is reconciled analytically, the two simulators produce nearly identical fidelity results across a broad parameter range.

Taken together, these findings validate the consistency and reliability of both timing and error models in QuISP and SeQUeNCe. Despite differences in modeling granularity and simulation strategies, both platforms faithfully reproduce expected quantum behaviors. For the purposes of this thesis, these results provide strong support for using QuISP as a platform for prototyping, simulating, and validating distributed quantum protocols within the QRSA software architecture and RuleSet-based control framework.

Limitation and future work

While the experiments discussed in this section demonstrate promising agreement between QuISP, SeQUeNCe, and analytical models, they represent only an initial validation step. Our current investigations are limited to small-scale networks, primarily linear chains and single-hop topologies, and focus on verifying timing models and fidelity predictions under controlled noise parameters. A natural next step is to apply QuISP to simulate larger-scale networks and study emergent behaviors, resource contention

across multiple connections, and the performance of routing and multiplexing protocols in more realistic settings.

It is important to emphasize that QuISP offers significantly more capabilities than demonstrated in this thesis. Built on the QRSA architecture and RuleSet-based control framework, QuISP provides modular tools for simulating resource allocation, routing, multiplexing, and AAA (authentication, authorization, and accounting) under heterogeneous network conditions. It supports a wide range of link types, including satellite-ground channels [229], memory–source–memory (MSM) and memory–interference–memory (MIM) links [145, 299], and is capable of modeling internetworking scenarios that span administrative domains [300]. However, these advanced features have not yet been applied to protocol validation or network-wide policy evaluation, which presents an important direction for future research.

Several other quantum network simulators have been developed, though many focus on different goals. Some emphasize hardware-level modeling [301, 302], others aim to integrate with specific applications or testbeds [29, 184, 303, 304], and some focus on entanglement distribution in linear repeater chains [186, 305] or specific network-layer issues such as routing and multi-path scheduling [256, 306, 307]. In contrast, QuISP is designed to support end-to-end architectural studies that encompass routing, multiplexing, AAA, and abstraction across link and network boundaries.

Unlike quantum error correction (QEC), where simplified noise models are supported by strong theoretical justification for fault-tolerant properties, quantum networking still lacks foundational results that explain when and why such models are appropriate [61, 131–133, 308]. As a result, the community tends to rely on more rigorous modeling approaches such as full density matrix simulations. We argue that further validation of simulators like QuISP, particularly through experimental testbeds and real-world deployments, is essential for increasing confidence in their predictive value. Building this empirical and theoretical foundation remains a key challenge for advancing simulation-based architecture design in quantum networking.

4.9 Discussion

Having presented the details of our architecture throughout the chapter, in this last section, we now reiterate its core principles and revisit the key design decisions introduced in section 4.1. This reflection highlights the rationale behind our architectural choices, the internal coherence of the overall design, and how the key components—RuleSets, QRNA, and QRSA—together provide a robust and extensible foundation for memory-based quantum networks. While this chapter has focused on one specific architecture grounded in RuleSets and recursive abstraction, it is not the only viable approach. Later, we broaden the discussion to consider alternative architectural paradigms and

other operational models of quantum networks that offer distinct trade-offs in generality, performance, and implementation complexity. This comparative discussion places our design choices within the broader landscape of quantum network architectures.

Revisiting the design decisions

Our architecture adopts a connection-oriented model in which connections are established through a two-pass setup protocol. These connections are managed by RuleSets distributed along the path between the Initiator and Responder nodes. This design treats Bell pair delivery as the fundamental network service, but it also leaves room for generalizing to the generation and distribution of multipartite entangled states, should a network operator choose to support such use cases (see section 4.4). This addresses the fundamental nature of the network's services and how connections are conceptualized and controlled.

RuleSets lie at the heart of our programming model, serving as distributed, event-driven control logic that defines local behavior and governs resource usage. They are generated based on user requests at the Responder and installed during connection setup, enabling localized yet coordinated behavior across the network. This allows connections to be customized and analyzable. At the application level, users specify high-level constraints, such as target fidelity or the number of Bell pairs, without dealing with low-level protocol details. These intents are translated by the Connection Manager into executable RuleSets. Furthermore, the Rule Engine, which allocates local resources to specific RuleSet instances (each corresponding to a single connection), enables multiplexing by tagging and isolating resources. This mechanism also lays the foundation for integrating fairness, access control, and scheduling policies across shared infrastructure—addressing the questions of how application requests are conveyed, how resources are multiplexed, and how security and AAA considerations can be incorporated.

In addition to defining connection semantics and control logic, our architecture clarifies node roles and facilitates scalable network composition through recursion. QRNA defines a recursive abstraction boundary that allows networks to be grouped and treated as virtual links at higher levels of hierarchy. This enables independent domains, each managed by different operators, to interconnect while preserving control over local behavior. Within each recursion level, the QRSA stack (especially the Connection Manager) rewrites RuleSets and coordinates connection setup through wrapper functions, supporting seamless internetworking across domains. Each recursion layer makes routing decisions based on its own topology and policy, enabling flexible and scalable routing strategies. At domain boundaries, RuleSets are translated to preserve end-to-end connection semantics, allowing interoperability across heterogeneous hardware and administrative zones. This modular and layered approach supports both architectural

consistency and extensibility—addressing the questions of defining network roles, supporting flexible routing strategies, enabling internetworking, and managing security and privacy at domain boundaries.

Taken together, these architectural mechanisms form a cohesive foundation for building programmable, distributed quantum networks that are scalable, analyzable, and technology-agnostic. The deliberate use of modularity, abstraction, and cross-domain compatibility ensures that the architecture can support a broad range of quantum networking applications and deployment scenarios.

Operation scheduling paradigms

The RuleSet-based architecture central to this thesis, as detailed in this chapter, operates on what can be best described as a dynamic scheduling paradigm. In this model, network activities are intrinsically event-driven, reacting to local information and resource availability, which fosters significant node autonomy in decision-making (albeit guided by the centrally-authored RuleSets). This approach offers considerable flexibility and responsiveness, mirroring the adaptive nature often sought in complex distributed systems.

To better contextualize this design choice and understand its advantages and trade-offs, it is instructive to examine it alongside other prominent paradigms for scheduling operations within a quantum network. These alternative approaches often differ fundamentally in how they manage the timing and coordination of key tasks—such as link-level entanglement generation, purification, and entanglement swapping—ranging from globally synchronized, discrete-time operations to more decentralized, continuous-time mechanisms. In the following, we will discuss three principal scheduling paradigms: the one-shot model, the time-slot based model, and will further elaborate on the characteristics of the dynamic model as exemplified by our work.

The *one-shot based paradigm*, explored in the work of Pant et al. [134] and Patil et al. [137], operates with operations occurring in predefined stages within synchronized slots. Its defining characteristic is that all unused entangled pairs (i.e., those not part of a successfully completed end-to-end entanglement by the slot's conclusion) are discarded at the end of each slot. This approach is particularly relevant for near-term quantum network implementations, especially where quantum memory lifetimes are short and cannot reliably hold qubits across multiple time slots. By ensuring each new operational round effectively starts with fresh quantum resources, this method minimizes the impact of memory decoherence on the final entanglement. Consequently, it tends to offer higher fidelity for successfully established end-to-end pairs, though often at the cost of slower overall throughput compared to approaches that preserve resources across slots. Furthermore, this operational model aligns well with assumptions for distributed

quantum computation at the datacenter scale, where protocols may rely on Bell pairs being generated and made available within specific time windows to perform remote operations between distinct quantum processing units or computers [220, 309–312].

In contrast, *the time-slot based paradigm*, as analyzed in the work of Cicconetti et al. [135] and Dhara et al. [136], also dictates that the network operates in globally coordinated time slots, with each slot typically divided into several predefined stages for sequential operations (e.g., link-level generation, purification, swapping, classical communication). However, unlike the one-shot model, entangled Bell pairs that are generated but not immediately consumed are retained in quantum memories and carried over to subsequent time slots. This approach necessitates quantum memories with longer coherence times capable of preserving qubit states reliably across these extended durations. The global coordination inherent in this paradigm often implies the need for a central decision-making unit to schedule and synchronize operations across the network. While such centralization can simplify the implementation of complex, multi-stage coordination strategies, the classical messaging overhead involved in communicating instructions and feedback between the central unit and distributed network nodes can become a significant bottleneck, potentially limiting performance and scalability, especially for wide-area networks. Nevertheless, for smaller or more controlled network segments, this paradigm can offer high pipeline efficiency.

Finally, the *dynamic paradigm*, which aligns with our RuleSet-driven architecture, schedules operations asynchronously. This approach is explored in the work of Zhan et al. [313], J. Li et al. [314], Z. Li et al. [315], and Andrade et al. [316]. In this model, actions are primarily event-driven and triggered by local conditions, such as resource availability or the arrival of messages. Link-level Bell pairs are generated whenever quantum memories become free and opportunities arise, and subsequent operations like purification or swapping occur opportunistically. This approach offers high flexibility and efficient memory utilization, embodying the adaptive and decentralized "spirit of the Internet." However, its asynchronous and highly state-dependent nature poses significant challenges for analytical performance modeling and prediction, especially at the multi-hop or network-wide scale.

Connection-oriented versus connectionless models

Two primary paradigms have emerged for managing entanglement distribution across quantum networks: the *connection-oriented model* and the *connectionless model*. Our proposed architecture adopts the connection-oriented model, as it offers strong guarantees in terms of coordination, fidelity, and scheduling flexibility. Understanding the fundamental distinctions between these two paradigms is crucial for appreciating the design rationale behind our framework.

The connection-oriented model [40, 135, 314, 315, 317], as detailed for our architecture in section 4.3, emphasizes establishing a confirmed path and securing resource commitments from all intermediate nodes before the attempt to distribute end-to-end Bell pairs begins. This pre-established agreement, orchestrated, for instance, by a Responder generating and distributing RuleSets, aims to enhance the reliability of entanglement distribution and provide more predictable end-to-end rate and state fidelity. This model is particularly well-suited for scenarios requiring non-local coordination, such as multi-hop entanglement purification and more complex scheduling.

In contrast, the connectionless model [33, 159, 315, 317] operates without prior resource reservation or explicit commitment from all nodes along an entire path. Instead, operations like Bell pair distribution typically proceed on a hop-by-hop basis, analogous to packet switching in classical networks. While this approach is often associated with certain third-generation repeater schemes designed for high-speed attempts, it could, in principle, be adapted for first or second-generation repeaters. If utilized with a RuleSet-like protocol, for example, it would likely require RuleSets to be more dynamically generated or updated, involving ongoing coordination between the current end-node of an extending Bell pair and the next hop node it is attempting to reach, rather than being entirely predetermined by a central Responder.

The hop-by-hop nature of the connectionless model generally restricts entanglement purification to occur only between adjacent nodes. This can make it challenging to achieve high end-to-end fidelity guarantees, as errors might accumulate along the path. If an operation within a segment fails, or if the fidelity of an intermediate entangled pair drops below a usable threshold and cannot be salvaged by link-level purification alone, the attempt for that path segment is aborted and may be retried. Consequently, connectionless approaches often impose more stringent requirements on the baseline error thresholds of individual quantum operations, the initial fidelity of elementary links, and the coherence times of any quantum memories involved. While many studies on connectionless models [315, 317] have focused on their potential advantages in terms of entanglement generation rates, comprehensive analyses of their end-to-end fidelity characteristics, especially across multi-hop paths with realistic noise models, are an ongoing area of investigation.

Chapter 5

INTEGRATING ALL-PHOTONIC REPEATERS INTO THE ARCHITECTURE

A key element that enables quantum repeaters to circumvent the fundamental repeaterless bound is their ability to store short-distance Bell pairs between adjacent nodes inside quantum memories, and later splice them together via entanglement swapping to extend entanglement across longer distances. This principle underlies both first- and second-generation repeaters. Other approaches—such as those proposed in third-generation repeaters—aim to significantly relax memory lifetime requirements, reducing them from the duration of an end-to-end round-trip to merely the time needed to receive and forward quantum states at each node, typically on the scale of link-level communication delays. These methods achieve this by leveraging fault-tolerant teleportation [156, 157], performing logical operations through surface code-style measurement-based computation [154], or transmitting photonic qubits encoded in fault-tolerant codes directly through the network in hop-by-hop manner decoding and re-encoding at every repeater node [44, 155].

Challenging the assumption that quantum memories are indispensable, Azuma, Tamaki, and Lo proposed a new type for quantum repeaters, the *all-photonic quantum repeaters*, built entirely from flying photons [2]. Their insight was that the use of memories arises primarily to deal with the probabilistic nature of entanglement generation and to facilitate the delayed operations required for high-fidelity entanglement swapping. By embedding these capabilities directly—near-deterministic link-level generation—into large photonic graph states, called the *repeater graph states (RGS)*, they outlined how repeaters could be realized without intermediate storage, instead relying on photonic loss-tolerant graph states and adaptive measurements.

In this chapter, we will revisit this all-photonic paradigm and focus on addressing the gap between conceptual designs and practical implementations [46]. We first review the original proposal by Azuma, Tamaki, and Lo, then present our contributions: a

Portions of this chapter have been adapted from the following publications:

- N. Benchasattabuse et al., "Engineering Challenges in All-Photonic Quantum Repeaters," *IEEE Network*, vol. 39, no. 1, pp. 132–139, Jan. 2025, DOI: 10.1109/MNET.2024.3411802.
- N. Benchasattabuse et al., "Architecture and Protocols for All-Photonic Quantum Repeaters," in *2024 IEEE International Conference on Quantum Computing and Engineering (QCE)*, vol. 01, Sep. 2024, pp. 1879–1889, DOI: 10.1109/QCE60285.2024.00217.
- N. Benchasattabuse et al., "Integrating Entanglement Purification into All-Photonic Quantum Repeaters," in *2025 IEEE International Conference on Quantum Computing and Engineering (QCE)*, to be published, Apr. 25, 2025, DOI: 10.48550/arXiv.2504.18121.

more modular approach to repeater graph state generation based on the introduction of half repeater graph states (half-RGS) [45, 47], which can be deterministically generated using quantum emitters. This design allows not only greater timing flexibility and modularity but also provides a natural way to hand off photonic entanglement to memory-equipped end nodes (addressing challenges raised in section 3.6). Leveraging this abstraction, we demonstrate how paths composed entirely of all-photonic repeaters can be treated as virtual link-level connections from the perspective of higher-level protocols. This integration point allows such paths to be seamlessly embedded into the recursive architecture introduced in chapter 4, despite their very different underlying hardware assumptions and operations. Finally, we address a long-standing question in the literature—how to support non-neighbor purification in an all-photonic segment—by showing that half-RGS offers the operational flexibility needed to realize purification and swapping even when limited to photonic-only infrastructure [48] in section 5.7.

5.1 Preliminaries to understanding all-photonic quantum repeaters

To understand the design and operation of all-photonic quantum repeaters, we begin by revisiting a key concept—the graph state formalism. While the basics of graph states were introduced in section 2.3, we will extend this discussion by introducing an additional measurement pattern, the XX measurement on adjacent qubits, and its corresponding graph transformation rule, both of which are central to the all-photonic quantum repeater protocol in this chapter.

Revisiting graph state transformation rules

Let us recall a basic transformation rule in the graph state formalism, the effect of a Z basis measurement on a qubit. In a graph state, measuring a qubit in the Z basis removes the corresponding vertex from the graph, and induces a Pauli Z^r operator as a side effect on each of the measured qubit's neighbors, where $r \in \{0, 1\}$ is the classical measurement outcome. An example of this Z measurement effect is depicted in Fig. 5.1 (bottom left).

To fully describe the operational mechanics of the all-photonic repeater scheme, however, we require an additional graph transformation rule—the *XX measurement rule*. This rule applies when two adjacent qubits in a graph state, say a and b, are each measured in the X basis, under the condition that these two qubits do not share any common neighbors. It is important to distinguish this sequence of two independent single-qubit X basis measurements from a joint measurement of the XX Pauli operator (a stabilizer measurement) often encountered in quantum error correction.

As illustrated in Fig. 5.1 (bottom right), performing these X basis measurements on adjacent qubits a and b transforms the graph such that the distinct neighborhoods $N_a \setminus \{b\}$

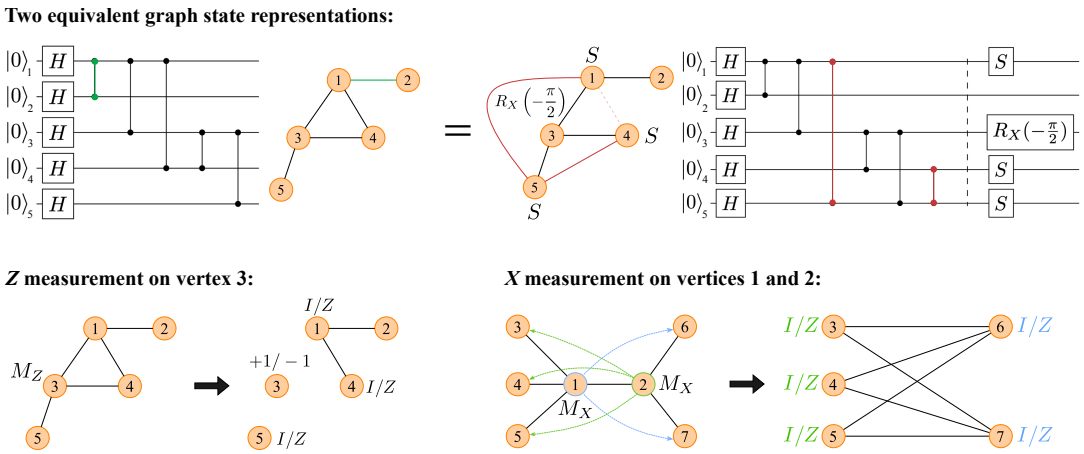


Figure 5.1: Top: Two equivalent graph state representations of the same quantum state. Vertices represent qubits, and edges correspond to controlled-phase gates. A green edge in the top left highlights one such gate. The top right graph is obtained by applying local complementation on vertex 3, replacing edges between its neighbors (e.g., deleting $(1, 4)$) with new ones (e.g., adding $(1, 5)$ and $(4, 5)$, shown in red). The states remain equivalent under local Clifford operations. Bottom: A Z measurement and its effect (left), and an XX measurement on two qubits in a different graph (right). Side effects on qubits 3, 4, and 5 depend on the measurement of qubit 2; those on qubits 6 and 7 depend on qubit 1. Blue and green arrows indicate these dependencies. (From [45].)

and $N_b \setminus \{a\}$ become fully interconnected, forming a complete bipartite graph between them, while a and b are removed. While such graphical transformation rules are powerful tools for bypassing the often cumbersome direct manipulation of stabilizer tableaus, we will derive this XX measurement rule explicitly using the stabilizer formalism. This derivation is particularly more convenient as X basis measurements involve multiple non-trivial steps to represent purely through local complementations or simple vertex and edge removal rules.

The derivation proceeds by analyzing the effect of the X basis measurements on a and b on the stabilizer generators of the initial graph state. We can categorize these generators as follows:

1. The stabilizer generator g_a associated with qubit a.
2. The stabilizer generator g_b associated with qubit b.
3. Generators g_j associated with qubits $j \in N_a \setminus \{b\}$ (neighbors of a, excluding b).
4. Generators g_k associated with qubits $k \in N_b \setminus \{a\}$ (neighbors of b, excluding a).
5. Generators associated with all other qubits outside of $N_a \cup N_b \cup \{a, b\}$.

Generators in the last category commute with the X basis measurements on a and b, thus remaining unaffected, and can be excluded from this specific derivation. For clarity, and

with slight abuse of notation, we denote the Pauli string for stabilizer generators of each group (omitting the last group) in table 5.1

Table 5.1: The stabilizer of graph states before XX measurement on adjacent qubits a and b. Operators acting on qubit groups: M_j, $N_a \setminus \{b\}$, a, b, $N_b \setminus \{a\}$, and M_k from left to right, with M_j (M_k) denoting qubits in the neighborhood of qubts in N_a (N_b) not including a (b).

Generator Type	**Pauli Operators on Qubit Groups**					
	M_j	$N_a \setminus \{b\}$	a	b	$N_b \setminus \{a\}$	M_k
g_a	I	Z_j^a	X_a	Z_b	I	I
g_b	I	I	Z_a	X_b	Z_k^b	I
g_j (for $j \in N_a \setminus \{b\}$)	Z_{M_j}	X_j^a	Z_a	I	I	I
g_k (for $k \in N_b \setminus \{a\}$)	I	I	I	Z_b	X_k^b	Z_{M_k}

To derive the effect of the XX measurement, we apply the stabilizer update procedure from section 2.2 sequentially: first for the measurement of $P_a = X_a$, then for $P_b = X_b$. Let the measurement outcomes be $s_a, s_b \in \{+1, -1\}$. After the measurements, the stabilizer group of the entire system will include $s_a X_a$ and $s_b X_b$. Our goal is to determine the updated stabilizer generators that describe the system post-measurement.

Starting with the measurement on qubit a, the system is projected into an eigenstate of X_a. We observe that g_b and all g_j anticommute with X_a. Following the rules from section 2.2, we update the stabilizer by multiplying each g_j with g_b, giving updated generators $g'_j = g_b \cdot g_j$. We then remove g_b from the stabilizer set. After this first measurement, the stabilizer becomes

$$\left\langle \begin{array}{c} Z_j^a X_a Z_b \\ s_a X_a \\ Z_{M_j} X_j^a I_a X_b Z_k^b \\ Z_b X_k^b Z_{M_k} \end{array} \right\rangle. \tag{5.1}$$

Here, each generator g_j associated with a neighbor $j \in N_a \setminus \{b\}$ has been updated, effectively transferring its entanglement to qubit b.

Next, we measure qubit b, obtaining outcome s_b for X_b. We choose g_a, which anticommutes with X_b, and multiply it into all g_k, yielding the updated stabilizer:

$$\left\langle \begin{array}{c} s_b X_b \\ s_a X_a \\ Z_{M_j} X_j^a I_a X_b Z_k^b \\ Z_j^a X_a X_k^b Z_{M_k} \end{array} \right\rangle, \tag{5.2}$$

To simplify, we now eliminate X_a and X_b from the remaining generators by multiplying the transformed g'_j and g'_k with $s_a X_a$ and $s_b X_b$, respectively. This results in:

$$\left\langle \begin{array}{c} s_b X_b \\ s_a X_a \\ s_b Z_{M_j} X_j^a Z_k^b \\ s_a Z_j^a X_k^b Z_{M_k} \end{array} \right\rangle, \tag{5.3}$$

The net effect of these two measurements is a new stabilizer set for the remaining unmeasured qubits. For each original neighbor $j \in N_a \setminus \{b\}$, its updated generator becomes:

$$g'_j = s_b X_j \left(\prod_{v \in N_j \setminus \{a\}} Z_v \right) \left(\prod_{k \in N_b \setminus \{a\}} Z_k \right). \tag{5.4}$$

Similarly, for each original neighbor $k \in N_b \setminus \{a\}$, its updated generator is:

$$g'_k = s_a X_k \left(\prod_{v \in N_k \setminus \{b\}} Z_v \right) \left(\prod_{j \in N_a \setminus \{b\}} Z_j \right). \tag{5.5}$$

Interpreting these in the graph state formalism (see eq. (2.18)), we see that each qubit j retains its original neighbors (excluding a) and gains new edges to all qubits in $N_b \setminus \{a\}$, as indicated by the $\prod_{k \in N_b \setminus \{a\}} Z_k$ term. Likewise, each k in $N_b \setminus \{a\}$ becomes connected to all $j \in N_a \setminus \{b\}$, as shown by the $\prod_{j \in N_a \setminus \{b\}} Z_j$ term. The resulting graph corresponds precisely to a complete bipartite connection between the two original neighborhoods (as shown in the bottom right of Fig. 5.1 as claimed), confirming the graphical transformation rule.

Having now extended the ruleset for graph state transformations and their associated effects under measurement, we focus on two key measurement sequences—single-qubit Z measurements and two-qubit XX measurements, to analyze how Bell pairs are distributed in the all-photonic quantum repeater scheme.

5.2 Original all-photonic quantum repeaters

We begin by reviewing the original proposal for all-photonic quantum repeaters by Azuma, Tamaki, and Lo [2]. The core component of this scheme is a highly entangled photonic graph state known as the *repeater graph state* (RGS). For brevity, we will refer to this all-photonic quantum repeater scheme as the *RGS scheme* throughout the rest of this thesis.

Using the terminology and node types defined in chapter 4, we can describe the RGS scheme as follows. The nodes responsible for generating repeater graph states are referred to as repeater graph state sources (RGSS). The intermediate measurement

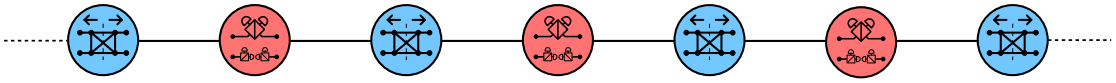


Figure 5.2: An illustration of a segment of a connection path in a network based on the RGS scheme. Blue nodes are RGS source nodes (RGSS), and red nodes are Advanced Bell State Analyzers (ABSAs). Note that end nodes are often omitted in diagrams of the original scheme, as it was primarily designed for applications like QKD where the end nodes can be simple measurement devices.

nodes, which must be capable of performing both Bell state measurements and adaptive single-qubit measurements, are known as advanced Bell state analyzers (ABSAs). An example of a connection path segment in the RGS scheme is shown in Fig. 5.2. It is important to note that the original scheme was designed with applications like QKD in mind, and thus often implicitly assumes the end nodes (e.g., Alice and Bob) are measurement-based terminals, which are typically not drawn in the diagrams.

We will now delve into the original RGS scheme in more detail, highlighting its key innovations, particularly its resilience to photon loss and operational errors. We will also describe the quantum error-correcting code it employs and briefly outline the conceptual and technological assumptions underlying the scheme.

Overview of the RGS scheme

A repeater graph state (RGS) consists of two sets of qubits, *inner* and *outer* qubits, each containing $2m$ qubits (as shown in Step 1 of Fig. 5.3). The inner qubits form a complete graph, and each inner qubit is connected to an outer qubit. We refer to m as the *number of arms* of the RGS. This complete graph among the inner qubits encodes the entanglement swapping, before photons are transmitted to establish link-level entanglement between adjacent neighbor nodes. This leads to the interpretation of the RGS scheme as a "time-reversed" version of conventional quantum repeaters, where entanglement swapping occurs *before* the link-level Bell pairs are generated.

To distribute a Bell pair between two end nodes (e.g., Alice and Bob), the RGS scheme employs a linear path made up of repeater graph state source (RGSS) and advanced Bell state analyzer (ABSA) nodes, analogous to a path consisting of memory–interference–memory (MIM) links in memory-based networks.

As illustrated in Fig. 5.2, each RGSS along the path generates an RGS. Each RGS is split in two: one half is sent to the left ABSA and the other to the right ABSA. Each ABSA receives the halves of *two* RGSs, one from each neighboring RGSS, and performs *three stages of adaptive measurement*.

In the first stage, the ABSA performs *rotated Bell state measurements (BSMs)* between each pair of outer qubits from the two RGS halves (Step 2 of Fig. 5.3). These rotated BSMs are equivalent to adding an edge between the two outer qubits and then performing

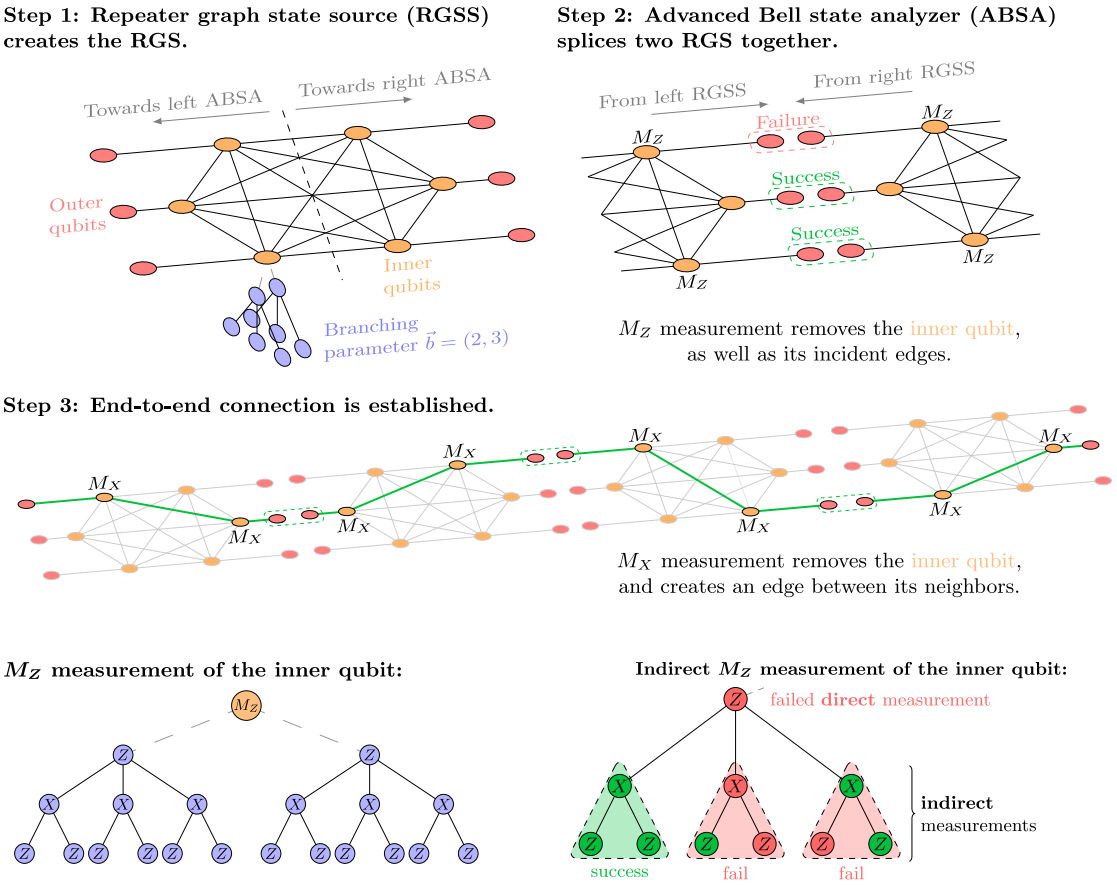


Figure 5.3: An overview of the RGS scheme. The three steps shown here have corresponding analogs in memory-based repeaters. Inner qubits play the role of quantum memories, while outer qubits act as flying qubits. Step 1 (at RGSS) mirrors entanglement swapping, though no specific inner qubit pairs are selected. Step 2 (at ABSA) generates link-level entanglement via Bell state measurements (BSMs) on outer qubit pairs. The Z and X basis measurements in Steps 2 and 3 effectively determine which inner qubit pairs are swapped. Logical Z measurements are indicated at the bottom. Z/X labels on blue qubits refer to physical measurement bases; logical X measurements swap these roles. The triangle structures in the bottom right show how indirect Z basis measurement is possible if photon loss occurs. (From [45].)

an XX measurement. This is repeated for all m pairs. Since BSMs implemented by linear optics are inherently probabilistic and photons may be lost in transit from RGSS to ABSA, the success of each measurement is uncertain.

If *at least one* BSM succeeds at an ABSA, the protocol proceeds to the second stage. Z basis measurements are performed on the inner qubits connected to outer qubits where the BSM failed. This is because a Z measurement effectively removes a vertex (and its edges) from the graph state. If *no* BSMs succeed at an ABSA node, the connection attempt fails and must be restarted. Inner qubits connected to successfully measured outer qubits are also Z measured, except for one surviving pair per ABSA. This procedure prunes the graph, yielding a *linear chain graph state* (as shown with green path in Step 3 of Fig. 5.3).

In the third and final stage (Step 3), the remaining inner qubits, now arranged in a linear chain, are measured in the X basis. Since only two inner qubits remain per ABSA, these measurements effectively implement XX measurements. The final result is a *Bell pair* (i.e., a two-qubit graph state) shared between the two end nodes, up to local Pauli corrections.

Having discussed the conceptual foundations of the RGS scheme, we now turn to how it mitigates photon loss and operational errors—features that qualify it as a third-generation quantum repeater, albeit without support for hop-by-hop, packet-switching-style connections.

Robustness of the RGS scheme: encoding and measurement of inner qubits

The RGS scheme incorporates several mechanisms to address photon loss during link-level entanglement generation and to mitigate operational errors. A primary strategy against photon loss in the Bell state measurements (BSMs) on the RGS's outer qubits is the use of multiple parallel "arms," a parameter denoted by m. Increasing m enhances the probability that at least one BSM succeeds at each advanced Bell state analyzer (ABSA), a necessary condition for establishing an end-to-end entangled connection.

Successfully measuring the outer qubits, however, is only one aspect of the challenge. The inner qubits of the RGS, which are also photonic and therefore susceptible to loss, must undergo successful measurement to complete the entanglement swapping protocol and prune the graph state into a final end-to-end Bell pair. To guarantee arrivals of these vital inner qubits, the RGS scheme typically employs the *loss-tolerant tree code encoding* [318]. In this encoding, each logical inner qubit is represented by a tree-like graph state composed of multiple physical photonic qubits. The specific structure of this tree, defined by its branching parameters $\vec{b} = (b_0, b_1, \ldots, b_{n-1})$, provides redundancy. This redundancy, in conjunction with the number of arms m, determines the overall loss and error tolerance of the RGS scheme. An illustrative example of such tree-encoded inner qubits can be seen in Fig. 5.3.

The logical state of these encoded inner qubits is determined by performing a pattern of single-qubit measurements on their constituent physical qubits. All physical qubits at the same level within a tree are measured in the same basis, with the measurement basis (e.g., X or Z) alternating between successive levels, as depicted in Fig. 5.3. For instance, to perform a logical measurement in the X basis (Z basis), we measure the first-level physical qubits in the X basis (Z basis), the second-level physical qubits in the Z basis (X basis), and so on, alternating the basis for each subsequent level. The logical outcome of the measurement is then inferred from the parity of a specific subset of these physical measurement outcomes. A logical Z basis measurement outcome is derived from the parity of eigenvalues from all physical qubits in the first level of the tree, whereas a

logical X basis outcome is determined by the parity of an X basis measurement on one first-level qubit combined with Z basis measurements of its neighbors at the second level.

Even if a physical qubit is lost, its measurement result can sometimes still be deduced through a technique known as *counterfactual or indirect measurement* [318]. This principle, rooted in the stabilizer formalism of graph states, allows the intended measurement outcome of a physical qubit a (which is part of a stabilizer generator g_i) to be inferred if all other physical qubits in the support of g_i have been successfully measured, either directly or indirectly. The inherent redundancy of the tree code thereby ensures that the logical outcome can often be consistently determined despite some physical qubit losses. For details of how logical operators are defined, readers are referred to the supplementary material of [2].

Beyond mitigating photon loss, the tree encoding also offers a degree of resilience against quantum operational errors that affect physical measurements. For instance, errors in physical measurement outcomes in the Z basis on a qubit at level k within the tree can often be corrected by performing a majority vote. This vote utilizes the indirect measurement results for that qubit, which are inferred from its neighboring qubits at level $k + 1$ and, in turn, their neighbors at level $k + 2$.

Considering the *logical* measurement outcomes, results from logical measurements in the X basis are amenable to correction through majority voting. Since the logical X outcome is determined by the parity of a measurement in the X basis on a single first-level physical qubit and the measurements in the Z basis of its neighbors in the second level; if no errors occur, any such valid grouping for parity calculation will yield the same logical result, allowing for error detection and correction via a majority vote across these possible derivations. In contrast, logical measurements in the Z basis, which depends on the combined parity of *all* physical measurements in the Z basis on the first-level qubits, are not directly correctable by a similar majority vote mechanism at the logical level. However, the individual physical measurements in the Z basis contributing to this logical outcome can still benefit from the aforementioned indirect measurement and majority voting techniques at the physical qubit level. Consequently, the overall error tolerance of the RGS scheme is a function of both the specific tree code structure and the prevailing photon loss probability.

Transmission order of photons in the RGS

The measurement basis selection of the inner qubits is dependent on the BSM outcome of the outer qubits. As long as the outer qubits arrive at the ABSA before their connected inner qubit, the measurement basis can be determined locally by the ABSA. The physical qubits composing the inner qubits can be sent in any order as long as it is known to

the ABSA beforehand, this is fixed and does not require any delay line to correct the ordering in our work as we will show the generation sequence in section 5.4. We note that even if the photons are lost, assuming that the photons are well separated temporally and time tagged, the ABSA can deterministically flag the loss event. This well-separated assumption is also commonly adopted in memory-based repeater schemes for multiplexing photons from multiple memories into a single fiber [40].

5.3 Half-RGS building block

While the original repeater graph state (RGS) connects its inner qubits in a complete graph, such full connectivity is not strictly necessary for entanglement swapping. Prior work has demonstrated that a complete bipartite structure (biclique), formed by linking two halves of the inner qubits, is sufficient to achieve the same functionality [319, 320], as illustrated in Fig. 5.4. This relaxation preserves the core capabilities of the RGS while potentially reducing the complexity of its generation [77, 321], and may also lower error rates given the same baseline fidelity of photon sources and detectors.

This insight motivates our first and most central contribution: the *half-Repeater Graph State (half-RGS)*—a modular, emitter-friendly building block for constructing biclique RGSs. The half-RGS is more than just a simplification, it is the foundation upon which nearly all contributions in this chapter are built. It directly addresses key engineering and architectural challenges in the RGS framework: it simplifies photonic generation procedures, enables clean timing synchronization, aligns naturally with end-node constraints, and allows direct interfacing with memory-based repeaters. Moreover, it introduces new possibilities for entanglement purification and modular scalability (section 5.7).

In this section, we describe the structure and capabilities of the half-RGS, and demonstrate how this modular abstraction leads to practical, scalable designs in both pure all-photonic and heterogeneous quantum networks—between all-photonic and memory-based repeaters. The modularity of the half-RGS building block offers significant advantages in addressing practical engineering challenges for RGS schemes, particularly in optimizing end-node resource requirements, simplifying timing synchronization, and facilitating the integration of all-photonic segments with memory-based repeater networks.

The half-RGS construct

The *half-RGS*, as illustrated in Fig. 5.4, is a multipartite entangled state derived from one half of a standard Repeater Graph State (RGS) with a key structural modification. A stationary quantum emitter or memory qubit, referred to as the *anchor qubit* (shown in purple), is introduced and becomes entangled with the photonic portion of that half, which consists of a set of inner qubits and their corresponding outer qubits. At a

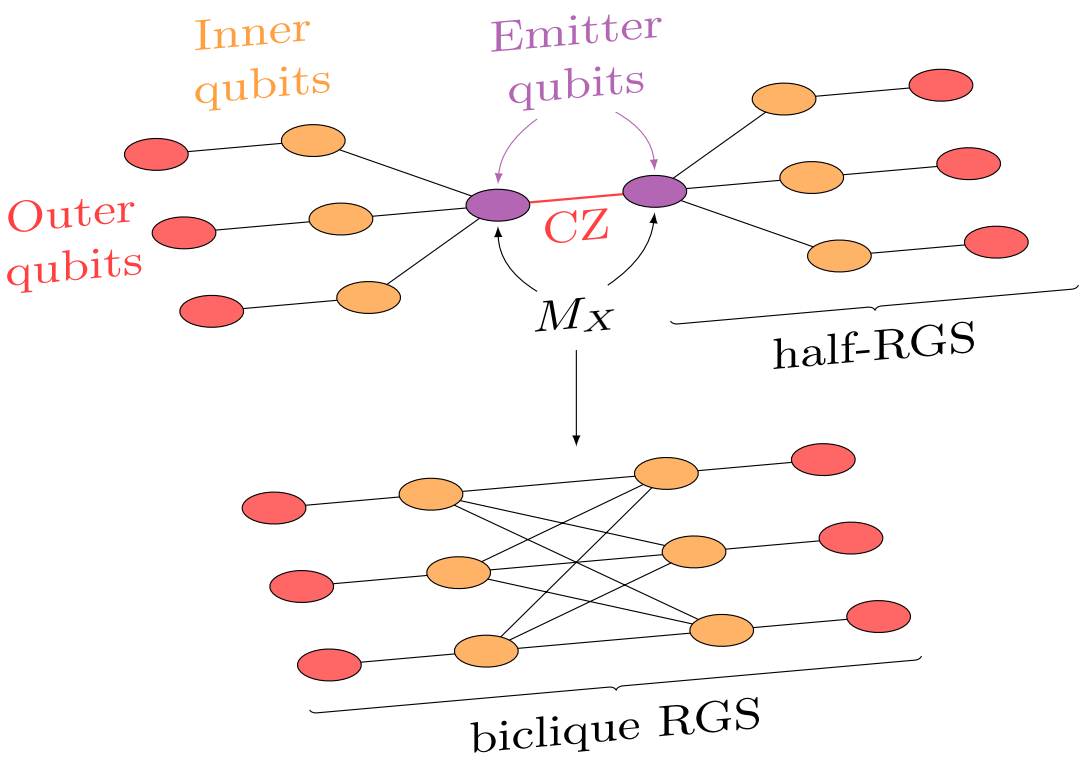


Figure 5.4: An example of a half-RGS (left and right structures) and the transformation of two such half-RGSs into a biclique RGS (center). The anchor emitter qubits (blue) of the two half-RGSs are joined via an application of a CZ gate followed by an XX measurement, resulting in the biclique RGS suitable for entanglement swapping. (From [45].)

given repeater graph state source (RGSS) node, two such half-RGSs can be generated independently and then deterministically fused by applying a CZ gate between their anchor qubits followed by an XX measurement (i.e., a rotated Bell measurement). This operation combines the two halves into a full biclique RGS, hence the name "half-RGS." As we will show next, this modular structure addresses several longstanding challenges associated with the broader RGS framework [46].

Improving resource requirements at end nodes

The integration of compute-capable end nodes into the RGS scheme, extending beyond simpler measurement-only nodes for QKD, has not been fully explored. These compute nodes require quantum memories to store entangled states while awaiting classical messages from distant ABSAs and RGSSs before Bell pairs are ready for use. One approach, explored in [322] (and depicted at the top of Fig. 5.5), necessitates that end nodes be equipped with at least m (the number of RGS arms) emissive quantum memories per trial. To sustain high trial rates, comparable to the RGS scheme's full potential, an additional rm memories must be held in reserve. This is because photons from previous trials would have already been emitted while the system awaits correction

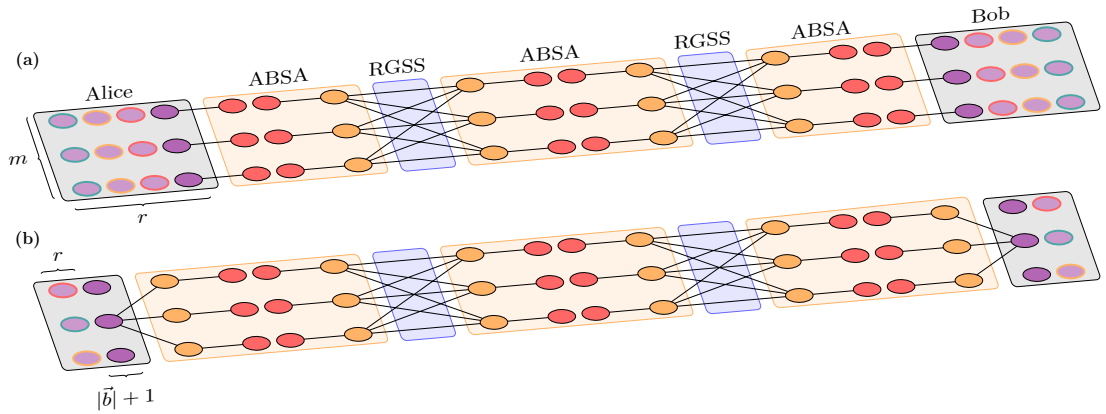


Figure 5.5: Architectures supporting the RGS scheme where the photonic states generated at end nodes are different. (a) The architecture proposed in [322], where end nodes require m emissive memories for each trial, along with rm idle memories that have already been utilized in prior trials and are awaiting messages from all ABSAs. (b) Proposed architecture, featuring end nodes equipped with half-RGS building blocks. In this setup, end nodes require $|\vec{b}|+1$ quantum emitters (where $|\vec{b}|$ represents the depth of the tree encoding plus one), while reserving r memories awaiting notification messages from prior trials. Purple circles with black borders represent the quantum emitters utilized in the current trial. Purple circles with colored borders in both (a) and (b) indicate idle memories awaiting messages from ABSAs, with the same colors denoting memories participating in the same trials. (From [45].)

messages, where r denotes the ratio of classical message travel time from the farthest ABSA to the RGS generation time. For instance, based on the analysis in [50] for a 1000 km separation between end nodes, each end node would require approximately 150 quantum memories to sustain high-rate operation, assuming $m = 15$ arms and a message delay ratio of $r = 10$.

In contrast, our proposed architecture leveraging the half-RGS building block (shown at the bottom of Fig. 5.5) streamlines this requirement. An end node would need only r quantum memories (for storing states from previous trials awaiting classical corrections) and $|\vec{b}|+1$ emitters (where $|\vec{b}|$ is the depth of the tree encoding) for generating its local half-RGS, resulting in a total of $r+|\vec{b}|+1$ combined memory and emitter units. Crucially, the r quantum memories do not need to be emissive if the state of the half-RGS anchor emitter can be swapped into them. Furthermore, the emitters themselves do not require coherence times as long as those of quantum memories, given their role is primarily the generation of photonic qubits for the half-RGS. This configuration can effectively double the entanglement trial rate compared to the rm emissive memories approach (or $2rm$ if multiplexed fiber is used to match our rate) discussed in [322], as illustrated in Fig. 5.5.

While this approach does imply that end nodes must be equipped with half-RGS generation capability, the hardware is identical to what is already required at repeater nodes,

making the assumption reasonable. A more end-node friendly approach would be that end nodes connected to a repeater bridging the RGS scheme and the port connecting to end nodes can be memory-based.

While this strategy implies that end nodes themselves must be equipped with Half-RGS generation capabilities, this requirement can be considered reasonable given that such hardware is functionally identical to that already necessary for the RGSS nodes within the all-photonic segments. Alternatively, for end nodes lacking this specific generation hardware, a more accommodating approach involves utilizing a memory-equipped repeater node to act as an interface. Such a repeater would bridge the all-photonic RGS segment to a conventional memory-based quantum link terminating at the end node, thereby allowing simpler, memory-based end nodes to connect to networks that leverage advanced RGS technology (see also section 5.3).

Simplifying timing synchronization.

For entanglement swapping via Bell state measurements (BSMs) to succeed, the involved photons must arrive at the measurement station within an indistinguishable time window. This requirement is universal across all quantum repeater schemes and stems from the physics of linear-optical two-photon interference, not from the design of any particular architecture.

What makes the RGS scheme uniquely challenging, however, is that the full RGS is typically generated by a single set of emitters [49, 50]. In this configuration, both halves of the photonic graph—destined for neighboring nodes on either side—are entangled with the same stationary qubits and must be emitted in a tightly coordinated sequence. For the outer photons to arrive simultaneously at their respective Bell state measurement stations (ABSAs), the emitter must coordinate timing with both neighboring nodes. This means that a single RGSS node must synchronize with both its left and right ABSAs, implicitly creating a dependency chain across the full connection path. Timing messages must propagate along the entire route, turning synchronization into a global coordination problem. This requirement imposes significant engineering complexity and hinders scalability [323].

Our half-RGS approach resolves this issue by introducing modularity at the level of RGS generation. Instead of creating both sides of the RGS from a single set of emitters, the biclique RGS is formed from two independently generated half-RGSs—each constructed by its own set of emitters (different QNICs) within the same RGSS node. This design localizes timing control as synchronization is now reduced to standard neighbor-to-neighbor coordination, just as in conventional memory-based repeater schemes. Timing messages are only exchanged between adjacent nodes and the local controller managing emitter operations, eliminating the need for global synchronization across the entire

connection path.

Integration with memory-based repeaters through virtualization of links

The integration of all-photonic repeaters—particularly the RGS scheme—into memory-based quantum repeater networks remains an underexplored but essential step toward scalable and heterogeneous quantum networks. By reframing an entire RGS segment as a single virtual link-level connection between two memory-equipped devices—rather than as a chain of network-level repeaters—we enable its seamless use within existing memory-based protocols, as illustrated in Fig. 5.6. This abstraction allows the Bell pairs generated over the RGS chain to serve as fundamental link-level resources, similar to those produced by memory-based links like MIM, MSM [119], or even Sneakernet [120].

Once generated, these Bell pairs can be directly handed off to the control stack (e.g., QRSA and RuleSets) at the terminating memory-equipped nodes, allowing seamless integration with existing entanglement management protocols. This abstraction aligns with the design principles of the future Quantum Internet—modularity, heterogeneity, and interoperability—by allowing RGS links to be treated identically to other link technologies within connection setup and resource scheduling processes, as discussed in chapter 4.

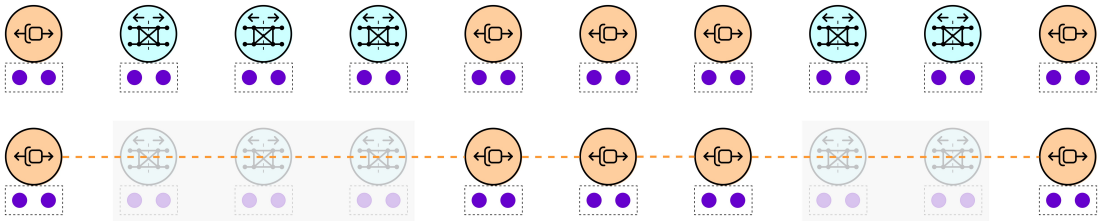


Figure 5.6: Segments constituting RGSS (repeater graph state source, in light blue) and ABSA (advanced Bell state analyzer, omitted) nodes are treated as virtual links to memory-equipped repeaters (orange). The RGS scheme link is invisible at the connection-level protocol, reducing the cost of connection setup and management. (From [45].)

This virtual link abstraction fits naturally within the QRNA framework described in chapter 4, where RGS-based subpaths can be treated as single hop-level links between memory-equipped nodes. At the boundaries of these subpaths, RuleSets and connection messages are translated into the lower-level RGS protocol—mirroring the internetworking process outlined in section 4.6. Conceptually, an RGS subnetwork may function like a router, due to its internal coordination, or like an optical switch (OSW), owing to its abstraction of link-level and tight real-time signaling requirement. While implementation specifics are left for future work, the key insight here is that encapsulating all-photonic repeater chains as virtual links is both technically viable and architecturally beneficial, paving the way for unified protocol designs across heterogeneous quantum

technologies.

5.4 Generation of half-RGS

Generating photonic RGS remains one of the most challenging aspects of implementing the RGS scheme. Numerous generation methods have been proposed and refined over time [2, 49, 324–326], alongside optimization techniques that reduce resource consumption and generation time [77, 321, 327].

In this work, we adopt the approach of Buterakos et al. [49], which enables the deterministic generation of photonic graph states via quantum emitters.

To support deterministic generation of half-RGS using emitter-based schemes, we begin by clarifying the operational assumptions and primitive transformations used in our construction.

Assumptions on the operations

We assume the following capabilities for the quantum emitters, summarized in Fig. 5.7. Emitters are arranged in a linear topology $(Q_a, Q_e, Q_0, Q_1, \ldots, Q_{|\vec{b}|-1})$, where Q_a is the anchor qubit and Q_e is the emitter for outer qubits. As a reminder, anchor qubits refer to emitters that are part of the half-RGS and serve to anchor photonic qubits (Fig. 5.4). Emitters support single-qubit Hadamard gates and controlled-phase gates between adjacent emitters. Photon emission is modeled as a controlled-Not operation to a newly initialized qubit in the $|0\rangle$ state (Fig. 5.7(a)). While we model this emission as deterministic for the purposes of the protocol logic, we acknowledge that no physical emitter can achieve perfect quantum efficiency. Any inefficiency in single-photon emission is treated as a form of initial photon loss, a type of error that the RGS scheme is inherently designed to tolerate.

In addition to the standard graph state manipulation rules, we use *push-out operations*, where an emitter is effectively replaced by its emitted photon and then disconnected from the rest of the graph [328, 329]. This is implemented by emitting a photon followed by a Hadamard gate, as shown in Fig. 5.7(b).

The Bell-state measurement (BSM) employed in our protocol is equivalent to applying a controlled-phase gate between the two qubits followed by an XX measurement. Due to limitations of linear optics, the BSM succeeds with probability 50%, which we model by treating odd-parity outcomes (i.e., 01 or 10) as success. This choice is arbitrary; any two of the four Bell basis outcomes would suffice.

With these assumptions in place, we now describe the full generation sequence for a half-RGS state with arbitrary branching parameters and multiple arms.

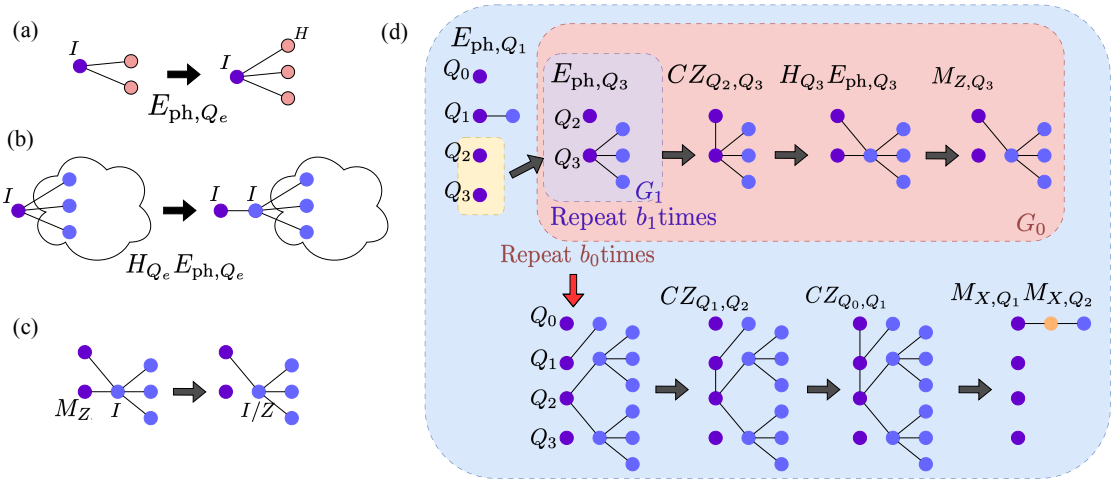


Figure 5.7: (a) Photon emission with side effects labeled. (b) Push-out operation: photon emission followed by a Hadamard gate, introducing no side effect. (c) Push-out followed by Z measurement, the main sequence used in half-RGS generation. (d) Generation sequence (adapted from [50]) for one arm of half-RGS with $\vec{b} = (2, 3)$. Repeat m times to construct m arms. (From [45].)

Sequences for generating the half-RGS

To generate a half-RGS with m arms and branching vector $\vec{b} = (b_0, b_1, \ldots, b_{n-1})$, anchored at Q_a, we use a sequence adapted from [49, 50], as also illustrated in Fig. 5.7(d), which is given by

$$\begin{aligned} &\left(M_{X,Q_e}M_{X,Q_0}CZ_{Q_a,Q_e}CZ_{Q_e,Q_0}G_0^{b_0}E_{\mathrm{ph},Q_e}\right)^m, \\ &\text{with } G_k = M_{Z,Q_{k+1}}H_{Q_{k+1}}E_{\mathrm{ph},Q_{k+1}}CZ_{Q_k,Q_{k+1}}G_{k+1}^{b_{k+1}} \\ &\text{and } G_{n-1} = E_{\mathrm{ph},Q_{n-1}}, \end{aligned} \tag{5.6}$$

where E_{ph,Q_i} denotes photon emission from emitter Q_i, and M_{P,Q_i} denotes measurement in basis P on emitter Q_i.

Joining two half-RGS states into a full RGS is achieved by applying a CZ gate between the two anchor qubits (Q_a) and performing an XX measurement. Although our generation sequence requires more emitters than some prior approaches, it simplifies both timing and resource management at ABSA nodes. In particular, photon emission proceeds in the same order as measurements, which reduces the need for photon storage and mitigates loss in optical fibers.

Compared to previous methods [50], our design uses twice as many emitters per RGSS and two additional emitters for generating the outer qubits. However, this cost is offset by significantly improved performance. Our scheme cuts the generation time in half and eliminates the need for optical delay lines or efficient photon storage at ABSA nodes, as required in [49, 50, 321, 322]. All photons are emitted in the same order as their corresponding measurements, ensuring minimal travel distance through fiber and reducing susceptibility to loss.

While this construction ensures deterministic generation and efficient timing, it introduces certain Clifford side effect operators during entangling gates and measurements. These side effects must be tracked and corrected to ensure the proper formation of Bell pairs, as discussed next.

Side effects created during the generation

The generation of half-RGS states and the subsequent measurements at ABSA nodes introduce Clifford side effects that must be carefully tracked and corrected to ensure proper state preparation. During generation, leaf nodes corresponding to the inner and outer qubits inherit an H side effect (Fig. 5.7(a)). These can be corrected either optically, by applying Hadamard gates to the emitted photons, or logically, by swapping the X and Z measurement bases at ABSA.

Beyond the H side effects, additional Z side effects may arise. In particular, qubits produced via a push-out operation followed by Z basis measurement of the emitter can acquire a Z side effect with 50% probability (Fig. 5.7(c)). Outer qubits may also pick up Z side effects from the XX measurement used to attach them to the inner qubits.

Further side effects are introduced when joining two half-RGS states to form a biclique RGS. This step involves a CZ gate and an XX measurement between the anchor qubits (Fig. 5.4), which can collectively toggle the Z side effect on all physical qubits in the first tree level (Fig. 2.1, bottom right). There are two ways to handle these effects: either (1) track a logical Z side effect on the inner qubit or (2) track the side effect status of each individual physical qubit. We adopt the second strategy—toggling the physical qubits—when describing the general protocol in section 5.5, and the first—tracking logical Z—when analyzing the protocol tailored to our half-RGS design in section 5.5.

5.5 The full RGS protocol

We now describe the communication protocol to realize the RGS scheme given our assumptions discussed above. First, we will describe a biclique RGS generation agnostic protocol outlining where information about measurement and side effects originate and how they should be passed to intermediate and end nodes. Later, in the following subsection, we will go into the generation specific with our proposed half-RGS interface leveraging graph state manipulation rules to argue about the calculations of Pauli frame corrections in arm-by-arm manner and hop-by-hop manners.

Protocol overview

A communication protocol for the RGS scheme must support fast generation and soft real-time processing. In particular, ABSAs must be able to determine measurement bases based solely on the arrival order of photons, without waiting for classical instructions from the RGSS. At the same time, the protocol must ensure that Pauli frame

corrections are correctly tracked—even as the number of measurement results scales with the size of the RGS—so that end nodes can deterministically correct the final entangled state.

In each trial, the RGSS sends an RGS to its adjacent ABSA, followed by a classical message detailing side effects associated with each physical qubit. These side effects—such as byproduct Pauli Z operators introduced during graph state generation—are determined at the RGSS regardless of the specific generation method used [2, 49, 321, 324, 325]. Importantly, in some architectures (e.g., full RGSs with complete graphs), certain generation methods may impose additional side effects on emitted photons. For example, if the emitter qubit is measured in the Y basis as in [50], then the first photon in a tree-encoded arm must be measured in the Y basis rather than X. This means ABSAs must be aware in advance of the exact form of side effects that may appear on each photon, in order to choose the correct measurement basis. Hence, side-effect information must be communicated ahead of time or synchronized on a per-trial basis between adjacent RGSS–ABSA pairs.

Once the RGS reaches the ABSA, Bell-state measurements (BSMs) are performed on the outer qubits. The outcomes of these BSMs determine the measurement bases of inner qubits, and the ABSA proceeds to measure them accordingly. To process these outcomes, each ABSA constructs a measurement tree per RGS arm, recording measurement outcomes and bases (or marking qubits as lost). This tree mirrors the logical tree encoding and enables the ABSA to determine the logical measurement result locally. The Z side effects from generation, along with the measurement tree data, can then be compressed into just two classical bits per ABSA per trial: one indicating whether the logical X result should be flipped, and another indicating whether decoding was successful. These bits are sent to the end nodes, which then apply the appropriate Pauli corrections to recover a deterministic Bell pair.

Generation-agnostic Pauli frame calculations

We now explain how Pauli frame corrections can be computed in a generation-agnostic manner—without relying on the internal steps of RGS creation. This tracking is a crucial but often overlooked part of RGS-based protocol design, as many works focus solely on the final entangled state up to local Clifford equivalence, sidestepping the operational details of how Pauli frame information is extracted and propagated during execution.

While photons are physically measured in the order they arrive at the ABSA, we can conceptually treat them as being measured in any order, provided each outer qubit is measured before its corresponding inner qubits. We now describe how to compute the resulting Pauli frame for the memory qubits at the end nodes. Fig. 5.8 outlines the procedure below.

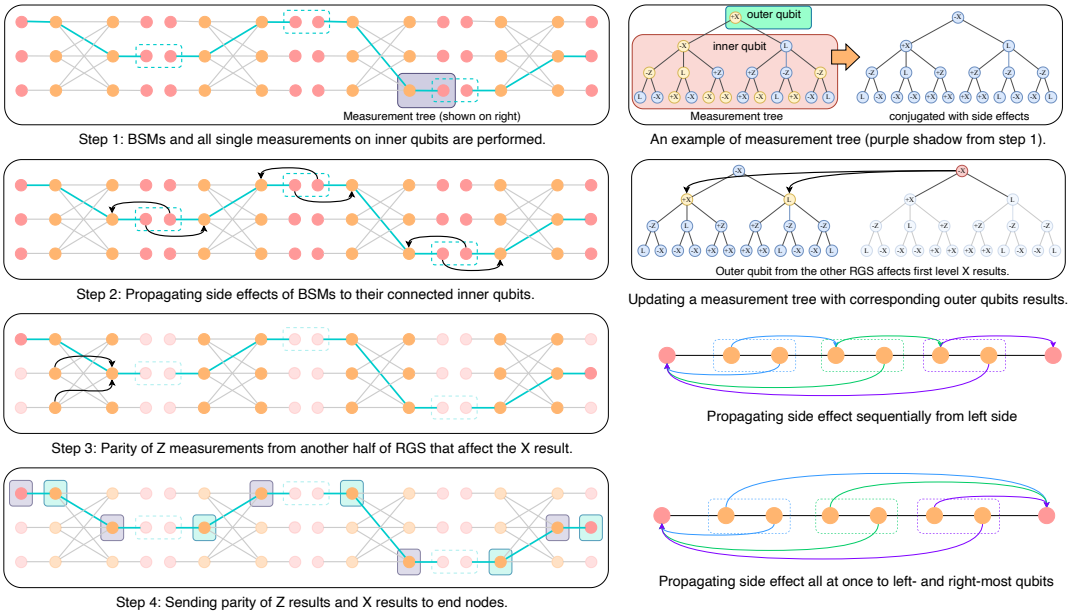


Figure 5.8: The tracking of side effect propagation at each step, the resolution of physical and logical measurement results, and the rationale behind the Pauli frame corrections applied at end nodes, as illustrated in Steps 1–4 on the left side of the figure. Examples of measurement trees and their updating during the Pauli frame calculation process are presented in the top two panels on the right side. Yellow-filled vertices represent qubits with a Z side effect. The sequential propagation of side effects, viewed hop-by-hop, is contrasted with an equivalent view of simultaneous propagation to the outermost qubits, depicted in the bottom two panels on the right side. (From [45].)

First, we assume that the success or failure of all BSMs on the outer qubits is known, so the measurement bases for all inner qubits are determined. Before we begin tracking state changes due to measurements, each ABSA constructs a measurement tree for each inner–outer qubit pair in every RGS arm. This tree includes the side effect information about all physical qubits, as provided by the RGSS and end nodes. The tree stores the measurement basis and result, or marks the qubit as lost. Measurement outcomes are denoted by + (for 0) or − (for 1), along with the basis (X or Z); loss events are marked by L. If a qubit has a Z side effect from generation (shown as a yellow-filled vertex in Fig. 5.8), then its X basis measurement result is flipped, while Z basis results and loss entries remain unchanged.

Next, we propagate the BSM results obtained from the outer qubit pair to their connected inner neighbors, as shown in Step 2 of Fig. 5.8. A BSM has the same effect as an XX measurement in Fig. 2.1 and may introduce a Z side effect. Successful BSM outcomes propagate across the RGSs to the inner qubits, indicated by the black arrows in Step 2. If the measurement outcome on the outer qubit from the left (right) is 1 (i.e., −X in the figure), then the measurement results of the first-level physical qubits in the right (left) tree are flipped. We do not propagate Z side effects to inner qubits whose outer neighbors were part of a failed or discarded BSM, since these inner qubits are measured

in the Z basis and Z side effects do not affect Z measurement results.

The results of Z measurements on the inner qubits of one half of the RGS affect the X outcomes of the other half; i.e., the two halves sent to different ABSAs. However, these measurements are performed at separate ABSAs, making their outcomes locally unavailable. This cross-arm dependency is illustrated in Step 3 of Fig. 5.8. Fortunately, as shown in the bottom two panels on the right side of the figure, the Z propagation from XX measurements can be handled at the end nodes via the parity product of measurement outcomes. This same logic applies to any Z side effect introduced by inner qubits measured in the Z basis. Each ABSA computes the parity of its local Z side effects (i.e., the product of the individual Pauli Z) needing corrections, and sends this partial parity to the end node, which combines them to determine the total parity.

Therefore, each ABSA needs to send only two bits of information to the end nodes per trial: one bit indicating whether a Z side effect propagates to that side (based on the parity), and one bit indicating whether the procedure was successful. The number of classical bits that each ABSA receives is equal to the number of photons it measured. We will discuss this total amount of classical bits to be sent in section 5.5

Half-RGS specific Pauli frame calculation

We now show how our proposed half-RGS building block, generated using quantum emitters, enables straightforward calculation of Pauli frame corrections and naturally leads to a simple graphical method for determining the end-to-end fidelity of the protocol (discuss in section 5.6). As in the generation-agnostic approach previously described, we begin with Step 1 of Fig. 5.8, but the subsequent steps differ.

With our construction of half-RGS blocks, all photons are generated in the same order they are sent and measured at an ABSA, eliminating the need for optical delay lines. Each arm of the half-RGS is created sequentially, which allows us to track and propagate Z side effects of each arm to the anchor qubit during the generation process. Importantly, the half-RGS anchor qubit never emits photons, and the only operation performed on it—namely, the CZ (used to join and form the biclique RGS at the final step)—commutes with Z side effects. Thus, this tracking approach is sound.

First, let us define the qubits involved in this one-hop picture, as shown in Fig. 5.9 (corresponding to the second-to-last step of Fig. 5.7, where tree-encoded physical qubits are represented as a single logical orange qubit). Let Q_a^L, Q_e^L, Q_0^L, Q_{out}^L, and Q_{in}^L denote the anchor qubit of the half-RGS, emitter for outer qubits, anchor of the inner logical qubits, outer qubit, and inner logical qubit, respectively, for the left side incoming to the ABSA. Similarly, Q_j^R denotes the corresponding qubits for the right incoming side.

Recall the side effects introduced by the generation sequence of the half-RGS as de-

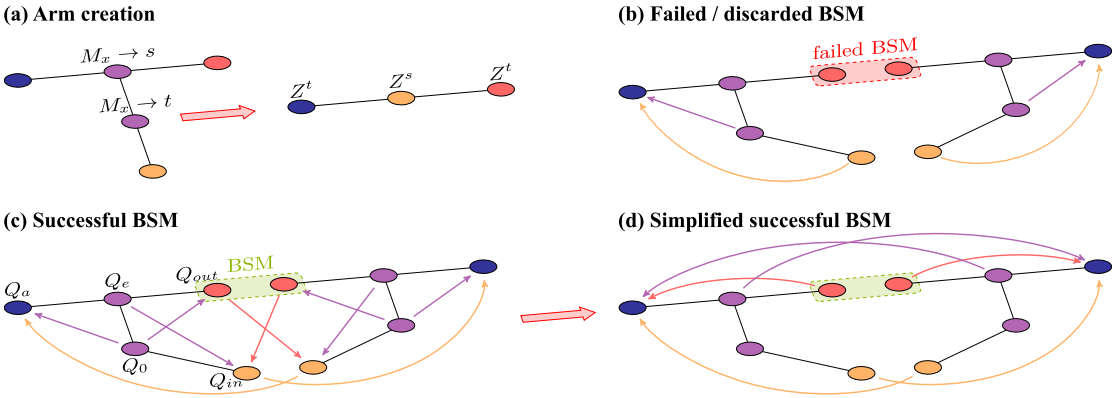


Figure 5.9: Side effect propagation for an arm of a one-hop RGS link from half-RGS. An arrow from a to b indicates that if the measurement result of a is 1, the result of b should be flipped. (a) Side effect propagation when creating an arm attached to anchor qubits via XX measurements on emitters joining inner and outer qubits. (b) Propagation for failed or discarded BSMs, where inner qubits are measured in the Z basis. (c) Side effect propagation for a successful BSM arm, with inner qubits measured in the X basis. (d) A simplified view of (c), where the tips of arrows are moved to anchor qubits by following the arrow chain.

scribed in section 5.4. Graphically (see Fig. 5.9), if a measurement in the RGS protocol introduces a side effect (flipping the result of another qubit's measurement) when the outcome is "1," we draw an arrow from the measured qubit to the affected one. The arrows from inner and outer qubits correspond to the view in the last step of Fig. 5.7, after Q_1 and Q_2 have been measured out. The arrows from Q_0 and Q_1 follow from the generation sequence in the same figure, where inner and outer qubits become part of the RGS arm attached to the anchor.

The figure depicts two measurement patterns: one where inner qubits are measured in the X basis and another where they are measured in the Z basis (left and right halves, respectively). From this graphical picture, we observe that a single flipped measurement result propagates through the chain of arrows and always affects an anchor qubit. This lets us simplify the entire effect of each measurement to a possible flip of the Z side effect on only the anchor qubits (lower half of Fig. 5.9).

By repeating this process for each arm of a one-hop RGS link between two half-RGSs, we can compute the full Pauli frame corrections. To generalize this to a multi-hop linear chain of RGS links, recall that at each RGSS node, the two halves of the half-RGS are joined to form a biclique RGS using a CZ gate followed by XX measurements. Because the tracked Z side effect in the one-hop picture commutes with CZ and flips X basis measurement outcomes, we simply flip the reported measurement result if the anchor has a Z side effect. These measurement results, which are then sent to the end nodes for post-processing, indicate the Pauli frame corrections—similar to Step 4 in Fig. 5.8.

One-Stage vs Two-Stage Pauli frame correction

An important aspect of RGS protocols is how classical information from measurements is processed to determine the Pauli frame corrections at the end nodes. Each measurement generates a classical bit and may introduce side effects that must be tracked. We now compare two strategies for processing and communicating this classical data: the conventional *One-Stage Correction Method*, and the *Two-Stage Correction Method* we propose (section 5.5).

In the conventional One-Stage Method, implied in most literature [2, 50], every intermediate node—including all Advanced Bell State Analyzers (ABSAs) and Repeater Graph State Sources (RGSSs)—forwards all measurement outcomes to the end nodes without performing any local post-processing. This includes results from every physical photon measured along the entire RGS chain and their side effects during the generation, potentially numbering in the hundreds or thousands depending on the encoding parameters (e.g., $m = 14$, $\vec{b} = (10, 5)$) per RGS. The end nodes then process this full dataset to reconstruct the logical qubit outcomes and calculate the required Pauli frame corrections. While conceptually straightforward and easier to implement, this approach incurs significant classical communication overhead and places a heavy processing burden on the end nodes, which must interpret the entire history of side effect propagation.

In contrast, our Two-Stage Method introduces local pre-processing at each ABSA. Rather than transmitting all raw photon measurement results, the ABSA collects and reduces this data—along with any necessary side-effect tracking from neighboring RGSSs—down to a small number of summary bits. Specifically, each ABSA computes and forwards only two bits: the effective parity of Z side effect for end nodes and a bit inidiciating the success of the procedure. These summary bits are then passed along the chain, and the end nodes use this compressed information to determine the final Pauli corrections.

This distributed correction strategy has two major advantages. First, the end nodes receive and process far fewer classical bits—only two bits per ABSA instead of all individual measurement results—resulting in a much lighter computational load. Second, the total classical communication volume across the network is significantly reduced. As illustrated in Fig. 5.10, the Two-Stage Method cuts down the number of bits the end nodes must process by up to three orders of magnitude compared to the One-Stage Method. The constant and relatively small processing load per ABSA, shown as the red dashed line, makes this a scalable and efficient approach.

By optimizing classical side-effect management in this way, our Two-Stage Method improves the practicality of RGS-based quantum repeater protocols, particularly for large-scale quantum networks where classical bandwidth and processing resources are

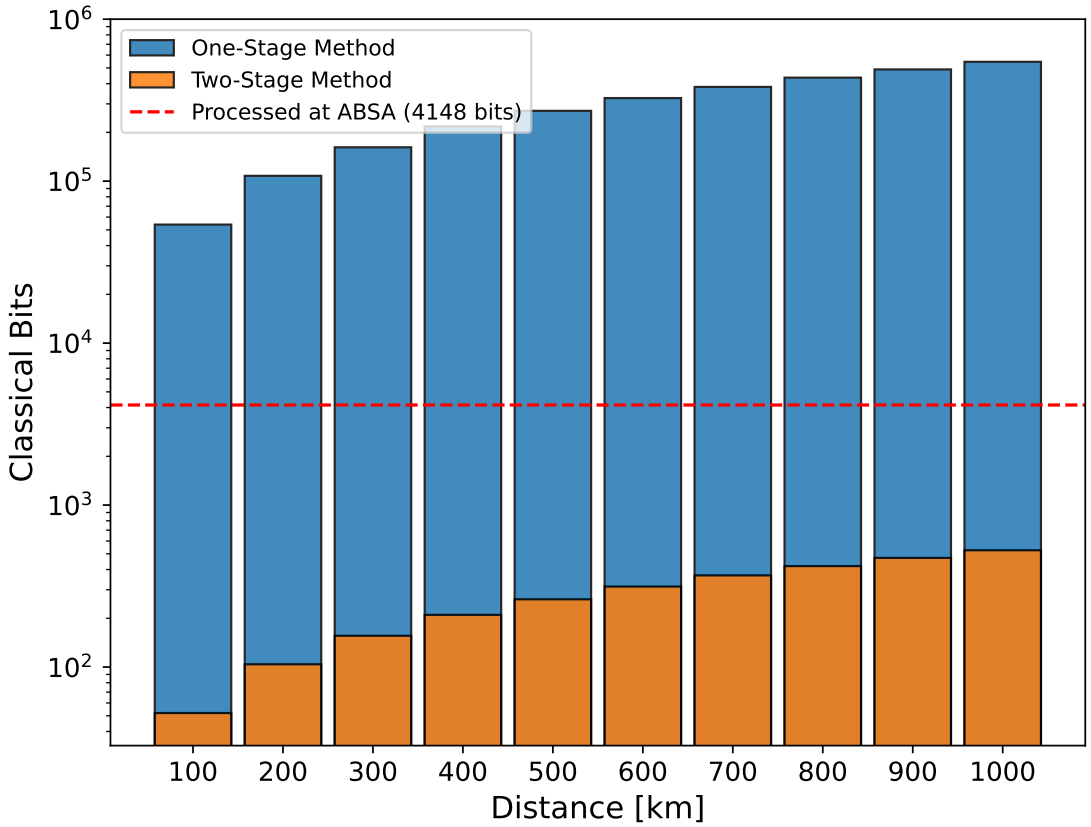


Figure 5.10: The number of classical bits to be processed (equivalent to the total number of photonic qubits involved in generating the RGS arms for one Bell pair attempt) versus the separation distance between two end nodes. This assumes a near-optimal RGS structure with $m = 14$ arms and inner qubit tree-encoding parameters $\vec{b} = (10, 5)$. Blue bars represent the classical bits processed by end nodes in the One-Stage Correction Method. Orange bars show the significantly reduced bits processed by end nodes in the Two-Stage Correction Method. The red dashed line indicates the constant, relatively small number of bits processed by each intermediate ABSA in the Two-Stage Method. (From [46].)

constrained.

Correctness of the protocol and the propagation rules

We validated the correctness of the two protocols, generation-agnostic and half-RGS building block, using the stabilizer tableau simulator Stim [330]. The code for this validation can be found at [331]. In the generation-agnostic simulation, each physical qubit forming an inner qubit is directly created from quantum circuits using Hadamard and CZ gates without measurements. To simulate the stochastic Z side effect on each physical qubit within an inner logical qubit, a Z gate is applied with 50% probability after state generation. Subsequently, inner–outer qubit arms are generated following the procedure in section 5.4 until half-RGSs are obtained and transformed into biclique RGSs. Once all RGSs are generated at each RGSS, along with the two half-RGSs from

the end nodes, BSMs are performed on all outer qubit pairs, and measurement bases are selected for all inner qubits. Finally, all physical qubits in the system are measured simultaneously in a single circuit layer.

Instead of directly simulating message passing, the correction procedure is executed step by step, as outlined in section 5.5. For each RGS arm, a measurement tree is created, followed by updating the measurement tree with side effect information and Z propagation from the BSM, and the decoding of the logical result. Finally, the 4 bits of information at each ABSA are processed, and correction operations are applied to the end nodes' memories accordingly.

For the half-RGS specific protocol—where biclique RGS is created via our proposed half-RGS building blocks—we simulate it hop-by-hop similarly to how we described it. This helps speed up the simulation as the total number of qubits during the simulation is bounded by only a single hop independent of how many hops we want to simulate. All the photons are generated exactly as described in section 5.4, photon-by-photon, confirming that our Pauli frame calculation and also the generation sequence create the right state.

To confirm that the final state between two end nodes is a two-vertex graph state without any side effect, we used Stim to examine the stabilizer generators of the final two qubits. We verified that the stabilizer generators are $X_a Z_b$ and $Z_a X_b$ as expected, where the subscripts a and b denote the qubit held by Alice and Bob, respectively. Simulating the photon loss event is achieved by probabilistically marking qubits as lost with some probability. If a qubit is marked as lost, we randomly apply Pauli X, Y, or Z before the measurement, and the results are then removed from the measurement tree, excluding their participation from the logical qubit decoding process.

5.6 Fidelity derivation of the end-to-end Bell pairs

The fidelity of end-to-end Bell pairs is a key performance metric in quantum networks, quantifying the closeness of the final shared quantum state to an ideal Bell state. In this section, we will derive this fidelity using the single-hop half-RGS architecture, guided by the Pauli frame propagation behavior illustrated in Fig. 5.9.

Error modeling assumptions

Our error model assumes that errors manifest as independent single-qubit Pauli error channels acting on each photonic qubit involved in the RGS. For simplicity, and in line with standard approaches [2, 50, 322], we consider quantum emitters, their associated gate operations, and measurement devices to be ideal and noiseless. This simplification is justified because noise from these components can be effectively approximated as an increased effective noise channel on the emitted photons themselves [332].

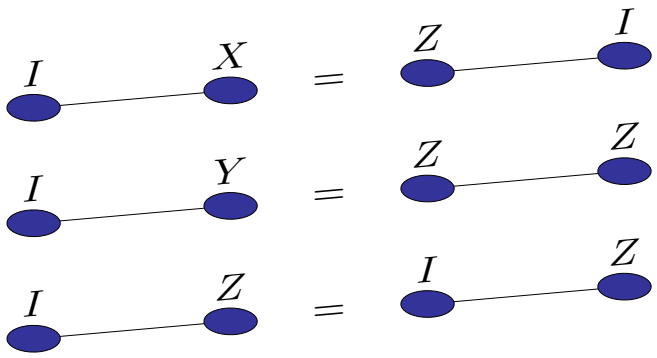


Figure 5.11: The equivalence descriptions of graph states with side effects. The graph states on the left side represent Bell pairs with X, Y, and Z Pauli errors, while the right side shows how we model errors in this work using graph state representations.

While standard approaches for modeling Bell pair errors typically apply Pauli errors (X, Y, or Z) to just one qubit—exploiting the symmetry of Bell states—we adopt a different viewpoint tailored to our protocol. Since our target state is a two-vertex graph state (equivalent to a Bell state), we can equivalently track only Z type side effects accumulating on each of the two end-node qubits. This modeling choice simplifies error propagation across hops without loss of generality.

Indeed, for a two-qubit graph state, these approaches are equivalent up to stabilizer transformations. As shown in Fig. 5.11, a Pauli X, Y, or Z error on Alice's qubit corresponds respectively to a Z error on Bob's, Z errors on both, or a Z error on Alice's qubit—each equivalent under multiplication by stabilizer generators $Z_a X_b$ or $X_a Z_b$. A key insight in our fidelity analysis is that the error probability for a single-hop half-RGS link can be decomposed into two contributions: (i) inner qubit contributions, which are local and uncorrelated between the two end nodes; and (ii) outer qubit contributions, which induce correlated errors across the link. The error on a memory qubit at an end node depends on whether the Pauli frame at its corresponding anchor is flipped. This occurs when an *odd* number of side-effect-inducing errors affect the qubits pointing to the anchor. An even number of such errors cancels out. Thus, the error probability can be computed by evaluating the parity of errors influencing each memory qubit.

Error propagation from outer qubit operations

We first analyze how errors arising from operations on the *outer qubits* contribute to the effective error on the Bell pair after link-level entanglement swapping. These outer qubits participate in Bell state measurements (BSMs) that connect adjacent half-RGS segments, analogous to elementary links in traditional repeater chains. For this analysis, we can conceptually consider all inner qubits to have been measured out first (as measurement order can be rearranged without affecting the final state apart from classical processing of outcomes), allowing us to isolate the BSMs on outer qubits.

Each outer qubit is subjected to an independent single-qubit Pauli error channel. Focusing on successful transmission events (i.e., no photon loss for these qubits), each resulting Bell pair formed by these outer qubits before a BSM can be described by a probabilistic mixture of Pauli errors. In our graph state framework, this is represented by a four-element probability vector $\mathbf{e}$ given by

$$\mathbf{e} = [w, x, y, z]^T \tag{5.7}$$

$$:= w\,|\psi_{II}\rangle\langle\psi_{II}| + x\,|\psi_{ZI}\rangle\langle\psi_{ZI}| + y\,|\psi_{ZZ}\rangle\langle\psi_{ZZ}| + z\,|\psi_{IZ}\rangle\langle\psi_{IZ}|\,, \tag{5.8}$$

where each $|\psi_P\rangle$ is a two-qubit graph state (Bell pair) affected by an effective Pauli error $P \in \{II, ZI, ZZ, IZ\}$, with w, x, y, z being their respective probabilities. When a BSM (specifically, the XX measurement following a CZ gate, as common in RGS protocols) is performed between two such Bell pairs, the output error vector is computed via the transformation given by

$$\mathrm{BSM}_{XX}\left(\begin{bmatrix} w_1 \\ x_1 \\ y_1 \\ z_1 \end{bmatrix}, \begin{bmatrix} w_2 \\ x_2 \\ y_2 \\ z_2 \end{bmatrix}\right) = \begin{bmatrix} w_1w_2 + x_1x_2 + y_1y_2 + z_1z_2 \\ w_1x_2 + x_1w_2 + z_1y_2 + y_1z_2 \\ w_1y_2 + y_1w_2 + x_1z_2 + z_1x_2 \\ w_1z_2 + z_1w_2 + x_1y_2 + y_1x_2 \end{bmatrix}. \tag{5.9}$$

To compute the error across the entire path, we can apply eq. (5.9) iteratively at each hop between the end nodes, ultimately yielding the final error probability vector.

For a special case of $N-1$ hops with the same depolarizing errors for every photons. We can derive the final error probability vector for an end-to-end Bell pair established across a chain of N initial elementary links. Each elementary link is assumed to start with an identical error vector $\mathbf{e}^{(0)} = [w_0, x_0, y_0, z_0]^T$. The end-to-end entanglement is created by performing $N-1$ successive BSMs, each extending the entanglement by one link.

Let $\mathbf{e}^{(k)} = [w^{(k)}, x^{(k)}, y^{(k)}, z^{(k)}]^T$ denote the error vector of an entangled pair that spans $k+1$ elementary links; i.e., the result of k BSM operations. The process is recursive; the k-th BSM operation combines the entangled pair from the previous $k-1$ swaps (with error vector $\mathbf{e}^{(k-1)}$) with a fresh elementary link (error vector $\mathbf{e}^{(0)}$). Using the BSM transformation rule defined previously, this can be expressed as a linear recurrence relation

$$\mathbf{e}^{(k)} = M \cdot \mathbf{e}^{(k-1)}, \tag{5.10}$$

where the transformation matrix M is determined by the components of the initial error vector $\mathbf{e}^{(0)}$ given by

$$M = \begin{pmatrix} w_0 & x_0 & y_0 & z_0 \\ x_0 & w_0 & z_0 & y_0 \\ y_0 & z_0 & w_0 & x_0 \\ z_0 & y_0 & x_0 & w_0 \end{pmatrix}. \tag{5.11}$$

The final state after $N-1$ swaps, which spans all N initial links, is therefore $\mathbf{e}^{(N-1)} = M^{N-1}\mathbf{e}^{(0)}$.

To compute M^{k_s} where $k_s = N-1$ is the number of swaps, we can diagonalize M as matrix M is real and symmetric, and its structure allows for straightforward determination of its eigenvalues. These eigenvalues are

$$\lambda_0 = w_0 + x_0 + y_0 + z_0 = 1 \quad \left(\text{as } \sum p_i = 1\right), \tag{5.12}$$

$$\lambda_1 = w_0 + x_0 - y_0 - z_0, \tag{5.13}$$

$$\lambda_2 = w_0 - x_0 + y_0 - z_0, \tag{5.14}$$

$$\lambda_3 = w_0 - x_0 - y_0 + z_0. \tag{5.15}$$

The initial state $\mathbf{e}^{(0)}$ can be decomposed in the eigenbasis of M. After $k_s = N-1$ applications of the transformation M, the components of the final error vector $\mathbf{e}^{(N-1)} = [w^{(N-1)}, x^{(N-1)}, y^{(N-1)}, z^{(N-1)}]^T$ are given by the standard solution for such iterated Pauli channels given by

$$\begin{aligned} w^{(N-1)} &= \frac{1}{4}\left[\lambda_0^N + \lambda_1^N + \lambda_2^N + \lambda_3^N\right] \\ &= \frac{1}{4}\left[1 + \lambda_1^N + \lambda_2^N + \lambda_3^N\right], \end{aligned} \tag{5.16}$$

$$x^{(N-1)} = \frac{1}{4}\left[1 + \lambda_1^N - \lambda_2^N - \lambda_3^N\right], \tag{5.17}$$

$$y^{(N-1)} = \frac{1}{4}\left[1 - \lambda_1^N + \lambda_2^N - \lambda_3^N\right], \tag{5.18}$$

$$z^{(N-1)} = \frac{1}{4}\left[1 - \lambda_1^N - \lambda_2^N + \lambda_3^N\right]. \tag{5.19}$$

In these equations, $\lambda_1, \lambda_2, \lambda_3$ are defined by eqs. (5.13) to (5.15) using the components w_0, x_0, y_0, z_0 of the initial error vector $\mathbf{e}^{(0)}$ of each elementary link, and N is the total number of elementary links in the chain. The fidelity of the final distributed Bell pair with respect to the ideal state $|\psi_{II}\rangle$ is given by $w^{(N-1)}$. This derivation quantifies how the initial Pauli error probabilities on elementary links propagate through a chain of $N-1$ entanglement swaps.

Error probability from inner qubits

The error contributions from inner qubits arise from measurements on their constituent physical photons and the subsequent impact on Pauli frame tracking. These contributions are effectively independent for Alice's and Bob's sides of the final entangled pair, allowing their respective error probabilities to be defined and calculated separately. To facilitate a compact derivation, we will assume an identical depolarizing channel affects all physical photonic qubits within the inner structures. Under this symmetric model, if p_{dep} is the single-photon depolarization probability, the effective single-photon

measurement error probability relevant for our analysis p_{ph} can be considered as $\frac{2}{3}p_{\text{dep}}$ since X and Y error would flip the results when measuring in the Z basis, and similarly for X and Y measurements. While this specific derivation assumes uniformity, the general approach can be extended to cases with non-identical or asymmetric Pauli error channels across different hops or within distinct tree code structures.

Let p_{ZI} and p_{IZ} denote the probability that a single-hop link's inner qubit measurements induce an effective Z error on the left (Alice's side) and right (Bob's side) memory (anchor qubit), respectively. Assuming there are N hops between the two end nodes and independent errors across these hops, the total end-to-end error probabilities for Alice's $\left(p_{ZI}^{\text{e2e}}\right)$ and Bob's $\left(p_{IZ}^{\text{e2e}}\right)$ memory qubits are given by summing over an odd number of such errors as

$$\begin{aligned} p_{ZI}^{\text{e2e}} &= \sum_{k \text{ odd}}^{N} \binom{N}{k} p_{ZI}^{k} (1 - p_{ZI})^{N-k} \\ &= \frac{1}{2}\left(1 - (1 - 2p_{ZI})^{N}\right), \end{aligned} \tag{5.20}$$

$$\begin{aligned} p_{IZ}^{\text{e2e}} &= \sum_{k \text{ odd}}^{N} \binom{N}{k} p_{IZ}^{k} (1 - p_{IZ})^{N-k} \\ &= \frac{1}{2}\left(1 - (1 - 2p_{IZ})^{N}\right), \end{aligned} \tag{5.21}$$

assuming $p_{ZI}, p_{IZ} < 0.5$, as to avoid a completely mixed state at the end.

The single-hop error probabilities, p_{ZI} and p_{IZ} (which are equal to a common $p_{\text{in-hop}}$ in the symmetric case), capture the likelihood that an odd number of physical errors within the m arms of an RGS segment's inner qubit measurements results in a net flip of the Pauli frame for that hop. This probability is given by

$$p_{\text{in-hop}} = p_{ZI} = p_{IZ} = \frac{1}{2}\left(1 - \left(1 - 2p_{\text{in}}^{X}\right)\left(1 - 2p_{\text{in}}^{Z}\right)^{m-1}\right), \tag{5.22}$$

where p_{in}^{P} denotes the probability that a logical measurement in basis $P \in \{\text{X}, \text{Z}\}$ on a tree encoded inner qubit is decoded incorrectly. The average error probability of logical measurements p_{in}^{P} can be derived from underlying physical error parameters (such as p_{ph} or p_{dep}), taking into account both photon loss probabilities and physical measurement error probabilities within the tree code structure, as detailed in analyses such as [50]. While eq. (5.22) provides the per-hop probability for identical links, in a general network where these probabilities might vary from hop to hop, the end-to-end calculation would involve a more generalized combination of these per-hop error events, though the fundamental principle of summing over configurations leading to an odd number of total Pauli frame flips remains.

Combined end-to-end fidelity derivation

Having separately characterized the error propagation from outer qubits (resulting in the error vector $\mathbf{e}_{\text{out}}^{(N-1)}$) and the effective Z type errors from inner qubit measurements (leading to independent Z error probabilities $p_A \equiv p_{ZI}^{\text{e2e}}$ and $p_B \equiv p_{IZ}^{\text{e2e}}$ on Alice's and Bob's sides, respectively, as given by eqs. (5.20) and (5.21)), we now combine these contributions to determine the overall error probability vector of the final Bell pair.

The effect of the independent Z errors from inner qubits can be described as a transition matrix (linear transformation) applied to the error vector $\mathbf{e}_{\text{out}}^{(N-1)}$ from the outer qubit chain. Let these inner qubit error probabilities be

$$w'_{\text{in}} = (1 - p_A)(1 - p_B) \quad \text{(Prob. of } II \text{ from inner channel)}, \tag{5.23}$$

$$x'_{\text{in}} = p_A(1 - p_B) \quad \text{(Prob. of } ZI \text{ from inner channel)}, \tag{5.24}$$

$$y'_{\text{in}} = p_A p_B \quad \text{(Prob. of } ZZ \text{ from inner channel)}, \tag{5.25}$$

$$z'_{\text{in}} = (1 - p_A)p_B \quad \text{(Prob. of } IZ \text{ from inner channel)}. \tag{5.26}$$

The final end-to-end error vector $\mathbf{e}_{\text{e2e}} = [w_{\text{e2e}}, x_{\text{e2e}}, y_{\text{e2e}}, z_{\text{e2e}}]^T$ is then obtained by applying the transformation M_{in} to $\mathbf{e}_{\text{out}}^{(N-1)}$ as $\mathbf{e}_{\text{e2e}} = M_{\text{in}} \cdot \mathbf{e}_{\text{out}}^{(N-1)}$, where the transition matrix M_{in} due to inner qubit errors is given by

$$M_{\text{in}} = \begin{pmatrix} w'_{\text{in}} & x'_{\text{in}} & y'_{\text{in}} & z'_{\text{in}} \\ x'_{\text{in}} & w'_{\text{in}} & z'_{\text{in}} & y'_{\text{in}} \\ y'_{\text{in}} & z'_{\text{in}} & w'_{\text{in}} & x'_{\text{in}} \\ z'_{\text{in}} & y'_{\text{in}} & x'_{\text{in}} & w'_{\text{in}} \end{pmatrix}. \tag{5.27}$$

This matrix M_{in} has the same structural form as the BSM transformation matrix M, with its components now being the probabilities derived from the inner qubit errors $(w'_{\text{in}}, x'_{\text{in}}, y'_{\text{in}}, z'_{\text{in}})$. Applying this transformation explicitly yields the components of the final error vector as

$$w_{\text{e2e}} = w'_{\text{in}} w_{\text{out}}^{(N-1)} + x'_{\text{in}} x_{\text{out}}^{(N-1)} + y'_{\text{in}} y_{\text{out}}^{(N-1)} + z'_{\text{in}} z_{\text{out}}^{(N-1)}, \tag{5.28}$$

$$x_{\text{e2e}} = x'_{\text{in}} w_{\text{out}}^{(N-1)} + w'_{\text{in}} x_{\text{out}}^{(N-1)} + z'_{\text{in}} y_{\text{out}}^{(N-1)} + y'_{\text{in}} z_{\text{out}}^{(N-1)}, \tag{5.29}$$

$$y_{\text{e2e}} = y'_{\text{in}} w_{\text{out}}^{(N-1)} + z'_{\text{in}} x_{\text{out}}^{(N-1)} + w'_{\text{in}} y_{\text{out}}^{(N-1)} + x'_{\text{in}} z_{\text{out}}^{(N-1)}, \tag{5.30}$$

$$z_{\text{e2e}} = z'_{\text{in}} w_{\text{out}}^{(N-1)} + y'_{\text{in}} x_{\text{out}}^{(N-1)} + x'_{\text{in}} y_{\text{out}}^{(N-1)} + w'_{\text{in}} z_{\text{out}}^{(N-1)}, \tag{5.31}$$

with the end-to-end fidelity given by $F_{\text{e2e}} = w_{\text{e2e}}$.

The graphical representation in Fig. 5.9 and these error expressions also inform efficient Monte Carlo simulation methods by tracking Z flips hop-by-hop. This aligns with techniques in QuISP, prior work [41, 189, 333], and Stim [330]. Our simulations using this approach show consistency with average error models like those in [50]. The code for this simulation can be found at [331].

5.7 Purification-enhanced RGS scheme

Entanglement purification is a powerful tool for improving the fidelity of long-distance quantum communication. However, integrating it beyond the link level into all-photonic architectures remains a significant challenge [2, 50]. In this section, we introduce a purification-enhanced framework for the RGS scheme that leverages our half-RGS building block to address this problem. Our approach uses optimistic purification protocols to achieve flexible purification between distant nodes without introducing memories or significantly increasing latency, opening the door to a new class of high-performance all-photonic repeaters.

Motivation for integrating purification into the RGS scheme

To understand the need for a new approach, we first consider conventional methods for purification in repeater networks. A natural strategy in RGS-based networks is to perform two-way purification at the end nodes once a long-distance entangled state has been established. While conceptually simple, this method suffers from increased latency due to one-way end-to-end classical communication and degradation from memory decoherence during the waiting period.

Alternatively, a measurement-based approach using purification-embedded graph states can potentially avoid quantum memories but is limited to purifying entanglement between adjacent nodes only [334]. This method still requires additional optical Bell-state measurements, which reduces success rates and lowers the overall generation rate. Such limitations highlight the fundamental difficulty of incorporating purification into RGS protocols, which are otherwise valued for their high loss tolerance and memory-free design.

We address this challenge by leveraging the half-RGS building block to enable purification directly on the generated half-RGS primitives. This is achieved using optimistic purification, which is executed immediately on the anchor qubits of multiple half-RGSs without awaiting heralding signals [51–53]. Success signals and Pauli frame corrections propagate only at the end of the process, preserving the fast rate and memoryless structure of the RGS scheme. This design allows for purification between non-neighboring nodes and supports flexible scheduling strategies—a capability previously limited to memory-based repeaters.

Crucially, the added overhead is modest, with generation time scaling proportionally with purification rounds while total latency remains effectively unchanged. These features enable RGS-based networks to benefit from a large body of prior work on purification scheduling and optimization, bringing established memory-based techniques into the all-photonic domain [282, 283, 286, 335]. The key enabler for this framework is the optimistic purification protocol, which we now review in more detail.

Optimistic purification

Optimistic entanglement purification, also known as blind purification, is a variant of the heralded approach where purification proceeds without waiting for confirmation that the initial Bell pairs were successfully established [51–53]. As illustrated in Fig. 5.12(b), multiple purification rounds can be executed sequentially, with a single round of classical messaging exchanged only at the very end to report the combined outcome.

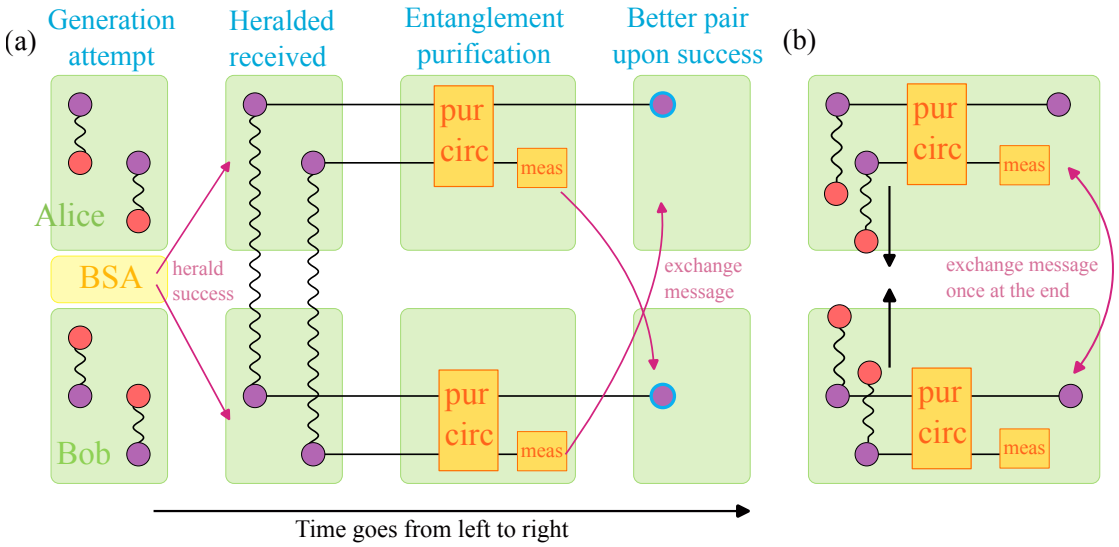


Figure 5.12: (a) Schematic of two-way heralded entanglement purification. The process begins with the probabilistic distribution of Bell pairs, followed by heralding signals from intermediate nodes confirming successful pair creation. A purification circuit is then applied, during which some Bell pairs are measured and consumed. The measurement results are exchanged between the two parties; if the outcomes agree (i.e., correlate or anti-correlate as expected), the remaining Bell pair is kept, resulting in improved fidelity. (b) Schematic of optimistic purification. In this approach, distribution and purification operations are performed immediately in sequence, without waiting for heralding confirmation. Classical outcomes are compared only once at the end, combining both the heralding of Bell pair generation and the results of purification in a single round of communication. (From [48].)

While this strategy may lower the overall success probability, it offers significant advantages in high-latency settings by minimizing memory idle time and the cumulative effects of decoherence [52, 53]. This trade-off can lead to higher output fidelity compared to traditional purification, even with fewer successful attempts.

The optimistic approach is particularly well-suited to scenarios with high or near-deterministic Bell pair generation rates, a key feature of the RGS framework. When pairs are created reliably, waiting for heralding signals offers little benefit and only increases latency. Having established the advantages of this protocol, we now detail how our half-RGS architecture is modified to incorporate it.

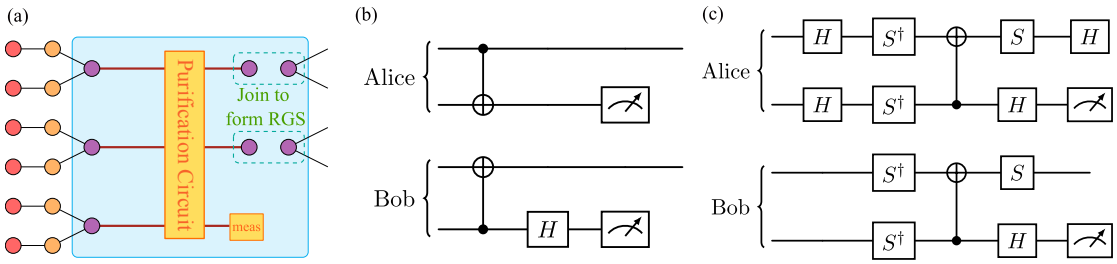


Figure 5.13: (a) Integration of entanglement purification into the RGS generation process. Instead of joining left and right half-RGSs immediately after their creation as in the original protocol [45], multiple half-RGSs are generated on each side. Their anchor emitters (purple qubits) serve as inputs to a purification circuit. After purification, the retained high-fidelity half-RGSs from the left and right are paired and joined to form complete RGSs. (Note that dark red lines connecting purple qubits in the blue shaded box represent the circuits or their evolution in time and do not represent entangled properties of graph states.) (b) Graph state purification circuits that measures the parity of stabilizer Z_aX_b where a represents the Alice's side and b for Bob [336, 337] Top (bottom) qubits of Alice and Bob are Bell pairs in the two-qubit graph state. The top qubits for both side are retained after the purification circuit. The purification circuit for X_aZ_b are similar, by exchanging the circuits that Alice and Bob locally performed. (c) Graph state purification circuit that measures the parity of Y_aY_b stabilizers. (From [48].)

Modification to the RGS generation process

The standard half-RGS generation process, detailed in section 5.4, concludes by immediately joining two half-RGSs. To integrate optimistic purification, we modify this procedure by delaying the joining step to insert a purification stage. Instead of creating a single half-RGS per side, we generate multiple half-RGSs, requiring additional anchor emitters. Once all half-RGSs are created, their anchor qubits are fed as inputs to a purification circuit. After purification, the retained high-fidelity half-RGSs from both sides are paired and joined. While this modification enables purification, it introduces new operational steps whose impact on resource overheads must be analyzed.

Overhead analysis of the purification-enhanced RGS scheme

We now analyze the time overhead of our purification-enhanced framework, focusing on the qubit coherence time required at the end nodes and the RGSS nodes. We evaluate how long these qubits must remain coherent to support applications requiring high-fidelity entanglement. For a direct comparison, we consider three scenarios: (1) the **raw RGS scheme** without purification; (2) a **baseline purification** approach where purification is applied only at the end nodes; and (3) our **proposed optimistic purification** framework.

First, in the *raw RGS scheme*, the memory coherence time $t_{\text{mem}}^{(\text{raw})}$ must span the half-RGS generation and the arrival of remote Pauli frame correction messages. Second, the *baseline purification* approach requires a significantly longer coherence time $t_{\text{mem}}^{(\text{p-base})}$ as it is dominated by the round-trip communication needed for heralding. Finally, our

proposed optimistic purification scheme has a memory coherence time of $t_{\text{mem}}^{(\text{p-opt})}$ that scales much more favorably than the baseline, as it avoids the heralding delay. This analysis reveals that our protocol significantly reduces the demands on quantum memory, a crucial advantage that sets the stage for evaluating its overall performance in terms of fidelity and rate.

5.8 Performance evaluation

Having established the architectural modifications and favorable overheads of our protocol, we now evaluate its performance and demonstrate its flexibility through numerical simulations. We showcase this flexibility by applying purification both at the link-level and between non-neighboring nodes. While a full optimization of purification scheduling is beyond the scope of this thesis, our analysis focuses on the fidelity of the distributed Bell pairs and the rate at which a minimum fidelity threshold is met, providing a clear benchmark of the scheme's capabilities.

Scope and assumptions

To ground our evaluation, we first define the scope of our simulation and the key assumptions made. We consider a repeater chain of 10 hops between two memory-equipped end nodes, Alice and Bob, connected by nine intermediate RGSS nodes. We model the outcome of half-RGS generation by applying effective errors directly to the anchor qubits, following the derivation in section 5.6. The half-RGS parameters are tuned to achieve a near-deterministic generation success rate (≥ 0.999). Our key assumptions are as follows:

- Nodes are spaced 2 km apart, with a channel loss of 0.2 dB/km.
- All other sources of photon loss, such as from coupling or emission failure, are considered negligible.
- Each half-RGS uses 18 arms and a tree encoding with branching parameters $(16, 14, 1)$ to ensure high success probability.
- All photons are subject to a uniform depolarizing noise channel.
- Quantum gates and measurements are assumed to be noiseless.
- End-node memories are assumed to be ideal, with no decoherence.

The justification for these ideal assumptions is twofold. First, operational noise can be effectively modeled as increased depolarizing error on the emitted photons. Second, recent advances in quantum error correction [162–164] make the availability of highly coherent logical memories a reasonable projection for future systems. With these assumptions defined, we can now specify the purification strategies used in our model.

Purification circuits and scheduling

We restrict our noise model to Pauli channels, a choice justified because most error-correcting codes convert general noise into an effective Pauli channel during syndrome measurement [133]. Both the baseline and our enhanced schemes use the same three 2-to-1 stabilizer-based purification circuits, which operate on two-qubit graph states (Bell pairs). These circuits, shown on the right in Fig. 5.13, measure the joint stabilizers Z_aX_b, X_aZ_b, and Y_aY_b [336, 337].

For the analysis, we adopt an error model where all errors are represented as Z-type Pauli operators that can affect either or both qubits of the Bell pair. This model simplifies the analysis of error propagation and aligns with the Pauli frame tracking used in our protocol. The purification circuits are sensitive to different error signatures; for instance, the ZX circuit detects Z errors on qubit b, the XZ circuit detects them on qubit a, and the YY circuit detects an odd number of total Z errors.

Our purification schedule for the baseline setting is inspired by the entanglement pumping protocol from [179]. The schedule applies these circuits iteratively in the sequence YY, ZX, YY, XZ. The rationale for starting with the YY circuit is that while the initial depolarizing noise is symmetric, the effective errors contributed by the inner qubits are biased towards independent Z errors on each end node (IZ and ZI), as shown in section 5.6. The YY circuit is therefore most effective at detecting these dominant error types in the early rounds.

Error vector formalism for purification

To evaluate the different purification schedules, we now define the mathematical rules for purification within the error vector formalism established in section 5.6. Each 2-to-1 purification circuit is characterized by its success probability and the corresponding transformation it applies to the error vectors of the two input Bell pairs, $\mathbf{e_1}$ and $\mathbf{e_2}$. The success probabilities for the three stabilizer measurement circuits (ZX, XZ, and YY) are given by

$$P_{ZX}(\mathbf{e_1}, \mathbf{e_2}) = (w_1 + x_1)(w_2 + x_2) + (z_1 + y_1)(z_2 + y_2), \tag{5.32}$$

$$P_{XZ}(\mathbf{e_1}, \mathbf{e_2}) = (w_1 + z_1)(w_2 + z_2) + (x_1 + y_1)(x_2 + y_2), \tag{5.33}$$

$$P_{YY}(\mathbf{e_1}, \mathbf{e_2}) = (w_1 + y_1)(w_2 + y_2) + (x_1 + z_1)(x_2 + z_2). \tag{5.34}$$

Conditioned on a successful outcome, the error vector of the retained Bell pair is given by one of the following transformations

$$\mathrm{Pur}_{ZX}(\mathbf{e_1}, \mathbf{e_2}) = \frac{1}{P_{ZX}(\mathbf{e_1}, \mathbf{e_2})} \begin{bmatrix} w_1w_2 + z_1z_2 \\ z_1w_2 + w_1z_2 \\ x_1y_2 + y_1x_2 \\ x_1x_2 + y_1y_2 \end{bmatrix}, \tag{5.35}$$

$$\mathrm{Pur}_{XZ}(\mathbf{e_1}, \mathbf{e_2}) = \frac{1}{P_{XZ}(\mathbf{e_1}, \mathbf{e_2})} \begin{bmatrix} w_1w_2 + x_1x_2 \\ z_1z_2 + y_1y_2 \\ z_1y_2 + y_1z_2 \\ x_1w_2 + w_1x_2 \end{bmatrix}, \tag{5.36}$$

$$\mathrm{Pur}_{YY}(\mathbf{e_1}, \mathbf{e_2}) = \frac{1}{P_{YY}(\mathbf{e_1}, \mathbf{e_2})} \begin{bmatrix} w_1w_2 + y_1y_2 \\ x_1z_2 + z_1x_2 \\ y_1w_2 + w_1y_2 \\ x_1x_2 + z_1z_2 \end{bmatrix}. \tag{5.37}$$

These new rules for purification, combined with the BSM transformation from eq. (5.9), form the complete analytical toolkit needed to model the resulting fidelity between baseline and our purification-enhanced scheme.

Evaluation methodology

With the analytical tools established, we now outline the methodology for our numerical simulation. Our simulation tracks the effective error vectors, conditioned on the successful generation of one-hop Bell pairs. The freedom to reorder measurements allows us to model each half-RGS as an effective two-qubit graph state between an anchor and an outer qubit, with noise from inner qubits applied directly to the anchors using the average logical error probability [2, 50]. The process to establish an end-to-end link involves first applying depolarizing noise to the outer qubits, then performing BSMs to connect the segments, and finally updating the anchor error vectors with the noise contributions from the inner qubits, per the rules in section 5.6.

After establishing the initial noisy end-to-end Bell pairs, we simulate the scheduled purification protocols for each of the three scenarios. For the baseline case, purification is applied to the final Bell pairs. For our enhanced scheme, purification is performed at intermediate stages according to the multi-stage schedule detailed in the next section. This consistent methodology ensures a direct and fair comparison of the end-to-end fidelities achieved by each framework.

5.9 Numerical results

To quantify the advantages of our architecture, we present a detailed performance comparison between our purification-enhanced scheme, the baseline approach, and

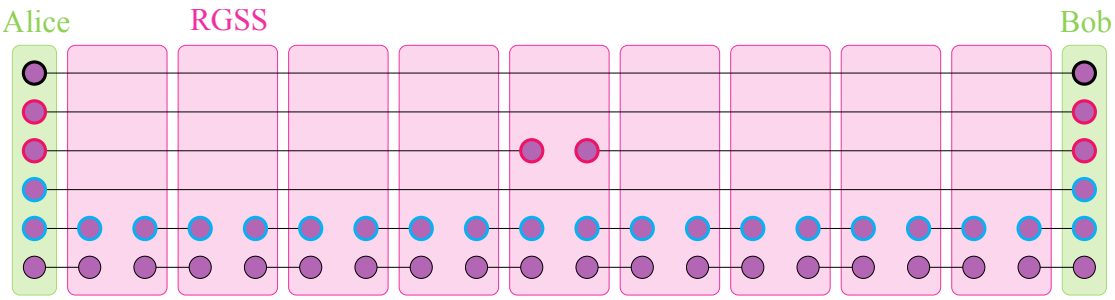


Figure 5.14: Illustration of flexible purification scheduling enabled by our purification-enhanced RGS framework. Using five half-RGSs per side at each hop (bottom line represents a copy), three end-to-end Bell pairs are created via different purification schedules, resulting in varying fidelities. These pairs then undergo further purification at the end nodes to yield a single high-fidelity Bell pair. The first is constructed by purifying (YY) each link-level connection and stitching the purified links across the full path (blue border). The second is formed by stitching link-level pairs into two copies of two 5-hop segments, purifying (YY) within each segment, and then connecting them (pink border). The third is created by directly stitching raw link-level pairs without intermediate purification (black thick border). These three pairs are then used in a pumping-like purification sequence: the first and second are purified (with the second as the sacrificial pair with ZX), followed by purification between the result and the third (with the third as the sacrificial pair with YY). The final output is a high-fidelity end-to-end Bell pair, significantly improved over the unpurified case upon success. (From [48].)

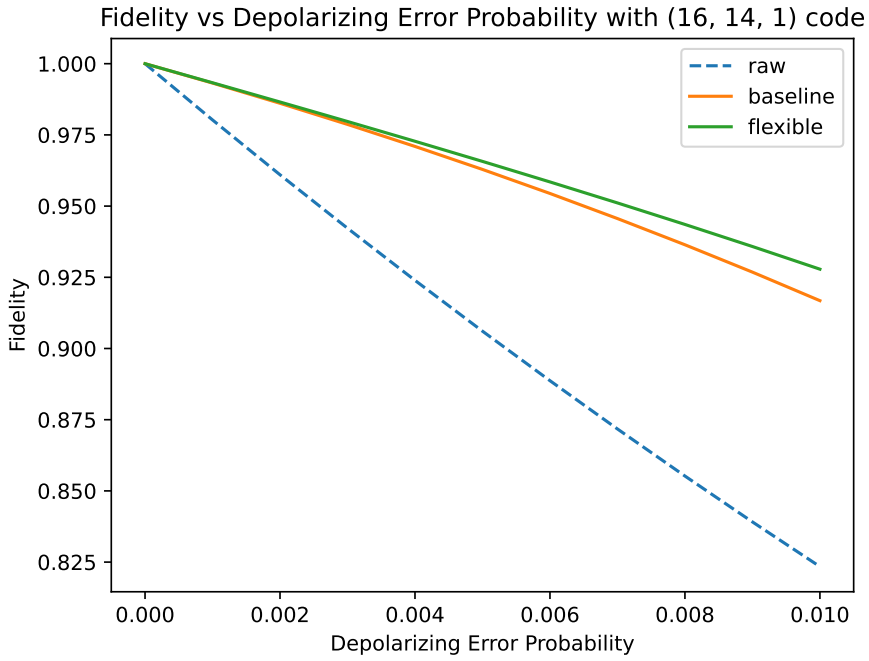


Figure 5.15: End-to-end fidelity comparison of repeater chain with ten hops between unpurified Bell pairs (dashed blue lines), baseline purification using direct end-to-end Bell pairs (orange), and our purification-enhanced scheme (green). Both of the purified schemes consume five half-RGSs per end-to-end Bell pair instead of one in the unpurified setting. (From [48].)

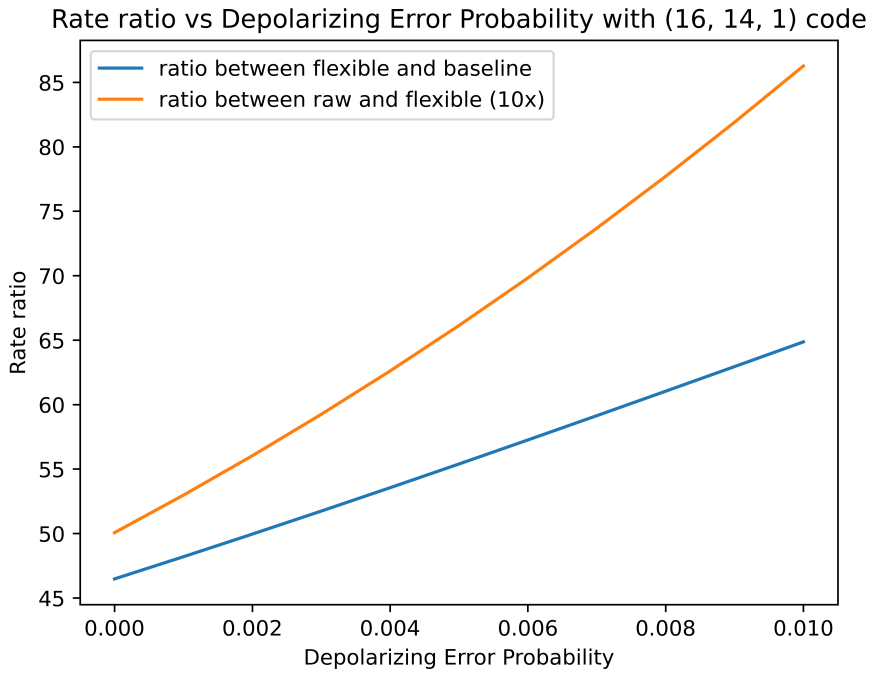


Figure 5.16: Comparison of Bell pair distribution rates for different purification strategies over a 40 km (10-hop) repeater chain. The blue curve shows the rate enhancement of the purification-enhanced strategy over a baseline pumping approach, achieving an improvement of 45 to 65 times, depending on the noise level. The orange curve, scaled by 10x for visibility, illustrates the overhead of purification by comparing its rate to that of a raw, unpurified link. This overhead corresponds to a rate reduction factor of 5 to 8.5. (From [48].)

the unpurified RGS protocol. For the purification-enhanced scheme, we implement a sophisticated multi-stage schedule, illustrated in Fig. 5.14. This strategy involves creating three distinct end-to-end Bell pairs using different intermediate purification schedules before combining them in a final purification sequence at the end nodes. This flexible, multi-layered approach exemplifies our framework's ability to strategically apply purification at the most effective points in the network.

The simulation results, presented in Figs. 5.15 and 5.16, clearly illustrate the critical trade-offs between fidelity and rate. As seen in Fig. 5.15, both purification schemes substantially improve fidelity over the unpurified case, consistently achieving values above 0.9 for the noise parameters considered ($0 \leq p_{dep} \leq 0.01$). Notably, our enhanced scheme maintains a consistent fidelity advantage over the baseline approach. This improvement is a direct consequence of its ability to suppress errors early in the repeater chain, thereby preventing noise from accumulating and propagating across the entire connection.

This performance advantage becomes even more pronounced when considering the distribution rate. As shown in Fig. 5.16, the purification-enhanced scheme outperforms the baseline strategy by a factor of 45 to 65 across the tested noise levels. This significant

gain arises because our optimistic protocol eliminates the need for multiple, long-distance classical communication rounds required by the baseline's schedule, pushing the exchange of all purification results to the very end.

Crucially, this comparison is based on equal resource consumption, as both purification schemes use five half-RGSs per side at each RGSS and end node. The performance gap is therefore attributable to the architectural design of our protocol. By exploiting the structure of repeater graph states, our framework embeds purification directly within the entanglement generation process. This approach enables multiple stages of error suppression that enhance fidelity while simultaneously reducing the time overhead from classical communication.

These results, obtained without a globally optimized schedule, suggest that further improvements are readily achievable through tailored scheduling strategies [176]. Nonetheless, they provide compelling evidence that the purification-enhanced RGS framework can simultaneously meet the dual goals of high fidelity and a high entanglement distribution rate. This result successfully overcomes a fundamental challenge in all-photonic quantum repeater design, marking a significant step toward their practical implementation.

In summary, by integrating optimistic purification directly into the generation process, our framework transforms the all-photonic repeater chain from a conceptual protocol into a robust, high-performance building block for larger quantum networks. The demonstrated ability to generate high-fidelity entanglement at high rates means that an entire RGS-based path can now be reliably treated as a single, abstracted virtual link. This capability is essential for creating scalable, heterogeneous networks, as it allows all-photonic segments to be seamlessly composed with memory-based systems under a unified control architecture, bridging a critical gap between hardware-level photonic protocols and the requirements of a future Quantum Internet.

5.10 Discussion and future outlook

In chapter 4, we established a comprehensive architecture for memory-based quantum networks built on the principles of the quantum recursive network architecture (QRNA) and managed by RuleSet-based protocols. A central tenet of that architecture is the ability to abstract entire network segments as virtual links, enabling scalable and heterogeneous internetworking. However, integrating memoryless, all-photonic repeaters into such a stateful, connection-oriented framework presents a significant challenge. In this chapter, we have addressed this challenge directly by developing a robust, high-performance RGS scheme. Our key contributions—the modular half-RGS building block and the integration of optimistic purification—provide the necessary tools to transform a volatile all-photonic path into a reliable architectural component.

Hardware progress and the experimental landscape

Recent experimental progress has begun to bridge the gap between theoretical all-photonic repeater protocols and their physical realization. Early proof-of-concept demonstrations of the RGS scheme have been achieved, for instance, over a single hop using a passive, specialized measurement apparatus [338] and in a two-arm configuration [339]. A critical challenge remains the deterministic generation of the large-scale photonic graph states required. Several hardware platforms are being pursued, with proposals for using arrays of quantum emitters [340] and promising demonstrations using semiconductor quantum dots as sources [341, 342].

The most resource-efficient theoretical proposals for deterministic graph state generation often rely on a single, highly-capable quantum emitter that sequentially emits entangled photons, a concept sometimes called a "photonic machine gun" [332, 343]. This approach has seen remarkable experimental success recently, with demonstrations in the optical domain using trapped atoms and quantum dots [344–350] and in the microwave domain with superconducting qubits [351, 352]. Among the most advanced experimental feats to date are the generation of reconfigurable 2D cluster states of up to 20 photons [353] and deterministically generated, redundantly encoded linear graph states that enable near-deterministic fusion operations with built-in error correction [354, 355].

Alternative all-photonic architectures

Beyond the RGS framework, a rich landscape of alternative all-photonic repeater architectures exists. These approaches can be broadly distinguished by their core operational strategy [161, 356], which has significant implications for their integration into a QRNA-style architecture.

The first category, like the RGS scheme, is based on pre-distributing entangled Bell pairs, a model that aligns naturally with the connection-oriented paradigm of chapter 4. This family includes protocols that refine the crucial BSM step with adaptive measurement strategies on logically encoded states [167]. Other variants in this family introduce short-term photonic memory by having inner qubits travel through spools of optical fiber, allowing for more complex operations like boosted BSMs or parallel transmission of RGS arms through multi-channel fibers [324, 357].

The second category [44, 358–361] operates via the direct, hop-by-hop transmission of encoded logical states, an approach akin to the packet-switching model of third-generation repeaters [161]. The core idea is to utilize a "decode-and-re-encode" strategy at each repeater station to correct for errors, a strategy that can be implemented with various optimized loss-tolerant codes [362, 363]. While this model potentially offers high speed, integrating such connectionless protocols into the stateful, connection-oriented architecture presented in chapter 4 is non-trivial. Our framework relies on

end-to-end RuleSets established via a two-pass setup, but a hop-by-hop protocol could be emulated by sequentially creating and tearing down a series of short-lived, single-hop connections. This would, however, introduce significant protocol overhead, especially for managing the rapid instantiation of RuleSets and coordinating handoffs at network boundaries.

Finally, other distinct approaches exist that rely on different physical encodings entirely. These include protocols based on bosonic GKP states [364, 365], coherent states which can naturally provide deterministic BSMs [366], or multipartite entangled states such as W-states [367], each representing a unique avenue of research.

Strategies for protocol enhancement

Further improvements to all-photonic repeaters can be pursued along several research directions, primarily by addressing key operational bottlenecks. A primary challenge in many schemes, including RGS, is the probabilistic nature of linear-optical Bell-state measurements (BSMs). Boosting the BSM success probability beyond the standard 50% limit is a major avenue for enhancement, as it would directly reduce the resource overhead of the virtual links presented to the higher-level network architecture. Several proposals exist to achieve this, using techniques like ancillary photons [368, 369], weak non-linearities [370], or squeezed states [371, 372]. However, these advanced techniques involve significant experimental complexity and are constrained by fundamental no-go theorems [373] and theoretical limits on loss tolerance [374].

Another strategy for enhancement is to improve the resource efficiency of the RGS protocol itself. Currently, only one Bell pair can be extracted from a successful RGS trial, even if multiple BSMs succeed at every hop. This limitation has motivated the proposal of a generalized RGS framework where the tree code for inner qubits is replaced with a classical low-density parity-check (LDPC) code [375]. For certain families of LDPC codes, this would enable the creation of multiple, locally-CNOT equivalent Bell pairs, thus allowing for the extraction of more than one end-to-end pair from a single trial. Ultimately, weighing the architectural trade-offs between these two enhancement strategies—one improving a fundamental operation and the other the protocol's core structure—will be crucial for developing the most practical and resource-efficient quantum repeaters.

Conclusion

In this chapter, we addressed several key architectural and operational challenges of the all-photonic repeater graph state (RGS) scheme. Our primary contribution is the introduction of the **half-RGS**, a modular construct that resolves fundamental issues related to monolithic graph state generation and global timing synchronization by decomposing

the problem into independent, localized operations. This building block provides a more practical and hardware-friendly foundation for constructing all-photonic repeater links.

Building upon this modular foundation, we then tackled the long-standing problem of incorporating entanglement purification into all-photonic repeater paths. The architectural flexibility afforded by the half-RGS enabled the seamless integration of optimistic purification directly into the generation process. As our numerical results demonstrated, this purification-enhanced framework achieves both high fidelity and significantly higher distribution rates—outperforming baseline approaches by a factor of 45 to 65—without incurring additional latency or memory requirements. This performance gain is a direct result of the protocol's efficient design, which eliminates the need for multiple long-distance classical communication rounds.

Ultimately, these contributions advance the all-photonic repeater chain from a theoretical protocol towards its realization as a robust, high-performance component for large-scale networks. The demonstrated ability to generate high-fidelity entanglement at high rates validates the abstraction of an entire RGS subpath as a well-behaved **virtual link**. This provides the crucial missing piece for integrating this entire class of technology into the scalable, recursive architecture laid out in chapter 4. The half-RGS, guided by optimistic purification, becomes the concrete physical-layer and link-layer implementations that can be controlled and composed by the higher-level Connection Management and RuleSet protocols, truly bridging the gap between hardware-specific paradigms and the architectural requirements of a universal Quantum Internet.

Chapter 6

AN APPLICATION OF DISTRIBUTED QUANTUM COMPUTATION: AN OPTIMIZATION ALGORITHM PERSPECTIVE

Having established an architectural framework for heterogeneous quantum networks in the preceding chapters, we now take a step back and turn to the applications that such a network is designed to support. It would be incomplete to discuss the architecture without considering the computations it must enable. A primary motivation for building a quantum network is to connect individual quantum processors into distributed quantum computer clusters capable of performing large-scale computations that are infeasible on a single machine. While truly large-scale algorithms like Shor's factoring algorithm [376] or complex chemical simulations may represent a distant goal, *combinatorial optimization* problems offer a compelling and industrially relevant use case for near- to medium-term distributed quantum computing.

This chapter bridges the analysis of quantum optimization algorithms with the practical requirements of the network architectures designed to run them. To build a network capable of executing these algorithms, it is essential to understand their runtimes and resource consumption. This knowledge is critical for creating realistic traffic models for network simulators like QuISP, informing resource allocation strategies, and ultimately understanding the demands that distributed computation will place on a future Quantum Internet.

To this end, this chapter presents two primary contributions. First, we explore potential advancements in optimization by developing an algorithm based on a generalization of amplitude amplification. This method uses a non-trivial phase shift oracle, which we term the *subdivided phase oracle* (SPO), to directly amplify states corresponding to high-quality solutions. Interestingly, we find that the SPO algorithm, with its repeated application of a problem-specific phase oracle and a general mixing operator, can be viewed as a specific instance of the QAOA framework [377].

Recognizing that our SPO algorithm is a member of this broader class, the second key

Portions of this chapter have been adapted from the following publications:

- N. Benchasattabuse et al., "Amplitude Amplification for Optimization via Subdivided Phase Oracle," in *2022 IEEE International Conference on Quantum Computing and Engineering (QCE)*, Sep. 2022, pp. 22–30, DOI: 10.1109/QCE53715.2022.00020.
- N. Benchasattabuse et al., "Lower bounds on the number of rounds of the quantum approximate optimization algorithm required for guaranteed approximation ratios," *Physical Review A*, vol. 111, no. 6, p. 062 411, Jun. 9, 2025, DOI: 10.1103/PhysRevA.111.062411.

contribution of this chapter is a detailed analysis of the general QAOA framework's performance, which has direct implications for network design. Specifically, we derive some of the first analytical lower bounds on the number of QAOA rounds required to achieve a guaranteed approximation ratio. This provides crucial, concrete data on algorithm runtimes needed for network planning.

This chapter is structured to unfold this narrative. We will first provide a detailed background on Grover's algorithm, QAOA, and the SPO. We will then present our main results on the lower bounds for QAOA rounds for different classes of problems and mixers. Finally, we will discuss the broader implications of these algorithmic runtimes for designing and managing resources in the quantum networks of the future.

6.1 Introduction: why study a heuristic optimization algorithm?

Solving optimization problems is a cornerstone computational task, with a wide range of applications across industry, science, and technology. A primary application for quantum computers is to tackle these problems, and the quantum approximate optimization algorithm [378] and its generalization to constrained optimization problems, the quantum alternating operator ansatz (QAOA) [379], have emerged as leading candidates. QAOA is a hybrid quantum-classical approach designed to find high-quality approximate solutions to combinatorial optimization problems. Its structure involves alternating between a phase separator, which encodes the problem's objective function into a Hamiltonian H_1, and a mixing operator, H_0. A classical optimizer then tunes the parameters for a fixed number of rounds, p, to guide the quantum state toward a low-energy solution that represents a high-quality answer.

While algorithms based on Grover's search can find exact solutions with a provable quadratic speedup, this advantage must be considered in context. It has been argued that a quadratic speedup alone may be insufficient to provide a practical advantage, as the immense overheads associated with fault-tolerant quantum error correction could nullify the gains over classical algorithms [380]. Furthermore, Grover-style algorithms solve optimization problems heuristically by treating the problem as a black box and forgoing any exploitable structure; they are the quantum equivalent of brute force and do not provide approximation guarantees. This has motivated the development of other approaches, such as QAOA, which are designed to find high-quality *approximate* solutions and are seen as a potential path to achieving practical performance benefits.

A common misconception is that variational heuristic algorithms like QAOA are merely a stop-gap for the noisy intermediate-scale quantum (NISQ) era [381], to be supplanted by algorithms with provable guarantees once fault-tolerant quantum computers (FTQCs) are available. This view is mistaken. In fact, running QAOA on problem sizes of industrial relevance will likely *require* fault-tolerant hardware [382]. Furthermore,

concerns that the classical optimization of QAOA's parameters may be a bottleneck for large-scale problems are being addressed by promising research into techniques like transfer learning [383]. Indeed, a growing body of numerical evidence supports QAOA as a leading candidate for the go-to optimization algorithm in future quantum computing scenarios, making it a key potential application for the quantum networks we envision [384–387].

Given the long-term relevance of QAOA, a deeper theoretical understanding of its performance is crucial. However, theoretical results regarding its runtime—specifically, the number of rounds p required to achieve a given approximation ratio—are still relatively sparse, especially beyond the low-depth regime. A key contribution of this thesis is to address this gap by deriving some of the first analytical lower bounds on p. We achieve this by leveraging a powerful connection between the digital, round-based structure of QAOA and the continuous-time evolution of quantum annealing, building on recent results that established bounds on annealing times. This analysis will provide crucial insights into how to plan for and manage the resources needed to run large-scale optimization tasks on the distributed quantum networks of the future.

6.2 Background on quantum optimization algorithms

To understand the computational workloads a quantum network must support, we first review several foundational algorithms for quantum optimization. This review will establish the key concepts, performance metrics, and operational principles that are relevant to the analysis in the subsequent sections of this chapter.

The combinatorial optimization problem

The core task we consider is the combinatorial optimization problem. We define this problem on bitstrings of length n, with an objective function $C(x)$ that maps a given bitstring x to a non-negative integer value. The goal is to find an input z that maximizes this objective function, yielding the value $C_{max} = \max_z C(z)$. For our analysis, it is also useful to define the average of the objective values, $C_{avg} = \frac{1}{N}\sum_z C(z)$, and their standard deviation, $\sigma_C = \sqrt{\sum_z \frac{(C(z)-C_{avg})^2}{N}}$. We will consider unconstrained problems where the search space F consists of all $N = 2^n$ possible bitstrings.

To approach this with a quantum algorithm, it is natural to translate the objective function into a corresponding problem Hamiltonian. We first define a diagonal cost Hamiltonian H_C whose eigenvalues are the objective function values

$$H_C = \sum_{z \in F} C(z) \left|z\right\rangle \left\langle z\right| . \tag{6.1}$$

A common way to evaluate the performance of a heuristic optimization algorithm is its

approximation ratio, $\lambda \in [0, 1]$, which we define as

$$\lambda = \frac{\langle H_C \rangle_p}{C_{max}} = \max_{|\psi_p\rangle} \frac{\langle \psi_p | H_C | \psi_p \rangle}{C_{max}}, \tag{6.2}$$

where $|\psi_p\rangle$ is the state prepared by p rounds of an algorithm. Since it is often more convenient to work with Hamiltonians whose ground state represents the solution, we define our primary problem Hamiltonian H_1 as

$$H_1 = \mathbb{1} C_{max} - H_C. \tag{6.3}$$

With this definition, the task of maximizing $C(x)$ is equivalent to finding the ground state of H_1.

Grover's algorithm and adaptive search

One quantum approach to optimization is to treat it as a search problem. Grover's search algorithm is a celebrated quantum algorithm that provides a provable quadratic speedup for finding a marked item in an unstructured database [388]. It can find one of t target inputs among N total possibilities in just $O(\sqrt{N/t})$ queries to an oracle, compared to the $O(N/t)$ required classically. It was later understood that Grover's algorithm is an optimal special case of a more general technique known as *amplitude amplification* [389]. While powerful for finding exact solutions, Grover's algorithm in its basic form does not provide approximation guarantees. However, it can be used as a heuristic optimizer in a strategy called *Grover Adaptive Search*, where the algorithm is run repeatedly to find solutions of increasing quality [390, 391].

Quantum annealing

One leading paradigm for quantum optimization is quantum annealing (QA). The protocol starts in an easy-to-prepare ground state of a simple "mixer" Hamiltonian, H_0. The system is then slowly evolved according to a time-dependent Hamiltonian

$$H(t) = [1 - g(t)]H_0 + g(t)H_1, \tag{6.4}$$

where H_1 is the problem Hamiltonian whose ground state encodes the optimal solution. According to the adiabatic theorem, if this evolution is sufficiently slow, the system remains in the instantaneous ground state throughout the process and ends in the desired solution state of H_1. The primary limitation of this approach is that for many hard problems, the required "annealing time" to ensure adiabaticity can become exponentially long, negating the potential speedup [392, 393].

The quantum approximate optimization algorithm

The quantum approximate optimization algorithm (QAOA) is a leading hybrid quantum-classical algorithm designed to find high-quality approximate solutions to combinatorial

optimization problems [378]. It is structured around p rounds of alternately applying two parameterized unitary operators: a *phase separator* $e^{-i\gamma H_1}$, which encodes the problem Hamiltonian, and a *mixing operator* $e^{-i\beta H_0}$, which drives transitions between different computational basis states [379]. The initial state $|\psi_0\rangle$ is typically the ground state of the mixer H_0. The state after p rounds is thus given by

$$|\psi(\vec{\beta},\vec{\gamma})\rangle_p = (e^{-i\beta_p H_0} e^{-i\gamma_p H_1}) \dots (e^{-i\beta_1 H_0} e^{-i\gamma_1 H_1}) |\psi_0\rangle . \tag{6.5}$$

The goal of the QAOA is to prepare a state that minimizes the energy expectation value $\langle H_1\rangle_p = \langle\psi(\vec{\beta},\vec{\gamma})|H_1|\psi(\vec{\beta},\vec{\gamma})\rangle_p$. This is achieved through a classical optimization loop that varies the angle parameters $\vec{\gamma}$ and $\vec{\beta}$ to find the minimum.

Conceptually, QAOA is deeply connected to Quantum Annealing (QA). In the limit of infinite rounds ($p \to \infty$), QAOA can be seen as a Trotterized discretization of a continuous-time annealing process. More specifically, the alternating application of the two Hamiltonians in QAOA corresponds to a "bang-bang" schedule in QA, where only one driving Hamiltonian is turned on at a time [394]. This connection is significant, as the bang-bang schedule is conjectured to be the optimal annealing schedule for many optimization problems [395].

On the other hand, the QAOA does not need to arise from the Trotterization or discretization of any annealing schedule; instead, it can be understood purely as a variational ansatz designed for optimization [396]. Research in the finite-p regime has shown that for certain problems, QAOA can achieve a constant approximation ratio with a constant number of rounds [378, 397, 398]. Other active areas of investigation include understanding the algorithm's expressibility (i.e., the subset of Hilbert space accessible by a given ansatz) [399–401], the phenomenon of parameter concentration [402, 403], and techniques for improving performance through better initialization or insights from counterdiabatic driving [404–406].

6.3 Amplitude amplification with subdivided phase oracle for optimization problems

A promising avenue for quantum optimization lies in extending the principles of amplitude amplification beyond a simple binary distinction between "good" and "bad" states. Recent work has investigated the use of oracles that apply non-trivial phase shifts, acting uniquely on each input based on a cost function [407–409]. Building on this, we introduce and analyze a specific implementation of this concept, which we term the *subdivided phase oracle* (SPO) [407]. This approach replaces the canonical Grover oracle in the amplitude amplification process, providing a new method for solving combinatorial optimization problems. While the word "oracle" typically refers to a black-box function that recognizes a solution, we use the term here to describe the phase-shifting operator to maintain consistency with the language of amplitude amplification and related literature.

Oracle construction

The core idea of the SPO is to embed the objective value $f(x)$ of each potential solution $|x\rangle$ directly into its quantum phase. Instead of applying a uniform phase shift to all "good" states, the SPO applies a unique phase to every computational basis state in proportion to its quality. The action of the phase oracle, O_ϕ, on a state $|x\rangle$ is formally defined as

$$O_\phi |x\rangle \rightarrow e^{i\phi(x)} |x\rangle , \tag{6.6}$$

where $\phi(x)$ is a function that maps the input x to a phase angle. This mapping is constructed from the objective function $f(x)$ via a simple linear relationship

$$\phi(x) = k f(x) + \theta, \tag{6.7}$$

where k is an adjustable real-valued parameter and θ represents an offset, typically corresponding to the minimum objective value. Since a constant phase offset applied to all states amounts to an irrelevant global phase, the θ term can be ignored. The action of the SPO is therefore fully characterized by the scaling parameter k, and we can write the operator as

$$O_\phi(k) |x\rangle \rightarrow e^{ikf(x)} |x\rangle . \tag{6.8}$$

The oracle's action is invariant under shifts in the objective values (e.g., making them all positive) and permutations of the input labels. For simplicity in our analysis, we can therefore assume without loss of generality that $f(x)$ is a non-decreasing function with a minimum value of $f(x_{min}) = 0$.

The SPO optimization algorithm

Our algorithm incorporates the SPO into the standard amplitude amplification framework, replacing the binary oracle but retaining the conventional diffusion operator. The full process is outlined in Algorithm 1.

Algorithm 1 Optimization via SPO-based Amplitude Amplification

1: Initialize state $|s\rangle = |0\rangle^{\otimes n}$
2: Prepare uniform superposition $|\Psi_0\rangle \leftarrow H^{\otimes n} |s\rangle$
3: **for** a predetermined number of iterations, j **do**
4: Apply the Subdivided Phase Oracle: $|\Psi'\rangle \leftarrow O_\phi(k) |\Psi_{j-1}\rangle$
5: Apply the Diffusion Operator: $|\Psi_j\rangle \leftarrow D |\Psi'\rangle$
6: Measure the final state $|\Psi_j\rangle$

The process begins by preparing an equal superposition of all possible states, $|\Psi_0\rangle$. Then, in each iteration, two operations are applied in sequence. First, the SPO, $O_\phi(k)$, rotates the phase of each computational basis state as described above. Second, the standard Grover diffusion operator, D, is applied, which performs an inversion about

the average of the amplitudes. The diffusion operator is given by

$$D = 2\,|\Psi_0\rangle\,\langle\Psi_0| - I, \tag{6.9}$$

where $|\Psi_0\rangle$ is the uniform superposition state and I is the identity operator. After a set number of iterations, the final state is measured in the computational basis, with a higher probability of yielding states with high objective values.

Amplification mechanism and visualization

To build intuition for how the SPO leads to optimization, it is helpful to visualize its effect on the state amplitudes. Unlike the standard Grover algorithm where all amplitudes remain real, the SPO introduces complex phases. We can represent the amplitude of each basis state $|x\rangle$ as a vector $\vec{\alpha}_x$ in the 2D complex plane. Initially, in the uniform superposition state $|\Psi_0\rangle$, all N amplitude vectors are identical, real, and have a magnitude of $1/\sqrt{N}$, as shown in Fig. 6.1(a).

The application of the SPO, $O_\phi(k)$, corresponds to rotating each of these vectors around the origin by an angle $k\,f(x)$, as seen in Fig. 6.1(b). States with higher objective values are rotated more. This step only changes the relative phases and does not by itself alter the measurement probabilities. The amplification occurs upon applying the diffusion operator, D, which can be understood as a geometric operation in the complex plane:

1. First, the mean of all N amplitude vectors, $\vec{\alpha}_{mean}$, is calculated.
2. Then, each individual amplitude vector $\vec{\alpha}_x$ is reflected about this mean.
3. The new amplitude vector, $\vec{\alpha}'_x$, is given by the transformation $\vec{\alpha}'_x = 2\vec{\alpha}_{mean} - \vec{\alpha}_x$.

This process is depicted in Fig. 6.1(c-f). The reflection preferentially increases the magnitude of vectors that are pointed in a direction roughly opposite to the mean vector. Since states with high objective values are rotated the most by the SPO, they are most likely to end up in this favorable position for amplification. The vector corresponding to the optimal solution (most saturated red) is significantly increased in magnitude, thus amplifying its measurement probability.

This visualization also makes the condition for successful amplification clear. The amplitude of the optimal state, $\vec{\alpha}_{best}$, will increase in magnitude only if it is pointing generally away from the mean vector, $\vec{\alpha}_{mean}$. More formally, significant amplification occurs when the angle between $\vec{\alpha}_{best}$ and $\vec{\alpha}_{mean}$ is in the range $[\pi/2, 3\pi/2]$. The choice of the parameter k in the SPO is what determines this crucial angle. Finding an optimal k is therefore essential for the algorithm's performance, as it ensures the optimal state is consistently positioned for amplification in each iteration.

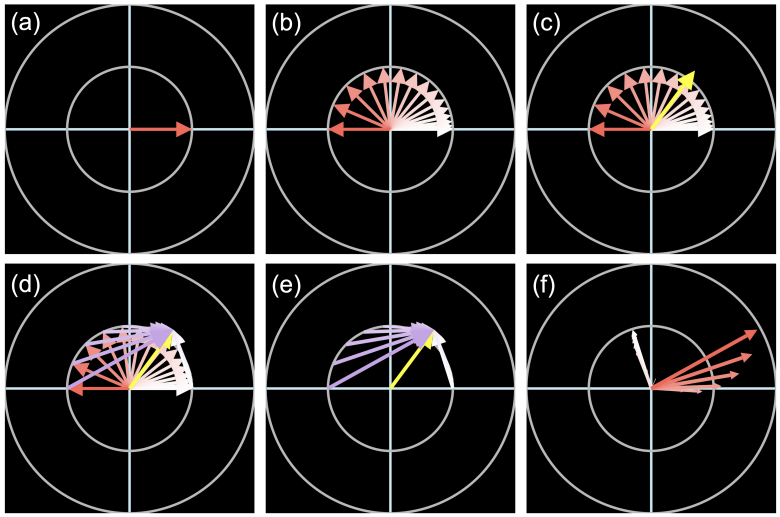


Figure 6.1: A 2D complex plot illustrating how the amplitudes of $N = 16$ states change during one iteration of amplitude amplification using an SPO with $f(x) = x^2$. The color of each vector, from white to red, denotes its objective value. (a) Initial amplitudes in a uniform superposition. (b) Amplitudes after the SPO applies a phase rotation proportional to each state's objective value. (c) The mean amplitude vector ($\vec{\alpha}_{mean}$) is calculated, shown here scaled by a factor of two (yellow arrow). (d, e) The diffusion operator reflects each amplitude vector about the mean. (f) The final amplitudes after one full iteration, showing a significant increase in the magnitude of the optimal state's amplitude. (From [57].)

6.4 Performance analysis for common objective distributions

Before presenting the simulation results, it is important to contextualize the problems being studied. The success of the Subdivided Phase Oracle (SPO) algorithm relies heavily on the distribution of the objective function values, as this distribution dictates the behavior of the mean amplitude vector during the amplification process. To explore the algorithm's performance, we therefore categorize our simulations based on the statistical properties of the objective function, focusing on three representative classes: symmetric (normal), asymmetric (skew normal), and heavy-tailed (exponential). These distributions are chosen not only for their analytical tractability but also because they serve as effective proxies for the types of objective landscapes found in real-world optimization problems.

Symmetric distributions: the normal distribution

We begin by considering objective values that follow a normal distribution, as described by the probability density function

$$\phi_{\text{norm}}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \tag{6.10}$$

where μ is the mean and σ is the standard deviation. This is a natural starting point, as many instances of NP-hard optimization problems, when formulated as Ising problems,

exhibit normally distributed objective values for large, random instances. This behavior arises from the central limit theorem, as the total cost is a sum of many local interaction terms sampled from a single distribution.

The key feature of the normal distribution is its symmetry. This symmetry makes it possible to choose a parameter k for the SPO such that the phase difference between the mean amplitude vector and the optimal solution's amplitude vector is exactly π, leading to highly effective amplification. Using the visualization technique presented in Fig. 6.1, one can see that an appropriate choice of k can ensure that both the mean vector and the optimal solution's vector remain on the real axis throughout the amplification process. However, this symmetry has a notable side effect: because the distribution is symmetric around its mean, the algorithm not only amplifies the probability of the best state $|N-1\rangle$ but also amplifies the probability of the worst state $|0\rangle$ to an equal degree.

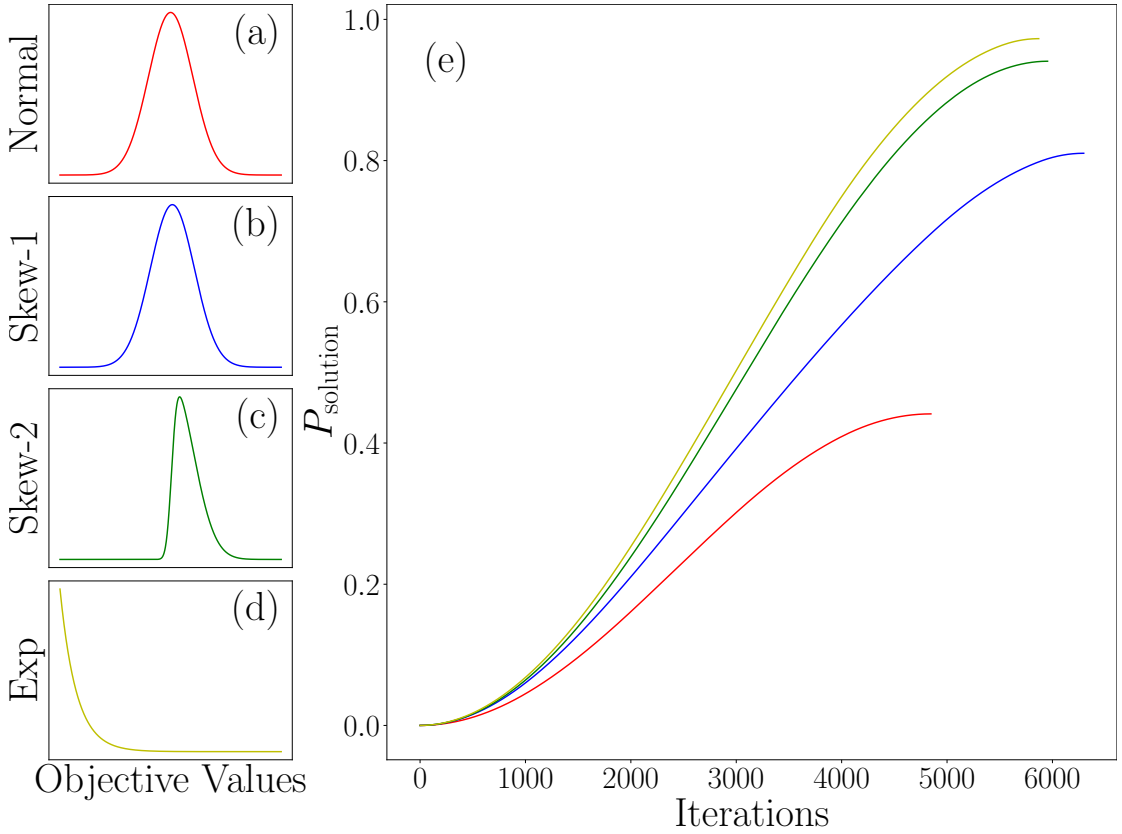


Figure 6.2: The probability of measuring the optimal solution, P_{solution}, as a function of the number of iterations for an optimization problem with search space $N \approx 2^{25}$. The objective values are sampled from four different distributions shown on the left: (a) Normal ($\sigma = 10$), (b) Skew Normal ($\alpha = 0.1$), (c) Skew Normal ($\alpha = 5$), and (d) Exponential. Plot (e) shows that the SPO algorithm successfully amplifies the solution probability to a significant degree for all distributions, though the maximum probability and required iterations vary. (From [57].)

Asymmetric distributions: skew normal and exponential

While the normal distribution provides a useful baseline, real-world optimization problems are unlikely to possess perfect symmetry. We therefore evaluate our algorithm's performance on two classes of asymmetric distributions. We find that not only do these asymmetries not pose an obstacle, they often prove to be advantageous for the algorithm's performance.

First, to capture a slight deviation from normality, we use the *skew normal distribution*, defined as

$$\phi_{\text{skew}}(x;\alpha) = 2\phi_{\text{norm}}(x)\Phi(x;\alpha), \tag{6.11}$$

where $\Phi(x;\alpha)$ is the cumulative distribution function of the standard normal distribution and the parameter α controls the degree and direction of the skew. As shown in Fig. 6.2(b) and (c), a non-zero α breaks the symmetry of the distribution. Second, we consider the *exponential distribution*,

$$\phi_{\text{exp}}(x;\lambda) = \lambda e^{-\lambda x}, \tag{6.12}$$

where λ is the rate parameter. This distribution is characterized by a long tail, representing problems where most solutions are poor, but a few are exceptionally good. For both of these asymmetric distributions, we observe that there exists an optimal parameter k that amplifies the probability of the optimal solution to a significant degree, as depicted in Fig. 6.2(e).

Common observations and performance analysis

By comparing the results across the three distributions, we can identify several common behaviors and key performance characteristics of the SPO algorithm. Fig. 6.2(e) shows how the probability of measuring the optimal solution, P_{solution}, evolves as the algorithm iterates. Starting from an initial probability of $P_{\text{solution}} \approx 1/N$ (here, $N = 2^{25}$), the algorithm successfully amplifies the solution in all cases. The specific distribution, however, affects both the number of iterations required and the maximum achievable probability, P^*_{solution}. For the symmetric normal distribution, $P^*_{\text{solution}} \approx 0.5$. In contrast, the two asymmetric distributions yield much stronger results: $P^*_{\text{solution}} = 0.94$ for the skew normal distribution and $P^*_{\text{solution}} = 0.97$ for the exponential distribution. This demonstrates that asymmetry in the objective landscape is beneficial, as it reduces the amplification of sub-optimal states and concentrates probability more effectively on the single best solution.

If the algorithm is run beyond the point of maximum amplification, P_{solution} begins to oscillate, reminiscent of the behavior in standard Grover search—exhibiting the souffle problem behavior. This brings us to the most critical factor for performance: the choice of the parameter k—obtaining the maximum possible amplification requires selecting an optimal k for a given objective distribution. The algorithm is extremely sensitive to this

choice, and deviating even slightly from the optimal value can cause the achievable amplification to decrease rapidly. Fig. 6.3 illustrates this high sensitivity, showing a sharp peak in performance at the optimal k for each of the three distributions. Finding this optimal k efficiently is a non-trivial task and a key challenge for practical implementations of this algorithm. Finally, a crucial observation is that for these broad distributions, the effectiveness of the algorithm—that is, the maximum achievable P^*_{solution}—is not affected by the size of the input space N. This is a highly desirable feature, signifying the potential applicability of the SPO approach to optimization problems with extremely large search spaces.

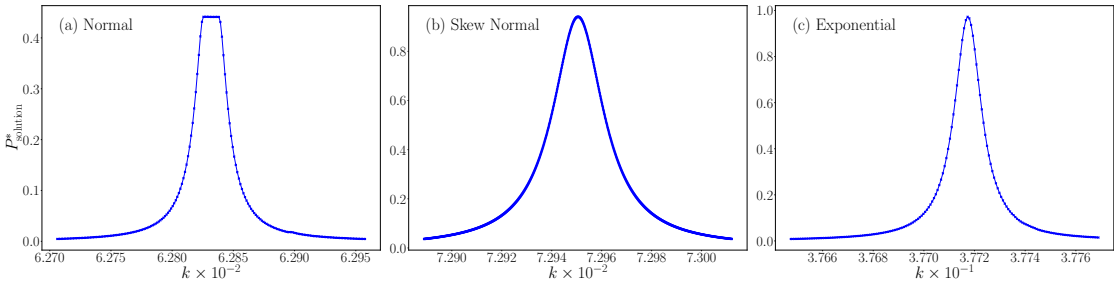


Figure 6.3: Sensitivity of the SPO algorithm's performance, measured by the maximum achievable solution probability P^*_{solution}, as a function of the scaling parameter k. The plots for (a) Normal, (b) Skew Normal, and (c) Exponential distributions all display an extremely sharp peak, demonstrating that performance rapidly degrades when deviating from the optimal k value for a given distribution. (From [57].)

Now that we have introduced the optimization algorithm based on subdivided phase oracle (SPO) and discussed its behavior under fixed-angle amplitude amplification, we turn to the question of formally bounding its runtime. Specifically, we aim to understand whether the fixed-angle strategy is inherently suboptimal, and whether allowing the angles k to vary across rounds—i.e., treating them as variational parameters—could enable performance that surpasses the Grover-like speedup.

From another perspective, the entire SPO algorithm can be viewed as a specific instance of the Quantum Approximate Optimization Algorithm (QAOA). The process of alternating a problem-dependent phase operator (O_ϕ) with a problem-independent mixing operator (the Grover diffusion operator) fits directly into the QAOA ansatz structure, where the diffusion operator acts as a "Grover-style" mixer [377]. This insight provides a powerful bridge between the two approaches and motivates a deeper analysis of the broader QAOA framework.

Recognizing that our SPO algorithm is a member of this more general class of protocols naturally leads to a crucial question: What are the performance limitations of the QAOA framework as a whole? While our results for the SPO are promising, they are specific to one type of mixer. To understand the full potential of QAOA for practical optimization, we must analyze its performance more broadly. A key unresolved issue in the field is

the lack of theoretical guarantees on the runtime of QAOA—specifically, the number of rounds p required to achieve a given solution quality.

6.5 Lower bounds on QAOA runtimes

While QAOA is a promising heuristic, a deeper theoretical understanding of its performance is crucial for evaluating its potential. A key performance metric is the number of rounds, p, required to achieve a desired solution quality, as this directly relates to the algorithm's time complexity. In this section, we present the second main contribution of this chapter: some of the first analytical lower bounds on p required to achieve a guaranteed approximation ratio, λ. We derive these bounds by establishing and exploiting a powerful connection between the digital, round-based structure of QAOA and the continuous-time evolution of quantum annealing.

The connection to quantum annealing

The link between QAOA and Quantum Annealing (QA) provides the foundation for our analysis. Recall that QA evolves a system under a time-dependent Hamiltonian $H(t) = (1 - g(t))H_0 + g(t)H_1$. The QAOA protocol, with its alternating application of unitaries $e^{-i\beta_j H_0}$ and $e^{-i\gamma_j H_1}$, can be viewed as a specific implementation of QA that uses a "bang-bang" schedule. In this type of schedule, the function $g(t)$ only takes the values 0 or 1, meaning only one Hamiltonian, H_0 or H_1, is active at any given time.

Under this interpretation, the sum of the absolute values of the QAOA angle parameters corresponds directly to the total annealing time, t_{anneal}, of the equivalent QA process:

$$t_{anneal} = \sum_{j=1}^{p} \left(|\beta_j| + |\gamma_j| \right). \tag{6.13}$$

This equivalence allows us to apply theoretical results from the study of general quantum annealing directly to the analysis of QAOA.

A general lower bound on QAOA rounds

Our derivation begins with a known lower bound on the annealing time required to reach a ground state with high probability, which holds regardless of the state's specific trajectory during the process.

Theorem 1 (from [59])**.** *Given two driving Hamiltonians with zero ground state energies, H_0 and H_1, a quantum annealing protocol that starts in the ground state of H_0 and ends at time t_f is bounded by*

$$t_f \geq \frac{\langle H_0 \rangle_{t_f} + \langle H_1 \rangle_0 - \langle H_1 \rangle_{t_f}}{\| [H_1, H_0] \|}, \tag{6.14}$$

where $\langle H \rangle_t$ is the expectation value with respect to the Hamiltonian H at time t and $\| \cdot \|$ denotes the spectral norm.

By combining Theorem 1 with the relationship in eq. (6.13), we can establish a lower bound on the sum of the QAOA angles.

Lemma 1. *For a QAOA protocol with p rounds driven by Hamiltonians H_0 and H_1 with zero ground state energies, starting from the ground state of H_0, the angle parameters $\vec{\gamma}$ and $\vec{\beta}$ are bounded by*

$$\sum_{j=1}^{p} \left(|\beta_j| + |\gamma_j|\right) \geq \frac{\langle H_0 \rangle_p + \langle H_1 \rangle_0 - \langle H_1 \rangle_p}{\|[H_1, H_0]\|}. \tag{6.15}$$

Proof. This is an immediate result from Theorem 1 and eq. (6.13). □

To translate this bound on angles into a bound on the number of rounds p, we introduce the reasonable assumption that the driving Hamiltonians are periodic. For the problem Hamiltonian H_1 built from an integer-valued objective function, a period of 2π arises naturally. Most common mixers, H_0, are also periodic. This allows us to bound the meaningful range of any individual angle, leading to our main result.

Theorem 2. *Given a classical objective function $C(x)$ for a maximization task, represented by the Hamiltonian H_C and encoded into the phase separator $H_1 = \mathbb{1}C_{max} - H_C$. Let the mixer be H_0, where all Hamiltonians are 2π periodic with zero ground state energies. If a QAOA protocol with p rounds reaches a state with approximation ratio λ, then*

$$p \geq \frac{\langle H_0 \rangle_p + \lambda C_{max} - C_{avg}}{4\pi \|[H_C, H_0]\|}. \tag{6.16}$$

Proof. From Lemma 1, when imposing the periodicity constraints, we get

$$\sum_{j=1}^{p} \left(|\beta_j| + |\gamma_j|\right) \geq \frac{\langle H_0 \rangle_p + \langle H_1 \rangle_0 - \langle H_1 \rangle_p}{\|[H_1, H_0]\|} \tag{6.17}$$

$$p(2\pi + 2\pi) \geq \frac{\langle H_0 \rangle_p + \langle H_1 \rangle_0 - \langle H_1 \rangle_p}{\|[H_1, H_0]\|} \tag{6.18}$$

$$p \geq \frac{\langle H_0 \rangle_p + \langle H_1 \rangle_0 - \langle H_1 \rangle_p}{4\pi \|[H_1, H_0]\|}, \tag{6.19}$$

where the second line comes from the fact that the two Hamiltonians are 2π periodic, thus, the largest meaningful β_j and γ_j are bounded by 2π. Since the phase separator H_1 of QAOA encodes the information of the objective value into the phase, we get

$$\langle H_1 \rangle_0 = C_{max} - C_{avg}, \tag{6.20}$$

$$\langle H_1 \rangle_p = C_{max} - \lambda C_{max}, \tag{6.21}$$

$$\langle H_1 \rangle_0 - \langle H_1 \rangle_p = \lambda C_{max} - C_{avg}. \tag{6.22}$$

This comes directly from the definitions of the approximation ratio and H_1 defined in eqs. (6.2) and (6.3). As for the denominator, we know that

$$[A + kI, B] = [A, B] \tag{6.23}$$

holds for any constant k and any matrix A, B. Because $H_1 = \mathbb{1}C_{max} - H_C$, we get $\|[H_1, H_0]\| = \|[H_C, H_0]\|$. □

To the best of our knowledge, this is the first result that analytically connects the number of rounds p to the approximation ratio λ for general QAOA on combinatorial optimization problems, independent of the mixer choice or problem structure. It provides a powerful tool for reasoning about the minimum circuit depth required for QAOA to achieve a meaningful result.

The effect of Hamiltonian rescaling

One might ask if this bound can be circumvented by simply rescaling the Hamiltonians, a common practice in quantum annealing to adjust energy ranges. If we rescale H_0 by $\alpha_0 > 0$ and H_1 by $\alpha_1 > 0$, the lower bound on p changes. However, for QAOA, the effect of rescaling is different than for continuous QA. Rescaling the Hamiltonians only changes their period, which means the angles β and γ can be rescaled accordingly without changing the number of rounds p. The number of rounds is therefore invariant under Hamiltonian rescaling.

This observation confirms that deriving the bounds from the 2π periodic case is logically sound. It also implies that we can make the bound tighter by choosing scaling factors that maximize the expression, giving a worst-case lower bound:

$$p \geq \max_{\alpha_0, \alpha_1} \frac{\alpha_0 \langle H_0 \rangle_p + \alpha_1(\langle H_1 \rangle_0 - \langle H_1 \rangle_p)}{2\pi(\alpha_0 + \alpha_1)\|[H_1, H_0]\|}. \tag{6.24}$$

This analysis shows that while the constant factor of the bound can be tightened, its asymptotic behavior cannot be arbitrarily changed by simple rescaling. Having established this general bound, we will now explore its implications for specific, concrete examples.

6.6 Lower bounds on QAOA with a Grover-style mixer

Having established a general lower bound on QAOA performance, we now turn our attention to a specific and important case: QAOA with a Grover-style mixer.

The Grover mixer and its general lower bound

One of the mixers that has been proposed for constrained optimization problems is the *Grover mixer* [377]. This operator is the parameterized version of the usual diffusion

operator in Grover's algorithm and amplitude amplification [388, 389]. It is a natural choice for many problems because it preserves the state within the subspace of feasible solutions and is invariant to permutations of the input. This means all states with the same objective value are treated equally, ensuring fair sampling. This style of QAOA is also studied from the perspective of amplitude amplification, in an attempt to decrease the cost of adaptive Grover searches and to generalize ground-state-finding techniques [57, 407–409] and also our SPO approach section 6.3. While the Grover mixer's ability to rapidly explore the state space provides better scaling for small problem sizes, it has been empirically shown to be potentially detrimental for large problems when using standard classical optimization loops [384, 385, 400, 410–414]. Nonetheless, characterizing its fundamental limitations is crucial, as finding optimal QAOA schedules can be computationally hard [415].

The Hamiltonian for the Grover mixer is defined as the projector onto the uniform superposition of all feasible states, $|\psi_0\rangle$,

$$H_{Grover} = \mathbb{1} - |\psi_0\rangle \langle\psi_0| . \tag{6.25}$$

For this specific mixer, our general lower bound from Theorem 2 simplifies significantly.

Theorem 3. *Given a classical objective function $C(x)$ for a maximization task, if a QAOA protocol with p rounds driven by the Grover mixer H_{Grover} reaches a state with approximation ratio λ, then*

$$p \geq \frac{1 - \left|\langle\psi_0|\psi_p\rangle\right|^2 + \lambda C_{max} - C_{avg}}{4\pi\sigma_C}, \tag{6.26}$$

where C_{max}, C_{avg}, and σ_C are the maximum, average, and standard deviation of the objective function, respectively.

Proof. The term $\langle H_{Grover}\rangle_p$ can be evaluated directly from eq. (6.25). Since H_C is diagonal in the computational basis and H_{Grover} is a rank-1 projector, the term $[H_C, H_{Grover}]$ evaluates to a projector of rank at most 2. There is a closed-form formula for calculating the spectral norm of a matrix of this form given in [416], which gives

$$\|[H_C, H_{Grover}]\| = \sigma_C. \tag{6.27}$$

Plugging these values into theorem 2 gives the lower bound as shown. □

This theorem reveals a crucial insight: the performance of QAOA with a Grover mixer depends only on the statistical distribution of the objective values, not on the combinatorial structure of the problem itself.

Application to k-local Hamiltonians

We can make this bound even more concrete for problems whose objective functions can be described by a k-local Hamiltonian. This is a natural assumption for many combinatorial optimization problems, where the cost function is a sum of terms that each act on at most k variables [417, 418]. Such a Hamiltonian can be written as

$$H_C = \mathbb{1}C_{avg} - \sum_{\nu=1}^{m} \alpha_\nu H_\nu, \tag{6.28}$$

where each H_ν is a product of at most k Pauli Z operators, and α_ν are real-valued weights. For this class of Hamiltonians, the standard deviation σ_C can be calculated directly from the coefficients as $\sigma_C^2 = \sum_{\nu=1}^{m} \alpha_\nu^2$. This leads to a more explicit lower bound.

Theorem 4. *Given a classical objective function $C(x)$ for a maximization task, represented by a k-local Hamiltonian $H_C = C_{avg}\mathbb{1} - \sum_{\nu=1}^{m} \alpha_\nu H_\nu$, where each H_ν is a product of Pauli Z with at most k terms, each α_ν is a real number denoting the weight of the term, and C_{avg} and C_{max} denote the average and the global maximum of $C(x)$. If a QAOA protocol with p rounds driven by the phase separator $H_1 = C_{max}\mathbb{1} - H_C$ and the mixer H_{Grover} that starts in the state $|\psi_0\rangle = |+\rangle^{\otimes n}$ reaches a state $|\psi_p\rangle$ with approximation ratio λ, then*

$$p \geq \frac{1 - \left|\langle\psi_0|\psi_p\rangle\right|^2 + \lambda C_{max} - C_{avg}}{4\pi\sqrt{\sum_{\nu=1}^{m} \alpha_\nu^2}}. \tag{6.29}$$

Proof. Since we have already established that C_{avg} is the coefficient of the trivial Pauli term, we now show that the standard deviation σ_C can be easily calculated from the problem Hamiltonian representation as well. Realizing that

$$\sigma_C^2 = \langle s|H_C^2|s\rangle - \langle s|H_C|s\rangle^2 \tag{6.30}$$

$$= \left(C_{avg}^2 + \sum_{\nu=1}^{m} \alpha_\nu^2\right) - C_{avg}^2 \tag{6.31}$$

$$= \sum_{\nu=1}^{m} \alpha_\nu^2, \tag{6.32}$$

where the first term of the second line is due to the fact that only when all the Pauli Z's cancel out into identity that we get non-zero contributions to the sum. □

This formulation is powerful because it allows us to estimate the lower bound on the required circuit depth directly from the coefficients of the problem Hamiltonian.

Implication for Max-Cut: a polynomial round requirement

The consequences of these bounds are most apparent when applied to a specific problem like Max-Cut. By applying Theorem 3 to a Max-Cut instance on a graph with $|E|$ edges, we can derive a direct corollary. We can form a corollary for Max-Cut to formally state this insight.

Corollary 1. *If a QAOA protocol with p rounds finds an approximate solution with approximation ratio λ to Max-Cut of a graph with $|E|$ edges driven by the objective value phase separator H_1 and the Grover-mixer H_{Grover} that starts in the state $|\psi_0\rangle = |+\rangle^{\otimes n}$, then*

$$p \geq \frac{1 - \left|\langle\psi_0|\psi_p\rangle\right|^2 + \lambda C_{max} - |E|/2}{2\pi\sqrt{|E|}}. \tag{6.33}$$

Proof. Applying the lower bound of Theorem 3 to an instance of Max-Cut with the number of edges denoted by $|E|$. Since there are 4 possible ways to assign a cut partition to an edge, it is easy to see that the expected cut value for each edge is $1/2$, and since any pair of edges are independent of each other, the expected cut values then equals $|E|/2$. We can use a similar line of argument to calculate the standard deviation of the cuts of a graph, which evaluates to $\sqrt{|E|}/2$. □

We know that for any bipartite graph, the maximum number of cut edges will be equal to the number of edges in the graph. Using the above corollary, we get

$$p \geq \frac{1 - \left|\langle\psi_0|\psi_p\rangle\right|^2 + \lambda|E| - |E|/2}{2\pi\sqrt{|E|}} \tag{6.34}$$

$$\geq \frac{(2\lambda - 1)}{4\pi}\sqrt{|E|}. \tag{6.35}$$

Since $|E|$ is at least linear in the number of vertices n when the graph is connected, this shows that there exist problem instances in the class of Max-Cut that QAOA with Grover-mixer cannot get any approximation ratio guarantee with constant rounds and would require the number of rounds to be at least a polynomial in n to obtain a constant approximation ratio. We note that this conclusion is different and cannot be compared directly to results in [419], where it is shown that constant rounds QAOA with transverse field mixer can obtain constant approximation ratio for a certain families of regular bipartite expander graphs due to different assumptions and choice of mixer.

Although we have only shown this formally with the Max-Cut problems, similar calculations can be made for most combinatorial optimization problems.

This is a significant finding, as it demonstrates that a constant-round QAOA with this mixer is not sufficient to solve many common optimization problems with a guarantee of quality.

Discussion on the tightness of the bound

A natural question is whether this polynomial lower bound is tight. While our analysis indicates a polynomial requirement, some empirical studies have suggested that finding optimal angles for QAOA on hard problem instances might require an exponential number of rounds [384, 385]. This apparent discrepancy may arise because the heuristic, round-by-round methods used to find angles in numerical simulations might not achieve the true, globally optimal parameters for a given p. Furthermore, intuition from the low-depth ($p = 1$) regime suggests that each round of this QAOA variant should only contribute a Grover-like progress, which does not fully capture the complex dynamics at higher depths [420]. Therefore, while our lower bound is rigorously proven, the possibility remains that this type of QAOA could offer a significant advantage over a simple Grover adaptive search, and determining its true upper-bound complexity remains an important open question.

6.7 Lower bounds applied to search problems

Having analyzed the lower bounds for general optimization problems, we now turn to the special case of search. A search problem on n qubits can be described as finding a bitstring z from a set $S = \{z|C_{search}(z) = 1\}$, where the objective function $C_{search}(x)$ maps inputs to $\{0, 1\}$. While using the hybrid QAOA framework to solve search is not practically efficient, analyzing this case provides a crucial test for the tightness and generality of our derived bounds. The well-established lower bound for unstructured search is $\Omega(\sqrt{N/m})$ queries to an oracle, where m is the number of marked items. Our goal is to see if our QAOA bounds can recover this fundamental limit.

This analysis is also motivated by a desire to understand the behavior of our bounds when the contribution from the mixer Hamiltonian, $\langle H_0 \rangle_p$, dominates, a scenario not fully explored in the previous section. We will first show that our energy-based bound (Theorem 2) is indeed tight for the Grover mixer. We will then show that for the transverse-field mixer, this same bound reveals a dependency on the problem's structure but is ultimately loose. Finally, we will derive and apply a new, more general bound based on state overlap that correctly captures the unstructured search limit for a broad class of mixers.

Performance with a Grover mixer: a tight bound

We begin by showing that the lower bound from Theorem 3 is tight when applied to search problems using the Grover mixer. Since Grover's algorithm can be viewed as a QAOA with fixed angles, our bound is expected to recover the well-known scaling.

Corollary 2. *If a QAOA protocol solves a search problem with p rounds, finding a marked state from m marked states, with success probability λ and is driven by*

the objective value phase separator H_1 (phase oracle with adjustable phase) and the Grover-mixer H_{Grover} that starts in the state $|\psi_0\rangle = |+\rangle^{\otimes n}$, then

$$p \geq \frac{\lambda}{2\pi}\sqrt{\frac{N-m}{m}} - \frac{1}{2\pi}\sqrt{\lambda(1-\lambda)}. \tag{6.36}$$

Proof. We can see that, by rewriting the states as linear combinations of solution and non-solution states, we get

$$\left|\langle\psi_0|\psi_p\rangle\right|^2 = \frac{1}{N}\left(\sqrt{\lambda m} + \sqrt{(1-\lambda)(N-m)}\right)^2, \tag{6.37}$$

$$C_{max} = 1, \text{ and } C_{avg} = m/N. \tag{6.38}$$

The standard deviation for the search cost function is

$$\sigma_C = \frac{\sqrt{m(N-m)}}{N}. \tag{6.39}$$

Plugging these into the lower bound in Theorem 3 gives this expression, thus completing the proof. □

As expected, the expression in eq. (6.36) recovers Grover's search scaling of order $\sqrt{N/m}$ when $m \ll N$. This result confirms that our energy-based lower bound is tight and consistent with the optimal query complexity for unstructured search when the appropriate mixer is used.

Performance with a transverse-field mixer: a structure-dependent bound

Next, we examine the search problem with the transverse-field (H_{TF}) mixer. Unlike the Grover mixer, H_{TF} is not invariant under input permutations, and its performance depends on the structure of the marked states. To illustrate this, we analyze two different search problems with the same number of solutions but different structures.

First, we consider a set of marked states S_{dist-3} where the Hamming distance between any two marked states is at least 3. For this problem, the lower bound evaluates to

$$p \geq \frac{n(1-2\sqrt{\lambda(1-\lambda)}) + 2\lambda - 2m/N}{4\pi\sqrt{n}}, \tag{6.40}$$

which is on the order of $\sqrt{n}$, regardless of the number of marked states m.

Second, we consider a set $S_{Hamming-k}$ where all states with a fixed Hamming weight k are marked. In this case, the lower bound becomes

$$p \geq \frac{n(1-2\sqrt{\lambda(1-\lambda)}) + 2\lambda - 2m/N}{4\pi\sqrt{2k(n-k)+n}}. \tag{6.41}$$

This bound depends on k; for instance, when $k = n/\log^2 n$, the lower bound is on the order of $\log n$. These two examples show that for the same number of marked states, the lower bound can yield different scaling ($\sqrt{n}$ vs. $\log n$), revealing that the energy-based bound for the transverse-field mixer is sensitive to the problem's structure. However, since neither recovers the expected $\Omega(\sqrt{N/m})$ worst-case scaling, we conclude that this bound is loose for the general case of unstructured search with H_{TF}.

A general and tighter bound via state overlap

To resolve the looseness of the energy-based bound for certain mixers, we can derive a new bound based on a different principle: the rate of change of state overlap. This leads to the following lemma.

Lemma 2. *Given a QAOA protocol driven by Hamiltonians H_0 and H_1, where H_1 is 2π periodic, and a projector P_0 that commutes with H_0, the number of rounds p is bounded by*

$$p \geq \frac{\left|\langle P_0\rangle_p - \langle P_0\rangle_0\right|}{2\pi \left\|[P_0, H_1]\right\|}. \tag{6.42}$$

Proof. Consider dynamics with a Hamiltonian

$$H(t) = (1 - g(t))H_0 + g(t)H_1, \tag{6.43}$$

where the value of $g(t) = \{0, 1\}$ in the case of QAOA. If P_0 is a projector that commutes with H_0, we have that

$$\left|\frac{d}{dt}\langle P_0\rangle_t\right| = \left|-i \operatorname{Tr}\left([H(t), \rho_t]P_0\right)\right| \tag{6.44}$$

$$= \left|\operatorname{Tr}\left(\rho_t[P_0, g(t)H_1]\right)\right| \tag{6.45}$$

$$\leq |g(t)|\left\|[P_0, H_1]\right\|, \tag{6.46}$$

where the last line arrived from the cyclic property of trace and Rayleigh quotient. Integrating over the duration t_f of the protocol gives

$$\left|\langle P_0\rangle_{t_f} - \langle P_0\rangle_0\right| = \left|\int_0^{t_f} \frac{d}{dt}\langle P_0\rangle_t \, dt\right| \tag{6.47}$$

$$\leq \left\|[P_0, H_1]\right\| \int_0^{t_f} |g(t)|dt \tag{6.48}$$

$$\leq \left\|[P_0, H_1]\right\| \sum_{j=1}^{p} |\gamma_j| \tag{6.49}$$

$$\leq \left\|[P_0, H_1]\right\| 2\pi p. \tag{6.50}$$

where we assumed $|\gamma_j| < 2\pi$ due to the periodicity. Then,

$$p \geq \frac{\left|\langle P_0\rangle_p - \langle P_0\rangle_0\right|}{2\pi \left\|[P_0, H_1]\right\|}, \tag{6.51}$$

completing the proof. □

While the numerator is always between 0 and 1, this bound proves to be much stronger and more general for search problems.

General lower bound for search with QAOA

By choosing $P_0 = |\psi_0\rangle \langle\psi_0|$ where $|\psi_0\rangle$ is the uniform superposition state, we can apply the above lemma to form a new, more general theorem for search.

Theorem 5. *If a QAOA protocol solves a search problem with p rounds, finding one of m marked states with success probability λ, and is driven by any mixer H_0 that has $|\psi_0\rangle = |+\rangle^{\otimes n}$ as its ground state, then*

$$p \geq \frac{\lambda(N-2m) + m - 2\sqrt{\lambda(1-\lambda)m(N-m)}}{2\pi\sqrt{m(N-m)}}. \tag{6.52}$$

Proof. Consider the lower bound expression in Lemma 2, when choosing $P_0 = |\psi_0\rangle \langle\psi_0|$ we see that

$$[P_0, H_1] = -[\mathbb{1} - P_0, H_1] = [H_1, H_{Grover}]. \tag{6.53}$$

Then using Lemma 2, we get

$$p \geq \frac{\left|\langle P_0\rangle_p - \langle P_0\rangle_0\right|}{2\pi \left\|[P_0, H_1]\right\|} = \frac{1 - \left|\langle\psi_0|\psi_p\rangle\right|^2}{2\pi\sigma_C}. \tag{6.54}$$

By plugging in the values from eq. (6.37) and eq. (6.39) into the above inequality, we get the expression as claimed, thus completing the proof. □

This powerful result successfully recovers the optimal Grover scaling for the unstructured search algorithm by providing a lower bound of $\Omega(\sqrt{N/m})$ that matches the known upper bound [388]. Crucially, it holds not only for the Grover mixer but also for the transverse-field mixer and any other mixer that begins in the standard uniform superposition state. This confirms that while the energy-based bound can reveal problem structure, this overlap-based bound correctly captures the fundamental query complexity for unstructured search within the QAOA framework. This result can also be adapted to apply to continuous-time quantum walk searches that start in a stationary distribution.

6.8 Implications for network architecture and resource management

The analysis of QAOA runtimes is not merely an exercise in algorithmic complexity theory; it provides fundamental, actionable insights that connect directly to the architectural challenges of designing heterogeneous quantum repeater networks. The performance characteristics of the applications that will run on a future Quantum Internet are the primary drivers of its architectural requirements. By establishing lower bounds on the number of rounds for a leading class of optimization algorithms, we can begin to quantitatively model the demands these applications will place on network resources.

Computation models, architecture, and traffic implications

The number of QAOA rounds p is a direct proxy for how long a distributed quantum computation will occupy network resources. This runtime determines the coherence time required at quantum memories, the operational lifetime of end-to-end entangled states, and the entanglement generation rate needed across elementary links to sustain the computation.

For industrially relevant problems, the number of algorithmic logical qubits often exceeds the capacity of a single device, motivating the distributed architecture central to this thesis. The resulting traffic model between nodes depends heavily on how qubits or subcircuits are partitioned across the network, which in turn is influenced by the architecture of the quantum devices themselves—including their physical qubit connectivity, the error-correcting codes employed, and the computational model used to realize inter-node operations.

These variables span different abstraction levels and can be managed individually (e.g., via hardware-specific compilers) or through a unified optimization framework that globally schedules physical qubit operations and Bell-pair usage. The best strategy depends on system scale and use case. Tightly coupled clusters, such as quantum datacenters, may benefit from holistic optimization. In contrast, wide-area networks or privacy-sensitive architectures like blind quantum computation require fundamentally different computational models that cannot be globally optimized due to the obfuscation of operational steps and the need for verification. Since nodes may not be fully trusted, protocols must enforce privacy and correctness, which introduces translation overheads and favors coarser abstractions—significantly altering resource demands.

To better understand these trade-offs, we now focus on the datacenter setting, where high trust and physical proximity simplify coordination and optimization. The choice of fault-tolerant computational model plays a crucial role in determining the timing, fidelity, and structure of Bell pairs needed for inter-node operations. Gate-based models typically use Bell pairs to implement nonlocal gates (via telegates) or to teleport qubits (teledata). Efficient partitioning and scheduling can significantly reduce entan-

glement consumption [421, 422]. By contrast, measurement-based models such as one-way MBQC [423], Pauli-based computation [424], lattice surgery [425–427], or CSS surgery [428] often require large graph or GHZ states as primitives, especially in topologies with dense connectivity.

An alternative approach is the *active volume* model [429], which combines architectural and computational perspectives. It can reduce overheads by orders of magnitude while requiring only Bell pairs, assuming a high-connectivity interconnect such as proposed in [430]. This paradigm fundamentally alters the traffic model, emphasizing fewer but more strategic Bell-pair uses over graph-state generation.

At a lower level of abstraction, code-specific implementations—such as blocklet-based surface code execution [431], qLDPC-based schemes [222], or distributed Floquet codes [221]—enable further optimization by aligning physical layout and scheduling with the error-correcting code structure. These models promise reduced resource consumption but demand precise control over how logical operations are mapped and executed. However, they currently lack efficient compilers capable of translating standard gate-based algorithms without incurring significant overheads, often quadratic in size [380]. New compilation efforts targeting these models are ongoing [74, 432–435].

Ultimately, accurate traffic modeling must incorporate the underlying error-correcting code (e.g., surface code vs. qLDPC [222]), the nature of communication primitives, and algorithm-specific patterns. Our findings show that QAOA runtimes span a broad spectrum: from constant-time $O(1)$ algorithms for easy instances, to polynomial-time $\text{poly}(n)$ scaling for structured problems, and even $O(\sqrt{N})$ runtimes for unstructured search-like tasks. This variability is crucial for classifying network traffic types. Short tasks generate bursty, lightweight entanglement demands. In contrast, long-running computations with polynomial complexity require sustained, high-rate delivery of high-fidelity Bell pairs over extended periods. Capturing this diversity in temporal and resource requirements is essential for building realistic traffic models in network simulators like QuISP. To accurately predict network performance, a simulator's traffic generator must be able to model this diversity of application demands, which arise directly from the algorithmic properties we have analyzed.

Resource allocation strategies informed by algorithmic runtimes

Understanding these algorithmic runtime profiles directly informs the architectural choices described in chapter 4, particularly with respect to resource allocation and connection setup strategies. The different runtime classes of QAOA translate naturally into different network operation modes.

Long-running applications with predictable runtimes and high entanglement demands (e.g., those with polynomial or square-root scaling) are ill-suited to a best-effort, con-

nectionless model. These tasks require end-to-end guarantees on fidelity and availability, motivating the use of a connection-oriented approach with dedicated resource reservation—akin to circuit switching. This is the paradigm implemented by our RuleSet protocol and two-pass connection setup mechanism, which allocate and manage network resources over the full duration of such computations.

Conversely, shorter and bursty workloads—such as constant-round algorithms—can be handled more efficiently with opportunistic resource allocation via dynamic, connectionless protocols or statistical multiplexing. This knowledge allows a network's Connection Manager (CM) within QRSA to make more intelligent decisions, authoring RuleSets that requrest the appropriate resource reservation strategy based on the anticipated runtime of the application.

These architectural insights are especially relevant for advanced scenarios like blind quantum computation, where a cluster of distributed quantum devices must coordinate on a large-scale computation while preserving privacy and correctness. Such applications impose stringent requirements on entanglement quality, reliability, and scheduling, all of which depend on robust error management protocols like those described in section 3.5. The runtime lower bounds derived in this work offer concrete, analytical tools to reason about total resource requirements and guide Quality of Service (QoS) decisions for these demanding use cases.

6.9 Conclusion

This chapter highlights the critical synergy between quantum algorithm analysis and quantum network design. We have demonstrated that understanding the performance and resource requirements of key applications, such as optimization algorithms like QAOA and the Subdivided Phase Oracle, is fundamental to designing practical and efficient quantum networks. The analytical lower bounds derived here are not just abstract results in complexity theory; they are essential engineering parameters that inform everything from traffic modeling to the choice of high-level architectural paradigms.

By quantifying the runtime demands of these algorithms, we provide the necessary tools for more effective traffic modeling, resource allocation, and overall architectural planning. This foundational understanding allows us to intelligently apply the principles of the Quantum Recursive Network Architecture (QRNA) and the RuleSet-based control plane. It enables a network architect to anticipate the needs of different computational workloads and provision resources accordingly, a critical capability for building the heterogeneous quantum repeater networks that will form the future Quantum Internet.

Chapter 7

DISCUSSION AND CONCLUSION

7.1 Summary of contributions

This thesis has addressed some of the most pressing engineering and architectural challenges facing the development of a scalable, heterogeneous Quantum Internet. By drawing on principles from classical network design and focusing on the practical requirements of real-world implementation, we have developed a cohesive framework that spans from high-level architecture to specific hardware-level protocols and their analysis. The main ideas and contributions presented can be summarized by four key pillars.

First, building upon foundational concepts from prior research, we introduced a comprehensive network architecture for memory-based quantum repeaters, grounded in the Quantum Recursive Network Architecture (QRNA) and managed by programmable, event-driven RuleSets. This framework, realized in the Quantum Routing Software Architecture (QRSA), provides a modular and extensible foundation for managing network resources, defining connection semantics, and enabling interoperability between diverse and independently operated quantum networks.

Second, to validate this architecture and explore protocol performance, we utilized and contributed to the development of the QuISP network simulator, a project that began prior to my joining the research group. We established its credibility through a rigorous cross-validation study against another independent simulator, demonstrating that despite different modeling philosophies, QuISP faithfully reproduces the expected physical behaviors of quantum networks. This work affirms that well-designed simulation is an indispensable tool for guiding the design and analysis of complex quantum network protocols and architectures.

My primary and novel contributions then extend this architectural model. Third, we extended our architectural framework to the domain of all-photonic repeaters. We introduced the half-Repeater Graph State (half-RGS) as a modular building block that resolves key engineering challenges in monolithic RGS generation and timing synchronization. Building on this, we developed a purification-enhanced RGS scheme, demonstrating for the first time a practical method for non-neighbor purification in a memory-less repeater chain. This work transforms the all-photonic segment into a robust, high-performance component that can be abstracted as a "virtual link" within the broader QRNA framework.

Finally, to ground our architectural work in the needs of real applications, I led the analysis of the performance of quantum optimization algorithms. We presented some of the first analytical lower bounds on the number of rounds required for the Quantum Approximate Optimization Algorithm (QAOA) to achieve a guaranteed approximation ratio. This analysis provides concrete data on the runtime and resource requirements of a key application class, enabling more realistic traffic modeling and informing resource management strategies for networks designed to support distributed quantum computation.

7.2 Revisiting the research questions

Having summarized the contributions, we now evaluate how this work addresses the primary research questions posed in the introduction of this thesis.

The first question asked: *How do we design a unified network architecture that supports heterogeneous quantum repeaters?* Chapters 4 and 5 of this thesis directly address this question. Chapter 4 laid out the QRNA and RuleSet framework as a candidate for a unified architecture, designed specifically to manage different technologies through abstraction. Chapter 5 then provided a crucial analysis by showing how all-photonic repeaters—a fundamentally different paradigm—can be successfully integrated into this memory-centric architecture via the half-RGS and virtual link abstractions. While full heterogeneity remains a challenge, this work provides a concrete and viable architectural blueprint for achieving it.

The second question was: *How can simulation guide the design and resource management of scalable quantum networks?* This thesis answers this question both by example and by providing tools. The cross-validation study in Chapter 4 established QuISP as a reliable tool for architectural exploration. Furthermore, the algorithmic analysis in Chapter 6 demonstrates how simulation can be used: by taking the runtime bounds of algorithms like QAOA, we can create realistic traffic models that allow simulators like QuISP to study network-wide phenomena like resource contention and scheduling policy performance, which would be intractable to analyze purely analytically.

The third question asked: *How can all-photonic quantum repeaters become practical and interoperable with memory-based designs?* Chapter 5 provides a partial but significant answer. By introducing the half-RGS to simplify generation and by solving the long-standing purification problem, we make the RGS scheme substantially more robust and practical for high-error environments. Furthermore, the "virtual link" abstraction provides the explicit mechanism for interoperability, defining how a memory-based control plane using RuleSets can manage and utilize a memory-less all-photonic subnetwork. The full realization of this vision still depends on experimental progress, but our work provides the necessary architectural and protocol-level solutions.

7.3 Future work and open questions

The contributions of this thesis lay the groundwork for numerous avenues of future research, ranging from near-term engineering projects to broader, more fundamental open questions for the field.

Near-term research directions

Several concrete projects can directly extend the work presented here:

1. Integrating Segment Routing into QRNA: A promising direction is to adapt segment routing concepts from classical networks to the QRNA framework. This could involve embedding segment identifiers corresponding to recursive virtual links, potentially streamlining route computation and enabling richer, programmable routing policies.

2. Expanding Simulation Capabilities in QuISP: Further work is needed to mature the QuISP simulator by expanding its error modeling and protocol support. Key areas for development include: (1) implementing real-time link monitoring, a responsibility of the Hardware Monitor that could draw on techniques from recent work [239–241]; (2) integrating routing algorithms that explicitly account for link error characteristics; (3) enabling adaptive routing to handle link failures; (4) incorporating the full range of node types defined in section 4.5; (5) adding support for quantum error-correcting codes to allow for the simulation of 2G, 3G, and all-photonic RGS repeaters; (6) implementing diverse multiplexing strategies; (7) completing the resource pool implementation to fully match the RuleSet specification; and (8) developing an advanced, configurable traffic modeling engine. While not exhaustive, this list represents a selection of the highest-priority areas for future development.

3. Bridging Simulation and Experiment: A crucial task is to develop methodologies for translating real-world experimental testbed parameters into QuISP simulations. Improving this pipeline is essential for using simulation to accurately predict or replicate the behavior of physical hardware.

4. Mapping Architecture to Hardware: A key engineering challenge is to map the abstract software components of the Quantum Routing Software Architecture (QRSA) onto concrete hardware platforms, defining the specific interfaces and control logic required for a practical implementation.

Broader open questions for the field

This research also highlights several deep, open questions that will shape the future of quantum networking:

1. In quantum error correction (QEC), a robust theoretical foundation exists for defining fault-tolerance paradigms, which establishes that analyzing Pauli channels is often sufficient for studying both phenomenological and circuit-level noise. However, the direct applicability of this foundation to quantum networking is less clear, especially since the dominant source of error is often photon loss. A key open question is therefore: Can a similarly robust theoretical framework be developed to enable efficient network simulations that accurately represent realistic noise models, including loss, without requiring full density-matrix simulation? Answering this is essential for building confidence in our near-term simulation and design tools.

2. This thesis has presented engineering advancements for all-photonic quantum repeaters, suggesting their operational characteristics can increasingly resemble those of memory-based networks. We have also proposed methods for managing network heterogeneity using the Quantum Recursive Network Architecture (QRNA) framework. To translate these insights into practice, a concrete research question arises: When comparing 2G and RGS-based repeaters, which technology is better suited for deployment near the network edge versus the backbone? This question requires considering realistic network topologies and the known robustness and fragility trade-offs of the classical Internet to guide optimal deployment strategies.

3. We have demonstrated that RGS-based segments can be successfully abstracted as virtual links within the QRNA framework. This raises a further architectural question: Are there other repeater models or paradigms that cannot be so readily abstracted and would thus necessitate fundamental modifications to the conceptual levels of QRNA? For instance, how might repeaters based on the Zeno effect, or those utilizing continuous-variable (CV) quantum information or other bosonic encodings, fit within or challenge this recursive architectural model? Fully understanding the integration requirements of these diverse repeater types is necessary to test the universality of the architecture.

4. A crucial challenge lies at the boundary between the network and its users: What is the optimal interface between the network-level operations managed by QRSA and RuleSets, and the application layer that consumes the generated entanglement? Defining appropriate "quantum socket" abstractions is critical. While recent work offers one vision for this interface by coupling application logic to specific hardware capabilities , it remains an open question whether this direction is the most promising path forward, or if more hardware-agnostic interfaces are required for a truly interoperable application ecosystem.

7.4 Perspective

Ultimately, the journey from isolated laboratory demonstrations to a globally interconnected Quantum Internet will not be defined by a single breakthrough technology. Instead, progress will come from our ability to compose a diverse ecosystem of specialized solutions—guided by the principle of openness and interoperability, much like the inclusive spirit of the early classical Internet. The central argument of this thesis is that designing for such heterogeneity is not a secondary concern to be addressed after scalable quantum systems arrive—it is the foundational engineering challenge of our time. Waiting for the experimental and system to settle is a risky strategy; history shows that once technologies are widely adopted, retrofitting for compatibility becomes prohibitively difficult. By demonstrating that memory-based and all-photonic repeater paradigms can be unified under a single, recursive architectural framework—and by grounding that framework in the concrete demands of emerging applications—this work offers a practical roadmap for an architecture that is modular, evolvable, and prepared for future innovation.

Realizing this vision will require sustained collaboration across the distinct cultures of theoretical computer science, quantum physics, and network engineering. The challenges remain formidable, but they are no longer merely conceptual; they are now the concrete engineering problems of interface design, protocol optimization, and resource management. While many of the core theoretical concepts are in place, progress will be driven as much by clever engineering as by new physics. As has been said in classical computing by Leiserson et al., there is still "plenty of room at the top" [436] for the clever optimizations and smarter implementations that can improve performance by orders of magnitude. Good engineering, done well, is what will transform today's feasibility into tomorrow's utility. The architectural principles and systems explored in this thesis are not the final answer, but they are a step toward it. They offer a foundation for future researchers, developers, and engineers to build on, refine, and extend—toward a future where the Quantum Internet is not just possible, but real.

BIBLIOGRAPHY

[1] K. Azuma et al., "Quantum repeaters: From quantum networks to the quantum internet," *Reviews of Modern Physics*, vol. 95, no. 4, p. 045 006, Dec. 20, 2023, DOI: 10.1103/RevModPhys.95.045006.

[2] K. Azuma, K. Tamaki, and H.-K. Lo, "All-photonic quantum repeaters," *Nature Communications*, vol. 6, no. 1, p. 6787, 1 Apr. 15, 2015, DOI: 10.1038/ncomms7787.

[3] C. H. Bennett and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing," in *International Conference on Computer System and Signal Processing*, India: IEEE, 1984, pp. 175–179.

[4] A. K. Ekert, "Quantum cryptography based on Bell's theorem," *Physical Review Letters*, vol. 67, no. 6, pp. 661–663, Aug. 5, 1991, DOI: 10.1103/PhysRevLett.67.661.

[5] C. L. Degen, F. Reinhard, and P. Cappellaro, "Quantum sensing," *Reviews of Modern Physics*, vol. 89, no. 3, p. 035 002, Jul. 25, 2017, DOI: 10.1103/RevModPhys.89.035002.

[6] T. J. Proctor, P. A. Knott, and J. A. Dunningham, "Networked quantum sensing," 2017, arXiv: 1702.04271 [math-ph, physics:quant-ph].

[7] A. Broadbent, J. Fitzsimons, and E. Kashefi, "Universal blind quantum computation," in *2009 50th Annual IEEE Symposium on Foundations of Computer Science*, Oct. 2009, pp. 517–526, DOI: 10.1109/FOCS.2009.36.

[8] T. Morimae and K. Fujii, "Blind quantum computation protocol in which Alice only makes measurements," *Physical Review A*, vol. 87, no. 5, p. 050 301, May 13, 2013, DOI: 10.1103/PhysRevA.87.050301.

[9] J. F. Fitzsimons, "Private quantum computation: An introduction to blind quantum computing and related protocols," *npj Quantum Information*, vol. 3, no. 1, pp. 1–11, Jun. 15, 2017, DOI: 10.1038/s41534-017-0025-3.

[10] S. Muralidharan, L. Li, J. Kim, N. Lütkenhaus, M. D. Lukin, and L. Jiang, "Optimal architectures for long distance quantum communication," *Scientific Reports*, vol. 6, no. 1, p. 20 463, 1 Feb. 15, 2016, DOI: 10.1038/srep20463.

[11] R. Valivarthi et al., "Quantum teleportation across a metropolitan fibre network," *Nature Photonics*, vol. 10, no. 10, pp. 676–680, Oct. 2016, DOI: 10.1038/nphoton.2016.180.

[12] Q.-C. Sun et al., "Quantum teleportation with independent sources and prior entanglement distribution over a network," *Nature Photonics*, vol. 10, no. 10, pp. 671–675, Oct. 2016, DOI: 10.1038/nphoton.2016.179.

[13] R. Valivarthi et al., "Teleportation Systems Toward a Quantum Internet," *PRX Quantum*, vol. 1, no. 2, p. 020 317, Dec. 4, 2020, DOI: 10.1103/PRXQuantum.1.020317.

[14] M. Pompili et al., "Realization of a multi-node quantum network of remote solid-state qubits," *Science*, vol. 372, no. 6539, pp. 259–264, Apr. 16, 2021, DOI: 10.1126/science.abg1919.

[15] M. Alshowkan et al., "Reconfigurable Quantum Local Area Network Over Deployed Fiber," *PRX Quantum*, vol. 2, no. 4, p. 040 304, Oct. 6, 2021, DOI: 10.1103/PRXQuantum.2.040304.

[16] S. L. N. Hermans, M. Pompili, H. K. C. Beukers, S. Baier, J. Borregaard, and R. Hanson, "Qubit teleportation between non-neighbouring nodes in a quantum network," *Nature*, vol. 605, no. 7911, pp. 663–668, May 26, 2022, DOI: 10.1038/s41586-022-04697-y.

[17] T. van Leent et al., "Entangling single atoms over 33 km telecom fibre," *Nature*, vol. 607, no. 7917, pp. 69–73, Jul. 2022, DOI: 10.1038/s41586-022-04764-4.

[18] J. Chung et al., "Design and Implementation of the Illinois Express Quantum Metropolitan Area Network," *IEEE Transactions on Quantum Engineering*, vol. 3, pp. 1–20, 2022, DOI: 10.1109/TQE.2022.3221029.

[19] V. Krutyanskiy et al., "Entanglement of Trapped-Ion Qubits Separated by 230 Meters," *Physical Review Letters*, vol. 130, no. 5, p. 050 803, Feb. 2, 2023, DOI: 10.1103/PhysRevLett.130.050803.

[20] S. I. Davis et al., "Teleportation of entanglement in the Fermilab Quantum Network," in *CLEO 2024*, Charlotte, North Carolina: Optica Publishing Group, 2024, FTu3F.7, DOI: 10.1364/CLEO_FS.2024.FTu3F.7.

[21] A. N. Craddock, A. Lazenby, G. B. Portmann, R. Sekelsky, M. Flament, and M. Namazi, "Automated Distribution of Polarization-Entangled Photons Using Deployed New York City Fibers," *PRX Quantum*, vol. 5, no. 3, p. 030 330, Aug. 9, 2024, DOI: 10.1103/PRXQuantum.5.030330.

[22] J.-L. Liu et al., "Creation of memory–memory entanglement in a metropolitan quantum network," *Nature*, vol. 629, no. 8012, pp. 579–585, May 2024, DOI: 10.1038/s41586-024-07308-0.

[23] C. M. Knaut et al., "Entanglement of nanophotonic quantum memory nodes in a telecom network," *Nature*, vol. 629, no. 8012, pp. 573–578, May 2024, DOI: 10.1038/s41586-024-07252-z.

[24] S. I. Davis et al., "Entanglement swapping systems toward a quantum internet," 2025, arXiv: 2503.18906 [quant-ph].

[25] A. Ruskuc et al., "Multiplexed entanglement of multi-emitter quantum network nodes," *Nature*, vol. 639, no. 8053, pp. 54–59, Mar. 2025, DOI: 10.1038/s41586-024-08537-z.

[26] M. Sena et al., "Robust High-Fidelity Quantum Entanglement Distribution over Large-Scale Metropolitan Fiber Networks with Co-propagating Classical Signals," 2025, arXiv: 2504.08927 [quant-ph].

[27] A. Dahlberg et al., "A link layer protocol for quantum networks," in *Proceedings of the ACM Special Interest Group on Data Communication*, ser. SIGCOMM '19, New York, NY, USA: Association for Computing Machinery, Aug. 19, 2019, pp. 159–173, DOI: 10.1145/3341302.3342070.

[28] J. Illiano, M. Caleffi, A. Manzalini, and A. S. Cacciapuoti, "Quantum Internet Protocol Stack: A Comprehensive Survey," *Computer Networks*, vol. 213, p. 109 092, Aug. 2022, DOI: 10.1016/j.comnet.2022.109092.

[29] A. Dahlberg et al., "NetQASM—a low-level instruction set architecture for hybrid quantum–classical programs in a quantum internet," *Quantum Science and Technology*, vol. 7, no. 3, p. 035 023, Jun. 2022, DOI: 10.1088/2058-9565/ac753f.

[30] S. Gauthier, G. Vardoyan, and S. Wehner, "An Architecture for Control of Entanglement Generation Switches in Quantum Networks," *IEEE Transactions on Quantum Engineering*, vol. 4, pp. 1–17, 2023, DOI: 10.1109/TQE.2023.3320047.

[31] S. Gauthier, G. Vardoyan, and S. Wehner, "A Control Architecture for Entanglement Generation Switches in Quantum Networks," in *Proceedings of the 1st Workshop on Quantum Networks and Distributed Quantum Computing*, ser. QuNet '23, New York, NY, USA: Association for Computing Machinery, Sep. 10, 2023, pp. 38–44, DOI: 10.1145/3610251.3610552.

[32] Y. Li, H. Zhang, C. Zhang, T. Huang, and F. R. Yu, "A Survey of Quantum Internet Protocols From a Layered Perspective," *IEEE Communications Surveys & Tutorials*, vol. 26, no. 3, pp. 1606–1634, 2024, DOI: 10.1109/COMST.2024.3361662.

[33] L. Bacciottini et al., "Leveraging Internet Principles to Build a Quantum Network," 2025, arXiv: 2410.08980 [quant-ph].

[34] T. R. Beauchamp, H. Jirovská, S. Gauthier, and S. Wehner, "A Modular Quantum Network Architecture for Integrating Network Scheduling with Local Program Execution," 2025, arXiv: 2503.12582 [quant-ph].

[35] V. Kumar, C. Cicconetti, M. Conti, and A. Passarella, "Quantum Internet: Technologies, Protocols, and Research Challenges," 2025, arXiv: 2502.01653 [quant-ph].

[36] R. Van Meter, J. Touch, and Horsman, "Recursive quantum repeater networks," *Progress in Informatics*, no. 8, p. 65, Mar. 2011, DOI: 10.2201/NiiPi.2011.8.8.

[37] J. D. Touch, Y.-S. Wang, and V. K. Pingali, "A Recursive Network Architecture," ISI, ISI-TR-2006-626, Oct. 20, 2006, [Online]. Available: https://www.strayalpha.com/pubs/isi-tr-626.pdf.

[38] J. Touch and V. Pingali, "The RNA Metaprotocol," in *2008 Proceedings of 17th International Conference on Computer Communications and Networks*, St. Thomas, US Virgin Islands: IEEE, Aug. 2008, pp. 1–6, DOI: 10.1109/icccn.2008.ecp.46.

[39] J. Touch et al., "A Dynamic Recursive Unified Internet Design (DRUID)," *Computer Networks*, vol. 55, no. 4, pp. 919–935, Mar. 2011, DOI: 10.1016/j.comnet.2010.12.016.

[40] R. Van Meter et al., "A Quantum Internet Architecture," in *2022 IEEE International Conference on Quantum Computing and Engineering (QCE)*, Broomfield, CO, USA: IEEE, Sep. 2022, pp. 341–352, DOI: 10.1109/QCE53715.2022.00055.

[41] R. Satoh et al., "QuISP: A Quantum Internet Simulation Package," in *2022 IEEE International Conference on Quantum Computing and Engineering (QCE)*, Sep. 2022, pp. 353–364, DOI: 10.1109/QCE53715.2022.00056.

[42] X. Wu et al., "SeQUeNCe: A customizable discrete-event simulator of quantum networks," *Quantum Science and Technology*, vol. 6, no. 4, p. 045 027, Sep. 2021, DOI: 10.1088/2058-9565/ac22f6.

[43] J. Chung et al., "Cross-Validating Quantum Network Simulators," in *IEEE INFOCOM 2025 - IEEE Conference on Computer Communications*, *Workshop on Quantum Networked Applications and Protocols*, to be published, Apr. 2, 2025, DOI: 10.48550/arXiv.2504.01290.

[44] J. Borregaard, H. Pichler, T. Schröder, M. D. Lukin, P. Lodahl, and A. S. Sørensen, "One-Way Quantum Repeater Based on Near-Deterministic Photon-Emitter Interfaces," *Physical Review X*, vol. 10, no. 2, p. 021 071, Jun. 30, 2020, DOI: 10.1103/PhysRevX.10.021071.

[45] N. Benchasattabuse, M. Hajdušek, and R. Van Meter, "Architecture and Protocols for All-Photonic Quantum Repeaters," in *2024 IEEE International Conference on Quantum Computing and Engineering (QCE)*, vol. 01, Sep. 2024, pp. 1879–1889, DOI: 10.1109/QCE60285.2024.00217.

[46] N. Benchasattabuse, M. Hajdušek, and R. Van Meter, "Engineering Challenges in All-Photonic Quantum Repeaters," *IEEE Network*, vol. 39, no. 1, pp. 132–139, Jan. 2025, DOI: 10.1109/MNET.2024.3411802.

[47] N. Benchasattabuse, M. Hajdušek, and R. Van Meter, "Bridging All-photonic and Memory-based Quantum Repeaters," submitted for publication, 2025.

[48] N. Benchasattabuse, M. Hajdušek, and R. Van Meter, "Integrating Entanglement Purification into All-Photonic Quantum Repeaters," in *2025 IEEE International Conference on Quantum Computing and Engineering (QCE)*, to be published, Apr. 25, 2025, DOI: 10.48550/arXiv.2504.18121.

[49] D. Buterakos, E. Barnes, and S. E. Economou, "Deterministic Generation of All-Photonic Quantum Repeaters from Solid-State Emitters," *Physical Review X*, vol. 7, no. 4, p. 041 023, Oct. 27, 2017, DOI: 10.1103/PhysRevX.7.041023.

[50] P. Hilaire, E. Barnes, and S. E. Economou, "Resource requirements for efficient quantum communication using all-photonic graph states generated from a few matter qubits," *Quantum*, vol. 5, p. 397, Feb. 15, 2021, DOI: 10.22331/q-2021-02-15-397.

[51] L. Hartmann, B. Kraus, H.-J. Briegel, and W. Dür, "Role of memory errors in quantum repeaters," *Physical Review A*, vol. 75, no. 3, p. 032 310, Mar. 8, 2007, DOI: 10.1103/PhysRevA.75.032310.

[52] M. Razavi, M. Piani, and N. Lütkenhaus, "Quantum repeaters with imperfect memories: Cost and scalability," *Physical Review A*, vol. 80, no. 3, p. 032 301, Sep. 1, 2009, DOI: 10.1103/PhysRevA.80.032301.

[53] M. Mobayenjarihani, G. Vardoyan, and D. Towsley, "Optimistic Entanglement Purification in Quantum Networks," in *2023 IEEE International Conference on Quantum Computing and Engineering (QCE)*, vol. 01, Sep. 2023, pp. 1143–1153, DOI: 10.1109/QCE57702.2023.00129.

[54] M. Caleffi, M. Amoretti, D. Ferrari, J. Illiano, A. Manzalini, and A. S. Cacciapuoti, "Distributed quantum computing: A survey," *Computer Networks*, vol. 254, p. 110 672, Dec. 1, 2024, DOI: 10.1016/j.comnet.2024.110672.

[55] D. Barral et al., "Review of Distributed Quantum Computing: From single QPU to High Performance Quantum Computing," *Computer Science Review*, vol. 57, p. 100 747, Aug. 2025, DOI: 10.1016/j.cosrev.2025.100747.

[56] A. Abbas et al., "Challenges and Opportunities in Quantum Optimization," *Nature Reviews Physics*, vol. 6, no. 12, pp. 718–735, Oct. 28, 2024, DOI: 10.1038/s42254-024-00770-9.

[57] N. Benchasattabuse, T. Satoh, M. Hajdušek, and R. Van Meter, "Amplitude Amplification for Optimization via Subdivided Phase Oracle," in *2022 IEEE International Conference on Quantum Computing and Engineering (QCE)*, Sep. 2022, pp. 22–30, DOI: 10.1109/QCE53715.2022.00020.

[58] N. Benchasattabuse, A. Bärtschi, L. P. García-Pintos, J. Golden, N. Lemons, and S. Eidenbenz, "Lower bounds on the number of rounds of the quantum approximate optimization algorithm required for guaranteed approximation ratios," *Physical Review A*, vol. 111, no. 6, p. 062 411, Jun. 9, 2025, DOI: 10.1103/PhysRevA.111.062411.

[59] L. P. García-Pintos, L. T. Brady, J. Bringewatt, and Y.-K. Liu, "Lower Bounds on Quantum Annealing Times," *Physical Review Letters*, vol. 130, no. 14, p. 140 601, Apr. 5, 2023, DOI: 10.1103/PhysRevLett.130.140601.

[60] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition*. Cambridge: Cambridge University Press, 2010, DOI: 10.1017/CBO9780511976667.

[61] D. Gottesman, "Stabilizer Codes and Quantum Error Correction," Ph.D. dissertation, California Institute of Technology, May 28, 1997, DOI: 10.48550/arXiv.quant-ph/9705052.

[62] S. Aaronson and D. Gottesman, "Improved simulation of stabilizer circuits," *Physical Review A*, vol. 70, no. 5, p. 052 328, Nov. 30, 2004, DOI: 10.1103/PhysRevA.70.052328.

[63] D. Gottesman, "The Heisenberg Representation of Quantum Computers," 1998, arXiv: quant-ph/9807006.

[64] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, "Teleporting an unknown quantum state via dual classical and Einstein-Podolsky-Rosen channels," *Physical Review Letters*, vol. 70, no. 13, pp. 1895–1899, Mar. 29, 1993, DOI: 10.1103/PhysRevLett.70.1895.

[65] E. Bernstein and U. Vazirani, "Quantum Complexity Theory," *SIAM Journal on Computing*, vol. 26, no. 5, pp. 1411–1473, Oct. 1997, DOI: 10.1137/S0097539796300921.

[66] D. Simon, "On the power of quantum computation," in *Proceedings 35th Annual Symposium on Foundations of Computer Science*, Nov. 1994, pp. 116–123, DOI: 10.1109/SFCS.1994.365701.

[67] D. Deutsch and R. Jozsa, "Rapid solution of problems by quantum computation," *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences*, vol. 439, no. 1907, pp. 553–558, Jan. 1997, DOI: 10.1098/rspa.1992.0167.

[68] M. Hein, J. Eisert, and H. J. Briegel, "Multiparty entanglement in graph states," *Physical Review A*, vol. 69, no. 6, p. 062 311, Jun. 9, 2004, DOI: 10.1103/PhysRevA.69.062311.

[69] M. Hein, W. Dür, J. Eisert, R. Raussendorf, M. V. den Nest, and H.-J. Briegel, "Entanglement in Graph States and its Applications," Feb. 11, 2006, DOI: 10.48550/arXiv.quant-ph/0602096.

[70] S. Anders and H. J. Briegel, "Fast simulation of stabilizer circuits using a graph state representation," *Physical Review A*, vol. 73, no. 2, p. 022 334, Feb. 21, 2006, DOI: 10.1103/PhysRevA.73.022334.

[71] A. Bouchet, "κ-Transformations, Local Complementations and Switching," in *Cycles and Rays*, G. Hahn, G. Sabidussi, and R. E. Woodrow, Eds., Dordrecht: Springer Netherlands, 1990, pp. 41–50, DOI: 10.1007/978-94-009-0517-7_5.

[72] R. Raussendorf and H. J. Briegel, "A One-Way Quantum Computer," *Physical Review Letters*, vol. 86, no. 22, pp. 5188–5191, May 28, 2001, DOI: 10.1103/PhysRevLett.86.5188.

[73] R. Raussendorf, D. E. Browne, and H. J. Briegel, "Measurement-based quantum computation on cluster states," *Physical Review A*, vol. 68, no. 2, p. 022 312, Aug. 25, 2003, DOI: 10.1103/PhysRevA.68.022312.

[74] S. Liu, N. Benchasattabuse, D. Q. Morgan, M. Hajdušek, S. J. Devitt, and R. Van Meter, "A Substrate Scheduler for Compiling Arbitrary Fault-Tolerant Graph States," in *2023 IEEE International Conference on Quantum Computing and Engineering (QCE)*, Bellevue, WA, USA: IEEE, Sep. 17, 2023, pp. 870–880, DOI: 10.1109/QCE57702.2023.00101.

[75] M. Krishnan Vijayan, A. Paler, J. Gavriel, C. R. Myers, P. P. Rohde, and S. J. Devitt, "Compilation of algorithm-specific graph states for quantum circuits," *Quantum Science and Technology*, vol. 9, no. 2, p. 025 005, Feb. 2024, DOI: 10.1088/2058-9565/ad1f39.

[76] A. Sen, K. Goodenough, and D. Towsley, "Multipartite Entanglement Distribution in Quantum Networks using Subgraph Complementations," 2024, arXiv: 2308.13700 [quant-ph].

[77] S. Ghanbari, J. Lin, B. MacLellan, L. Robichaud, P. Roztocki, and H.-K. Lo, "Optimization of deterministic photonic-graph-state generation via local operations," *Physical Review A*, vol. 110, no. 5, p. 052 605, Nov. 15, 2024, DOI: 10.1103/PhysRevA.110.052605.

[78] R. Van Meter, *Quantum Networking* (Networks and Telecommunications Series). Hoboken, NJ: Wiley, 2014, 1 p., DOI: 10.1002/9781118648919.

[79] M. Sasaki, "Quantum networks: Where should we be heading?" *Quantum Science and Technology*, vol. 2, no. 2, p. 020 501, Jun. 1, 2017, DOI: 10.1088/2058-9565/aa6994.

[80] C. Simon, "Towards a global quantum network," *Nature Photonics*, vol. 11, no. 11, pp. 678–680, Nov. 2017, DOI: 10.1038/s41566-017-0032-0.

[81] W. J. Munro, N. L. Piparo, J. Dias, M. Hanks, and K. Nemoto, "Designing tomorrow's quantum internet," *AVS Quantum Science*, vol. 4, no. 2, p. 020 503, Jun. 8, 2022, DOI: 10.1116/5.0092069.

[82] C. H. Bennett and G. Brassard, "Quantum cryptography: Public key distribution and coin tossing," *Theoretical Computer Science*, Theoretical Aspects of Quantum Cryptography – Celebrating 30 Years of BB84, vol. 560, pp. 7–11, Dec. 4, 2014, DOI: 10.1016/j.tcs.2014.05.025.

[83] F. Xu, X. Ma, Q. Zhang, H.-K. Lo, and J.-W. Pan, "Secure quantum key distribution with realistic devices," *Reviews of Modern Physics*, vol. 92, no. 2, p. 025 002, May 26, 2020, DOI: 10.1103/RevModPhys.92.025002.

[84] M. Peev et al., "The SECOQC quantum key distribution network in Vienna," *New Journal of Physics*, vol. 11, no. 7, p. 075 001, Jul. 2009, DOI: 10.1088/1367-2630/11/7/075001.

[85] M. Sasaki et al., "Field test of quantum key distribution in the Tokyo QKD Network," *Optics Express*, vol. 19, no. 11, pp. 10 387–10 409, May 23, 2011, DOI: 10.1364/OE.19.010387.

[86] Y.-A. Chen et al., "An integrated space-to-ground quantum communication network over 4,600 kilometres," *Nature*, vol. 589, no. 7841, pp. 214–219, Jan. 2021, DOI: 10.1038/s41586-020-03093-8.

[87] M. Pittaluga et al., "600-km repeater-like quantum communications with dual-band stabilization," *Nature Photonics*, vol. 15, no. 7, pp. 530–535, Jul. 2021, DOI: 10.1038/s41566-021-00811-0.

[88] M. Pittaluga et al., "Long-distance coherent quantum communications in deployed telecom networks," *Nature*, vol. 640, no. 8060, pp. 911–917, Apr. 2025, DOI: 10.1038/s41586-025-08801-w.

[89] D. Gottesman, T. Jennewein, and S. Croke, "Longer-Baseline Telescopes Using Quantum Repeaters," *Physical Review Letters*, vol. 109, no. 7, p. 070 503, Aug. 16, 2012, DOI: 10.1103/PhysRevLett.109.070503.

[90] E. T. Khabiboulline, J. Borregaard, K. De Greve, and M. D. Lukin, "Optical Interferometry with Quantum Networks," *Physical Review Letters*, vol. 123, no. 7, p. 070 504, Aug. 15, 2019, DOI: 10.1103/PhysRevLett.123.070504.

[91] R. Jozsa, D. S. Abrams, J. P. Dowling, and C. P. Williams, "Quantum Clock Synchronization Based on Shared Prior Entanglement," *Physical Review Letters*, vol. 85, no. 9, pp. 2010–2013, Aug. 28, 2000, DOI: 10.1103/PhysRevLett.85.2010.

[92] P. Kómár et al., "A quantum network of clocks," *Nature Physics*, vol. 10, no. 8, pp. 582–587, Aug. 2014, DOI: 10.1038/nphys3000.

[93] B. C. Nichol et al., "An elementary quantum network of entangled optical atomic clocks," *Nature*, vol. 609, no. 7928, pp. 689–694, Sep. 2022, DOI: 10.1038/s41586-022-05088-z.

[94] J.-P. Chen et al., "Quantum Key Distribution over 658 km Fiber with Distributed Vibration Sensing," *Physical Review Letters*, vol. 128, no. 18, p. 180 502, May 2, 2022, DOI: 10.1103/PhysRevLett.128.180502.

[95] K. Kanneworff et al., "Towards experimental demonstration of quantum position verification using true single photons," 2025, arXiv: 2502.04125 [quant-ph].

[96] V. Giovannetti, S. Lloyd, and L. Maccone, "Advances in quantum metrology," *Nature Photonics*, vol. 5, no. 4, pp. 222–229, Apr. 2011, DOI: 10.1038/nphoton.2011.35.

[97] T. J. Proctor, P. A. Knott, and J. A. Dunningham, "Multiparameter Estimation in Networked Quantum Sensors," *Physical Review Letters*, vol. 120, no. 8, p. 080 501, Feb. 21, 2018, DOI: 10.1103/PhysRevLett.120.080501.

[98] W. Ge, K. Jacobs, Z. Eldredge, A. V. Gorshkov, and M. Foss-Feig, "Distributed Quantum Metrology with Linear Networks and Separable Inputs," *Physical Review Letters*, vol. 121, no. 4, p. 043 604, Jul. 25, 2018, DOI: 10.1103/PhysRevLett.121.043604.

[99] M. Hillery, V. Bužek, and A. Berthiaume, "Quantum secret sharing," *Physical Review A*, vol. 59, no. 3, pp. 1829–1834, Mar. 1, 1999, DOI: 10.1103/PhysRevA.59.1829.

[100] R. Cleve, D. Gottesman, and H.-K. Lo, "How to Share a Quantum Secret," *Physical Review Letters*, vol. 83, no. 3, pp. 648–651, Jul. 19, 1999, DOI: 10.1103/PhysRevLett.83.648.

[101] A. Broadbent and R. Islam, "Quantum Encryption with Certified Deletion," in *Theory of Cryptography*, R. Pass and K. Pietrzak, Eds., Cham: Springer International Publishing, 2020, pp. 92–122, DOI: 10.1007/978-3-030-64381-2_4.

[102] T. Morimae and T. Yamakawa, "Quantum commitments and signatures without one-way functions," in vol. 13507, 2022, pp. 269–295, DOI: 10.1007/978-3-031-15802-5_10.

[103] S. Gunn, N. Ju, F. Ma, and M. Zhandry, "Commitments to Quantum States," in *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, Orlando FL USA: ACM, Jun. 2, 2023, pp. 1579–1588, DOI: 10.1145/3564246.3585198.

[104] S. Tani, H. Kobayashi, and K. Matsumoto, "Exact Quantum Algorithms for the Leader Election Problem," in *STACS 2005*, V. Diekert and B. Durand, Eds., Berlin, Heidelberg: Springer, 2005, pp. 581–592, DOI: 10.1007/978-3-540-31856-9_48.

[105] M. Ben-Or and A. Hassidim, "Fast quantum byzantine agreement," in *Proceedings of the Thirty-Seventh Annual ACM Symposium on Theory of Computing*, ser. STOC '05, New York, NY, USA: Association for Computing Machinery, May 22, 2005, pp. 481–485, DOI: 10.1145/1060590.1060662.

[106] K. Chen and H.-K. Lo, "Multi-partite quantum cryptographic protocols with noisy GHZ states," *Quantum Information and Computation*, vol. 7, no. 8, pp. 689–715, Nov. 2007, DOI: 10.26421/QIC7.8-1.

[107] R. Augusiak and P. Horodecki, "Multipartite secret key distillation and bound entanglement," *Physical Review A*, vol. 80, no. 4, p. 042 307, Oct. 7, 2009, DOI: 10.1103/PhysRevA.80.042307.

[108] S. Tani, H. Kobayashi, and K. Matsumoto, "Exact Quantum Algorithms for the Leader Election Problem," *ACM Trans. Comput. Theory*, vol. 4, no. 1, 1:1–1:24, Mar. 2012, DOI: 10.1145/2141938.2141939.

[109] G. Murta, F. Grasselli, H. Kampermann, and D. Bruß, "Quantum Conference Key Agreement: A Review," *Advanced Quantum Technologies*, vol. 3, no. 11, p. 2 000 025, 2020, DOI: 10.1002/qute.202000025.

[110] R. Cleve and H. Buhrman, "Substituting quantum entanglement for communication," *Physical Review A*, vol. 56, no. 2, pp. 1201–1204, Aug. 1, 1997, DOI: 10.1103/PhysRevA.56.1201.

[111] J. I. Cirac, A. K. Ekert, S. F. Huelga, and C. Macchiavello, "Distributed quantum computation over noisy channels," *Physical Review A*, vol. 59, no. 6, pp. 4249–4254, Jun. 1, 1999, DOI: 10.1103/PhysRevA.59.4249.

[112] D. Gottesman and I. L. Chuang, "Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations," *Nature*, vol. 402, no. 6760, pp. 390–393, Nov. 1999, DOI: 10.1038/46503.

[113] Y. Dulek, C. Schaffner, and F. Speelman, "Quantum homomorphic encryption for polynomial-sized circuits," *Theory of Computing*, vol. 14, no. 1, pp. 1–45, 2018, DOI: 10.4086/toc.2018.v014a007.

[114] U. Mahadev, "Classical Homomorphic Encryption for Quantum Circuits," *SIAM Journal on Computing*, vol. 52, no. 6, FOCS18–189, Dec. 31, 2023, DOI: 10.1137/18M1231055.

[115] D. Aharonov, M. Ben-Or, E. Eban, and U. Mahadev, "Interactive Proofs for Quantum Computations," 2017, arXiv: 1704.04487 [quant-ph].

[116] D. Dieks, "Communication by EPR devices," *Physics Letters A*, vol. 92, no. 6, pp. 271–272, Nov. 1982, DOI: 10.1016/0375-9601(82)90084-6.

[117] W. K. Wootters and W. H. Zurek, "A single quantum cannot be cloned," *Nature*, vol. 299, no. 5886, pp. 802–803, Oct. 1982, DOI: 10.1038/299802a0.

[118] A. Peres, "How the no-cloning theorem got its name," *Fortschritte der Physik*, vol. 51, no. 4–5, pp. 458–461, May 7, 2003, DOI: 10.1002/prop.200310062.

[119] C. Jones, D. Kim, M. T. Rakher, P. G. Kwiat, and T. D. Ladd, "Design and analysis of communication protocols for quantum repeater networks," *New Journal of Physics*, vol. 18, no. 8, p. 083 015, Aug. 2016, DOI: 10.1088/1367-2630/18/8/083015.

[120] S. J. Devitt, A. D. Greentree, A. M. Stephens, and R. Van Meter, "High-speed quantum networking by ship," *Scientific Reports*, vol. 6, no. 1, p. 36 163, 1 Nov. 2, 2016, DOI: 10.1038/srep36163.

[121] S. Khatri, "Policies for elementary links in a quantum network," *Quantum*, vol. 5, p. 537, Sep. 7, 2021, DOI: 10.22331/q-2021-09-07-537.

[122] S. Pirandola, R. Laurenza, C. Ottaviani, and L. Banchi, "Fundamental limits of repeaterless quantum communications," *Nature Communications*, vol. 8, no. 1, p. 15 043, Apr. 26, 2017, DOI: 10.1038/ncomms15043.

[123] M. Takeoka, S. Guha, and M. M. Wilde, "Fundamental rate-loss tradeoff for optical quantum key distribution," *Nature Communications*, vol. 5, no. 1, p. 5235, Oct. 24, 2014, DOI: 10.1038/ncomms6235.

[124] W. Dür, H.-J. Briegel, J. I. Cirac, and P. Zoller, "Quantum repeaters based on entanglement purification," *Physical Review A*, vol. 59, no. 1, pp. 169–181, Jan. 1, 1999, DOI: 10.1103/PhysRevA.59.169.

[125] S. Guha et al., "Rate-loss analysis of an efficient quantum repeater architecture," *Physical Review A*, vol. 92, no. 2, p. 022 357, Aug. 31, 2015, DOI: 10.1103/PhysRevA.92.022357.

[126] M. Żukowski, A. Zeilinger, M. A. Horne, and A. K. Ekert, ""Event-ready-detectors" Bell experiment via entanglement swapping," *Physical Review Letters*, vol. 71, no. 26, pp. 4287–4290, Dec. 27, 1993, DOI: 10.1103/PhysRevLett.71.4287.

[127] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, "Purification of Noisy Entanglement and Faithful Teleportation via Noisy Channels," *Physical Review Letters*, vol. 76, no. 5, pp. 722–725, Jan. 29, 1996, DOI: 10.1103/PhysRevLett.76.722.

[128] D. Deutsch, A. Ekert, R. Jozsa, C. Macchiavello, S. Popescu, and A. Sanpera, "Quantum Privacy Amplification and the Security of Quantum Cryptography over Noisy Channels," *Physical Review Letters*, vol. 77, no. 13, pp. 2818–2821, Sep. 23, 1996, DOI: 10.1103/PhysRevLett.77.2818.

[129] H.-J. Briegel, W. Dür, J. I. Cirac, and P. Zoller, "Quantum Repeaters: The Role of Imperfect Local Operations in Quantum Communication," *Physical Review Letters*, vol. 81, no. 26, pp. 5932–5935, Dec. 28, 1998, DOI: 10.1103/PhysRevLett.81.5932.

[130] W. Dür and H. J. Briegel, "Entanglement purification and quantum error correction," *Reports on Progress in Physics*, vol. 70, no. 8, pp. 1381–1424, Aug. 1, 2007, DOI: 10.1088/0034-4885/70/8/R03.

[131] S. J. Devitt, K. Nemoto, and W. J. Munro, "Quantum Error Correction for Beginners," *Reports on Progress in Physics*, vol. 76, no. 7, p. 076 001, Jul. 1, 2013, DOI: 10.1088/0034-4885/76/7/076001.

[132] J. Roffe, "Quantum Error Correction: An Introductory Guide," *Contemporary Physics*, vol. 60, no. 3, pp. 226–245, Jul. 3, 2019, DOI: 10.1080/00107514.2019.1667078.

[133] D. Gottesman, "Surviving as a Quantum Computer in a Classical World," 2025, Accessed: May 1, 2025. [Online]. Available: https://www.cs.umd.edu/class/spring2024/cmsc858G/QECCbook-2024-ch1-15.pdf.

[134] M. Pant et al., "Routing entanglement in the quantum internet," *npj Quantum Information*, vol. 5, no. 1, p. 25, Dec. 2019, DOI: 10.1038/s41534-019-0139-x.

[135] C. Cicconetti, M. Conti, and A. Passarella, "Request Scheduling in Quantum Networks," *IEEE Transactions on Quantum Engineering*, vol. 2, pp. 2–17, 2021, DOI: 10.1109/TQE.2021.3090532.

[136] P. Dhara, A. Patil, H. Krovi, and S. Guha, "Subexponential rate versus distance with time-multiplexed quantum repeaters," *Physical Review A*, vol. 104, no. 5, p. 052 612, Nov. 19, 2021, DOI: 10.1103/PhysRevA.104.052612.

[137] A. Patil, M. Pant, D. Englund, D. Towsley, and S. Guha, "Entanglement generation in a quantum network at distance-independent rate," *npj Quantum Information*, vol. 8, no. 1, pp. 1–9, 1 May 6, 2022, DOI: 10.1038/s41534-022-00536-0.

[138] X.-L. Feng, Z.-M. Zhang, X.-D. Li, S.-Q. Gong, and Z.-Z. Xu, "Entangling Distant Atoms by Interference of Polarized Photons," *Physical Review Letters*, vol. 90, no. 21, p. 217 902, May 28, 2003, DOI: 10.1103/PhysRevLett.90.217902.

[139] L.-M. Duan and H. J. Kimble, "Efficient Engineering of Multiatom Entanglement through Single-Photon Detections," *Physical Review Letters*, vol. 90, no. 25, p. 253 601, Jun. 23, 2003, DOI: 10.1103/PhysRevLett.90.253601.

[140] C. Simon and W. T. M. Irvine, "Robust Long-Distance Entanglement and a Loophole-Free Bell Test with Ions and Photons," *Physical Review Letters*, vol. 91, no. 11, p. 110 405, Sep. 12, 2003, DOI: 10.1103/PhysRevLett.91.110405.

[141] C. Jones, K. D. Greve, and Y. Yamamoto, "A high-speed optical link to entangle quantum dots," 2013, arXiv: 1310.4609 [quant-ph].

[142] Y. Huang, F. Salces–Carcoba, R. X. Adhikari, A. H. Safavi-Naeini, and L. Jiang, "Vacuum Beam Guide for Large Scale Quantum Networks," *Physical Review Letters*, vol. 133, no. 2, p. 020 801, Jul. 9, 2024, DOI: 10.1103/PhysRevLett.133.020801.

[143] W. J. Munro, K. A. Harrison, A. M. Stephens, S. J. Devitt, and K. Nemoto, "From quantum multiplexing to high-performance quantum networking," *Nature Photonics*, vol. 4, no. 11, pp. 792–796, Nov. 2010, DOI: 10.1038/nphoton.2010.213.

[144] A. Zang, A. Kolar, J. Chung, M. Suchara, T. Zhong, and R. Kettimuthu, "Simulation of Entanglement Generation between Absorptive Quantum Memories," in *2022 IEEE International Conference on Quantum Computing and Engineering (QCE)*, Sep. 2022, pp. 617–623, DOI: 10.1109/QCE53715.2022.00084.

[145] K. S. Soon et al., "An Implementation and Analysis of a Practical Quantum Link Architecture Utilizing Entangled Photon Sources," in *2024 International Conference on Quantum Communications, Networking, and Computing (QCNC)*, Jul. 2024, pp. 25–32, DOI: 10.1109/QCNC62729.2024.00014.

[146] B. P. Abbott et al., "LIGO: The Laser Interferometer Gravitational-Wave Observatory," *Reports on Progress in Physics*, vol. 72, no. 7, p. 076 901, Jun. 2009, DOI: 10.1088/0034-4885/72/7/076901.

[147] L. Jiang, J. M. Taylor, K. Nemoto, W. J. Munro, R. Van Meter, and M. D. Lukin, "Quantum repeater with encoding," *Physical Review A*, vol. 79, no. 3, p. 032 325, Mar. 20, 2009, DOI: 10.1103/PhysRevA.79.032325.

[148] N. Sangouard, R. Dubessy, and C. Simon, "Quantum repeaters based on single trapped ions," *Physical Review A*, vol. 79, no. 4, p. 042 340, Apr. 29, 2009, DOI: 10.1103/PhysRevA.79.042340.

[149] Y. Li, S. D. Barrett, T. M. Stace, and S. C. Benjamin, "Long range failure-tolerant entanglement distribution," *New Journal of Physics*, vol. 15, no. 2, p. 023 012, Feb. 2013, DOI: 10.1088/1367-2630/15/2/023012.

[150] M. Zwerger, H. J. Briegel, and W. Dür, "Hybrid architecture for encoded measurement-based quantum computation," *Scientific Reports*, vol. 4, no. 1, p. 5364, Jun. 20, 2014, DOI: 10.1038/srep05364.

[151] P. Mazurek et al., "Long-distance quantum communication over noisy networks without long-time quantum memory," *Physical Review A*, vol. 90, no. 6, p. 062 311, Dec. 4, 2014, DOI: 10.1103/PhysRevA.90.062311.

[152] F. Kimiaee Asadi, N. Lauk, S. Wein, N. Sinclair, C. O'Brien, and C. Simon, "Quantum repeaters with individual rare-earth ions at telecommunication wavelengths," *Quantum*, vol. 2, p. 93, Sep. 13, 2018, DOI: 10.22331/q-2018-09-13-93.

[153] F. Kimiaee Asadi, S. C. Wein, and C. Simon, "Protocols for long-distance quantum communication with single 167Er ions," *Quantum Science and Technology*, vol. 5, no. 4, p. 045 015, Sep. 2020, DOI: 10.1088/2058-9565/abae7c.

[154] A. G. Fowler, D. S. Wang, C. D. Hill, T. D. Ladd, R. Van Meter, and L. C. L. Hollenberg, "Surface Code Quantum Communication," *Physical Review Letters*, vol. 104, no. 18, p. 180 503, May 6, 2010, DOI: 10.1103/PhysRevLett.104.180503.

[155] W. J. Munro, A. M. Stephens, S. J. Devitt, K. A. Harrison, and K. Nemoto, "Quantum communication without the necessity of quantum memories," *Nature Photonics*, vol. 6, no. 11, pp. 777–781, Nov. 2012, DOI: 10.1038/nphoton.2012.243.

[156] S. Muralidharan, J. Kim, N. Lütkenhaus, M. D. Lukin, and L. Jiang, "Ultrafast and Fault-Tolerant Quantum Communication across Long Distances," *Physical Review Letters*, vol. 112, no. 25, p. 250 501, Jun. 27, 2014, DOI: 10.1103/PhysRevLett.112.250501.

[157] S. Muralidharan, C.-L. Zou, L. Li, J. Wen, and L. Jiang, "Overcoming erasure errors with multilevel systems," *New Journal of Physics*, vol. 19, no. 1, p. 013 026, Jan. 20, 2017, DOI: 10.1088/1367-2630/aa573a.

[158] F. Rozpędek, K. Noh, Q. Xu, S. Guha, and L. Jiang, "Quantum repeaters based on concatenated bosonic and discrete-variable quantum codes," *npj Quantum Information*, vol. 7, no. 1, pp. 1–12, Jun. 22, 2021, DOI: 10.1038/s41534-021-00438-7.

[159] S. DiAdamo, B. Qi, G. Miller, R. Kompella, and A. Shabani, "Packet switching in quantum networks: A path to the quantum Internet," *Physical Review Research*, vol. 4, no. 4, p. 043 064, Oct. 28, 2022, DOI: 10.1103/PhysRevResearch.4.043064.

[160] M. Razavi, *An Introduction to Quantum Communications Networks: Or, How Shall We Communicate in the Quantum Era?* Morgan & Claypool Publishers, May 1, 2018, DOI: 10.1088/978-1-6817-4653-1.

[161] W. J. Munro, K. Azuma, K. Tamaki, and K. Nemoto, "Inside Quantum Repeaters," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 21, no. 3, pp. 78–90, May 2015, DOI: 10.1109/JSTQE.2015.2392076.

[162] R. Acharya et al., "Suppressing quantum errors by scaling a surface code logical qubit," *Nature*, vol. 614, no. 7949, pp. 676–681, Feb. 2023, DOI: 10.1038/s41586-022-05434-1.

[163] D. Bluvstein et al., "Logical quantum processor based on reconfigurable atom arrays," *Nature*, Dec. 6, 2023, DOI: 10.1038/s41586-023-06927-3.

[164] A. Paetznick et al., "Demonstration of logical qubits and repeated error correction with better-than-physical error rates," 2024, arXiv: 2404.02280 [quant-ph].

[165] J. Ramette, J. Sinclair, N. P. Breuckmann, and V. Vuletić, "Fault-tolerant connection of error-corrected qubits with noisy links," *npj Quantum Information*, vol. 10, no. 1, pp. 1–6, Jun. 10, 2024, DOI: 10.1038/s41534-024-00855-4.

[166] J. Sinclair, J. Ramette, B. Grinkemeyer, D. Bluvstein, M. D. Lukin, and V. Vuletić, "Fault-tolerant optical interconnects for neutral-atom arrays," *Physical Review Research*, vol. 7, no. 1, p. 013 313, Mar. 25, 2025, DOI: 10.1103/PhysRevResearch.7.013313.

[167] P. Hilaire, E. Barnes, S. E. Economou, and F. Grosshans, "Error-correcting entanglement swapping using a practical logical photon encoding," *Physical Review A*, vol. 104, no. 5, p. 052 623, Nov. 29, 2021, DOI: 10.1103/PhysRevA.104.052623.

[168] N. Gisin, N. Linden, S. Massar, and S. Popescu, "Error filtration and entanglement purification for quantum communication," *Physical Review A*, vol. 72, no. 1, p. 012 338, Jul. 28, 2005, DOI: 10.1103/PhysRevA.72.012338.

[169] C. Pattison, G. Baranes, J. P. Bonilla Ataides, M. D. Lukin, and H. Zhou, "Constant-Rate Entanglement Distillation for Fast Quantum Interconnects," in *Proceedings of the 52nd Annual International Symposium on Computer Architecture*, Tokyo Japan: ACM, Jun. 21, 2025, pp. 257–270, DOI: 10.1145/3695053.3731069.

[170] Y. Shi, A. Patil, and S. Guha, "Stabilizer Entanglement Distillation and Efficient Fault-Tolerant Encoders," *PRX Quantum*, vol. 6, no. 1, p. 010 339, Mar. 4, 2025, DOI: 10.1103/PRXQuantum.6.010339.

[171] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, "Mixed-state entanglement and quantum error correction," *Physical Review A*, vol. 54, no. 5, pp. 3824–3851, Nov. 1, 1996, DOI: 10.1103/PhysRevA.54.3824.

[172] M. Zwerger, H. J. Briegel, and W. Dür, "Robustness of hashing protocols for entanglement purification," *Physical Review A*, vol. 90, no. 1, p. 012 314, Jul. 10, 2014, DOI: 10.1103/PhysRevA.90.012314.

[173] N. P. Breuckmann and J. N. Eberhardt, "Quantum Low-Density Parity-Check Codes," *PRX Quantum*, vol. 2, no. 4, p. 040 101, Oct. 11, 2021, DOI: 10.1103/PRXQuantum.2.040101.

[174] J. P. B. Ataides et al., "Constant-Overhead Fault-Tolerant Bell-Pair Distillation using High-Rate Codes," 2025, arXiv: 2502.09542 [quant-ph].

[175] L. Jiang, J. M. Taylor, N. Khaneja, and M. D. Lukin, "Optimal approach to quantum communication using dynamic programming," *Proceedings of the National Academy of Sciences*, vol. 104, no. 44, pp. 17 291–17 296, Oct. 30, 2007, DOI: 10.1073/pnas.0703284104.

[176] S. Krastanov, V. V. Albert, and L. Jiang, "Optimized Entanglement Purification," *Quantum*, vol. 3, p. 123, Feb. 18, 2019, DOI: 10.22331/q-2019-02-18-123.

[177] E. Shchukin and P. van Loock, "Optimal Entanglement Swapping in Quantum Repeaters," *Physical Review Letters*, vol. 128, no. 15, p. 150 502, Apr. 13, 2022, DOI: 10.1103/PhysRevLett.128.150502.

[178] S. Jansen, K. Goodenough, S. de Bone, D. Gijswijt, and D. Elkouss, "Enumerating all bilocal Clifford distillation protocols through symmetry reduction," *Quantum*, vol. 6, p. 715, May 19, 2022, DOI: 10.22331/q-2022-05-19-715.

[179] C. Gidney, "Tetrationally Compact Entanglement Purification," 2023, arXiv: 2311.10971 [quant-ph].

[180] A. Zang, X. Chen, E. Chitambar, M. Suchara, and T. Zhong, "No-Go Theorems for Universal Entanglement Purification," *Physical Review Letters*, vol. 134, no. 19, p. 190 803, May 13, 2025, DOI: 10.1103/PhysRevLett.134.190803.

[181] A. Zang, X.-A. Chen, E. Chitambar, M. Suchara, and T. Zhong, "Entanglement Purification in Quantum Networks: Guaranteed Improvement and Optimal Time," 2025, arXiv: 2505.02286 [quant-ph].

[182] L. Aparicio, R. Van Meter, and H. Esaki, "Protocol design for quantum repeater networks," in *Proceedings of the 7th Asian Internet Engineering Conference*, Bangkok Thailand: ACM, Nov. 9, 2011, pp. 73–80, DOI: 10.1145/2089016.2089029.

[183] B. Ru, W. K. G. Seah, and A. C. Valera, "Synchronization Control-Plane Protocol for Quantum Link Layer," 2024, arXiv: 2409.07049 [cs].

[184] S. Oslovich, B. van der Vecht, and S. Wehner, "Compilation strategies for quantum network programs using Qoala," 2025, arXiv: 2505.06162 [quant-ph].

[185] A. Pirker and W. Dür, "A quantum network stack and protocols for reliable entanglement-based networks," *New Journal of Physics*, vol. 21, no. 3, p. 033 003, Mar. 2019, DOI: 10.1088/1367-2630/ab05f7.

[186] K. Fang, J. Zhao, X. Li, Y. Li, and R. Duan, "Quantum NETwork: From theory to practice," *Science China Information Sciences*, vol. 66, no. 8, p. 180 509, Jul. 5, 2023, DOI: 10.1007/s11432-023-3773-4.

[187] C. Delle Donne et al., "An operating system for executing applications on quantum network nodes," *Nature*, vol. 639, no. 8054, pp. 321–328, Mar. 2025, DOI: 10.1038/s41586-025-08704-w.

[188] B. van der Vecht, A. T. Yücel, H. Jirovská, and S. Wehner, "Qoala: An Application Execution Environment for Quantum Internet Nodes," 2025, arXiv: 2502.17296 [quant-ph].

[189] T. Matsuo, C. Durand, and R. Van Meter, "Quantum link bootstrapping using a RuleSet-based communication protocol," *Physical Review A*, vol. 100, no. 5, p. 052 320, Nov. 15, 2019, DOI: 10.1103/PhysRevA.100.052320.

[190] T. Matsuo, "Simulation of a Dynamic, RuleSet-based Quantum Network," M.S. thesis, Keio University, Aug. 26, 2019, DOI: 10.48550/arXiv.1908.10758.

[191] S. Suzuki and R. Van Meter, "Classification of Quantum Repeater Attacks," in *Proceedings 2015 Workshop on Security of Emerging Networking Technologies*, San Diego, CA: Internet Society, 2015, DOI: 10.14722/sent.2015.23003.

[192] T. Satoh, S. Nagayama, S. Suzuki, T. Matsuo, M. Hajdušek, and R. Van Meter, "Attacking the Quantum Internet," *IEEE Transactions on Quantum Engineering*, vol. 2, pp. 1–17, 2021, DOI: 10.1109/TQE.2021.3094983.

[193] B. Blakely, J. Chung, A. Poczatek, R. Syed, and R. Kettimuthu, "Toward a Quantum Information System Cybersecurity Taxonomy and Testbed: Exploiting a Unique Opportunity for Early Impact," 2024, arXiv: 2404.12465 [cs].

[194] M. B. Plenio and S. Virmani, "An introduction to entanglement measures," *Quantum Info. Comput.*, vol. 7, no. 1, pp. 1–51, Jan. 1, 2007.

[195] L. Amico, R. Fazio, A. Osterloh, and V. Vedral, "Entanglement in many-body systems," *Reviews of Modern Physics*, vol. 80, no. 2, pp. 517–576, May 6, 2008, DOI: 10.1103/RevModPhys.80.517.

[196] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, "Quantum entanglement," *Reviews of Modern Physics*, vol. 81, no. 2, pp. 865–942, Jun. 17, 2009, DOI: 10.1103/RevModPhys.81.865.

[197] M. Hajdušek and V. Vedral, "Entanglement in pure and thermal cluster states," *New Journal of Physics*, vol. 12, no. 5, p. 053 015, May 2010, DOI: 10.1088/1367-2630/12/5/053015.

[198] M. Hajdušek and M. Murao, "Direct evaluation of pure graph state entanglement," *New Journal of Physics*, vol. 15, no. 1, p. 013 039, Jan. 2013, DOI: 10.1088/1367-2630/15/1/013039.

[199] M. Ma, Y. Li, and J. Shang, "Multipartite entanglement measures: A review," *Fundamental Research*, S2667325824002590, Jun. 2024, DOI: 10.1016/j.fmre.2024.03.031.

[200] W. Dür, G. Vidal, and J. I. Cirac, "Three qubits can be entangled in two inequivalent ways," *Physical Review A*, vol. 62, no. 6, p. 062 314, Nov. 14, 2000, DOI: 10.1103/PhysRevA.62.062314.

[201] E. O. Ilo-Okeke, L. Tessler, J. P. Dowling, and T. Byrnes, "Remote quantum clock synchronization without synchronized clocks," *npj Quantum Information*, vol. 4, no. 1, pp. 1–5, 1 Aug. 15, 2018, DOI: 10.1038/s41534-018-0090-2.

[202] T. Satoh and R. Van Meter, "Quantum Sockets for the NISQ Era Quantum Internet," 2018. [Online]. Available: https://www.ngc.is.ritsumei.ac.jp/~ger/static/AQIS18/OnlineBooklet/158.pdf.

[203] R. Van Meter, T. Satoh, S. Nagayama, T. Matsuo, and S. Suzuki, "Optimizing Timing of High-Success-Probability Quantum Repeaters," 2017, arXiv: 1701.04586 [quant-ph].

[204] N. McKeown et al., "OpenFlow: Enabling innovation in campus networks," *ACM SIGCOMM Computer Communication Review*, vol. 38, no. 2, pp. 69–74, Mar. 31, 2008, DOI: 10.1145/1355734.1355746.

[205] P. Bosshart et al., "P4: Programming protocol-independent packet processors," *ACM SIGCOMM Computer Communication Review*, vol. 44, no. 3, pp. 87–95, Jul. 28, 2014, DOI: 10.1145/2656877.2656890.

[206] R. Van Meter and T. Matsuo, "Connection Setup in a Quantum Network," Internet Engineering Task Force, Internet Draft draft-van-meter-qirg-quantum-connection-setup-01, Sep. 11, 2019, 12 pp., [Online]. Available: https://datatracker.ietf.org/doc/draft-van-meter-qirg-quantum-connection-setup-01.

[207] C. Meignant, D. Markham, and F. Grosshans, "Distributing Graph States Over Arbitrary Quantum Networks," *Physical Review A*, vol. 100, no. 5, p. 052 333, Nov. 27, 2019, DOI: 10.1103/PhysRevA.100.052333.

[208] Z. Huang, G. K. Brennen, and Y. Ouyang, "Imaging Stars with Quantum Error Correction," *Physical Review Letters*, vol. 129, no. 21, p. 210 502, Nov. 16, 2022, DOI: 10.1103/PhysRevLett.129.210502.

[209] A. P. Vepsäläinen et al., "Impact of ionizing radiation on superconducting qubit coherence," *Nature*, vol. 584, no. 7822, pp. 551–556, Aug. 2020, DOI: 10.1038/s41586-020-2619-8.

[210] J. M. Martinis, "Saving superconducting quantum processors from decay and correlated errors generated by gamma and cosmic rays," *npj Quantum Information*, vol. 7, no. 1, p. 90, Jun. 2021, DOI: 10.1038/s41534-021-00431-0.

[211] Q. Xu et al., "Distributed Quantum Error Correction for Chip-Level Catastrophic Errors," *Physical Review Letters*, vol. 129, no. 24, p. 240 502, Dec. 2022, DOI: 10.1103/PhysRevLett.129.240502.

[212] B. O. Sane, R. V. Meter, and M. Hajdusek, "Fight or Flight: Cosmic Ray-Induced Phonons and the Quantum Surface Code*," in *2023 IEEE International Conference on Quantum Computing and Engineering (QCE)*, vol. 01, Sep. 2023, pp. 1378–1388, DOI: 10.1109/QCE57702.2023.00156.

[213] X. Wu et al., *Mitigating cosmic ray-like correlated events with a modular quantum processor*, May 2025, DOI: 10.48550/arXiv.2505.15919.

[214] X. Li et al., "Cosmic-ray-induced correlated errors in superconducting qubit array," *Nature Communications*, vol. 16, no. 1, p. 4677, May 2025, DOI: 10.1038/s41467-025-59778-z.

[215] A. Ambainis, H. Buhrman, Y. Dodis, and H. Rohrig, "Multiparty quantum coin flipping," in *Proceedings. 19th IEEE Annual Conference on Computational Complexity, 2004.*, Jun. 2004, pp. 250–259, DOI: 10.1109/CCC.2004.1313848.

[216] D. Mayers, "Unconditional security in Quantum Cryptography," 2004, arXiv: quant-ph/9802025.

[217] M. Christandl and S. Wehner, "Quantum Anonymous Transmissions," in *Advances in Cryptology - ASIACRYPT 2005*, B. Roy, Ed., Berlin, Heidelberg: Springer, 2005, pp. 217–235, DOI: 10.1007/11593447_12.

[218] M. A. Taherkhani, K. Navi, and R. V. Meter, "Resource-aware system architecture model for implementation of quantum aided Byzantine agreement on quantum repeater networks," *Quantum Science and Technology*, vol. 3, no. 1, p. 014 011, Dec. 2017, DOI: 10.1088/2058-9565/aa9bb1.

[219] J. Kim et al., "A Fault-Tolerant Million Qubit-Scale Distributed Quantum Computer," in *Proceedings of the 29th ACM International Conference on Architectural Support for Programming Languages and Operating Systems, Volume 2*, La Jolla CA USA: ACM, Apr. 2024, pp. 1–19, DOI: 10.1145/3620665.3640388.

[220] H. Shapourian et al., "Quantum Data Center Infrastructures: A Scalable Architectural Design Perspective," 2025, arXiv: 2501.05598 [quant-ph].

[221] E. Sutcliffe, B. Jonnadula, C. L. Gall, A. E. Moylett, and C. M. Westoby, *Distributed quantum error correction based on hyperbolic Floquet codes*, Jan. 2025, DOI: 10.48550/arXiv.2501.14029.

[222] T. J. Yoder et al., *Tour de gross: A modular quantum computer based on bivariate bicycle codes*, Jun. 2025, DOI: 10.48550/arXiv.2506.03094.

[223] A. Fischer and D. Towsley, "Distributing Graph States Across Quantum Networks," in *2021 IEEE International Conference on Quantum Computing and Engineering (QCE)*, Oct. 2021, pp. 324–333, DOI: 10.1109/QCE52317.2021.00049.

[224] L. Bugalho, B. C. Coutinho, F. A. Monteiro, and Y. Omar, "Distributing Multipartite Entanglement over Noisy Quantum Networks," *Quantum*, vol. 7, p. 920, Feb. 9, 2023, DOI: 10.22331/q-2023-02-09-920.

[225] X. Fan, C. Zhan, H. Gupta, and C. R. Ramakrishnan, "Optimized Distribution of Entanglement Graph States in Quantum Networks," *IEEE Transactions on Quantum Engineering*, vol. 6, pp. 1–17, 2025, DOI: 10.1109/TQE.2025.3552006.

[226] J. Yin et al., "Satellite-based entanglement distribution over 1200 kilometers," *Science*, vol. 356, no. 6343, pp. 1140–1144, Jun. 2017, DOI: 10.1126/science.aan3211.

[227] S. Khatri, A. J. Brady, R. A. Desporte, M. P. Bart, and J. P. Dowling, "Spooky action at a global distance: Analysis of space-based entanglement distribution for the quantum internet," *npj Quantum Information*, vol. 7, no. 1, p. 4, Jan. 2021, DOI: 10.1038/s41534-020-00327-5.

[228] S. Haldar, I. Agullo, A. J. Brady, A. Lamas-Linares, W. C. Proctor, and J. E. Troupe, "Towards global time distribution via satellite-based sources of entangled photons," *Physical Review A*, vol. 107, no. 2, p. 022 615, Feb. 2023, DOI: 10.1103/PhysRevA.107.022615.

[229] P. Fittipaldi, K. Teramoto, N. Benchasattabuse, M. Hajdušek, R. Van Meter, and F. Grosshans, "Entanglement Swapping in Orbit: A Satellite Quantum Link Case Study," in *2024 IEEE International Conference on Quantum Computing and Engineering (QCE)*, vol. 01, Sep. 2024, pp. 1924–1930, DOI: 10.1109/QCE60285.2024.00222.

[230] M. Koyama et al., "Optimal Switching Networks for Paired-Egress Bell State Analyzer Pools," in *2024 IEEE International Conference on Quantum Computing and Engineering (QCE)*, vol. 01, Sep. 2024, pp. 1897–1907, DOI: 10.1109/QCE60285.2024.00219.

[231] B. Kar and P. Kumar, "Routing Protocols for Quantum Networks: Overview and Challenges," 2023, arXiv: 2305.00708 [quant-ph].

[232] A. Abane, M. Cubeddu, V. S. Mai, and A. Battou, "Entanglement Routing in Quantum Networks: A Comprehensive Survey," *IEEE Transactions on Quantum Engineering*, vol. 6, pp. 1–39, 2025, DOI: 10.1109/TQE.2025.3541123.

[233] C. Di Franco and D. Ballester, "Optimal path for a quantum teleportation protocol in entangled networks," *Physical Review A*, vol. 85, no. 1, p. 010 303, Jan. 9, 2012, DOI: 10.1103/PhysRevA.85.010303.

[234] R. Van Meter, T. Satoh, T. D. Ladd, W. J. Munro, and K. Nemoto, "Path Selection for Quantum Repeater Networks," *Networking Science*, vol. 3, no. 1–4, pp. 82–95, Dec. 2013, DOI: 10.1007/s13119-013-0026-2.

[235] D. F. V. James, P. G. Kwiat, W. J. Munro, and A. G. White, "Measurement of qubits," *Physical Review A*, vol. 64, no. 5, p. 052 312, Oct. 16, 2001, DOI: 10.1103/PhysRevA.64.052312.

[236] G. M. D'Ariano, M. G. A. Paris, and M. F. Sacchi, "Quantum Tomography," in ser. Advances in Imaging and Electron Physics, vol. 128, Elsevier, 2003, pp. 205–308, DOI: 10.1016/S1076-5670(03)80065-4.

[237] J. B. Altepeter, D. F. James, and P. G. Kwiat, "Qubit Quantum State Tomography," in *Quantum State Estimation*, M. Paris and J. Řeháček, Eds., Berlin, Heidelberg: Springer, 2004, pp. 113–145, DOI: 10.1007/978-3-540-44481-7_4.

[238] J. Eisert et al., "Quantum certification and benchmarking," *Nature Reviews Physics*, vol. 2, no. 7, pp. 382–390, Jul. 2020, DOI: 10.1038/s42254-020-0186-4.

[239] A. G. Maity, J. C. A. Casapao, N. Benchasattabuse, M. Hajdusek, R. Van Meter, and D. Elkouss, "Noise estimation in an entanglement distillation protocol," *ACM SIGMETRICS Performance Evaluation Review*, vol. 51, no. 2, pp. 66–68, Sep. 28, 2023, DOI: 10.1145/3626570.3626594.

[240] J. C. A. Casapao, A. G. Maity, N. Benchasattabuse, M. Hajdušek, R. Van Meter, and D. Elkouss, "Disti-Mator: An entanglement distillation-based state estimator," *Communications Physics*, 2025, to be published, DOI: 10.48550/arXiv.2406.13937.

[241] J. C. A. Casapao et al., "A Double Selection Entanglement Distillation-Based State Estimator," in *2025 IEEE International Conference on Quantum Computing and Engineering (QCE)*, to be published, May 30, 2025, DOI: 10.48550/arXiv.2505.24280.

[242] H.-Y. Huang, "Learning in the quantum universe," Ph.D. dissertation, California Institute of Technology, Aug. 17, 2023, DOI: 10.7907/fgpv-3112.

[243] M. Caleffi, "Optimal Routing for Quantum Networks," *IEEE Access*, vol. 5, pp. 22 299–22 312, 2017, DOI: 10.1109/ACCESS.2017.2763325.

[244] K. Chakraborty, D. Elkouss, B. Rijsman, and S. Wehner, "Entanglement Distribution in a Quantum Network: A Multicommodity Flow-Based Approach," *IEEE Transactions on Quantum Engineering*, vol. 1, pp. 1–21, 2020, DOI: 10.1109/TQE.2020.3028172.

[245] Y. Zhao, G. Zhao, and C. Qiao, "E2E Fidelity Aware Routing and Purification for Throughput Maximization in Quantum Networks," in *IEEE INFOCOM 2022 - IEEE Conference on Computer Communications*, May 2022, pp. 480–489, DOI: 10.1109/INFOCOM48880.2022.9796814.

[246] Y. Zeng, J. Zhang, J. Liu, Z. Liu, and Y. Yang, "Multi-Entanglement Routing Design over Quantum Networks," in *IEEE INFOCOM 2022 - IEEE Conference on Computer Communications*, May 2022, pp. 510–519, DOI: 10.1109/INFOCOM48880.2022.9796810.

[247] L. Gyongyosi and S. Imre, "Decentralized base-graph routing for the quantum internet," *Physical Review A*, vol. 98, no. 2, p. 022 310, Aug. 10, 2018, DOI: 10.1103/PhysRevA.98.022310.

[248] K. Chakraborty, F. Rozpedek, A. Dahlberg, and S. Wehner, "Distributed Routing in a Quantum Internet," 2019, arXiv: 1907.11630 [quant-ph].

[249] L. Gyongyosi and S. Imre, "Entanglement-Gradient Routing for Quantum Networks," *Scientific Reports*, vol. 7, no. 1, p. 14 255, Oct. 27, 2017, DOI: 10.1038/s41598-017-14394-w.

[250] L. Le and T. N. Nguyen, "DQRA: Deep Quantum Routing Agent for Entanglement Routing in Quantum Networks," *IEEE Transactions on Quantum Engineering*, vol. 3, pp. 1–12, 2022, DOI: 10.1109/TQE.2022.3148667.

[251] A. Acín, J. I. Cirac, and M. Lewenstein, "Entanglement percolation in quantum networks," *Nature Physics*, vol. 3, no. 4, pp. 256–259, 4 Apr. 2007, DOI: 10.1038/nphys549.

[252] S. Pirandola, "End-to-end capacities of a quantum communication network," *Communications Physics*, vol. 2, no. 1, pp. 1–10, 1 May 17, 2019, DOI: 10.1038/s42005-019-0147-3.

[253] E. Sutcliffe and A. Beghelli, "Multiuser Entanglement Distribution in Quantum Networks Using Multipath Routing," *IEEE Transactions on Quantum Engineering*, vol. 4, pp. 1–15, 2023, DOI: 10.1109/TQE.2023.3329714.

[254] J. Halder, E. Matus, and G. Fettweis, "On the Concurrent Multipath Entanglement Distribution in Quantum Networks," in *GLOBECOM 2024 - 2024 IEEE Global Communications Conference*, Dec. 2024, pp. 2791–2796, DOI: 10.1109/GLOBECOM52923.2024.10901810.

[255] N. Lo Piparo, W. J. Munro, and K. Nemoto, "Quantum aggregation with temporal delay," *Physical Review A*, vol. 110, no. 3, p. 032 613, Sep. 13, 2024, DOI: 10.1103/PhysRevA.110.032613.

[256] H. Leone, N. R. Miller, D. Singh, N. K. Langford, and P. P. Rohde, "Cost vector analysis & multi-path entanglement routing in quantum networks," 2024, arXiv: 2105.00418 [quant-ph].

[257] K. Azuma, "Networking quantum networks with minimum cost aggregation," *npj Quantum Information*, vol. 11, no. 1, pp. 1–7, Mar. 21, 2025, DOI: 10.1038/s41534-025-01000-5.

[258] D. Lago-Rivera, S. Grandi, J. V. Rakonjac, A. Seri, and H. de Riedmatten, "Telecom-heralded entanglement between multimode solid-state quantum memories," *Nature*, vol. 594, no. 7861, pp. 37–40, 7861 Jun. 2021, DOI: 10.1038/s41586-021-03481-8.

[259] M. J. Clark et al., "Entanglement distribution quantum networking within deployed telecommunications fibre-optic infrastructure," in *Quantum Technology: Driving Commercialisation of an Enabling Science III*, Jan. 11, 2023, p. 22, DOI: 10.1117/12.2645095.

[260] T. Chakraborty et al., "Towards a spectrally multiplexed quantum repeater," *npj Quantum Information*, vol. 11, no. 1, pp. 1–6, Jan. 10, 2025, DOI: 10.1038/s41534-024-00946-2.

[261] Y.-R. Fan et al., "Quantum entanglement network enabled by a state-multiplexing quantum light source," *Light: Science & Applications*, vol. 14, no. 1, p. 189, May 12, 2025, DOI: 10.1038/s41377-025-01805-1.

[262] A. Khodadad Kashi and M. Kues, "Frequency-bin-encoded entanglement-based quantum key distribution in a reconfigurable frequency-multiplexed network," *Light: Science & Applications*, vol. 14, no. 1, p. 49, Jan. 16, 2025, DOI: 10.1038/s41377-024-01696-8.

[263] O. A. Collins, S. D. Jenkins, A. Kuzmich, and T. A. B. Kennedy, "Multiplexed Memory-Insensitive Quantum Repeaters," *Physical Review Letters*, vol. 98, no. 6, p. 060 502, Feb. 7, 2007, DOI: 10.1103/PhysRevLett.98.060502.

[264] L. Aparicio, "Design and Evaluation of Communication Protocols for Quantum Repeater Networks," M.S. thesis, Tokyo University, Tokyo, Aug. 17, 2011, 73 pp., [Online]. Available: https://aqua.sfc.wide.ad.jp/publications/Aparicio-master-thesis.pdf.

[265] L. Aparicio and R. Van Meter, "Multiplexing schemes for quantum repeater networks," presented at the SPIE Optical Engineering + Applications, R. E. Meyers, Y. Shih, and K. S. Deacon, Eds., San Diego, California, USA, Sep. 8, 2011, p. 816 308, DOI: 10.1117/12.893272.

[266] A. A. Barakabitze, A. Ahmad, R. Mijumbi, and A. Hines, "5G network slicing using SDN and NFV: A survey of taxonomy, architectures and future challenges," *Computer Networks*, vol. 167, p. 106 984, Feb. 11, 2020, DOI: 10.1016/j.comnet.2019.106984.

[267] P. Pathumsoot, T. Matsuo, T. Satoh, M. Hajdušek, S. Suwanna, and R. Van Meter, "Modeling of measurement-based quantum network coding on a superconducting quantum processor," *Physical Review A*, vol. 101, no. 5, p. 052 301, May 1, 2020, DOI: 10.1103/PhysRevA.101.052301.

[268] Z. Xiao et al., "A Connectionless Entanglement Distribution Protocol Design in Quantum Networks," *IEEE Network*, vol. 38, no. 1, pp. 131–139, Jan. 2024, DOI: 10.1109/MNET.2023.3321044.

[269] S.-M. Huang, T.-M. Hsu, J.-J. Du, J.-J. Kuo, and C.-Y. Wang, "Near-Optimal Swapping and Purifying Strategy for All-Optical-Switching Entanglement Routing," in *GLOBECOM 2024 - 2024 IEEE Global Communications Conference*, Dec. 2024, pp. 2803–2808, DOI: 10.1109/GLOBECOM52923.2024.10901022.

[270] Z. Li et al., "NarrowGap: Reducing Bottlenecks for End-to-End Entanglement Distribution in Quantum Networks," *IEEE Transactions on Networking*, vol. 33, no. 1, pp. 162–177, Feb. 2025, DOI: 10.1109/TNET.2024.3476342.

[271] V. Fajardo, J. Arkko, J. A. Loughney, and G. Zorn, "Diameter Base Protocol," Internet Engineering Task Force, Request for Comments RFC 6733, Oct. 2012, 152 pp., DOI: 10.17487/RFC6733.

[272] H. Zhou, K. Lv, L. Huang, and X. Ma, "Quantum Network: Security Assessment and Key Management," *IEEE/ACM Transactions on Networking*, vol. 30, no. 3, pp. 1328–1339, Jun. 2022, DOI: 10.1109/TNET.2021.3136943.

[273] C. Jiang, X.-L. Hu, Z.-W. Yu, and X.-B. Wang, "Side-channel security of practical quantum key distribution," *Physical Review Research*, vol. 6, no. 1, p. 013 266, Mar. 11, 2024, DOI: 10.1103/PhysRevResearch.6.013266.

[274] C.-H. F. Fung, B. Qi, K. Tamaki, and H.-K. Lo, "Phase-remapping attack in practical quantum-key-distribution systems," *Physical Review A*, vol. 75, no. 3, p. 032 314, Mar. 12, 2007, DOI: 10.1103/PhysRevA.75.032314.

[275] N. Gisin, S. Fasel, B. Kraus, H. Zbinden, and G. Ribordy, "Trojan Horse attacks on Quantum Key Distribution systems," *Physical Review A*, vol. 73, no. 2, p. 022 320, Feb. 13, 2006, DOI: 10.1103/PhysRevA.73.022320.

[276] I. S. Sushchev, D. S. Bulavkin, K. E. Bugai, A. S. Sidelnikova, and D. A. Dvoretskiy, "Trojan-horse attack on a real-world quantum key distribution system: Theoretical and experimental security analysis," *Physical Review Applied*, vol. 22, no. 3, p. 034 032, Sep. 12, 2024, DOI: 10.1103/PhysRevApplied.22.034032.

[277] N. Lütkenhaus and M. Jahma, "Quantum key distribution with realistic states: Photon-number statistics in the photon-number splitting attack," *New Journal of Physics*, vol. 4, pp. 44–44, Jul. 12, 2002, DOI: 10.1088/1367-2630/4/1/344.

[278] N. Lütkenhaus and A. J. Shields, "Focus on Quantum Cryptography: Theory and Practice," *New Journal of Physics*, vol. 11, no. 4, p. 045 005, Apr. 2009, DOI: 10.1088/1367-2630/11/4/045005.

[279] W. Kozlowski, A. Dahlberg, and S. Wehner, "Designing a Quantum Network Protocol," *Proceedings of the 16th International Conference on emerging Networking EXperiments and Technologies*, pp. 1–16, Nov. 23, 2020, DOI: 10.1145/3386367.3431293.

[280] S. Nagayama, B.-S. Choi, S. Devitt, S. Suzuki, and R. Van Meter, "Interoperability in encoded quantum repeater networks," *Physical Review A*, vol. 93, no. 4, p. 042 338, Apr. 26, 2016, DOI: 10.1103/PhysRevA.93.042338.

[281] R. Van Meter, T. D. Ladd, W. J. Munro, and K. Nemoto, "System Design for a Long-Line Quantum Repeater," *IEEE/ACM Transactions on Networking*, vol. 17, no. 3, pp. 1002–1013, Jun. 2009, DOI: 10.1109/TNET.2008.927260.

[282] P. Pathumsoot et al., "Hybrid Error-Management Strategies in Quantum Repeater Networks," 2023, arXiv: 2303.10295 [quant-ph].

[283] P. Pathumsoot et al., "Boosting end-to-end entanglement fidelity in quantum repeater networks via hybridized strategies," in *2024 International Conference on Quantum Communications, Networking, and Computing (QCNC)*, Jul. 1, 2024, pp. 49–56, DOI: 10.1109/QCNC62729.2024.00054.

[284] A. Zang, J. Chung, R. Kettimuthu, M. Suchara, and T. Zhong, "Analytical Performance Estimations for Quantum Repeater Network Scenarios," in *2024 IEEE International Conference on Quantum Computing and Engineering (QCE)*, Sep. 15, 2024, pp. 1960–1966, DOI: 10.1109/QCE60285.2024.00226.

[285] L. Chen and Z. Jia, "On Optimum Entanglement Purification Scheduling in Quantum Networks," *IEEE Journal on Selected Areas in Communications*, vol. 42, no. 7, pp. 1779–1792, Jul. 2024, DOI: 10.1109/JSAC.2024.3380080.

[286] S. Haldar et al., "Reducing classical communication costs in multiplexed quantum repeaters using hardware-aware quasi-local policies," *Communications Physics*, vol. 8, no. 1, pp. 1–20, Apr. 1, 2025, DOI: 10.1038/s42005-025-02029-w.

[287] D. Cai, A. Wielosz, and S. Wei, "Evolve carrier ethernet architecture with SDN and segment routing," in *Proceeding of IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks 2014*, Jun. 2014, pp. 1–6, DOI: 10.1109/WoWMoM.2014.6918980.

[288] R. Carpa, O. Glück, and L. Lefevre, "Segment routing based traffic engineering for energy efficient backbone networks," in *2014 IEEE International Conference on Advanced Networks and Telecommuncations Systems (ANTS)*, Dec. 2014, pp. 1–6, DOI: 10.1109/ANTS.2014.7057272.

[289] C. Filsfils, S. Previdi, L. Ginsberg, B. Decraene, S. Litkowski, and R. Shakir, "Segment Routing Architecture," Internet Engineering Task Force, Request for Comments RFC 8402, Jul. 2018, 32 pp., DOI: 10.17487/RFC8402.

[290] X. Xiao and L. Ni, "Internet QoS: A big picture," *IEEE Network*, vol. 13, no. 2, pp. 8–18, Mar. 1999, DOI: 10.1109/65.768484.

[291] D. D. Clark, J. Wroclawski, K. R. Sollins, and R. Braden, "Tussle in cyberspace: Defining tomorrow's internet," *SIGCOMM Comput. Commun. Rev.*, vol. 32, no. 4, pp. 347–356, Aug. 2002, DOI: 10.1145/964725.633059.

[292] *Sfc-aqua/RuleSetSpec*, sfc-aqua, May 31, 2024, [Online]. Available: https://github.com/sfc-aqua/RuleSetSpec.

[293] T. Wagner, H. Kampermann, D. Bruß, and M. Kliesch, "Optimal noise estimation from syndrome statistics of quantum codes," *Physical Review Research*, vol. 3, no. 1, p. 013 292, Mar. 31, 2021, DOI: 10.1103/PhysRevResearch.3.013292.

[294] T. Wagner, H. Kampermann, D. Bruß, and M. Kliesch, "Pauli channels can be estimated from syndrome measurements in quantum error correction," *Quantum*, vol. 6, p. 809, Sep. 19, 2022, DOI: 10.22331/q-2022-09-19-809.

[295] X. Wu et al., "Simulations of Photonic Quantum Networks for Performance Analysis and Experiment Design," in *2019 IEEE/ACM Workshop on Photonics-Optics Technology Oriented Networking, Information and Computing Systems (PHOTONICS)*, Nov. 2019, pp. 28–35, DOI: 10.1109/PHOTONICS49561.2019.00010.

[296] X. Wu, A. Kolar, J. Chung, D. Jin, M. Suchara, and R. Kettimuthu, "Parallel Simulation of Quantum Networks with Distributed Quantum State Management," *ACM Transactions on Modeling and Computer Simulation*, vol. 34, no. 2, pp. 1–28, Apr. 30, 2024, DOI: 10.1145/3634701.

[297] A. Varga and R. Hornig, "AN OVERVIEW OF THE OMNeT++ SIMULATION ENVIRONMENT," in *Proceedings of the First International ICST Conference on Simulation Tools and Techniques for Communications Networks and Systems*, Marseille, France: ICST, 2008, DOI: 10.4108/ICST.SIMUTOOLS2008.3027.

[298] A. Varga, "OMNeT++," in *Modeling and Tools for Network Simulation*, K. Wehrle, M. Güneş, and J. Gross, Eds., Berlin, Heidelberg: Springer, 2010, pp. 35–59, DOI: 10.1007/978-3-642-12331-3_3.

[299] K. S. Soon, N. Benchasattabuse, M. Hajdušek, K. Teramoto, S. Nagayama, and R. Van Meter, "Performance of Quantum Networks Using Heterogeneous Link Architectures," in *2024 IEEE International Conference on Quantum Computing and Engineering (QCE)*, vol. 01, Sep. 2024, pp. 1914–1923, DOI: 10.1109/QCE60285.2024.00221.

[300] K. Teramoto, M. Hajdušek, T. Sasaki, R. Van Meter, and S. Nagayama, "RuleSet-based Recursive Quantum Internetworking," in *Proceedings of the 1st Workshop on Quantum Networks and Distributed Quantum Computing*, New York NY USA: ACM, Sep. 10, 2023, pp. 25–30, DOI: 10.1145/3610251.3610556.

[301] B. Bartlett, "A distributed simulation framework for quantum networks and channels," 2018, arXiv: 1808.07047 [quant-ph].

[302] T. Coopmans et al., "NetSquid, a NETwork Simulator for QUantum Information using Discrete events," *Communications Physics*, vol. 4, no. 1, p. 164, Jul. 16, 2021, DOI: 10.1038/s42005-021-00647-8.

[303] *QuTech-Delft/squidasm*, QuTech, May 13, 2025, [Online]. Available: https://github.com/QuTech-Delft/squidasm.

[304] *QuTech-Delft/qne-adk*, QuTech, May 11, 2025, [Online]. Available: https://github.com/QuTech-Delft/qne-adk.

[305] J. Wallnöfer, F. Hahn, F. Wiesner, N. Walk, and J. Eisert, "Faithfully Simulating Near-Term Quantum Repeaters," *PRX Quantum*, vol. 5, no. 1, p. 010 351, Mar. 27, 2024, DOI: 10.1103/PRXQuantum.5.010351.

[306] A. Dahlberg and S. Wehner, "SimulaQron—a simulator for developing quantum internet software," *Quantum Science and Technology*, vol. 4, no. 1, p. 015 001, Sep. 2018, DOI: 10.1088/2058-9565/aad56e.

[307] S. Diadamo, J. Nötzel, B. Zanger, and M. M. Beşe, "QuNetSim: A Software Framework for Quantum Networks," *IEEE Transactions on Quantum Engineering*, vol. 2, pp. 1–12, 2021, DOI: 10.1109/TQE.2021.3092395.

[308] D. Gottesman, "Opportunities and Challenges in Fault-Tolerant Quantum Computation," 2022, arXiv: 2210.15844 [quant-ph].

[309] H. Jnane, B. Undseth, Z. Cai, S. C. Benjamin, and B. Koczor, "Multicore Quantum Computing," *Physical Review Applied*, vol. 18, no. 4, p. 044 064, Oct. 26, 2022, DOI: 10.1103/PhysRevApplied.18.044064.

[310] H. Leone, S. Srikara, P. P. Rohde, and S. Devitt, "Upper bounds for the clock speeds of fault-tolerant distributed quantum computation using satellites to supply entangled photon pairs," *Physical Review Research*, vol. 5, no. 4, p. 043 302, Dec. 29, 2023, DOI: 10.1103/PhysRevResearch.5.043302.

[311] H. Leone, T. Le, S. Srikara, and S. Devitt, "Resource overheads and attainable rates for trapped-ion lattice surgery," *Physical Review Research*, vol. 7, no. 2, p. 023 088, Apr. 24, 2025, DOI: 10.1103/PhysRevResearch.7.023088.

[312] S. Pouryousef, R. Nejabati, D. Towsley, R. Kompella, and E. Kaur, "Network-Aware Scheduling for Remote Gate Execution in Quantum Data Centers," 2025, arXiv: 2504.20176 [quant-ph].

[313] C. Zhan, J. Chung, A. Zang, A. Kolar, and R. Kettimuthu, "Design and Simulation of the Adaptive Continuous Entanglement Generation Protocol," 2025, arXiv: 2502.01964 [cs].

[314] J. Li, Q. Jia, K. Xue, D. S. L. Wei, and N. Yu, "A Connection-Oriented Entanglement Distribution Design in Quantum Networks," *IEEE Transactions on Quantum Engineering*, vol. 3, pp. 1–13, 2022, DOI: 10.1109/TQE.2022.3176375.

[315] Z. Li, K. Xue, J. Li, N. Yu, D. S. L. Wei, and R. Li, "Connection-Oriented and Connectionless Remote Entanglement Distribution Strategies in Quantum Networks," *IEEE Network*, vol. 36, no. 6, pp. 150–156, Nov. 2022, DOI: 10.1109/MNET.107.2100483.

[316] M. G. de Andrade et al., "On the Analysis of Quantum Repeater Chains with Sequential Swaps," 2024, arXiv: 2405.18252 [quant-ph].

[317] H. Zhang, Y. Li, C. Zhang, and T. Huang, "Connection-oriented and Connectionless Quantum Internet Considering Quantum Repeaters," 2022, arXiv: 2208.03930 [quant-ph].

[318] M. Varnava, D. E. Browne, and T. Rudolph, "Loss Tolerance in One-Way Quantum Computation via Counterfactual Error Correction," *Physical Review Letters*, vol. 97, no. 12, p. 120 501, Sep. 20, 2006, DOI: 10.1103/PhysRevLett.97.120501.

[319] A. Russo, E. Barnes, and S. E. Economou, "Photonic graph state generation from quantum dots and color centers for quantum communications," *Physical Review B*, vol. 98, no. 8, p. 085 303, Aug. 14, 2018, DOI: 10.1103/PhysRevB.98.085303.

[320] I. Tzitrin, "Local equivalence of complete bipartite and repeater graph states," *Physical Review A*, vol. 98, no. 3, p. 032 305, Sep. 6, 2018, DOI: 10.1103/PhysRevA.98.032305.

[321] B. Li, S. E. Economou, and E. Barnes, "Photonic resource state generation from a minimal number of quantum emitters," *npj Quantum Information*, vol. 8, no. 1, pp. 1–7, 1 Feb. 3, 2022, DOI: 10.1038/s41534-022-00522-6.

[322] Y. Zhan, P. Hilaire, E. Barnes, S. E. Economou, and S. Sun, "Performance analysis of quantum repeaters enabled by deterministically generated photonic graph states," *Quantum*, vol. 7, p. 924, Feb. 16, 2023, DOI: 10.22331/q-2023-02-16-924.

[323] Y. Mori et al., "Scalable Timing Coordination of Bell State Analyzers in Quantum Networks," in *2024 IEEE International Conference on Quantum Computing and Engineering (QCE)*, Montreal, QC, Canada: IEEE, Sep. 15, 2024, pp. 1890–1896, DOI: 10.1109/QCE60285.2024.00218.

[324] M. Pant, H. Krovi, D. Englund, and S. Guha, "Rate-distance tradeoff and resource costs for all-optical quantum repeaters," *Physical Review A*, vol. 95, no. 1, p. 012 304, Jan. 4, 2017, DOI: 10.1103/PhysRevA.95.012304.

[325] Y. Zhan and S. Sun, "Deterministic Generation of Loss-Tolerant Photonic Cluster States with a Single Quantum Emitter," *Physical Review Letters*, vol. 125, no. 22, p. 223 601, Nov. 24, 2020, DOI: 10.1103/PhysRevLett.125.223601.

[326] H. Shapourian and A. Shabani, "Modular architectures to deterministically generate graph states," *Quantum*, vol. 7, p. 935, Mar. 2, 2023, DOI: 10.22331/q-2023-03-02-935.

[327] E. Kaur, A. Patil, and S. Guha, "Resource-efficient and loss-aware photonic graph state preparation using an array of quantum emitters, and application to all-photonic quantum repeaters," 2024, arXiv: 2402.00731 [quant-ph].

[328] S. E. Economou, N. Lindner, and T. Rudolph, "Optically Generated 2-Dimensional Photonic Cluster State from Coupled Quantum Dots," *Physical Review Letters*, vol. 105, no. 9, p. 093 601, Aug. 24, 2010, DOI: 10.1103/PhysRevLett.105.093601.

[329] A. Russo, E. Barnes, and S. E. Economou, "Generation of arbitrary all-photonic graph states from quantum emitters," *New Journal of Physics*, vol. 21, no. 5, p. 055 002, May 1, 2019, DOI: 10.1088/1367-2630/ab193d.

[330] C. Gidney, "Stim: A fast stabilizer circuit simulator," *Quantum*, vol. 5, p. 497, Jul. 6, 2021, DOI: 10.22331/q-2021-07-06-497.

[331] N. Benchasattabuse, *Naphann/repeater-graph-state-protocol-based-on-half-RGS*, Mar. 2025, [Online]. Available: https://github.com/Naphann/repeater-graph-state-protocol-based-on-half-RGS.

[332] N. H. Lindner and T. Rudolph, "Proposal for Pulsed On-Demand Sources of Photonic Cluster State Strings," *Physical Review Letters*, vol. 103, no. 11, p. 113 602, Sep. 8, 2009, DOI: 10.1103/PhysRevLett.103.113602.

[333] V. L. Addala, S. Ge, and S. Krastanov, "Faster-than-Clifford simulations of entanglement purification circuits and their full-stack optimization," *npj Quantum Information*, vol. 11, no. 1, pp. 1–10, Jan. 23, 2025, DOI: 10.1038/s41534-024-00948-0.

[334] M. Zwerger, W. Dür, and H. J. Briegel, "Measurement-based quantum repeaters," *Physical Review A*, vol. 85, no. 6, p. 062 326, Jun. 27, 2012, DOI: 10.1103/PhysRevA.85.062326.

[335] Z. Wang et al., "An Efficient Scheduling Scheme of Swapping and Purification Operations for End-to-End Entanglement Distribution in Quantum Networks," *IEEE Transactions on Network Science and Engineering*, vol. 11, no. 1, pp. 380–391, Jan. 2024, DOI: 10.1109/TNSE.2023.3299177.

[336] W. Dür, H. Aschauer, and H.-J. Briegel, "Multiparticle entanglement purification for graph states," *Physical Review Letters*, vol. 91, no. 10, p. 107 903, Sep. 5, 2003, DOI: 10.1103/PhysRevLett.91.107903.

[337] C. Kruszynska, A. Miyake, H. J. Briegel, and W. Dür, "Entanglement purification protocols for all graph states," *Physical Review A*, vol. 74, no. 5, p. 052 316, Nov. 10, 2006, DOI: 10.1103/PhysRevA.74.052316.

[338] Y. Hasegawa et al., "Experimental time-reversed adaptive Bell measurement towards all-photonic quantum repeaters," *Nature Communications*, vol. 10, no. 1, p. 378, 1 Jan. 28, 2019, DOI: 10.1038/s41467-018-08099-5.

[339] Z.-D. Li et al., "Experimental quantum repeater without quantum memory," *Nature Photonics*, vol. 13, no. 9, pp. 644–648, 9 Sep. 2019, DOI: 10.1038/s41566-019-0468-5.

[340] K. Tiurev et al., "High-fidelity multiphoton-entangled cluster state with solid-state quantum emitters in photonic nanostructures," *Physical Review A*, vol. 105, no. 3, p. L030601, Mar. 2, 2022, DOI: 10.1103/PhysRevA.105.L030601.

[341] D. Istrati et al., "Sequential generation of linear cluster states from a single photon emitter," *Nature Communications*, vol. 11, no. 1, p. 5501, Oct. 30, 2020, DOI: 10.1038/s41467-020-19341-4.

[342] S. Chen et al., "Heralded Three-Photon Entanglement from a Single-Photon Source on a Photonic Chip," *Physical Review Letters*, vol. 132, no. 13, p. 130 603, Mar. 28, 2024, DOI: 10.1103/PhysRevLett.132.130603.

[343] A. Reiserer and G. Rempe, "Cavity-based quantum networks with single atoms and optical photons," *Reviews of Modern Physics*, vol. 87, no. 4, pp. 1379–1418, Dec. 1, 2015, DOI: 10.1103/RevModPhys.87.1379.

[344] C.-W. Yang, Y. Yu, J. Li, B. Jing, X.-H. Bao, and J.-W. Pan, "Sequential generation of multiphoton entanglement with a Rydberg superatom," *Nature Photonics*, vol. 16, no. 9, pp. 658–661, Sep. 2022, DOI: 10.1038/s41566-022-01054-3.

[345] P. Thomas, L. Ruscio, O. Morin, and G. Rempe, "Efficient generation of entangled multiphoton graph states from a single atom," *Nature*, vol. 608, no. 7924, pp. 677–681, Aug. 2022, DOI: 10.1038/s41586-022-04987-5.

[346] D. Cogan, Z.-E. Su, O. Kenneth, and D. Gershoni, "Deterministic generation of indistinguishable photons in a cluster state," *Nature Photonics*, vol. 17, no. 4, pp. 324–329, Apr. 2023, DOI: 10.1038/s41566-022-01152-2.

[347] N. Coste et al., "High-rate entanglement between a semiconductor spin and indistinguishable photons," *Nature Photonics*, vol. 17, no. 7, pp. 582–587, Jul. 2023, DOI: 10.1038/s41566-023-01186-0.

[348] P. Thomas, L. Ruscio, O. Morin, and G. Rempe, "Fusion of deterministically generated photonic graph states," *Nature*, vol. 629, no. 8012, pp. 567–572, May 2024, DOI: 10.1038/s41586-024-07357-5.

[349] Z.-E. Su et al., "Continuous and deterministic all-photonic cluster state of indistinguishable photons," *Reports on Progress in Physics*, vol. 87, no. 7, p. 077 601, Jul. 2024, DOI: 10.1088/1361-6633/ad4c93.

[350] Y. Meng et al., "Deterministic photon source of genuine three-qubit entanglement," *Nature Communications*, vol. 15, no. 1, p. 7774, Sep. 5, 2024, DOI: 10.1038/s41467-024-52086-y.

[351] V. S. Ferreira, G. Kim, A. Butler, H. Pichler, and O. Painter, "Deterministic generation of multidimensional photonic cluster states with a single quantum emitter," *Nature Physics*, vol. 20, no. 5, pp. 865–870, May 2024, DOI: 10.1038/s41567-024-02408-0.

[352] J. O'Sullivan et al., "Deterministic generation of two-dimensional multi-photon cluster states," *Nature Communications*, vol. 16, no. 1, p. 5505, Jul. 1, 2025, DOI: 10.1038/s41467-025-60472-3.

[353] H. Huet et al., "Deterministic and reconfigurable graph state generation with a single solid-state quantum emitter," *Nature Communications*, vol. 16, no. 1, p. 4337, May 9, 2025, DOI: 10.1038/s41467-025-59693-3.

[354] P. Hilaire, L. Vidro, H. S. Eisenberg, and S. E. Economou, "Near-deterministic hybrid generation of arbitrary photonic graph states using a single quantum emitter and linear optics," *Quantum*, vol. 7, p. 992, Apr. 27, 2023, DOI: 10.22331/q-2023-04-27-992.

[355] L. A. Pettersson, A. S. Sørensen, and S. Paesani, "Deterministic Generation of Concatenated Graph Codes from Quantum Emitters," *PRX Quantum*, vol. 6, no. 1, p. 010 305, Jan. 8, 2025, DOI: 10.1103/PRXQuantum.6.010305.

[356] M. Zwerger, H. J. Briegel, and W. Dür, "Measurement-based quantum communication," *Applied Physics B*, vol. 122, no. 3, p. 50, Mar. 2016, DOI: 10.1007/s00340-015-6285-8.

[357] A. Patil and S. Guha, "An Improved Design for All-Photonic Quantum Repeaters," 2024, arXiv: 2405.11768 [quant-ph].

[358] F. Ewert, M. Bergmann, and P. Van Loock, "Ultrafast Long-Distance Quantum Communication with Static Linear Optics," *Physical Review Letters*, vol. 117, no. 21, p. 210 501, Nov. 14, 2016, DOI: 10.1103/PhysRevLett.117.210501.

[359] F. Ewert and P. van Loock, "Ultrafast fault-tolerant long-distance quantum communication with static linear optics," *Physical Review A*, vol. 95, no. 1, p. 012 327, Jan. 25, 2017, DOI: 10.1103/PhysRevA.95.012327.

[360] S.-W. Lee, T. C. Ralph, and H. Jeong, "Fundamental building block for all-optical scalable quantum networks," *Physical Review A*, vol. 100, no. 5, p. 052 303, Nov. 4, 2019, DOI: 10.1103/PhysRevA.100.052303.

[361] D. Niu, Y. Zhang, A. Shabani, and H. Shapourian, "All-photonic one-way quantum repeaters with measurement-based error correction," *npj Quantum Information*, vol. 9, no. 1, pp. 1–9, 1 Oct. 21, 2023, DOI: 10.1038/s41534-023-00775-9.

[362] T. C. Ralph, A. J. F. Hayes, and A. Gilchrist, "Loss-Tolerant Optical Qubits," *Physical Review Letters*, vol. 95, no. 10, p. 100 501, Aug. 30, 2005, DOI: 10.1103/PhysRevLett.95.100501.

[363] T. J. Bell, L. A. Pettersson, and S. Paesani, "Optimizing Graph Codes for Measurement-Based Loss Tolerance," *PRX Quantum*, vol. 4, no. 2, p. 020 328, May 17, 2023, DOI: 10.1103/PRXQuantum.4.020328.

[364] V. V. Albert et al., "Performance and structure of single-mode bosonic codes," *Physical Review A*, vol. 97, no. 3, p. 032 346, Mar. 30, 2018, DOI: 10.1103/PhysRevA.97.032346.

[365] K. Fukui, R. N. Alexander, and P. van Loock, "All-optical long-distance quantum communication with Gottesman-Kitaev-Preskill qubits," *Physical Review Research*, vol. 3, no. 3, p. 033 118, Aug. 5, 2021, DOI: 10.1103/PhysRevResearch.3.033118.

[366] F. Rozpędek, K. P. Seshadreesan, P. Polakos, L. Jiang, and S. Guha, "All-photonic Gottesman-Kitaev-Preskill–qubit repeater using analog-information-assisted multiplexed entanglement ranking," *Physical Review Research*, vol. 5, no. 4, p. 043 056, Oct. 18, 2023, DOI: 10.1103/PhysRevResearch.5.043056.

[367] J. Miguel-Ramiro, F. Riera-Sàbat, and W. Dür, "Quantum Repeater for *W* States," *PRX Quantum*, vol. 4, no. 4, p. 040 323, Nov. 8, 2023, DOI: 10.1103/PRXQuantum.4.040323.

[368] W. P. Grice, "Arbitrarily complete Bell-state measurement using only linear optical elements," *Physical Review A*, vol. 84, no. 4, p. 042 331, Oct. 19, 2011, DOI: 10.1103/PhysRevA.84.042331.

[369] F. Ewert and P. Van Loock, "3 / 4 -Efficient Bell Measurement with Passive Linear Optics and Unentangled Ancillae," *Physical Review Letters*, vol. 113, no. 14, p. 140 403, Sep. 30, 2014, DOI: 10.1103/PhysRevLett.113.140403.

[370] S. D. Barrett, P. Kok, K. Nemoto, R. G. Beausoleil, W. J. Munro, and T. P. Spiller, "Symmetry analyzer for nondestructive Bell-state detection using weak nonlinearities," *Physical Review A*, vol. 71, no. 6, p. 060 302, Jun. 7, 2005, DOI: 10.1103/PhysRevA.71.060302.

[371] H. A. Zaidi, C. Dawson, P. van Loock, and T. Rudolph, "Near-deterministic creation of universal cluster states with probabilistic Bell measurements and three-qubit resource states," *Physical Review A*, vol. 91, no. 4, p. 042 301, Apr. 1, 2015, DOI: 10.1103/PhysRevA.91.042301.

[372] T. Kilmer and S. Guha, "Boosting linear-optical Bell measurement success probability with predetection squeezing and imperfect photon-number-resolving detectors," *Physical Review A*, vol. 99, no. 3, p. 032 302, Mar. 1, 2019, DOI: 10.1103/PhysRevA.99.032302.

[373] N. Lütkenhaus, J. Calsamiglia, and K.-A. Suominen, "Bell measurements for teleportation," *Physical Review A*, vol. 59, no. 5, pp. 3295–3300, May 1, 1999, DOI: 10.1103/PhysRevA.59.3295.

[374] P. Hilaire, Y. Castor, E. Barnes, S. E. Economou, and F. Grosshans, "Linear Optical Logical Bell State Measurements with Optimal Loss-Tolerance Threshold," *PRX Quantum*, vol. 4, no. 4, p. 040 322, Nov. 6, 2023, DOI: 10.1103/PRXQuantum.4.040322.

[375] B. Li, K. Goodenough, F. Rozpędek, and L. Jiang, "Generalized Quantum Repeater Graph States," *Physical Review Letters*, vol. 134, no. 19, p. 190 801, May 12, 2025, DOI: 10.1103/PhysRevLett.134.190801.

[376] P. W. Shor, "Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer," *SIAM Review*, vol. 41, no. 2, pp. 303–332, Jun. 1, 1999, DOI: 10.1137/S0036144598347011.

[377] A. Bärtschi and S. Eidenbenz, "Grover Mixers for QAOA: Shifting Complexity from Mixer Design to State Preparation," in *2020 IEEE International Conference on Quantum Computing and Engineering (QCE)*, Oct. 2020, pp. 72–82, DOI: 10.1109/QCE49297.2020.00020.

[378] E. Farhi, J. Goldstone, and S. Gutmann, "A Quantum Approximate Optimization Algorithm," 2014, arXiv: 1411.4028 [quant-ph].

[379] S. Hadfield, Z. Wang, B. O'Gorman, E. G. Rieffel, D. Venturelli, and R. Biswas, "From the Quantum Approximate Optimization Algorithm to a Quantum Alternating Operator Ansatz," *Algorithms*, vol. 12, no. 2, p. 34, Feb. 12, 2019, DOI: 10.3390/a12020034.

[380] R. Babbush, J. McClean, M. Newman, C. Gidney, S. Boixo, and H. Neven, "Focus beyond quadratic speedups for error-corrected quantum advantage," *PRX Quantum*, vol. 2, no. 1, p. 010 103, Mar. 29, 2021, DOI: 10.1103/PRXQuantum.2.010103.

[381] J. Preskill, "Quantum Computing in the NISQ era and beyond," *Quantum*, vol. 2, p. 79, Aug. 2018, DOI: 10.22331/q-2018-08-06-79.

[382] B. Augustino et al., "Strategies for running the QAOA at hundreds of qubits," 2024, arXiv: 2410.03015 [quant-ph].

[383] R. Shaydulin, P. C. Lotshaw, J. Larson, J. Ostrowski, and T. S. Humble, "Parameter Transfer for Quantum Approximate Optimization of Weighted MaxCut," *ACM Transactions on Quantum Computing*, vol. 4, no. 3, pp. 1–15, Sep. 30, 2023, DOI: 10.1145/3584706.

[384] J. Golden, A. Bärtschi, D. O'Malley, and S. Eidenbenz, "Numerical Evidence for Exponential Speed-Up of QAOA over Unstructured Search for Approximate Constrained Optimization," in *2023 IEEE International Conference on Quantum Computing and Engineering (QCE)*, vol. 01, Sep. 2023, pp. 496–505, DOI: 10.1109/QCE57702.2023.00063.

[385] J. Golden, A. Bärtschi, D. O'Malley, and S. Eidenbenz, "The Quantum Alternating Operator Ansatz for Satisfiability Problems," in *2023 IEEE International Conference on Quantum Computing and Engineering (QCE)*, vol. 01, Sep. 2023, pp. 307–312, DOI: 10.1109/QCE57702.2023.00042.

[386] R. Shaydulin et al., "Evidence of scaling advantage for the quantum approximate optimization algorithm on a classically intractable problem," *Science Advances*, vol. 10, no. 22, eadm6761, May 29, 2024, DOI: 10.1126/sciadv.adm6761.

[387] S. Boulebnane and A. Montanaro, "Solving Boolean Satisfiability Problems With The Quantum Approximate Optimization Algorithm," *PRX Quantum*, vol. 5, no. 3, p. 030 348, Sep. 10, 2024, DOI: 10.1103/PRXQuantum.5.030348.

[388] L. K. Grover, "A fast quantum mechanical algorithm for database search," in *Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing*, ser. STOC '96, New York, NY, USA: Association for Computing Machinery, Jul. 1, 1996, pp. 212–219, DOI: 10.1145/237814.237866.

[389] G. Brassard, P. Høyer, M. Mosca, and A. Tapp, "Quantum amplitude amplification and estimation," in *Contemporary Mathematics*, S. J. Lomonaco and H. E. Brandt, Eds., vol. 305, Providence, Rhode Island: American Mathematical Society, 2002, pp. 53–74, DOI: 10.1090/conm/305/05215.

[390] C. Durr and P. Hoyer, "A Quantum Algorithm for Finding the Minimum," 1999, arXiv: quant-ph/9607014.

[391] A. Gilliam, S. Woerner, and C. Gonciulea, "Grover Adaptive Search for Constrained Polynomial Binary Optimization," *Quantum*, vol. 5, p. 428, Apr. 8, 2021, DOI: 10.22331/q-2021-04-08-428.

[392] T. Albash and D. A. Lidar, "Adiabatic quantum computation," *Reviews of Modern Physics*, vol. 90, no. 1, p. 015 002, Jan. 2018, DOI: 10.1103/RevModPhys.90.015002.

[393] M. Bernaschi, I. González-Adalid Pemartín, V. Martín-Mayor, and G. Parisi, "The quantum transition of the two-dimensional Ising spin glass," *Nature*, vol. 631, no. 8022, pp. 749–754, Jul. 2024, DOI: 10.1038/s41586-024-07647-y.

[394] L. Innocenti, G. D. Chiara, M. Paternostro, and R. Puebla, "Ultrafast critical ground state preparation via bang–bang protocols," *New Journal of Physics*, vol. 22, no. 9, p. 093 050, Sep. 2020, DOI: 10.1088/1367-2630/abb1df.

[395] Z.-C. Yang, A. Rahmani, A. Shabani, H. Neven, and C. Chamon, "Optimizing Variational Quantum Algorithms Using Pontryagin's Minimum Principle," *Physical Review X*, vol. 7, no. 2, p. 021 027, May 18, 2017, DOI: 10.1103/PhysRevX.7.021027.

[396] M. Cerezo et al., "Variational quantum algorithms," *Nature Reviews Physics*, vol. 3, no. 9, pp. 625–644, Sep. 2021, DOI: 10.1038/s42254-021-00348-9.

[397] E. Farhi, J. Goldstone, and S. Gutmann, "A Quantum Approximate Optimization Algorithm Applied to a Bounded Occurrence Constraint Problem," 2015, arXiv: 1412.6062 [quant-ph].

[398] J. Wurtz and P. Love, "MaxCut quantum approximate optimization algorithm performance guarantees for p > 1," *Physical Review A*, vol. 103, no. 4, p. 042 612, Apr. 27, 2021, DOI: 10.1103/PhysRevA.103.042612.

[399] M. Y. Niu, S. Lu, and I. L. Chuang, "Optimizing QAOA: Success Probability and Runtime Dependence on Circuit Depth," 2019, arXiv: 1905.12134 [quant-ph].

[400] V. Akshay, H. Philathong, M. E. S. Morales, and J. D. Biamonte, "Reachability Deficits in Quantum Approximate Optimization," *Physical Review Letters*, vol. 124, no. 9, p. 090 504, Mar. 5, 2020, DOI: 10.1103/PhysRevLett.124.090504.

[401] V. Akshay et al., "Circuit depth scaling for quantum approximate optimization," *Physical Review A*, vol. 106, no. 4, p. 042 438, Oct. 25, 2022, DOI: 10.1103/PhysRevA.106.042438.

[402] J. Basso, D. Gamarnik, S. Mei, and L. Zhou, "Performance and limitations of the QAOA at constant levels on large sparse hypergraphs and spin glass models," in *2022 IEEE 63rd Annual Symposium on Foundations of Computer Science (FOCS)*, Oct. 2022, pp. 335–343, DOI: 10.1109/FOCS54457.2022.00039.

[403] A. Anshu and T. Metger, "Concentration bounds for quantum states and limitations on the QAOA from polynomial approximations," *Quantum*, vol. 7, p. 999, May 11, 2023, DOI: 10.22331/q-2023-05-11-999.

[404] S. H. Sack and M. Serbyn, "Quantum annealing initialization of the quantum approximate optimization algorithm," *Quantum*, vol. 5, p. 491, Jul. 1, 2021, DOI: 10.22331/q-2021-07-01-491.

[405] M. Cain, E. Farhi, S. Gutmann, D. Ranard, and E. Tang, "The QAOA gets stuck starting from a good classical string," 2022, arXiv: 2207.05089 [quant-ph].

[406] J. Wurtz and P. J. Love, "Counterdiabaticity and the quantum approximate optimization algorithm," *Quantum*, vol. 6, p. 635, Jan. 27, 2022, DOI: 10.22331/q-2022-01-27-635.

[407] T. Satoh, Y. Ohkura, and R. Van Meter, "Subdivided Phase Oracle for NISQ Search Algorithms," *IEEE Transactions on Quantum Engineering*, vol. 1, pp. 1–15, 2020, DOI: 10.1109/TQE.2020.3012068.

[408] D. Koch, M. Cutugno, S. Karlson, S. Patel, L. Wessing, and P. M. Alsing, "Gaussian Amplitude Amplification for Quantum Pathfinding," *Entropy*, vol. 24, no. 7, p. 963, 7 Jul. 2022, DOI: 10.3390/e24070963.

[409] P. Shyamsundar, "Non-Boolean quantum amplitude amplification and quantum mean estimation," *Quantum Information Processing*, vol. 22, no. 12, p. 423, Nov. 30, 2023, DOI: 10.1007/s11128-023-04146-3.

[410] S. Marsh and J. B. Wang, "A quantum walk-assisted approximate algorithm for bounded NP optimisation problems," *Quantum Information Processing*, vol. 18, no. 3, p. 61, Jan. 16, 2019, DOI: 10.1007/s11128-019-2171-3.

[411] E. Farhi, D. Gamarnik, and S. Gutmann, "The Quantum Approximate Optimization Algorithm Needs to See the Whole Graph: A Typical Case," 2020, arXiv: 2004.09002 [quant-ph].

[412] E. Farhi, D. Gamarnik, and S. Gutmann, "The Quantum Approximate Optimization Algorithm Needs to See the Whole Graph: Worst Case Examples," 2020, arXiv: 2005.08747 [quant-ph].

[413] S. Marsh and J. B. Wang, "Combinatorial optimization via highly efficient quantum walks," *Physical Review Research*, vol. 2, no. 2, p. 023 302, Jun. 8, 2020, DOI: 10.1103/PhysRevResearch.2.023302.

[414] J. Golden, A. Bärtschi, D. O'Malley, and S. Eidenbenz, "Threshold-Based Quantum Optimization," in *2021 IEEE International Conference on Quantum Computing and Engineering (QCE)*, Oct. 2021, pp. 137–147, DOI: 10.1109/QCE52317.2021.00030.

[415] L. Bittel, S. Gharibian, and M. Kliesch, "The Optimal Depth of Variational Quantum Algorithms Is QCMA-Hard to Approximate," in *38th Computational Complexity Conference (CCC 2023)*, A. Ta-Shma, Ed., ser. Leibniz International Proceedings in Informatics (LIPIcs), vol. 264, Dagstuhl, Germany: Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2023, 34:1–34:24, DOI: 10.4230/LIPIcs.CCC.2023.34.

[416] A. İpek, "ON EIGENVALUE, SINGULAR VALUE AND NORMS OF A REAL SKEW-SYMMETRIC MATRIX," *Electronic Journal of Mathematical Analysis and Applications*, vol. 11, no. 1, pp. 111–115, Jan. 1, 2023, DOI: 10.21608/ejmaa.2023.284575.

[417] A. Lucas, "Ising formulations of many NP problems," *Frontiers in Physics*, vol. 2, Feb. 12, 2014, DOI: 10.3389/fphy.2014.00005.

[418] É. D. Taillard, *Design of Heuristic Algorithms for Hard Optimization: With Python Codes for the Travelling Salesman Problem* (Graduate Texts in Operations Research). Cham: Springer International Publishing, 2023, DOI: 10.1007/978-3-031-13714-3.

[419] S. Bravyi, A. Kliesch, R. Koenig, and E. Tang, "Obstacles to State Preparation and Variational Optimization from Symmetry Protection," *Physical Review Letters*, vol. 125, no. 26, p. 260 505, Dec. 24, 2020, DOI: 10.1103/PhysRevLett.125.260505.

[420] J. R. McClean et al., "Low-Depth Mechanisms for Quantum Optimization," *PRX Quantum*, vol. 2, no. 3, p. 030 312, Jul. 21, 2021, DOI: 10.1103/PRXQuantum.2.030312.

[421] R. D. Van Meter, "Architecture of a Quantum Multicomputer Optimized for Shor's Factoring Algorithm," Ph.D. dissertation, Keio University, Jul. 11, 2006, DOI: 10.48550/arXiv.quant-ph/0607065.

[422] M. G. Davis, J. Chung, D. Englund, and R. Kettimuthu, "Towards Distributed Quantum Computing by Qubit and Gate Graph Partitioning Techniques," in *2023 IEEE International Conference on Quantum Computing and Engineering (QCE)*, vol. 01, Sep. 2023, pp. 161–167, DOI: 10.1109/QCE57702.2023.00026.

[423] B. J. Brown and S. Roberts, "Universal fault-tolerant measurement-based quantum computation," *Physical Review Research*, vol. 2, no. 3, p. 033 305, Aug. 25, 2020, DOI: 10.1103/PhysRevResearch.2.033305.

[424] D. Litinski, "A Game of Surface Codes: Large-Scale Quantum Computing with Lattice Surgery," *Quantum*, vol. 3, p. 128, Mar. 5, 2019, DOI: 10.22331/q-2019-03-05-128.

[425] D. Horsman, A. G. Fowler, S. Devitt, and R. Van Meter, "Surface code quantum computing by lattice surgery," *New Journal of Physics*, vol. 14, no. 12, p. 123 011, Dec. 1, 2012, DOI: 10.1088/1367-2630/14/12/123011.

[426] D. Litinski and F. von Oppen, "Lattice Surgery with a Twist: Simplifying Clifford Gates of Surface Codes," *Quantum*, vol. 2, p. 62, May 4, 2018, DOI: 10.22331/q-2018-05-04-62.

[427] A. G. Fowler and C. Gidney, "Low overhead quantum computation using lattice surgery," 2019, arXiv: 1808.06709 [quant-ph].

[428] C. Poirson, J. Roffe, and R. I. Booth, "Engineering CSS surgery: Compiling any CNOT in any code," 2025, arXiv: 2505.01370 [quant-ph].

[429] D. Litinski and N. Nickerson, "Active volume: An architecture for efficient fault-tolerant quantum computers with limited non-local connections," 2022, arXiv: 2211.15465 [quant-ph].

[430] D. Sakuma et al., "An Optical Interconnect for Modular Quantum Computers," 2024, arXiv: 2412.09299 [quant-ph].

[431] D. Litinski, "Blocklet concatenation: Low-overhead fault-tolerant protocols for fusion-based quantum computation," 2025, arXiv: 2506.13619 [quant-ph].

[432] S. N. Saadatmand et al., "Fault-tolerant resource estimation using graph-state compilation on a modular superconducting architecture," 2024, arXiv: 2406.06015 [quant-ph].

[433] A. Robertson, H. Gao, and Y. R. Sanders, "A Resource Allocating Compiler for Lattice Surgery," 2025, arXiv: 2506.04620 [quant-ph].

[434] S. N. Saadatmand et al., "Superconducting qubits at the utility scale: The potential and limitations of modularity," 2025, arXiv: 2406.06015 [quant-ph].

[435] G. Bowen, A. Caesura, S. Devitt, and M. K. Vijayan, "Design and efficiency in graph-state computation," 2025, arXiv: 2502.18985 [quant-ph].

[436] C. E. Leiserson et al., "There's plenty of room at the Top: What will drive computer performance after Moore's law?" *Science*, vol. 368, no. 6495, eaam9744, Jun. 5, 2020, DOI: 10.1126/science.aam9744.

Appendix A

LIST OF PAPERS AND PRESENTATIONS

A.1 Peer-Reviewed Journals

1. N. Benchasattabuse, A. Bärtschi, L. P. García-Pintos, J. Golden, N. Lemons, and S. Eidenbenz, "Lower bounds on the number of rounds of the quantum approximate optimization algorithm required for guaranteed approximation ratios," *Physical Review A*, vol. 111, no. 6, p. 062 411, Jun. 9, 2025, DOI: 10.1103/PhysRevA.111.062411.

2. N. Benchasattabuse, M. Hajdušek, and R. Van Meter, "Engineering Challenges in All-Photonic Quantum Repeaters," *IEEE Network*, vol. 39, no. 1, pp. 132–139, Jan. 2025, DOI: 10.1109/MNET.2024.3411802.

3. J. C. A. Casapao, A. G. Maity, N. Benchasattabuse, M. Hajdušek, R. Van Meter, and D. Elkouss, "Disti-Mator: An entanglement distillation-based state estimator," *Communications Physics*, 2025, to be published, DOI: 10.48550/arXiv.2406.13937.

4. K. Suksen, N. Benchasattabuse, and P. Chongstitvatana, "Compact Genetic Algorithm with Quantum-Assisted Feasibility Enforcement," *ECTI Transactions on Computer and Information Technology (ECTI-CIT)*, vol. 16, no. 4, pp. 422–435, Oct. 11, 2022, DOI: 10.37936/ecti-cit.2022164.247821.

A.2 Conference Papers (Peer-Reviewed)

1. N. Benchasattabuse, M. Hajdušek, and R. Van Meter, "Integrating Entanglement Purification into All-Photonic Quantum Repeaters," in *2025 IEEE International Conference on Quantum Computing and Engineering (QCE)*, to be published, Apr. 25, 2025, DOI: 10.48550/arXiv.2504.18121.

2. J. C. A. Casapao et al., "A Double Selection Entanglement Distillation-Based State Estimator," in *2025 IEEE International Conference on Quantum Computing and Engineering (QCE)*, to be published, May 30, 2025, DOI: 10.48550/arXiv.2505.24280.

3. J. Chung et al., "Cross-Validating Quantum Network Simulators," in *IEEE INFOCOM 2025 - IEEE Conference on Computer Communications, Workshop on Quantum Networked Applications and Protocols*, to be published, Apr. 2, 2025, DOI: 10.48550/arXiv.2504.01290.

4. N. Patamawisut, N. Benchasattabuse, M. Hajdušek, and R. Van Meter, "Quantum Circuit Design for Decoded Quantum Interferometry," in *2025 IEEE International Conference on Quantum Computing and Engineering (QCE)*, to be published, Apr. 25, 2025, DOI: 10.48550/arXiv.2504.18334.

5. H. Yokomori, M. Koyama, N. Benchasattabuse, M. Hajdušek, S. Nagayama, and R. Van Meter, "Evaluation of Distimation's Real-world Performance on a Superconducting Quantum Computer," in *2025 IEEE International Conference on Quantum Computing and Engineering (QCE)*, to be published, Apr. 25, 2025, DOI: 10.48550/arXiv.2504.18141.

6. N. Benchasattabuse, M. Hajdušek, and R. Van Meter, "Architecture and Protocols for All-Photonic Quantum Repeaters," in *2024 IEEE International Conference on Quantum Computing and Engineering (QCE)*, vol. 01, Sep. 2024, pp. 1879–1889, DOI: 10.1109/QCE60285.2024.00217. **(2nd Best Paper)**

7. P. Fittipaldi, K. Teramoto, N. Benchasattabuse, M. Hajdušek, R. Van Meter, and F. Grosshans, "Entanglement Swapping in Orbit: A Satellite Quantum Link Case Study," in *2024 IEEE International Conference on Quantum Computing and Engineering (QCE)*, vol. 01, Sep. 2024, pp. 1924–1930, DOI: 10.1109/QCE60285.2024.00222.

8. M. Koyama et al., "Optimal Switching Networks for Paired-Egress Bell State Analyzer Pools," in *2024 IEEE International Conference on Quantum Computing and Engineering (QCE)*, vol. 01, Sep. 2024, pp. 1897–1907, DOI: 10.1109/QCE60285.2024.00219.

9. P. Pathumsoot et al., "Boosting end-to-end entanglement fidelity in quantum repeater networks via hybridized strategies," in *2024 International Conference on Quantum Communications, Networking, and Computing (QCNC)*, Jul. 1, 2024, pp. 49–56, DOI: 10.1109/QCNC62729.2024.00054.

10. K. S. Soon, N. Benchasattabuse, M. Hajdušek, K. Teramoto, S. Nagayama, and R. Van Meter, "Performance of Quantum Networks Using Heterogeneous Link Architectures," in *2024 IEEE International Conference on Quantum Computing and Engineering (QCE)*, vol. 01, Sep. 2024, pp. 1914–1923, DOI: 10.1109/QCE60285.2024.00221.

11. K. S. Soon et al., "An Implementation and Analysis of a Practical Quantum Link Architecture Utilizing Entangled Photon Sources," in *2024 International Conference on Quantum Communications, Networking, and Computing (QCNC)*, Jul. 2024, pp. 25–32, DOI: 10.1109/QCNC62729.2024.00014.

12. S. Liu, N. Benchasattabuse, D. Q. Morgan, M. Hajdušek, S. J. Devitt, and R. Van Meter, "A Substrate Scheduler for Compiling Arbitrary Fault-Tolerant Graph States," in *2023 IEEE International Conference on Quantum Computing and Engineering (QCE)*, Bellevue, WA, USA: IEEE, Sep. 17, 2023, pp. 870–880, DOI: 10.1109/QCE57702.2023.00101.

13. A. G. Maity, J. C. A. Casapao, N. Benchasattabuse, M. Hajdusek, R. Van Meter, and D. Elkouss, "Noise estimation in an entanglement distillation protocol," *ACM SIGMETRICS Performance Evaluation Review*, vol. 51, no. 2, pp. 66–68, Sep. 28, 2023, DOI: 10.1145/3626570.3626594.

14. N. Benchasattabuse, T. Satoh, M. Hajdušek, and R. Van Meter, "Amplitude Amplification for Optimization via Subdivided Phase Oracle," in *2022 IEEE International Conference on Quantum Computing and Engineering (QCE)*, Sep. 2022, pp. 22–30, DOI: 10.1109/QCE53715.2022.00020.

15. R. Satoh et al., "QuISP: A Quantum Internet Simulation Package," in *2022 IEEE International Conference on Quantum Computing and Engineering (QCE)*, Sep. 2022, pp. 353–364, DOI: 10.1109/QCE53715.2022.00056. **(Best Paper)**

16. R. Van Meter et al., "A Quantum Internet Architecture," in *2022 IEEE International Conference on Quantum Computing and Engineering (QCE)*, Broomfield, CO, USA: IEEE, Sep. 2022, pp. 341–352, DOI: 10.1109/QCE53715.2022.00055.

17. P. Pathumsoot, N. Benchasattabuse, R. Satoh, M. Hajdušek, R. Van Meter, and S. Suwanna, "Optimizing Link-Level Entanglement Generation in Quantum Networks with Unequal Link Lengths," in *2021 25th International Computer Science and Engineering Conference (ICSEC)*, Nov. 2021, pp. 179–184, DOI: 10.1109/ICSEC53205.2021.9684634.

18. K. Suksen, N. Benchasattabuse, and P. Chongstitvatana, "Experiments on Quantum Circuits for Compact Genetic Algorithm," in *2021 25th International Computer Science and Engineering Conference (ICSEC)*, Nov. 2021, pp. 185–190, DOI: 10.1109/ICSEC53205.2021.9684643.

A.3 Posters with Proceedings (Peer-Reviewed)

1. K. S. Soon, F. Kobayashi, N. Benchasattabuse, and S. Nagayama, "Surface Code Entanglement Ejection," in *2025 IEEE International Conference on Quantum Computing and Engineering (QCE)*, to be published, Albuquerque, NM, USA: IEEE, Aug. 31, 2025.

2. N. Benchasattabuse, M. Hajdušek, and R. Van Meter, "Protocols for All-Photonic Quantum Repeaters," in *2023 IEEE International Conference on Quantum Computing and Engineering (QCE)*, Bellevue, WA, USA: IEEE, Sep. 17, 2023, pp. 314–315, DOI: 10.1109/QCE57702.2023.10259.

A.4 Posters without Proceedings (Peer-Reviewed)

1. "Lower Bounds on Number of QAOA Rounds Required for Guaranteed Approximation Ratios," Quantum Information Processing Conference (QIP), Raleigh, NC, 25 February 2025.

2. "Flexible Repeater Graph States for Non-uniform End-to-end Paths," Workshop for Quantum Repeaters and Networks (WQRN3), Chicago, IL, 21 August 2022.

A.5 Manuscripts under Review and In Preparation

1. N. Benchasattabuse, M. Hajdušek, and R. Van Meter, "Bridging All-photonic and Memory-based Quantum Repeaters," submitted for publication, 2025.

2. S. Miyashita et al., "Digital quantum simulator for the time-dependent Dirac equation using discrete-time quantum walks," 2023, arXiv: 2305.19568 [quant-ph].

3. P. Pathumsoot et al., "Hybrid Error-Management Strategies in Quantum Repeater Networks," 2023, arXiv: 2303.10295 [quant-ph].

A.6 Talks

1. "Architecture and Protocols for All-photonic Quantum Repeaters," Invited talk at Mahidol University Summer School on Quantum Computing and Quantum Communication 2025 (SCQC2025), 27 June 2025.

2. "Towards a heterogeneous quantum network: bridging all-photonic and memory-based quantum repeaters," Quantum Internet Task Force Workshop, 11 June 2025. (online)

3. "Architecture and Protocols for All-photonic Quantum Repeaters," Invited talk at Mahidol University special seminar Celebrating 100 Years of Quantum Physics, 21 March 2025.

4. "Architecture and Protocols for All-photonic Quantum Repeaters," IEEE International Conference on Quantum Computing and Engineering (QCE24), 15 September 2024.

5. "Lower Bound on Number of QAOA Round Required for Guaranteed Approximation Ratios," Keio University Yagami Campus, 29 February 2024.

6. "Lower Bound on Number of QAOA Round Required for Guaranteed Approximation Ratios," Invited talk at Mahidol University, 8 January 2024.

7. "Lower Bound on Number of QAOA Round Required for Guaranteed Approximation Ratios," Japanese-French Quantum Information workshop (JFQI), 15 December 2023.

8. "Amplitude Amplification for Optimization via Subdivided Phase Oracle," IEEE International Conference on Quantum Computing and Engineering (QCE22), 19 September 2022.

9. "Distributing Multipartite Entangled State over Quantum Internet," International Congress on Science, Technology and Technology-based Innovation (STT47), 5 October 2021.