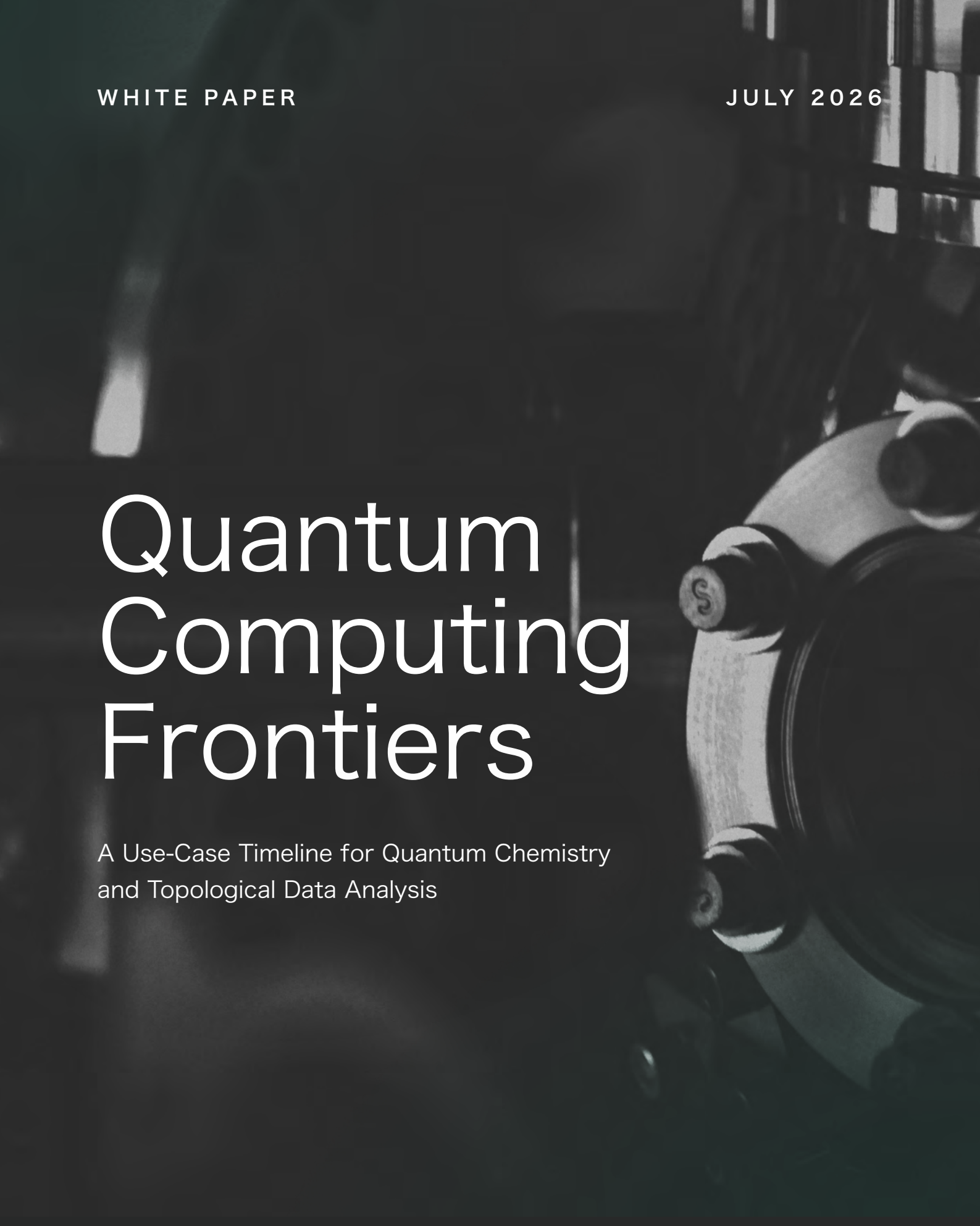
WHITE PAPER

JULY 2026



Quantum Computing Frontiers

A Use-Case Timeline for Quantum Chemistry and Topological Data Analysis

SoftBank

QUANTINUUM

Contents

Executive Summary

Quantum computing is entering a decisive phase in which progress in hardware performance must be translated into clearly defined, economically meaningful use cases. The central question is no longer whether quantum computing may deliver value in principle, but rather which problem classes become executable at which stage of hardware maturity, and how that technical progression can be aligned with viable commercial deployment.

This white paper presents the outcomes of a joint research initiative between SoftBank Corp. and Quantinuum, conducted under a shared long-term vision of realizing quantum AI data centers —hybrid computational infrastructures in which quantum processors operate alongside AI and HPC systems. The core objective of this work is to construct a structured *use-case timeline* that connects algorithmic resource requirements to Quantinuum's hardware roadmap.

Two representative application domains are analyzed: quantum chemistry and topological data analysis (TDA). These domains were selected because they combine structural computational challenges, industrial relevance, and scalability characteristics that are well aligned with quantum computational paradigms.

In quantum chemistry, feasibility is expressed in terms of system size measured by the number of spin orbitals. The study evaluates phase-estimation-based algorithms and, uniquely, incorporates error-corrected circuit execution at the algorithmic level using the Steane [[7,1,3]] code. Rather than demonstrating break-even behavior for isolated logical gates, this work evaluates break-even fidelity for entire algorithm-inspired circuits, marking a significant step toward realistic fault-tolerant quantum computation. The resulting timeline indicates a gradual expansion from small proof-of-concept systems to medium-scale excited-state calculations and, ultimately, chemically accurate simulations in the large-scale fault-tolerant regime.

In TDA, feasibility is characterized through graph size and structured parameters in complete k-partite graphs. Emphasis is placed on Noisy Intermediate scale quantum (NISQ)-executable circuits and early-stage applications such as International Revenue Share Fraud (IRSF) detection in telecommunications networks. Because improvements in anomaly detection precision can directly translate into measurable financial impact, TDA may represent a shorter commercialization pathway relative to quantum chemistry.

By aligning technical feasibility for both application domains against the roadmap of each successive generation of Quantinuum's quantum computers—Helios (current generation), Sol (anticipated in 2027), Apollo (anticipated in 2029), and Lumos, it's planned large-scale fault-tolerant quantum computing (FTQC) anticipated in the 2030s—the study provides a generation-indexed use case timeline of executable workloads. Importantly, this roadmap is framed as an assumption-driven envelope rather than a deterministic forecast, explicitly acknowledging uncertainties in qubit scaling, physical error rates, and error-correction overhead.

Beyond technical feasibility, this white paper could potentially help future enterprise customers of a quantum AI data center to connect executable workloads to value creation mechanisms and indicative market opportunities.

Taken together, the use-case timeline developed in this study serves as both a technical roadmap and a strategic governance framework. It aligns hardware evolution, algorithmic capability, market formation, and long-term commercial value under a coherent vision of quantum-enabled computational infrastructure.

1 Introduction

1.1 Background and Motivation

In what is often referred to as the second quantum revolution, quantum computing has witnessed remarkable and sustained progress. Initial expectations have been clearly surpassed, and the pace of advancement continues to accelerate.

Quantinuum's third-generation quantum computer, *Helios*, exemplifies this rapid development. As one of the most advanced quantum systems currently available, Helios achieves the highest accuracy of any commercial system in the world today[1] and enables the execution of quantum circuits that extend beyond the reach of classical simulation. Such milestones mark a transition from proof-of-principle demonstrations to increasingly ambitious computational experiments.

However, hardware progress alone is not sufficient to unlock practical value. To leverage these capabilities for modeling and solving real-world problems, advances in quantum computing hardware must be complemented by equally strong developments in quantum algorithms and system-level integration. This requirement is particularly important because quantum computing relies on fundamentally different computational principles compared to established classical paradigms such as high-performance computing (HPC) and artificial intelligence (AI).

A central challenge therefore emerges: which real-world problems can benefit from quantum computation, at what scale, and under what hardware conditions? Answering this question requires a structured and quantitative approach that connects algorithmic development, resource estimation, hardware capabilities, and application needs.

1.2 Purpose of this White Paper

This white paper presents the results of a joint research effort between SoftBank Corp. and Quantinuum, conducted under the long-term vision of realizing quantum AI data centers. Within this vision, quantum computers are considered as computational resources that operate alongside AI and HPC systems in future hybrid infrastructures.

The primary objective of this document is to construct a *use-case timeline*. Rather than discussing quantum computing in abstract terms, we aim to clarify which problem classes, at what scale and accuracy requirements, and under which generation of quantum hardware may become practically achievable.

To this end, the document details the development of quantum computational methods and experiments for two major use cases: quantum chemistry and topological data analysis (TDA). Beyond scalable quantum algorithm design and the quantum error correction layers required to execute them, particular attention is given to the interplay between quantum computing, HPC, and AI, as well as to the integration of quantum processors within modern computational infrastructures.

[1] Based on two-qubit gate fidelity as of December 31, 2025

1.3 Target Domains

This study focuses on two representative domains:

· Quantum chemistry, particularly excited-state dynamics and photochemical reactions
· Topological data analysis (TDA)

Although these domains differ in application area, they share several important characteristics:

(1) They exhibit structural computational challenges in classical approaches, where resource requirements grow rapidly with problem size.

(2) They involve high-dimensional, strongly correlated, or non-linear structures that are naturally aligned with quantum computational frameworks.

(3) They maintain clear connections to industrially relevant applications.

Quantum chemistry is closely linked to advanced material design, photo-responsive molecules, and energy-related technologies that underpin a sustainable society. In contrast, TDA provides tools for analyzing structural patterns in large-scale graph data, with potential relevance to high-reliability networks and anomaly detection. Together, these domains offer complementary perspectives on how quantum computing may create value across different sectors.

1.3.1 Quantum Chemistry

Quantum chemistry plays a central role in understanding and designing molecular systems whose behavior is governed by electronic structure. Of particular importance are excited-state dynamics and photochemical reactions, which underpin photo-responsive materials, molecular switches, and energy-conversion processes.

Modeling such phenomena presents fundamental challenges for classical computing. Excited states often involve strong electronic correlation, near-degeneracies, and conical intersections, where multiple electronic surfaces come close in energy and interact non-trivially. Accurately describing these effects typically requires computational methods whose cost scales steeply with system size, making high-precision simulations of larger molecular systems increasingly impractical even on advanced HPC platforms.

Quantum computing offers a conceptually different approach by directly representing and manipulating quantum states. In principle, algorithms such as quantum phase estimation enable access to energy spectra and phase information that are central to excited-state chemistry. Within this study, quantum chemistry is therefore investigated as a domain where quantum algorithms may progressively extend the tractable range of molecular complexity as hardware capabilities mature.

1.3.2 Topological Data Analysis

Topological Data Analysis (TDA) comprises mathematical methods rooted in topology and geometry that extract structural information from high-dimensional data. The core idea is to interpret data as occupying an underlying manifold and to then characterize the manifold's global and local properties through scaffolding constructions such as simplicial complexes and discrete Laplacian operators.

Two prominent branches are persistent homology, which captures multi-scale topological features such as connected components and holes, and Laplacian-based approaches, which analyze flows and spectral properties on high-dimensional structures. These methods provide descriptive and exploratory tools that complement both unsupervised and supervised machine learning.

Despite their expressive power, state-of-the-art classical TDA implementations face severe scaling limitations. The combinatorial growth of simplicial complexes leads to exponential increases in computational cost, often requiring aggressive sampling or dimensional truncation. This makes large-scale or high-resolution analysis difficult in practice.

Because industrial and network data are frequently high-dimensional, correlated, noisy, and partially unlabeled, TDA provides a natural analytical framework. Applications such as anomaly, reliability, or fraud detection benefit from understanding the geometric structure and distributional shifts

of system states, rather than relying solely on models trained on historical norms. In this context, quantum-enhanced TDA is explored as a potential pathway to overcome classical scaling barriers.

1.4 Approach of this Study

The approach adopted in this white paper is structured and quantitative. For each selected domain, we first define representative use cases and identify the corresponding quantum algorithms that would be required to address them.

Next, we characterize the relevant problem sizes and complexity parameters, and perform resource estimations in terms of logical qubit counts, quantum gate counts, and circuit depth. These estimates are then evaluated under explicit assumptions about quantum error correction schemes, including code overheads and target logical error rates.

The resulting resource requirements are compared with projected hardware roadmaps, including expected numbers of available qubits and achievable error rates at different stages of technological development. By aligning algorithmic resource demands with hardware capabilities, we construct a timeline that answers the question: *which tasks are expected to become feasible, and at approximately what stage of hardware maturity?*

In addition to this forward-looking analysis, a distinguishing feature of this study is the execution of proof-of-concept (PoC) demonstrations on currently available quantum hardware as of 2026. For tasks that are predicted to be executable with present-day systems, we implement and run concrete programs in order to validate feasibility and to benchmark performance under realistic experimental conditions.

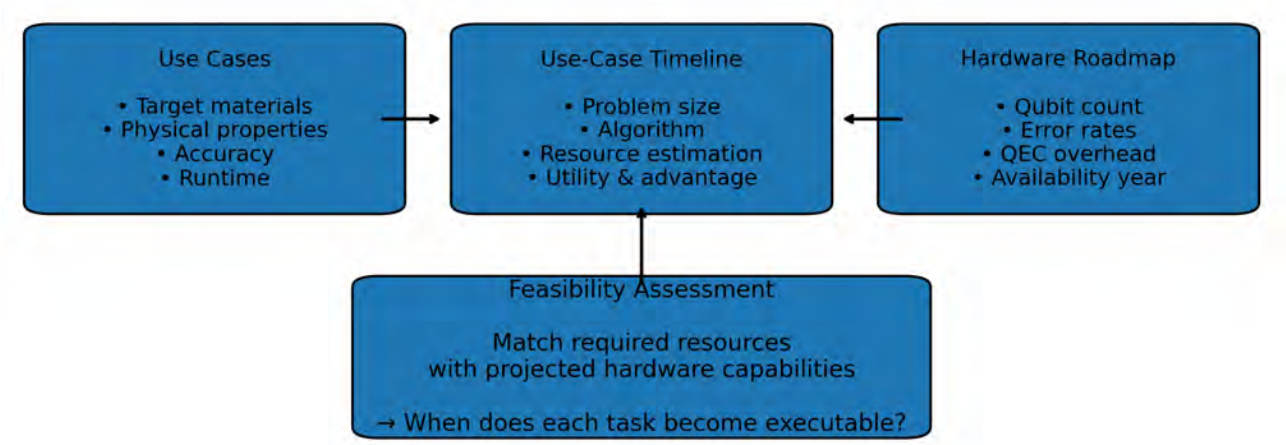


Figure 1.1: Conceptual framework for constructing a use-case timeline

1.5 Organization of this White Paper

The remainder of this document is organized as follows:

- Chapter 2 introduces the evaluation framework and the integrated roadmap used throughout the study.
- Chapter 3 examines quantum chemistry use cases and their technological outlook.
- Chapter 4 analyzes quantum-enhanced approaches in topological data analysis.
- Chapter 5 synthesizes both domains into a consolidated use-case timeline and discusses strategic implications.

Technical details—including algorithmic formulations, circuit constructions, resource estimations, and experimental configurations—are provided in the Appendices. The main chapters focus on the interpretation of results and their positioning along the proposed timeline.

2 Methodology and Evaluation Framework

2.1 Overall Approach

In this study, we deliberately move beyond purely theoretical complexity analyses or abstract algorithmic discussions. Instead, we place strong emphasis on explicitly constructing concrete quantum circuits and, whenever possible, executing them on simulators and available quantum hardware. In other words, we prioritize algorithms that can be written down at the circuit level, implemented under realistic gate-set constraints, and evaluated in the presence of noise.

This implementation-driven approach is intended to narrow the gap between theoretical proposals and practical realization. By grounding the analysis in executable circuits, we aim to provide a more credible and actionable answer to the central question of this white paper: which tasks can become feasible, and when?

2.2 Evaluation Strategy for Quantum Chemistry

For quantum chemistry, we focus on algorithms based on Quantum Phase Estimation (QPE). Comprehensive resource estimations for large-scale quantum chemistry calculations using QPE—covering logical qubit counts, gate counts, and error-correction overhead—have already been investigated in a previously published white paper. Accordingly, we do not repeat those detailed resource analyses here, but instead build upon those established results.

The distinctive contribution of the present study lies in the experimental investigation of error-corrected quantum circuits at the algorithmic level. Specifically, we implement proof-of-concept (PoC) demonstrations in which quantum error correction is incorporated directly into benchmark circuits inspired by QPE.

As an initial and conceptually clear code suitable for small-scale demonstrations, we employ the Steane code [[7,1,3]]. While it remains uncertain whether this code will play a central role in long-term large-scale architectures, it provides a well-understood framework for early-stage fault-tolerant experimentation. Logical qubits are encoded using this code, and benchmark circuits representative of phase-estimation-style computations are implemented at the logical level.

In this work, we define *break-even* as the point at which the overall execution fidelity of an error-corrected logical circuit exceeds the fidelity obtained by executing the same algorithm directly on physical qubits without error correction. This definition differs fundamentally from many previous demonstrations of quantum error correction. Earlier break-even studies have typically focused on metrics such as extended qubit lifetime, the fidelity of a single logical gate, or the preparation of states using only a small number of logical operations. By contrast, our evaluation targets the fidelity of an entire, algorithm-inspired quantum circuit. In this sense, the break-even condition is assessed at the level of a realistic computational workload rather than at the level of isolated operations.

Furthermore, our implementation assumes a fully fault-tolerant setting. Controlled rotation gates appearing in phase estimation must be decomposed into a universal fault-tolerant gate set, typically expressed in terms of Clifford gates (such as the Hadamard gate) and the non-Clifford T gate. Since the T gate cannot be implemented transversally in many codes, it requires ancillary resources such as magic-state distillation and is therefore significantly more costly. While some previous experimental studies have effectively treated rotation gates as physical-level operations—leading to partially fault-tolerant demonstrations—we explicitly decompose these rotations and assume a fully fault-tolerant implementation. To keep resource requirements within a demonstrable range, benchmark circuits are

carefully designed to limit the number of required T gates while preserving the essential structure of the algorithm.

2.3 Evaluation Strategy for Topological Data Analysis

In contrast to quantum chemistry, our investigation of Topological Data Analysis (TDA) concentrates on feasibility within the NISQ regime and does not assume quantum error correction for every stage of the timeline. We analyze quantum algorithms for Laplacian-moment-type computations under realistic constraints on qubit counts, circuit depth, and shot requirements, and implement small-scale instances on available hardware.

To enable controlled and systematic resource estimation, we do not rely on arbitrary real-world graphs, whose intrinsic complexity is difficult to tune. Instead, we adopt complete k-partite graphs as structured benchmark instances. Prior research has shown that computing Betti numbers for such graphs (without, of course, using analytic results) can become classically challenging. Moreover, the complexity of these graphs can be adjusted in a transparent manner through two parameters: the number of clusters k and the number of nodes within each cluster. This parametric control allows us to scale problem difficulty in a systematic and reproducible way.

While theoretical analyses often relate computational difficulty to the magnitude of the spectral gap (i.e., the smallest non-zero eigenvalue of the Laplacian), this quantity is generally not directly controllable in advance. For this reason, we do not use the Laplacian gap as a primary complexity parameter in this study. Instead, we rely on the structurally tunable parameters of the k-partite construction to characterize and scale problem instances.

2.4 Positioning of This Chapter

Taken together, this methodology integrates theoretical resource considerations, explicit circuit construction, NISQ-level demonstrations, and fault-tolerant proof-of-concept experiments. The framework established in this chapter serves as the common foundation for the domain-specific analyses presented in Chapters 3 and 4, and ultimately supports the construction of the consolidated use-case timeline in Chapter 5.

3 Quantum Computing in Quantum Chemistry

3.1 Industrial Relevance of Quantum Chemistry

Quantum chemistry is a high-impact arena for quantum computing because many industrial materials depend on phenomena that are too complex to model reliably at scale with classical methods. As hardware matures, quantum simulations promise predictive accuracy for light-driven reactions in large, realistic systems. Unlocking design insights that are currently out of reach, quantum chemistry turns R&D from trial-and-error into model-driven engineering.

Optical switching materials are a flagship example. These molecules power reconfigurable photonics, smart windows, augmented-reality displays, high-density data storage, and responsive coatings. Their performance-speed, energy efficiency, contrast, and durability hinges on ultrafast electronic transitions that determine whether a photon triggers the "right" structural change or funnels energy into side reactions and degradation reactions. In real devices, substitutions, heavy-atom effects, and host environments all shift these transitions in ways that classical multireference methods struggle to capture at industrial scales. Quantum simulations at scale allows map and tune these critical transitions in chemically realistic settings. This helps to accelerate discovery by screening switch architectures in silico before synthesis. Furthermore, raise hit rates by targeting structures with the desired switching thresholds, fast write/erase times, and long cycle life. Co-optimize devices by co-designing molecules with hosts and optical stacks to improve contrast and reduce power. Finally, accelerated quantum-accelerated R&D cycles enable early go/no-go decisions allowing for a significant de-risking of complex materials-science investments.

A practical roadmap starts with near-term pilots demonstrating a end-to-end quantum error corrected workflows implementing small, yet scalable quantum phase estimation circuits. Later work gradually extends the size of quantum computing experiments while leveraging increasingly higher fidelity quantum error corrected logical qubits. Finally, increased simulation sizes enable also vastly more complex models culminating, on fault-tolerant hardware, in end-to-end predictive pipelines from molecular design to device specification. As such, optical switching materials offer a credible path to first, tangible quantum advantage as well as a durable source of competitive differentiation.

3.2 Target Application Domains

Optical switches are materials and devices that reversibly modulate light with light, heat, or electrical stimuli. Such switches are foundational enablers for multiple growth markets. Their value stems from fast response, low energy per operation, compact footprints, and compatibility with integrated photonics.

Telecom & Datacom: In backbone networks and data centers, optical switches power wavelength routing, protection switching, and dynamic bandwidth allocation. They enable flexible, low-latency reconfiguration with minimal optical - electrical-optical conversions, improving throughput and energy efficiency.

On-Chip Interconnects & HPC: In silicon photonics, switches provide nanosecond-to-picosecond path selection for memory and compute disaggregation, photonic AI accelerators, and chiplet fabrics which reduce congestion and energy per bit while scaling bandwidth density.

Displays, AR/VR & Spatial Light Modulation: Ultrafast, high-contrast switching underpins holographic and near-eye displays, dynamic beam steering, and pixel-level dimming, improving brightness, color fidelity, and refresh rates for immersive systems.

Smart Infrastructure: Photoresponsive coatings and switchable windows deliver adaptive tinting, privacy on demand, and glare control, contributing to building energy savings and user comfort without bulky mechanics.
Sensing, Imaging & Lidar: Optical shutters and modulators enable time-gated imaging, background suppression, hyperspectral selection, and rapid beam control-boosting signal-to-noise and range precision in autonomous systems and scientific instruments.
Reconfigurable Photonics & Neuromorphic Computing: Nonvolatile or low-power switches program photonic circuits, optical FPGAs, and synaptic weights in analog photonic processors, supporting agile, edge-deployed AI with high throughput per watt.
Security & Data Storage: Switchable optical states support anti-counterfeiting features, optical encryption, and multilayer/holographic storage, increasing data density and tamper resistance.

Across these domains, business value concentrates in lower total cost of ownership (via energy and component reductions), new performance envelopes (speed, contrast, endurance), and faster product cycles enabled by model-driven quantum computational co-design of materials, devices, and systems.

Furthermore, the techniques developed for studying optical switches can be readily applied to other excited-state chemical reactions, including those in batteries and solar cells, because their potential energy surfaces share many common features. For example, Fig. 3.1 shows the potential energy curve along the double-bond rotation of an ethylene molecule. The reaction starts by ab-

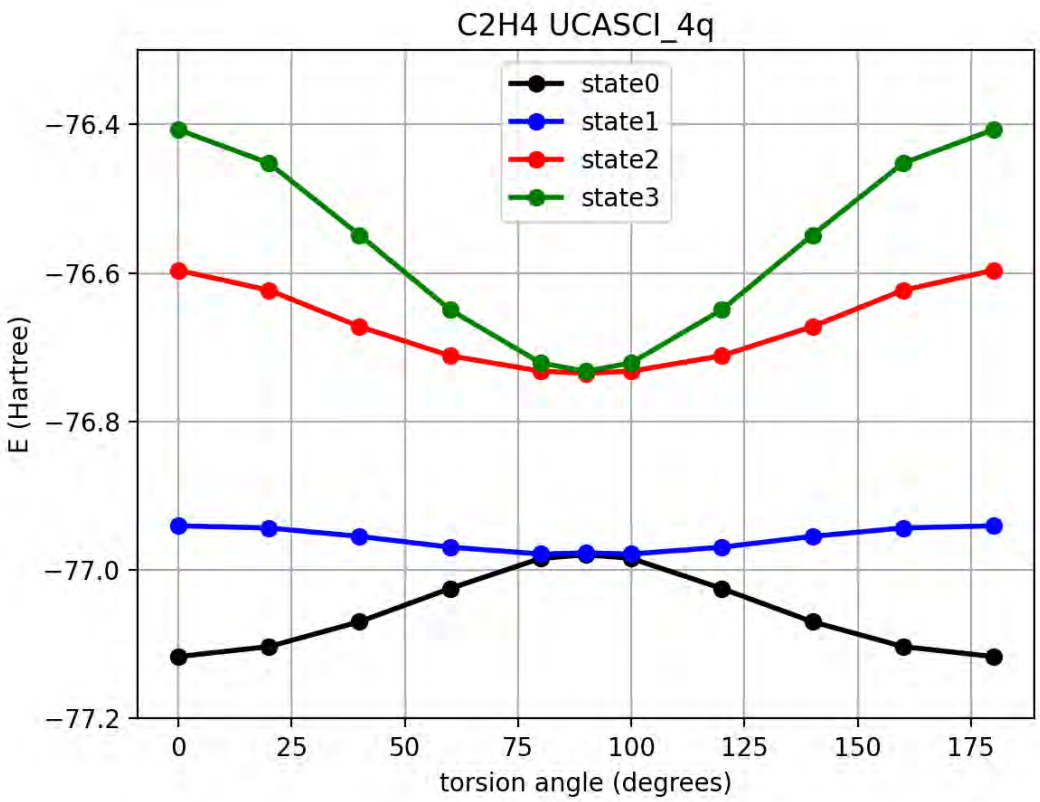


Figure 3.1: Potential energy curves of ethylene obtained from classical electronic structure calculations using the STO-3G basis set. To capture the essential degrees of freedom governing the conical intersection (CI), we adopt a minimal active space of (2e,2o), corresponding to two electrons in the π and π^* orbitals.

sorbing a photon to reach an excited state at the equilibrium geometry (torsion angle = 0 degrees), then follows the excited-state potential energy surface, passing through the conical intersection (torsion angle = 90 degrees) to return to the ground state. The industry-relevant chemical reactions are way more complicated than those of the ethylene molecule. See App. A.1 for a classical analysis of

photochromic molecules.

Excited-state calculations are generally more challenging than ground-state calculations due to stronger configuration interactions, i.e., involving more spin orbitals. This leads to an exponential increase in the memory requirements of classical computers, which quantum computers may mitigate. Once the mapping of chemistry information to the quantum computers is complete, we can use a scalable quantum algorithm as described below.

3.3 Quantum Algorithms for Quantum Chemistry

Quantum chemistry is a leading near-term domain for quantum advantage. This is due to the ability of quantum computing to solve electronic structure problems that describe many industrially critical phenomena around strong electron correlation, multiple near-degenerate states, and couplings with atomic degrees of freedom. High-accuracy electronic structure models like CASSCF/CASPT2 or MRCI become prohibitive for classical high-performance computing as problem sizes and model capabilities grow. However, such models and problems are precisely where predictive power matters most for real computational materials science R&D. On future fault-tolerant hardware, quantum algorithms promise scalable, accurate and reliable solution of ground and excited state properties, accurate energy gaps, and electrodynamic responses. Such predictive capabilities are central to design-grade simulation rather than post-experimental rationalization.

Optical switching materials provide a compelling exemplar where quantum simulation can move decisively beyond classical reach. Their function-fast, reversible, and fatigue-resistant photoisomerization depends on traversing conical intersections that funnel population between electronic states on sub-picosecond timescales. Engineering device - level figures of merit for materials design requires controlling the conical intersection topography (location, slope, and seam connectivity) and nonadiabatic couplings in chemically realistic settings. These scenarios demand large active spaces, multiple state manifolds, and explicit environment modeling across hundreds of atoms.

Quantum computational pipelines can natively target these bottlenecks. Quantum Phase Estimation (QPE) is a quantum subroutine that provides an exponential speedup for energy estimation. Given the desired precision ε, the QPE circuit generally grows $O(\frac{1}{\varepsilon})$ while the shot overhead remains constant. In other words, we need to double the circuit to double the accuracy. Given the practical limitations on the fidelity of physical quantum operations, QPE requires QEC to exploit its full potential.

We take a top-down approach to the end-to-end QPE, as in Ref. [1], where we first build the end-to-end pipeline and then improve its components step-by-step. This complements the bottom-up approach, which is a common practice among QEC researchers.

3.4 Molecular Hamiltonian

In this work, we consider the ethylene molecule (C_2H_4) as a prototypical and minimal photochemical model system exhibiting a conical intersection (CI) [2]. Ethylene is one of the simplest π-bonded unsaturated molecules, and it is well established that torsion about the C=C double bond, together with pyramidalization, leads to a degeneracy between electronic potential energy surfaces, giving rise to a conical intersection. Thus, it serves as a simple model system for photoinduced chemical reactions.

To capture the essential degrees of freedom governing the CI, we adopt a minimal active space of (2e,2o), i.e., two electrons in the π and π^* orbitals. We employ the minimal basis set (STO-3G) throughout this work. Fig. 3.1 shows the potential energy curves of ethylene obtained from classical electronic structure calculations. Near a torsion angle of approximately 90 degrees, the structure corresponding to the conical intersection emerges. Our objective is to evaluate the energy difference between the ground state and the first excited state at this 90-degree geometry.

First, the electronic Hamiltonian at the 90-degree structure is computed using classical electronic structure methods. The resulting second-quantized Hamiltonian is mapped to a qubit Hamiltonian

via the Jordan–Wigner transformation. We then reduce the number of qubits by exploiting particle-number and spin symmetries through tapering. As a result, the problem is reduced to an effective two-qubit Hamiltonian given by

$$H = h_1 Z_1 + h_2 Z_2 + h_3 Y_1 Y_2 + h_4 Z_1 Z_2 + h_5 I, \tag{3.1}$$

where $\{X_i, Y_i, Z_i\}$ denote the Pauli operators acting on the i-th qubit, I is the identity operator, and $\{h_i\}$ are real coefficients. The Hamiltonian coefficients are below in atomic units [3].

$$(h_1, h_2, h_3, h_4, h_5) = (-0.000302, -0.000302, -0.122188, 0.001280, -76.856020). \tag{3.2}$$

These parameters are evaluated with `InQuanto` v4.2 [4] and its extension to `PySCF` v2.4 [5] by performing the Hartree–Fock method.

This two-qubit model serves as a minimal effective Hamiltonian describing the energy gap between the ground and first excited states in the vicinity of the CI. It therefore provides a suitable framework for evaluating the performance and error resilience of quantum phase difference estimation.

3.5 Quantum Phase Difference Estimation

Let $\{|\phi_i\rangle\}_{i=0}^{N_{\text{eig}}-1}$ be the eigenstates of the unitary $U = e^{-iHt}$,

$$U\,|\phi_i\rangle = e^{i\phi_i}\,|\phi_i\rangle\,. \tag{3.3}$$

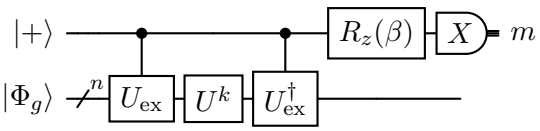
We consider two quantum states expressed in the eigenbasis as

$$|\Phi_g\rangle = \sum_{i=0}^{N_{\text{eig}}-1} c_i\,|\phi_i\rangle\,, \tag{3.4}$$

$$|\Phi_e\rangle = \sum_{j=0}^{N_{\text{eig}}-1} d_j\,|\phi_j\rangle\,. \tag{3.5}$$

The goal of phase difference estimation is to determine a subset of eigenphase differences $\Delta\phi_{ij} := \phi_i - \phi_j$. In the present work, we use information-theoretic sampling method [6] to estimate the phase difference. The measurement statistics of the ancilla qubit contain contributions weighted by $|c_i|^2|d_j|^2$.

The Quantum phase difference estimation(QPDE) [7–9] uses the following circuit:



(3.6)

This circuit has 5 key components:

(1) Preparation of $|\Phi_g\rangle$ and $|+\rangle$.

(2) Application of a controlled excitation operation that prepares $|\Phi_e\rangle$ conditioned on the ancilla qubit (ctrl-$U\text{ex}$).

(3) Application of U^k on the system register.

(4) Application of a phase-shift gate, $R_Z(\beta) = e^{-i\frac{\beta}{2}Z}$.

(5) Measurement of the ancilla qubit in the X-basis.

The parameters $\{k_l\}$ and $\{\beta_l\}$ are randomly sampled from $k \in \{1, 2, \ldots, k_{\max}\}$ and $\beta \in \{0, \frac{\pi}{2}\}$, respectively, where $k_{\max}$ controls the precision of QPDE, and determines the maximum depth of the circuit (3.6). The probability of measuring $m \in \{0, 1\}$ from a single ideal execution of the circuit (3.6) is expressed as,

$$P(m|\{c_i, d_j, \Delta\phi_{ij}\}; k, \beta) = \frac{1}{2}\left[1 + \sum_{i,j} |c_i|^2 |d_j|^2 \cos\left(k\Delta\phi_{ij} + \beta - m\pi\right)\right]. \tag{3.7}$$

We run the circuit (3.6) $N_{\rm s}$ times and collect the measurement outcomes $\{m_l\}_{l=1}^{N_s}$.

Provided $\{m_l\}_{l=1}^{N_{\rm s}}$, we calculate the following distribution over $\tilde{\phi} \in [-\pi, \pi)$,

$$Q(\tilde{\phi} \mid \{m_l\}; \{k_l, \beta_l\}) := \prod_{l=1}^{N_{\rm s}} Q(m_l \mid \tilde{\phi}; k_l, \beta_l) \tag{3.8}$$

$$\text{with} \quad Q(m_l \mid \tilde{\phi}; k_l, \beta_l) = \frac{1 + \cos(k_l\tilde{\phi} + \beta_l - m_l\pi)}{2}. \tag{3.9}$$

Eq. (3.9) represents the probability of obtaining a measurement outcome m_l if the input state were an eigenstate pair associated with a phase difference $\tilde{\phi}$. The resulting distribution is peaked around $\{\Delta\phi_{ij}\}_{i,j}$, and the dominant peak corresponds to $\Delta\phi_{i^*j^*}$. Hence, the target phase difference can be estimated by maximizing the likelihood function,

$$\arg\max_{\tilde{\phi}} \; Q(\tilde{\phi} \mid \{m_l\}; \{k_l, \beta_l\}). \tag{3.10}$$

For phase-difference estimation, we approximate the time evolution operator $U^k = e^{-iHkt}$ up to a global phase by a first-order Lie–Trotter product formula

$$U^k := \left(e^{-ih_1 Z_1 t} e^{-ih_2 Z_2 t} e^{-ih_3 Y_1 Y_2 t}\right)^k \times e^{-ih_4 Z_1 Z_2 kt} \tag{3.11}$$

$$= u^k v. \tag{3.12}$$

Here, we define u and v as below.

$$u := e^{-ih_1 Z_1 t} e^{-ih_2 Z_2 t} e^{-ih_3 Y_1 Y_2 t}. \tag{3.13}$$

$$v := e^{-ih_4 Z_1 Z_2 kt}. \tag{3.14}$$

Because v commutes with u, the Lie–Trotter decomposition is unnecessary.

In this work, the evolution time is set to $t = \pi/(16h_1)$, leading to $e^{-ih_1 Z_1 t} = e^{-i\frac{\pi}{16}Z_1} = \sqrt{T_1}$ and $e^{-ih_2 Z_2 t} = e^{-i\frac{\pi}{16}Z_2} = \sqrt{T_2}$. Moreover, we approximate $h_3 t/\pi$ and $h_4 t/\pi$ with a 5-bit binary fraction. With this approximation, we obtain $e^{-ih_3 Y_1 Y_2 t} = e^{-0.5000i\pi Y_1 Y_2} = R_{Y_1 Y_2}(\pi/2)$ and $e^{-ih_4 Z_1 Z_2 kt} = e^{-1.5000ik\pi Z_1 Z_2} = R_{Z_1 Z_2}(3k\pi/2)$. These rotations correspond to Clifford operations and can therefore be implemented using only Clifford gates.

The input state $|\Phi_g\rangle$ is

$$|\Phi_g\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle). \tag{3.15}$$

To generate an approximate excited state, we apply the unitary operator

$$U_{\rm ex} = X_2 Z_1. \tag{3.16}$$

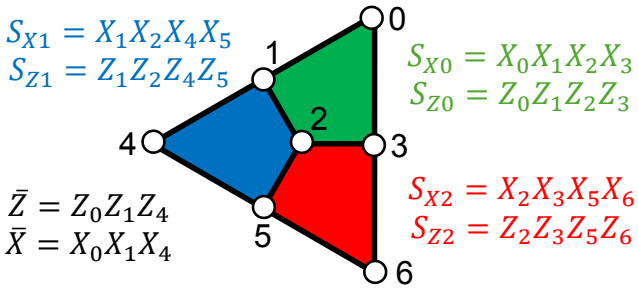


Figure 3.2: Convention of the stabilizer generators and logical operators for the $[\![7,1,3]\!]$ color code in the present work. Physical qubits are indicated by white circles. Logical operators and stabilizer generators are defined as tensor products of single-qubit Pauli operators. For each cell, the four qubits at the vertices define both X-type and Z-type stabilizer generators.

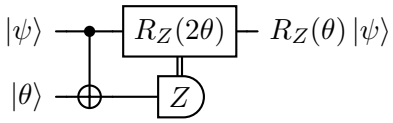
The action of $U_{\rm ex}$ on the ground-state reference yields

$$|\Phi_e\rangle = U_{\rm ex}\,|\Phi_g\rangle \tag{3.17}$$

$$= \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle). \tag{3.18}$$

3.6 Error correction with $[\![7,1,3]\!]$ code

The $[\![7,1,3]\!]$ color code [10, 11] (Steane code) takes 7 physical qubits to encode 1 logical qubit to distance 3, enabling the correction of any single-qubit errors. The generating set of stabilizers and logical operators are given in Fig. 3.2 with the overhead bar indicating the logical operation. In the Steane code, all the Clifford operations can be implemented transversally, whereas the non-Clifford gate, e.g., an arbitrary-angle single-qubit rotation, requires heavy lifting. We use the recursive gate teleportation gadget expressed as: [1]



(3.19)

Gate teleportation is understood by considering the state before the measurement,

$$\frac{1}{\sqrt{2}}R_Z(\theta)\,|\psi\rangle\otimes|0\rangle + \frac{1}{\sqrt{2}}R_Z(-\theta)\,|\psi\rangle\otimes|1\rangle\,. \tag{3.20}$$

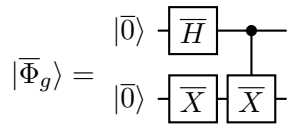
Measuring "0" in the second register (bottom register in circuit (3.19)) results in the desired state, and otherwise, the rotation angle is applied with an opposite sign. Each case happens with probability $1/2$. If the latter happens, this protocol with the rotation angle 2θ is applied to correct the angle to θ. This procedure is repeated until "0" is measured to indicate a successful application of $R_Z(\theta)$. Provided a rotation angle represented by an n_b-bit binary fraction, i.e. $\theta/\pi = b_0 + b_1/2 + \ldots b_{n_b}/2^{n_b-1}$ with $b_i \in \{0,1\}$, the protocol succeeds after at most $n_b - 2$ recursions. This is because $R_Z(2^{n_b-2}\theta) = R_Z(\pi(b_{n_b-1} + b_{n_b}/2))$ is a Clifford gate, which can be applied transversely, and thus does not require gate teleportation.

The encoded QPDE circuit consists of logical state preparation, logical gate operations, QEC cycles, and final measurements. While fully fault-tolerant (FT) circuit constructions suppress logical errors to second order in the physical error rate, they typically require a large circuit overhead. Instead, we adopt a partial fault-tolerant design that introduces minimal error-detection components to reduce the logical error rate while keeping the circuit overhead small. In particular, arbitrary-angle logical rotations $\overline{R}_Z(\theta)$ are implemented using recursive gate teleportation (RGT) with resource states and quantum error detection (QED). This approach avoids the costly procedures of magic-state dis-

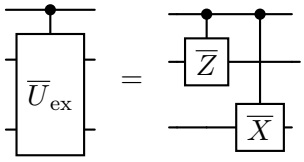
tillation and Clifford+T synthesis required for fully FT implementations, providing an efficient method for realizing logical rotations with arbitrary angles in the encoded QPDE circuit [12].

In the following, we describe the specific circuit implementation used in this work. All logical gates are given in App. A.2. The logical circuit for preparing $|\overline{\Phi}_g\rangle$ is compiled to

$$|\overline{\Phi}_g\rangle = \tag{3.21}$$



The logical circuit component $\overline{\mathrm{ctrl}-U_{\mathrm{ex}}}$ is compiled to



(3.22)

To encode $\overline{U}^k = \overline{u}^k\overline{v}$, we decompose the unitary operator U in Eq. (3.11) into the Clifford+R_Z gate set. Each logical $\overline{R}_Z$ gate is implemented using the RGT gadget $\overline{R}_Z^{\mathrm{RGT}}$.

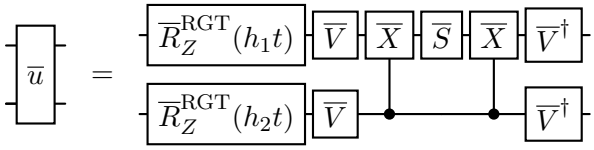
The operation $u = e^{-ih_1Z_1t}e^{-ih_2Z_2t}e^{-ih_3Y_1Y_2t}$ consists of single-qubit Z rotations and a two-qubit Y_1Y_2 rotation. With the chosen evolution time, the single-qubit terms become

$$e^{-ih_1Z_1t} = \sqrt{T_1}, \qquad e^{-ih_2Z_2t} = \sqrt{T_2}, \tag{3.23}$$

while the two-qubit interaction reduces to

$$e^{-ih_3Y_1Y_2t} = R_{Y_1Y_2}(\pi/2). \tag{3.24}$$

Since the rotation angle is $\pi/2$, this operation belongs to the Clifford group. The non-Clifford Z rotations are implemented using the RGT gadget $\overline{R}_Z^{\mathrm{RGT}}$. The circuit representation of $\overline{u}$ is shown in (3.25).



(3.25)

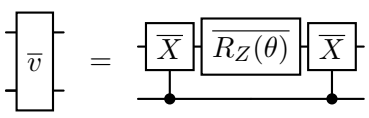
The operation $v = e^{-ih_4Z_1Z_2kt}$ corresponds to a two-qubit Z_1Z_2 rotation

$$e^{-ih_4Z_1Z_2kt} = R_{Z_1Z_2}(3k\pi/2). \tag{3.26}$$

For the values of k considered in this work, the resulting rotation angles are multiples of $\pi/2$,

$$\theta \in \{\pi/2,\ \pi,\ 3\pi/2,\ 2\pi\}. \tag{3.27}$$

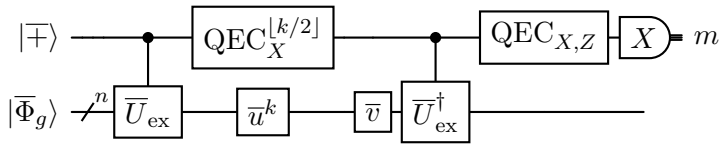
Since these angles correspond to Clifford rotations, the operation belongs to the Clifford group and can therefore be implemented using only Clifford gates such as S, $S^\dagger$ and Z. The circuit representation of $\overline{v}$ is shown in (3.28).



(3.28)

We insert Steane QEC gadgets as follows. After every two applications of $\overline{u}$, we apply an X-type

quantum error correction gadget (QEC_X) to the ancilla qubit. Thus, during the implementation of $\overline{u}^k$, the gadget QEC_X is applied $\lfloor k/2 \rfloor$ times. In addition, after the application of $\overline{U}_{\mathrm{ex}}^\dagger$, both X- and Z-type QEC gadgets ($\mathrm{QEC}_{X,Z}$) are applied to the ancilla. The overall QPDE circuit is given by (3.29).

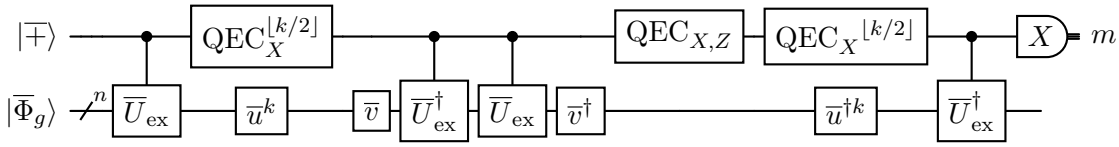


(3.29)

3.7 Calibration

To evaluate the accuracy of the circuit execution, we employ a mirror benchmark circuit. In contrast to the calibration procedure used in [1], where the rotation angle β of $R_Z(\beta)$ can be tuned such that the measurement outcome is always zero, the present setup contains multiple phase differences. Therefore, the calibration cannot be performed by simply adjusting the rotation angle.

Instead, we use a mirror circuit that ideally returns the system to the initial state in the absence of noise. The circuit used for benchmarking is shown in (3.30). Thus, the measurement outcome is expected to be 0 with unit probability in the noiseless case. With hardware noise, however, the probability is reduced.



(3.30)

To model the effect of decoherence, we introduce a parameter q and describe the probability of obtaining the correct measurement outcome as

$$P^{(\mathrm{dec})}(q) = 1 - \frac{q}{2}. \tag{3.31}$$

Since the mirror benchmark runs the time evolution both forward and backward, a circuit labeled by k applies the evolution block u a total of $2k$ times. We therefore model the decoherence parameter as

$$q(k) = 1 - e^{-a(2k)-b}, \tag{3.32}$$

where a approximately represents the decay rate of the signal per U application.

Using this model, we analyze the encoded circuits both with and without Steane QEC gadgets, expecting that it captures the dominant influence of physical and logical errors on the circuit outcome.

3.8 Experiment of Circuit Calibration

To evaluate the utility of the Steane QEC gadgets and to compare the accuracy with and without encoding, we benchmark the circuit fidelity under the following three conditions: (i) with Steane QEC gadgets, (ii) without Steane QEC gadgets, and (iii) unencoded logical circuit. For each condition, we fit the estimated success probability with $P^{(\mathrm{dec})}$ to provide a quantitative estimate of the circuit infidelity.

3.8.1 Experimental Setup

Devices.

The QEC-encoded circuits are executed on Quantinuum System Model H2-2 [13, 14], Quantinuum Helios-1 [14, 15], and Quantinuum Helios-1E [14, 15]. The logical circuits are executed on Quantinuum H2-2E [13, 14]. Details of the device noise parameters for each system are summarized in App. A.3.

Dynamical Decoupling.
To suppress coherent memory errors during idle periods, Dynamical Decoupling (DD) [16] is applied in all experiments. On System Model H2-2 and H2-2E, DD pulses are inserted automatically by the hardware compiler: whenever a qubit remains idle for a duration exceeding a threshold time, the compiler opportunistically inserts DD pulses to decouple accumulated coherent errors [17]. The threshold time was set to 0.03 seconds in all H2-2 and H2-2E experiments. On Helios-1 and Helios-1E, DD pulses were inserted manually at the logical circuit level, as automatic DD compilation is not yet available on that system. The timing is set to align with the state preparations.

Sampling settings for the calibration circuit.
For each experimental configuration and exponent k in the unitary U^k ($k = 1$–5), we collect 500 shots of the QPDE calibration circuit shown in (3.30) to estimate the probability of obtaining the zero-measurement outcome $m = 0$. For experiments performed on Quantinuum Helios-1, we report the actual number of completed shots as well as the number of unitary exponents k considered.

3.8.2 Results
The calibration results are shown in Fig. (3.3). Data points without a suffix correspond to circuits with Steane QEC gadgets; those labeled "(NoQEC)" correspond to circuits without QEC gadgets; and those labeled "(L)" correspond to the logical circuit. The dashed curves are fits to Eq. (3.32), and the extracted fitting parameters a and b are summarized in Tab. 3.1.

3.8.3 Discussion
Comparison between Helios-1E and H2-2.
As discussed in [1], the most influential noise channel is the memory error in the present type of experiments. There are two main strategies to suppress this: reducing idling time and applying DD. Helios has a real-time engine that can dynamically schedule the ion transport. Our circuit has many repeat-until-success (RUS) operations, and the Helios stack can literally execute them, whereas the H2 stack redundantly transports the ions if the RUS completes before its upper bound. On the other hand, in the Helios stack, automatic DD pulse insertion is not available at the time these experiments are conducted, and thus coherent memory noise suppression relies on manual implementation. The difference between H2 and Helios generations are interpretted from these two aspects, among others. Fig. 3.3 suggests that these two effects are competing with each other to show similar q. We expect that once automatic DD compilation becomes available on Helios-1E, its lower physical noise rate should translate into improved circuit fidelity compared to H2-2.

Effect of Steane QEC gadgets.
On H2-2, a comparison between the with QEC gadgets and without QEC gadgets reveals a crossover behavior as a function of circuit depth k. For $k \geq 3$, the encoded circuit with QEC gadgets exhibits a smaller decoherence effect than the circuit without QEC, indicating that error correction provides a net benefit at larger circuit depths. In contrast, for $k = 1$ and $k = 2$, the without QEC circuit outperforms the encoded one, as the overhead associated with the QEC gadgets exceeds the benefit of error suppression at shallow depths. On Helios-1E, there is a marginal difference between the with-QEC and without-QEC cases. It may be due to the remaining coherent error, which causes a phase-flip error with a probability that is quadratically dependent on the idling time.

Comparison with the unencoded logical circuit.
While the goal of our top-down approach here is to implement the end-to-end chemistry workflow for benchmarking, it would be worth comparing the encoded and unencoded cases to clarify the current status. At present, as in the other similar work in literature [1, 18], the QEC-encoded implementation does not achieve the performance of the logical circuit. This is primarily due to the accumulation of errors beyond the correction capability of the current distance-3 code, particularly memory noise. During the encoding, decoding, and error-correction processes, the ions experience extended idle

periods. Memory errors accumulated over these intervals increase the logical error rate beyond the physical baseline.

Recall that break-even has been demonstrated at the level of individual QEC gadgets on current devices [19–24], realizing effective error correction at the full-circuit level, including many non-Clifford operations, remains challenging. Achieving break-even at the application circuit level is therefore a key objective for future work, which will likely require a combination of techniques to suppress memory errors and the use of higher-distance error-correcting codes, together with hardware advancements.

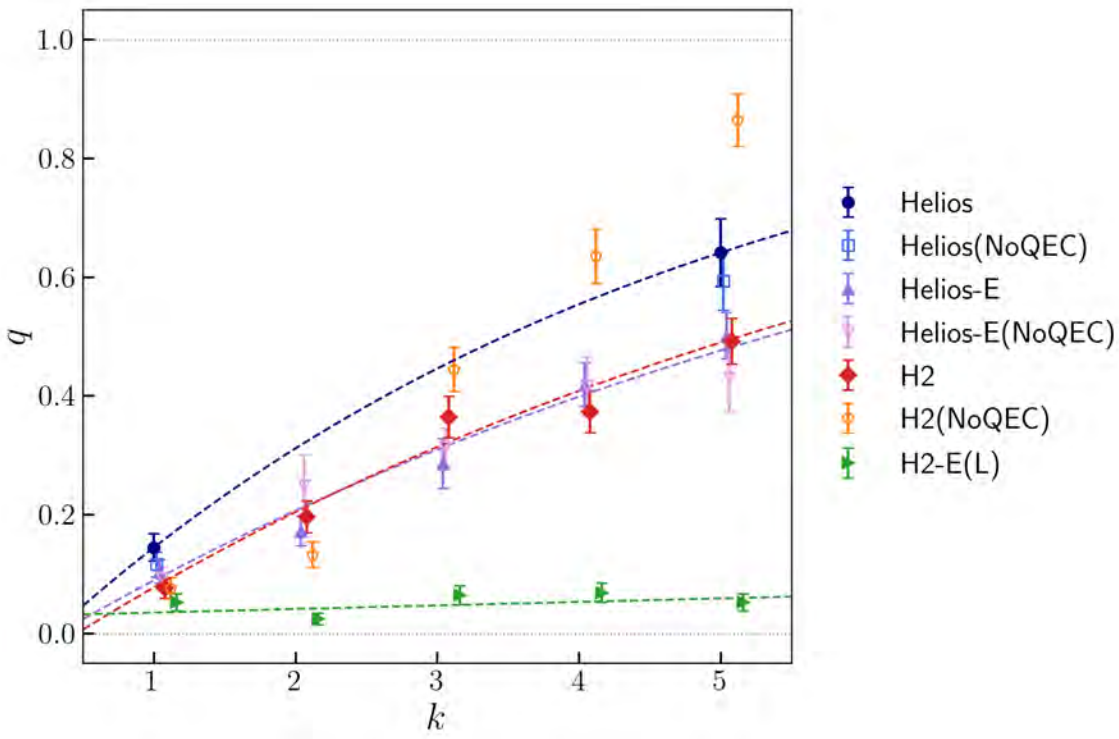


Figure 3.3: Decoherence parameter q as a function of circuit depth k, obtained by fitting the probability of the zero-measurement outcome from the QPDE benchmark circuits to $P^{(\mathrm{dec})}$ (Eq. (3.31)). Data points without a suffix correspond to circuits with Steane QEC gadgets; "(NoQEC)" denotes circuits without QEC gadgets; "(L)" denotes the logical circuit. Dashed curves show fits to Eq. (3.32). In the ideal (noiseless) case, $q = 0$.

Device	a	b
Helios	$0.1086(201)$	$-0.0613(520)$
Helios-E	$0.0692(71)$	$-0.0452(303)$
H2	$0.0741(67)$	$-0.0680(270)$
H2-E (L)	$0.0031(23)$	$0.0293(141)$

Table 3.1: Fitting parameters a and b in Eq. (3.32) extracted from the calibration data shown in Fig. 3.3.

3.9 Experiment of QPDE

We proceed to the phase difference estimation following Sec. 3.5. To assess the performance of QPDE under realistic noise conditions, we perform experiments using the noise-aware likelihood model calibrated in Sec. 3.8.

3.9.1 Experimental Setup

Devices.

The QEC-encoded circuits are executed on Quantinuum System Model H2-2 [13, 14]. Details of the device noise parameters are summarized in App. A.3.

Sampling strategy.

We randomly draw $N_s/10$ samples of (k, β) with a maximum circuit depth $k_{\max} = 10$. For each sampled pair (k_l, β_l), we execute the circuit 10 shots, resulting in a total of N_s measurement outcomes.

Noise-aware likelihood.

To incorporate hardware noise effects in the classical post-processing, we use the noise-aware distribution

$$Q(m_l|\tilde{\phi}; k_l, \beta_l) = \frac{1 + e^{-ak_l - b}\cos(k_l\tilde{\phi} + \beta_l - m_l\pi)}{2}, \tag{3.33}$$

instead of Eq. (3.9). Here, the parameters a and b are obtained from the circuit calibration in Sec. 3.8.

Estimation procedure.

From the experimentally obtained measurement outcomes $\{m_l\}_{l=1}^{N_s}$, we construct the posterior distribution over the phase $\tilde{\phi}$ using Eqs. (3.8) and (3.33). To estimate the phase and its uncertainty, we employ a bootstrap procedure: we draw N_s samples with replacement from the dataset, compute the phase estimate $\phi^{(i)}$ using Eq. (3.10), and repeat this procedure R times to obtain $\{\phi^{(i)}\}_{i=1}^{R}$. The phase is then estimated by the circular mean $\mathrm{Arg}(\sum_i e^{i\phi^{(i)}}/R)$, and the uncertainty is quantified using the Holevo variance.

Target phase.

The target quantity considered in this work is the energy difference at the conical intersection. Let E_{g} and E_{e} denote the ground- and excited-state energies, respectively. We define the energy difference as $\Delta E := E_{\mathrm{g}} - E_{\mathrm{e}}$, and the corresponding phase difference as $\Delta\phi := -\Delta E\, t$.

3.9.2 Results

The experimentally obtained distributions over the phase $\tilde{\phi}$ for different numbers of samples N_s are shown in Fig. 3.4. As N_s increases, the dominant peak of the distribution becomes sharper, concentrating around the target phase $\Delta\phi$. This demonstrates that increasing the number of samples improves the precision of the phase difference estimation, and consequently the accuracy of the energy gap estimation.

With $R = 1000$ bootstrap samples and $N_s = 1400$ circuit samples, the energy diference is estimated as $\Delta E = -0.0025561(24)$ Ha, which agrees with the approximation energy differnce $\Delta E_{\mathrm{APPROX}} = -0.002559$ Ha with the statistical uncertainty of σ.

3.10 Numerical simulations with tunable noise parameters

We investigate the impact of hardware noise on the partially fault-tolerant QPDE circuit through numerical simulations with tunable noise parameters. All simulations in this section are performed using the H2 emulator.

The noise sources considered in the simulations are summarized in Tab. 3.2. The two-qubit gate fault is modeled by an anisotropic two-qubit Pauli channel. Readout faults flip a measurement outcome from “1” to “0” with probability $p_{\mathrm{r},1\to0}$ and from “0” to “1” with probability $p_{\mathrm{r},0\to1}$. Memory noise acting on idling or transported qubits is modeled by the dephasing channel

$$\rho \mapsto e^{-ift_{\mathrm{m}}Z/2}\big((1 - gt_{\mathrm{m}})\rho + gt_{\mathrm{m}}Z\rho Z\big)e^{ift_{\mathrm{m}}Z/2}, \tag{3.34}$$

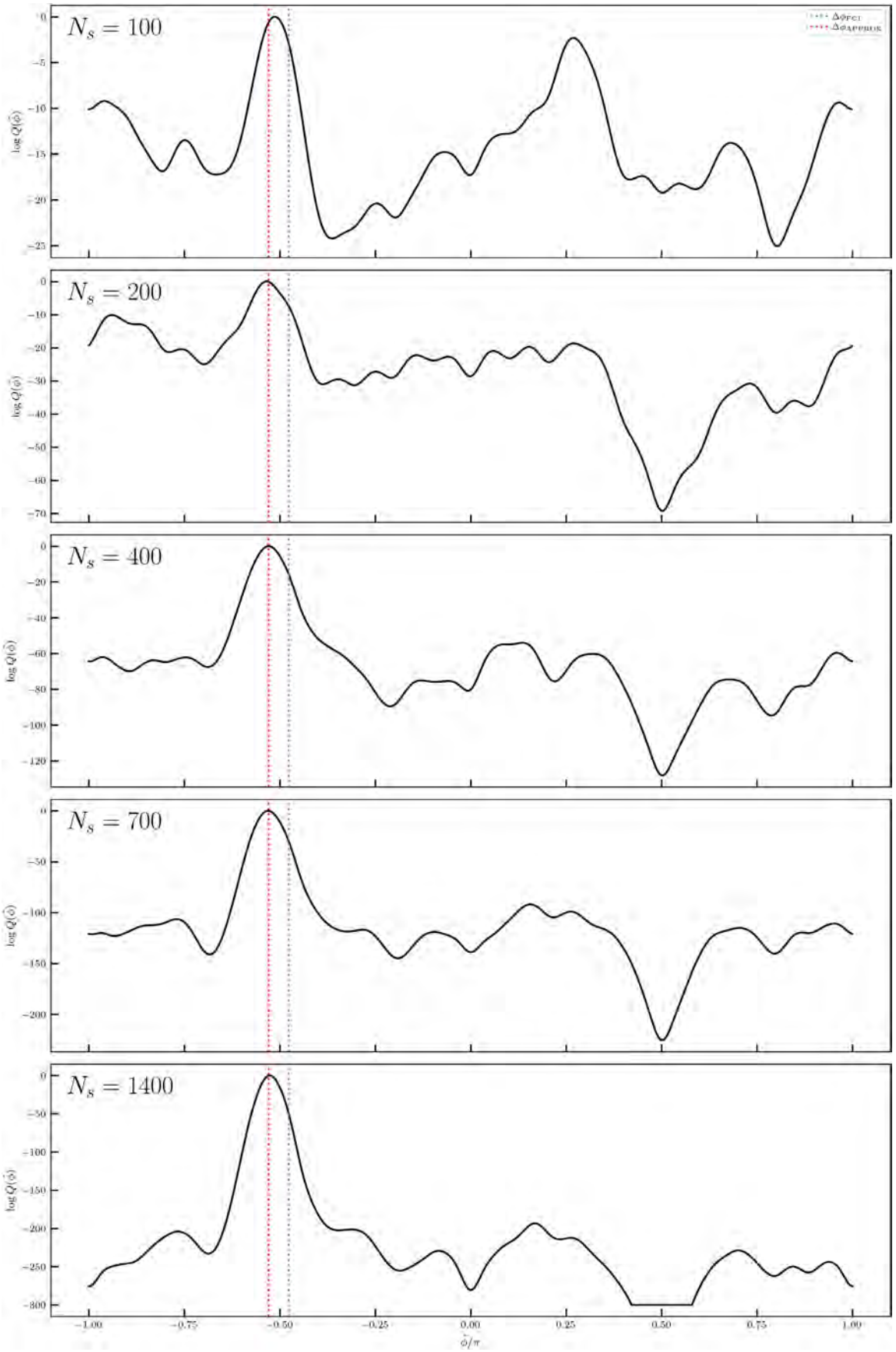


Figure 3.4: The distributions [Eqs. (3.8) and (3.33)] obtained by experimentally collecting the measurement outcomes of the circuit (3.29) for different numbers of samples N_s. The target phase $\Delta\phi$ computed from the Hamiltonian is shown as a gray dashed line, while the target phase including Trotter error is shown as a red dashed line.

where t_{m} denotes the idling or transport time. The parameters f and g characterize the coherent and incoherent components of memory noise, respectively. Further details of the noise model are given in [25].

To identify the dominant noise source in the partially fault-tolerant setup, we simulate the encoded QPDE circuits while activating each noise component separately (Tab. 3.2). The H2-2 quantum computer is emulated with the parameters $p_2 = 1.29 \times 10^{-3}$, $p_{\mathrm{r},0\to1} = 0.9 \times 10^{-3}$, $p_{\mathrm{r},1\to0} = 1.8 \times 10^{-3}$, $f = 4.3 \times 10^{-2}$ [rad· s^{-1}], and $g = 2.8 \times 10^{-3}$ [s^{-1}] [25].

Tab. 3.3 shows that incoherent memory noise and gate readout error are the dominant error source under the present setup. The encoded circuit involves long idling periods caused by nested conditional operations in the RGTs and QEC gadgets, which enhances the impact of memory noise. In contrast, coherent memory noise remains well suppressed even at the current value $f = 4.3 \times 10^{-2}\,\mathrm{rad}\cdot\mathrm{s}^{-1}$. This suppression is attributed to the use of dynamical decoupling (DD), which mitigates the accumulation of coherent phase errors. Fig. 3.5 illustrates the decoherence parameter q obtained from the $k = 3$ and $k = 5$ mirror benchmark circuits for each noise model.

Noise model		Parameter
Two-qubit gate fault		p_2
Readout fault	(“1”→“0”)	$p_{\mathrm{r},1\to0}$
	(“0”→“1”)	$p_{\mathrm{r},0\to1}$
Memory noise	coherent	f
	incoherent	g

Table 3.2: Noise sources included in the numerical simulations and their associated parameters.

Noise type	$q(k=3)$	$q(k=5)$
Gate & readout error	0.136	0.224
Coherent memory noise	0.120	0.116
Incoherent memory noise	0.104	0.160

Table 3.3: Simulation results of the encoded circuit with different noise sources activated individually. The first row uses $p_2 = 1.29 \times 10^{-3}$, $p_{\mathrm{r},0\to1} = 0.9 \times 10^{-3}$, and $p_{\mathrm{r},1\to0} = 1.8 \times 10^{-3}$. The second and third rows use the coherent memory noise parameter $f = 4.3 \times 10^{-2}$ [rad· s^{-1}] and the incoherent memory noise parameter $g = 2.8 \times 10^{-3}$ [s^{-1}], respectively. All other noise sources are turned off in each setup.

3.11 Discussion and Outlook

This work builds on the QEC-assisted QPE pipeline [1], implemented on the Quantinuum H2 trapped-ion processor with the $[\![7,1,3]\!]$ color code (Steane code). We introduced two update points: (1) replacing QPE with iterative quantum phase difference estimation (IQPDE), and (2) expanding the logical system register from one to two qubits. Three implementations were benchmarked—logical with QEC, logical without QEC (NoQEC), and unencoded physical—to evaluate QEC effectiveness in this setting.

The logical QEC implementation achieved higher fidelity than NoQEC, confirming that QEC gadgets improve circuit performance. This is consistent with the calibration results of the original work [1]. However, the physical implementation outperformed the logical one, and break-even with respect to the physical baseline was not reached as in the previous work [1, 18].

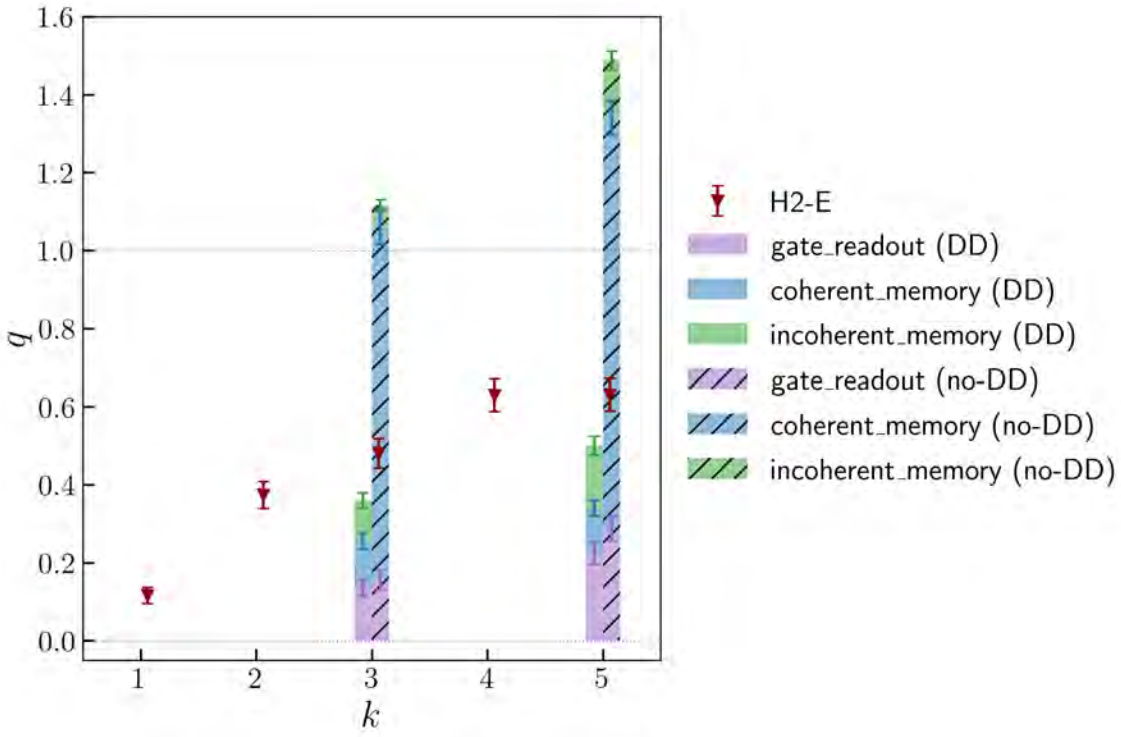


Figure 3.5: Decoherence parameter q obtained from the $k = 3$ and $k = 5$ mirror benchmark circuits for different noise models using the H2 emulator. The dark red markers represent the H2 emulator backend with dynamical decoupling enabled. (DD) and (noDD) denote the noise model results with each noise channel activated, with and without dynamical decoupling, respectively.

This contrasts with existing QEC demonstrations, where break-even has been widely achieved at the component level. On the same trapped-ion platform, physical-level fidelity has been surpassed for logical CNOT and SPAM [19], Bell state preparation [20], cat state preparation [21], and logical qubit teleportation [22]. The same trend holds across other qubit modalities: the Google Willow superconducting processor benchmarked surface-code memory and demonstrated that the logical error rate is lower than that of the best physical qubit [23], and the Harvard/QuEra neutral-atom group demonstrated logical advantage in a 48-logical-qubit sampling circuit [24]. In these cases, break-even is confined to low-depth circuits or component benchmarks. Break-even has been established at the component level, but has not been demonstrated at the algorithmic level, mainly due to the heavy-lifting non-Clifford operations (arbitrary-angle single-qubit rotation in the present work). The present work benchmarks QEC performance at the algorithm level using IQPDE, an algorithm for energy gap estimation.

The dominant factor behind this shortfall is the accumulation of errors beyond the corrective capacity of the code, particularly memory errors. During encoding, decoding, and error correction, ions experience extended idle periods. Memory errors accumulating during these intervals drive the logical error rate above the physical baseline.

This work extends QEC-assisted energy estimation [1] to energy gap estimation, providing a reference point for the practical utility and current limitations of QEC at the algorithm level. Revisiting this experiment on next-generation hardware will serve as a natural benchmark for progress toward algorithm-level fault tolerance.

We note that the present chemistry pipeline, along with the use cases (e.g., an optical switch), will remain virtually unchanged even after fault-tolerant quantum computing is achieved. The target molecular system will grow larger as the number of logical qubits we can use increases. The number of logical qubits will remain modest until we reach sufficient logical fidelity for chemistry analysis, as the additional qubits would be allocated to further protection of the logical qubits in the early stage. We will continue developing the software components to keep pace with advances across

all relevant software/hardware layers. In particular, the development of the error correction code is rapidly proceeding. Our pipeline is designed to incorporate such a new QEC code without changing the overall workflow. In the later stage, most likely in the Apollo generation, we start expanding the system size to tackle the industry-relevant problems.

4 Quantum Computing in Topological Data Analysis

4.1 Foundations of Topological Data Analysis

Topological Data Analysis (TDA) is an approach to data analysis guided by the principle: *data has shape, shape has meaning, and meaning produces value*. Rather than treating data merely as collections of numerical values, TDA interprets data as objects possessing an intrinsic "shape," and seeks to extract essential structural information from this geometric and topological perspective. The central feature of TDA is the use of *topological invariants*. These quantities remain unchanged under continuous deformations such as stretching or bending, and are therefore robust against noise and small perturbations in the data, when accommodated for in the 'discretization' step (for example, as is done for persistence diagrams [26]). This robustness arises because topological invariants capture global correlations within a dataset that are invariant to certain local transformations. They are insensitive to local coordinate transformations or minor fluctuations, and instead describe the fundamental large-scale structure of the data. At the same time, precisely because they encode global structure, topological descriptors are 'coarse-grained' properties. They do not preserve detailed local information, and therefore may overlook fine-grained structural differences. This dual nature—robust yet coarse—constitutes both a strength and a limitation of TDA.

TDA has been applied to a variety of use-cases across many scientific disciplines. One historically representative example is the identification of glassy states [27–29]. Structurally, glass and liquid states can appear remarkably similar at the level of local molecular configurations, making them difficult to distinguish using conventional measures. However, when TDA is applied to such systems, differences in topological invariants become apparent, enabling discrimination between these states. This example illustrates how TDA can detect structural distinctions that are not readily visible through traditional local or statistical analyses (see [28, 29]).

Early developments in TDA primarily focused on unsupervised, exploratory tasks in which structural differences were identified through comparisons of topological invariants (see [30]). In this setting, the goal was to uncover intrinsic organization within the data without relying on labeled examples. More recently, TDA has evolved toward integration with modern machine learning frameworks. Two major directions have emerged:

- incorporating topological invariants directly into learning objectives, for example as loss terms or regularizers (see [31–34]), and
- vectorizing topological summaries (such as persistence-based descriptors) and using them as input features in supervised learning models (see [35–39]).

In this work, we focus on the second of these approaches.

4.2 Limitations of Classical TDA and Quantum Advantage

A natural starting point for investigating the usefulness of TDA primitives for real-world data is looking at Betti numbers of clique complexes as simple global descriptors of network/graph data. Betti numbers are topological invariants that summarize global structural properties such as the number of connected components and higher-dimensional holes. These quantities capture correlations that are often invisible to purely local statistics but can be very computationally expensive to calculate, particularly in high dimensions.

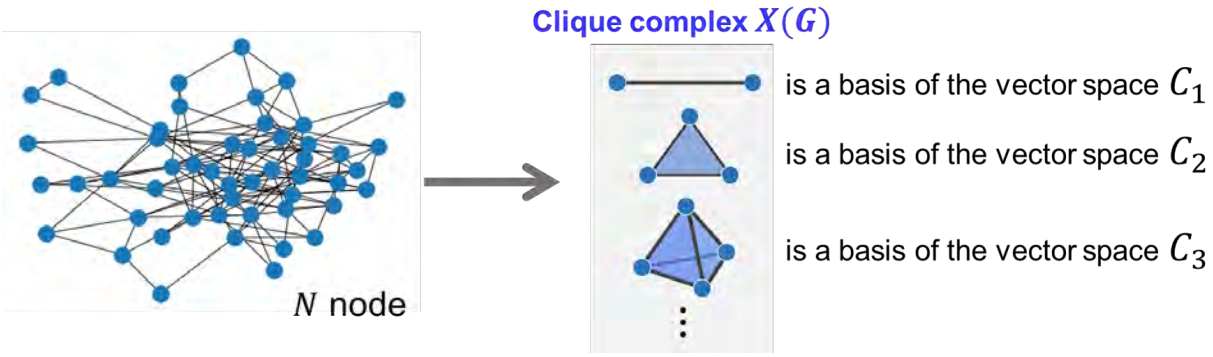


Figure 4.1: Conceptual diagram of a clique complex

4.2.1 Exact Computation and Its Limitations

To illustrate why exact Betti-number computation does not scale well classically, we consider a highly structured example with analytical properties: the complete k-partite graph. This follows the example set by Berry et al. [40]. A graph $G = (V, E)$ is called a *complete k-partite graph* if its vertex set admits a decomposition

$$V = V_1 \sqcup V_2 \sqcup \cdots \sqcup V_k,$$

such that no edges exist within each partition V_i, and every pair of vertices belonging to different partitions is connected by an edge. For simplicity, we assume equal partition sizes, $|V_i| = n \quad (i = 1, \ldots, k)$.

Given a graph G, its *clique complex* $X(G)$ is the simplicial complex whose simplices correspond to cliques (complete subgraphs) of G. The vertices of $X(G)$ are the vertices of G, and a clique of size $(r + 1)$ corresponds to an r-simplex. In a complete k-partite graph, any clique is formed by selecting at most one vertex from each partition, which makes the combinatorial structure of $X(G)$ explicit (see Fig. 4.1).

Let S_r denote the set of r-simplices of the clique complex of a complete k-partite graph, where $r \in [0, k - 1]$. The number of r-simplices is

$$|S_r| = \binom{k}{r+1} n^{r+1}.$$

In particular, $|S_{k-1}| = n^k$, so the size of the clique complex scales as $\mathcal{O}(n^k)$.

Betti numbers are computed via linear-algebraic operations on chain groups. For each dimension r, one considers the chain group $C_r = \mathrm{span}(S_r)$, where each simplex is formally associated with an independent unit vector. Betti numbers are then defined in terms of the rank of the associated boundary operators. Thus, the computational bottleneck is governed directly by the dimension of C_r and the resulting matrix sizes (see the left panel of Fig. 4.2).

In practice, many algorithmic improvements exist, including sparse-matrix techniques and optimized reduction procedures. In particular, `flagser` [41] is a highly optimized implementation specialized for computing Betti numbers of clique complexes of graphs and is among the most efficient classical tools currently available. Nevertheless, the fundamental combinatorial growth in the number of simplices remains unavoidable (see the right panel of Fig. 4.2).

As k increases, memory usage grows rapidly and computation time increases sharply. Beyond a certain threshold, the computation becomes infeasible due to memory exhaustion or excessive runtime. Therefore, even in this structurally simple setting—where simplex counts are known exactly—classical explicit computation of Betti numbers faces intrinsic scalability limitations.

Moreover, from a complexity-theoretic perspective, computing Betti numbers exactly is known to be #P-hard in general [42]. This further confirms that large-scale exact evaluation is fundamentally intractable.

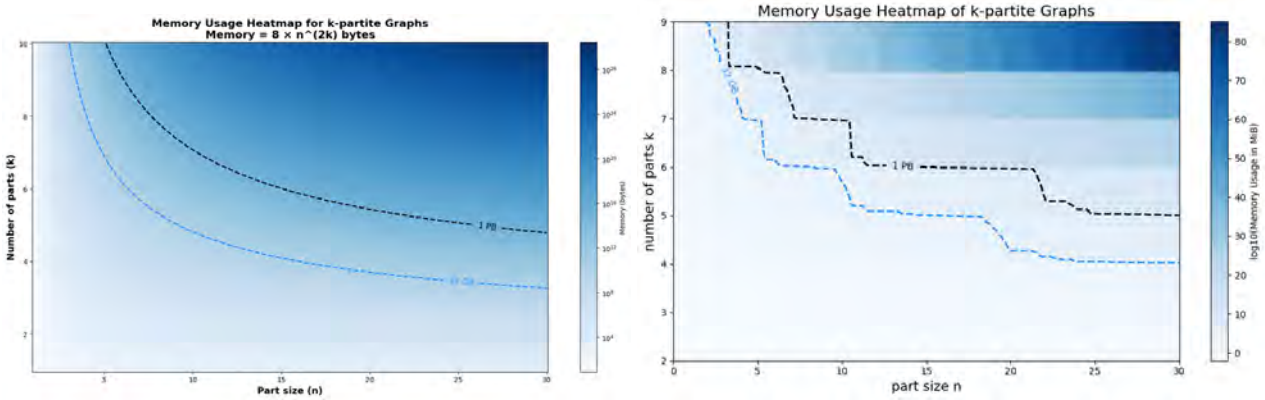


Figure 4.2: (Left) The dimension of the highest-dimensional vector space in the clique complex of a k-partite graph is n^k. A matrix acting on this space therefore has n^{2k} entries. If stored densely, the required memory usage would be $8n^{2k}$ bytes (assuming double-precision representation). Note that in actual computations the matrix is not stored in dense form. (Right) Memory usage when computing all Betti numbers of the clique complex of a k-partite graph using `flagser` [41]. Values exceeding 8GB are extrapolated and should therefore not be regarded as exact measurements.

4.2.2 Approximation Problems

Given the hardness of exact computation, we consider approximation. In particular, we discuss both multiplicative and additive approximation schemes.

Multiplicative Approximation. For a Betti number β_k, an estimate $\hat{\beta}_k$ is called an ε-multiplicative approximation if

$$(1-\varepsilon)\beta_k \;\leq\; \hat{\beta}_k \;\leq\; (1+\varepsilon)\beta_k.$$

However, multiplicative approximation of Betti numbers is NP-hard in general [42].

Additive Approximation. Instead, we consider the additive approximation of the normalized Betti number

$$\overline{\beta}_k = \frac{\beta_k}{\dim C_k}.$$

An estimate $\tilde{\beta}_k$ is called an ε-additive approximation if

$$\left|\tilde{\beta}_k - \overline{\beta}_k\right| \leq \varepsilon.$$

The computational complexity of this additive approximation problem for Betti numbers remains an open problem. However, since homology decision problems for clique complexes are known to be QMA-hard, it is conjectured that this problem is also QMA-hard [43, 44]. In other words, even additive approximation is unlikely to be efficiently solvable, even with quantum computation.

To proceed, we introduce the k-th Hodge Laplacian. Let C_k denote the k-chain space and $\partial_k : C_k \to C_{k-1}$ the boundary operator. The k-th Hodge Laplacian is defined as

$$\Delta_k = \partial_{k+1}\partial_{k+1}^\dagger + \partial_k^\dagger \partial_k.$$

A key property of this operator is that the multiplicity of the zero eigenvalue coincides with the Betti number β_k. Thus, computing Betti numbers can be viewed as a spectral problem.

Since counting exact zero eigenvalues is computationally difficult, we instead introduce the notion of *quasi-Betti number* [43]. For a threshold λ, we define quasi-Betti number as the number of eigenvalues smaller than λ. For sufficiently small λ, this provides a good approximation of the Betti number.

In this way, we relax the problem from counting exact zero modes to estimating the number of low-lying eigenmodes under additive error. This problem is known to be DQC1-hard for general simplicial complexes [43]. While the precise complexity for clique complexes remains open, it is expected to be similarly hard.

Therefore, instead of counting strictly zero eigenvalues, we consider the normalized number of eigenmodes within a small low-eigenvalue window. Under this formulation, the problem is expected to become DQC1-hard, making it intractable for classical computation while remaining potentially efficiently solvable by quantum algorithms.

4.2.3 Normalized Laplacian moments (LM)

Laplacian moments play a central role in the practical implementation of this problem. They provide a natural framework for realizing quasi-Betti quantities within a quantum algorithm. The quasi-Betti number is defined via a spectral threshold λ, effectively counting eigenvalues below λ. This corresponds to a discontinuous function of the spectrum, which is not directly amenable to efficient quantum implementation. To overcome this difficulty, we replace the discontinuous threshold function with a polynomial approximation. Laplacian moments arise precisely from this approximation, enabling an efficient and physically realizable quantum procedure.

We define the normalized k-th Hodge Laplacian as

$$\overline{\Delta}_k = \frac{1}{\Lambda}\Big(\partial_{k+1}\partial_{k+1}^{\dagger} + \partial_k^{\dagger}\partial_k\Big),$$

where Λ is chosen so that $\mathrm{Spec}(\overline{\Delta}_k) \subset [0,1]$.

The normalized Laplacian moment is defined as

$$\overline{T}_k^{(d)} = \frac{1}{\dim C_k}\mathrm{Tr}\big(I - \overline{\Delta}_k\big)^d.$$

Let $\{\lambda_i\}$ be the eigenvalues of $\overline{\Delta}_k$. Since the number of zero eigenvalues equals β_k, we obtain

$$\overline{T}_k^{(d)} = \overline{\beta}_k + \frac{1}{\dim C_k}\sum_{\lambda_i>0}(1-\lambda_i)^d.$$

Hence,

$$\lim_{d\to\infty}\overline{T}_k^{(d)} = \overline{\beta}_k.$$

In the following, we demonstrate that quantum advantage emerges in estimating such approximate Betti numbers under additive error, by analyzing the sample complexity of concrete algorithms. Detailed derivations and simulation results are presented in Ref. [45].

Even though we move to Laplacian moments for calculation purposes, in that we can approximate the normalized Betti number as d increases, this project investigates the usefulness of the vector of Laplacian moments for values of d from one to some manageable number. This move naturally allows the Laplacian moments to contain information that is beyond just the invariant topological information and includes geometric information present in the moments of the Laplacian (see subsection 4.4.2). Thus, not only does the Laplacian moment approach address the exact computational limitation outlined above but also the coarseness limitation introduced in Section 4.1.

Classical Monte Carlo Complexity To approximate the normalised Betti number within accuracy ε, one requires

$$d \geq \frac{\log(2/\varepsilon)}{\delta} \quad \Rightarrow \quad (1-\delta)^d \leq \frac{\varepsilon}{2},$$

where δ denotes the spectral gap.

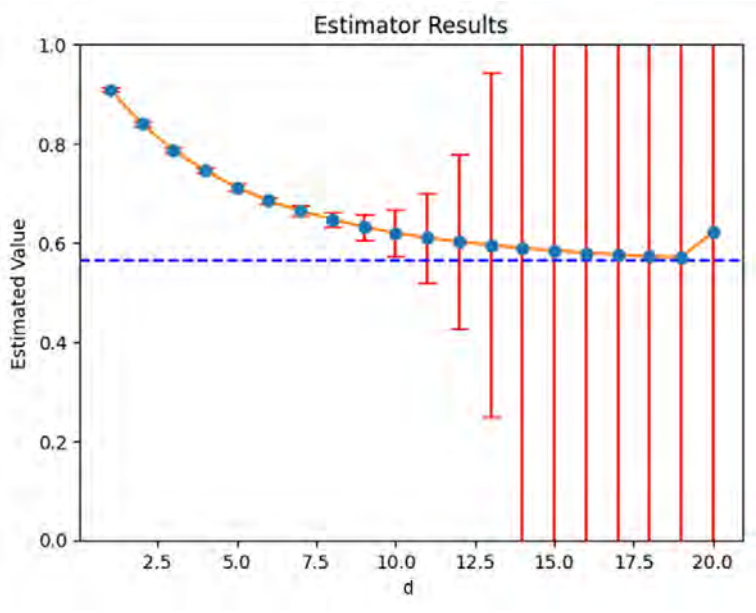


Figure 4.3: As shown in Section 4.2.1, a 6-partite graph with 11 nodes per partition lies at the boundary of feasibility for supercomputer-based exact computation. This figure shows the result of computing the same 6-partite graph (with 11 nodes per partition) using the classical sampling algorithm for Laplacian moments. The error bars in this plot are derived from the Hoeffding bound used by Apers et al. [46]. We observe that the empirical variance for this algorithm does not attain this bound in this case. For more on this question see [47]

In classical trace estimation, the number of required samples scales as

$$\sim \frac{2^{d+1}\log(2/\eta)}{\varepsilon^2},$$

and total matrix queries scale as

$$\sim \frac{2^{d+1}\log(2/\eta)}{\varepsilon^2} \times d.$$

Thus,

$$\text{Classical complexity: } \mathcal{O}(2^d).$$

While classical Monte Carlo simulation suffers from an exponential increase in the number of required samples, it can still provide approximate results, unlike exact computation, which becomes infeasible at scale.

Quantum Sampling Complexity We conclude by introducing the sampling cost of the quantum algorithm; a detailed description of the algorithm itself will be provided in the next section. In contrast, quantum sampling requires

$$\sim \frac{\log(2/\eta)}{\varepsilon^2}$$

samples, independent of d, and total query complexity

$$\sim \frac{\log(2/\eta)}{\varepsilon^2} \times d.$$

Hence,

$$\text{Quantum complexity: } \mathrm{poly}(d).$$

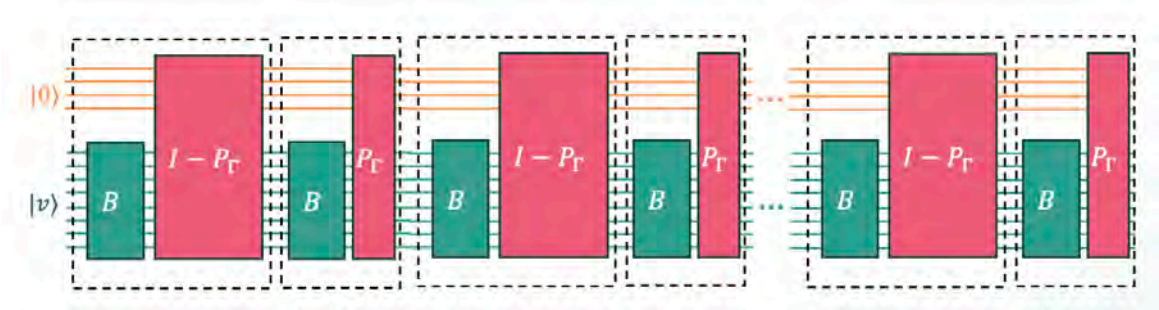


Figure 4.4: Circuit diagram demonstrating the general structure of the circuit for estimating $\overline{T}^{(d)}_{v,k}$. Each dotted box represents a "block".

4.3 Quantum circuits and resource estimates for computing Laplacian moments

In this section, we discuss a quantum algorithm used to generate the normalized Laplacian moments and provide resource estimates for three regimes in which this algorithm could be run on quantum hardware. The algorithm we focus on here is based on Algorithm 6 described in Akhalwaya et al. [45].

4.3.1 Algorithm structure

The algorithm we consider produces a vector of estimates for $\overline{T}^{(d)}_k$ for each $d = 1, 2, \ldots D$ where D is a number given as input to the algorithm. Similar to the classical algorithm of Apers et al. [46], the algorithm produces these estimates using a core subroutine which estimates, for an input k-clique v, the quantity $\overline{T}^{(d)}_{v,k} = \langle v| (I - \overline{\Delta}_k)^d |v\rangle$ for each d in the chosen range. This subroutine is leveraged to compute normalized Laplacian moments by observing that $\overline{T}^{(d)}_k = \mathbb{E}_v[\langle v| (I - \overline{\Delta}_k)^d |v\rangle]$ where v is a k-clique of G chosen uniformly at random.

Sampling simplices As in [46], we will assume access to an efficient sampling algorithm for the k-cliques of our input graph. Such an assumption is empirically reasonable on the small graph sizes we deal with in this investigation. Furthermore, the proof of Cade and Crichigno that quasi-Betti number estimation on general simplicial complexes is DQC1-hard holds in the presence of an efficient oracle for sampling from the simplices [43].

Circuits for $\overline{T}^{(d)}_{v,k}$ Given a k-clique v of G we run the two circuit blocks C_{D_0} and C_{D_1} which block-encode matrices D_0 and D_1 such that $I - \overline{\Delta}_k = D_0 D_1^\dagger$. The query counts discussed in the previous section refer to the number of uses of either of these circuits. These circuit blocks, seen in Figure 4.4, each consist of a unitary B (the Hermitian boundary operator) and a projection P_{Γ_G} or $I - P_{\Gamma_G}$ which consists of entangling operations between the main register and some ancilla qubits followed by measurements and post-selection on a particular set of results. The output of the algorithm is a random variable τ_v which is the number of blocks that succeed in a row (that is the post-selection succeeds up until the first failure). It is shown in [45] that $\mathbb{E}[\tau_v \geq d] = \overline{T}^{(d)}_{v,k}$. For more detail on circuits, see Appendix B.

4.3.2 Quantum resource estimates

Given the descriptions of the components in the quantum Laplacian moments algorithm, we can estimate the quantum resources required to run a single shot of this algorithm on given graph inputs.

These resource estimates allow us to better understand the scope for using quantum computers to explore Laplacian moments on graphs which are beyond the reach of existing classical methods. To this end, we look at resource counts in three regimes: the first corresponding to validating this algorithm's performance on hardware, the second towards demonstrating quantum advantage with

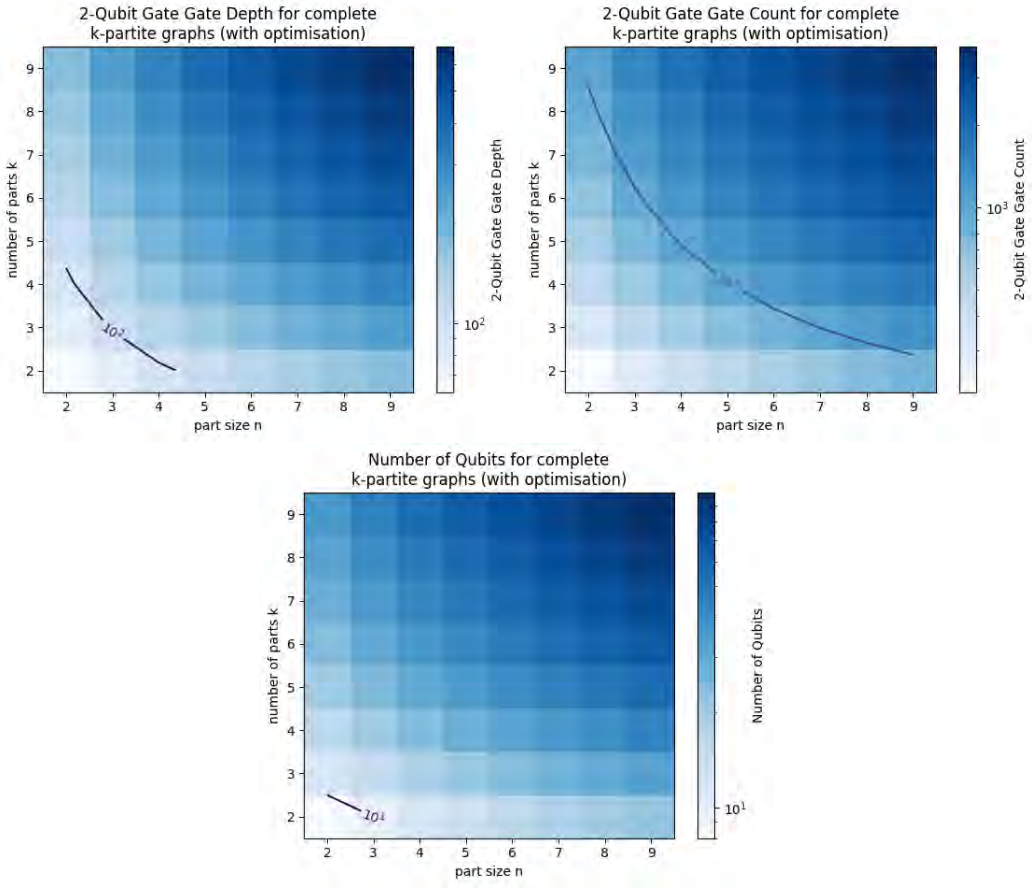


Figure 4.5: Resources required for running one shot of Laplacian moments algorithm (up to $\overline{T}_k^{(15)}$) on complete k-partite graphs. These circuits in question are optimised to exploit this k-partite structure.

the algorithm and the third corresponding to searching for advantage in real-world data on early fault tolerant quantum computers.

Regime 1: Validating the algorithm on complete k-partite graphs On the example of complete k-partite graphs which are defined above and studied by Berry et al. [40], we can optimise the projections to get resource counts that are amenable to running on current hardware. The ground truth for the Laplacian's is known exactly for these graphs so running the algorithm here would not be a demonstration of quantum advantage. However this extra structure allows us to test the functioning of the algorithm in the presence of noise and develop noise mitigation strategies. In Figure 4.5, we show that running the algorithm for any complete k-partite graph on up to 9 parts for up to 15 steps will take less than 4000 two-qubit gates. In Figure 4.6, we show the results of the output of such circuits running for the complete k-partite graph $K(5,4)$ on the Quantinuum Helios emulator.

Regime 2: Exploring early targets for quantum advantage As identified in Figure 4.2, graphs with small vertex counts but large clique counts represent a limitation for current classical TDA methods. These graphs are natural targets for exploring early quantum advantage for this algorithm. In Figure 4.7, we provide resource counts for running 30 steps of the algorithm on the same range of graphs in Figure 4.2. These resource counts do not make use of the optimisations for complete k-partite graphs in Figure 4.5. This means that such resources will be sufficient to perform the Laplacian moments algorithm on a wide range of generic graphs that are built from perturbing complete k-partite graphs. This class of graphs is studied by Berry et al. [40] as a promising target for quantum advantage in the related problem of estimating the normalised Betti number.

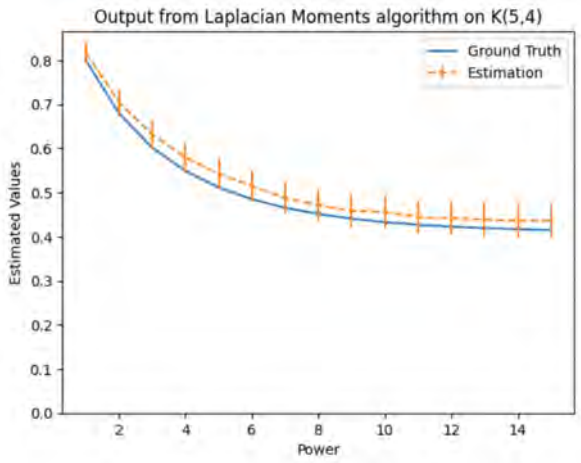


Figure 4.6: Results of 1000 shots of the quantum algorithm on a complete k-partite graph $K(5,4)$ on 20 vertices, run on a noisy emulator of Quantinuum Helios.

Regime 3: Towards advantage on real-world data Moving towards more general application of the quantum Laplacian moments algorithm, we consider the regime of graphs arising from real-world data. To this end we take the range of graph sizes occurring as 2-hop ego-graphs in the data from Section 4.6. This range is empircally observed to be from 10 to 2×10^4 vertices for the majority of ego-graphs. While we didn't find the high clique counts needed to make classical methods infeasible, we expect that target data sets which do require the quantum Laplacian moments algorithm can be found (see for example Section 4.5) and may be good targets for commercial advantage in the early fault tolerant era. In Figure 4.8, we provide T gate counts and depths for performing 50 blocks of the algorithm on graphs whose number of complement edges scales proportionally to n^2 and complement degree scales proportional to n. We use GRIDSYNTH (whole circuit accuracy 10^{-10}) for converting the rotations in the circuit's Hermitian boundary operator B to Clifford +T.

4.4 Quantum-Enhanced TDA and Quantum-HPC-AI Workflows

4.4.1 Baseline

We consider graph-structured data as the primary input. Graphs provide a broadly applicable representation: many data modalities can be interpreted geometrically through graphs (e.g., images can be viewed as grid graphs). Adopting a graph viewpoint therefore enables a natural geometric interpretation of data.

On top of this representation, we employ graph neural networks (GNNs), which respect graph structure when aggregating information. Conceptually, GNNs can be viewed as a generalization of convolutional neural networks (CNNs): they learn by repeatedly aggregating and transforming information along edges.

However, the fundamental aggregation mechanism in standard message-passing GNNs is local. Each node can access information only within a limited neighborhood determined by the number of network layers, and even within this hop-limited region, the representation is dominated by local interactions. As a result, structurally global information within that neighborhood is not necessarily captured explicitly. This motivates augmenting GNNs with complementary descriptors that summarize more global structural properties, such as topological invariants.

We therefore propose to compute TDA-derived descriptors and use them as additional node features for GNN training. Concretely, for each node we construct its *ego-graph*, i.e., the induced subgraph consisting of the node and all nodes within a fixed hop distance (together with all edges among them). We then apply TDA to this local subgraph, enabling the resulting descriptors to be attached as node-level features and passed downstream to the GNN.

As discussed previously, exact TDA computation can become computationally prohibitive. We thus focus on additive-approximation approaches implemented via Laplacian moments. Since nor-

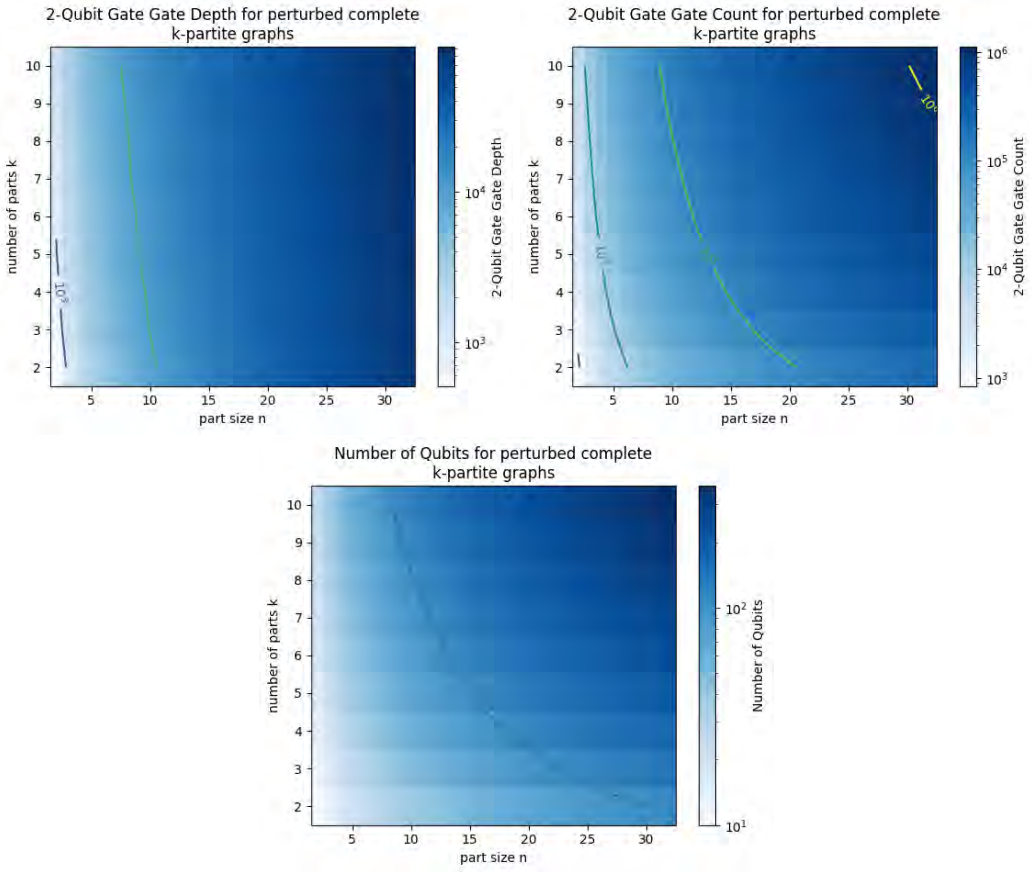


Figure 4.7: Resources required for running one shot of the Laplacian moments algorithm (for 30 steps) on generic graphs formed by perturbing complete k-partite graphs.

malized Betti numbers arise only in the infinite-depth limit of Laplacian moments, we do not attempt to compute the limit directly. Instead, we provide the GNN with a finite sequence of Laplacian moments,

$$\left(T^{(1)}, T^{(2)}, T^{(3)}, \ldots, T^{(d)}\right),$$

as structured features.

In summary, our baseline Quantum–HPC–AI workflow is as follows: (i) given a graph, construct an ego-graph for each node; (ii) compute a sequence of Laplacian moments for each ego-graph using quantum computation; (iii) attach these moment sequences as node features; and (iv) train the downstream GNN on HPC infrastructure.

This division of labor is natural: quantum computation is used for geometric/topological feature extraction, while classical AI and HPC handle large-scale model training and deployment.

Finally, it is important to emphasize the motivation for this hybrid design. Direct quantum machine learning approaches face well-known challenges such as barren plateaus and data-loading bottlenecks. In contrast, extracting quantum-derived features and leveraging them within strong classical machine learning pipelines is a highly promising path. The present baseline follows exactly this philosophy: quantum resources are used to produce features that enrich the classical model.

4.4.2 Beyond Topological Data Analysis

We now highlight an important perspective briefly raised in Section 4.2.3: Laplacian moments are useful beyond their asymptotic connection to topology.

The k-th Hodge Laplacian is defined as

$$\Delta_k = \partial_{k+1}\partial_{k+1}^{\dagger} + \partial_k^{\dagger}\partial_k.$$

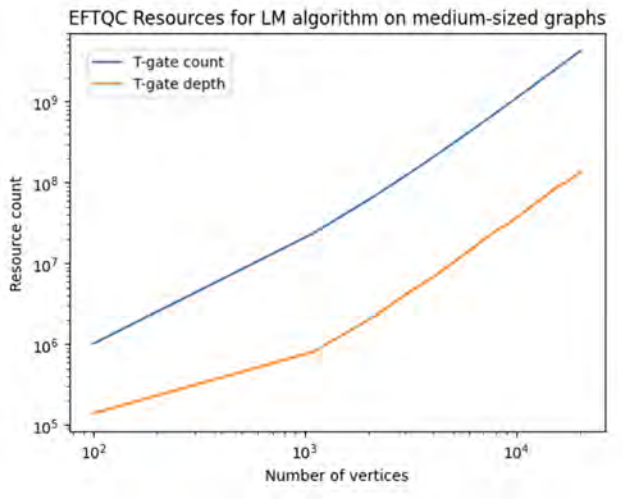


Figure 4.8: T-gate resources required to run one shot of the Laplacian moments algorithm (up to 50 steps) on graphs of varying sizes assuming the number of complement edges grows as $n^2/20$ and the max complement degree grows as $n/20$.

This operator can be interpreted as a diffusion operator on k-simplices. The first term mediates interactions through $(k+1)$-simplices, while the second term mediates interactions through $(k-1)$-simplices. In this sense, Δ_k governs how information propagates across the simplicial complex at dimension k.

Under this diffusion viewpoint, $(I-\Delta_k)^d$ corresponds to d-hop diffusion on k-simplices, and the Laplacian moment

$$T_k^{(d)} = \mathrm{Tr}(I-\Delta_k)^d$$

can be interpreted as the total self-correlation of a d-step diffusion process. The role of d is therefore transparent: small d emphasizes local structure, while larger d progressively incorporates broader and more global correlations.

In the limit $d \to \infty$, only the zero-eigenvalue subspace survives, and the Laplacian moments converge to (normalized) Betti numbers. This shows that Betti numbers represent the most global correlation information accessible through this diffusion process.

Crucially, finite-d Laplacian moments retain intermediate-scale correlation information that is lost in purely topological summaries. Betti numbers are robust but can be overly coarse: distinct structures may share identical Betti numbers. By contrast, finite-d moments interpolate between local and global organization, providing descriptors sensitive to mesoscopic structural changes. In Appendix B.2, we explore an application of this Laplacian feature to the study of heterophily in networks.

This is particularly relevant for real-world network applications such as communication networks, where anomalous or fraudulent behavior may not alter global topological invariants, yet can modify intermediate-scale correlation structure. Finite-d Laplacian moments can capture such changes and therefore serve as informative features for anomaly or fraud detection.

Taken together, this framework is not only a candidate route toward quantum advantage, but also a practically useful feature-augmentation strategy that introduces additional structural information beyond what standard classical TDA pipelines typically provide.

4.5 Synthetic ML Dataset with likely advantage for Quantum over GNNs

Before we apply the new pipeline introduced in the previous section to real-world data, we first provide a proof-of-concept example to show that likely classically-inaccessible Laplacian moments can provide an advantage over GNNs. In this section, we demonstrate a (highly engineered) synthetic dataset that shows the maximal possible benefit that Laplacian moments could have.

In Figure 4.9 we have constructed a toy dataset where the labels are associated with substruc-

Figure 4.9: Synthetic dataset where the “fraud” and “non-fraud” labels accompany different WL indistinguishable substructures.

tures that are indistinguishable by Weisfeiler-Leman hashes and therefore vanilla GNNs. These same substructures are distinguishable by high-order Laplacian moments. Indeed we run the full ML pipeline and achieve 100% classification accuracy with Laplacian moments and 50% without, the best possible situation. Furthermore, by virtue of the construction we can move the valuable information into an arbitrary Laplacian order of high power.

The details of the construction are as follows. Our goal is to produce two graphs that are indistinguishable by Weisfeiler-Leman but distinguishable by Laplacian moments. Graphs that satisfy the first condition are well studied in the literature. We simply borrow one such example which also differs by homology and then push the homology to arbitrary dimension. One of the most famous examples is the circulant graph with $n = km$ nodes with $r = (m-1)/2$ skip connections for skips $j \bmod n, j \in [1, r]$, versus the disconnected graph consisting of k cliques of size-m. When $k = 2$ and $m = 3$, we have Figure 4.10 consisting of two triangles versus one hexagon. These two graphs are indistinguishable by Weisfeiler-Leman because nodes in both graphs have the same one-hop neighbourhood structure. They are however distinguishable by their topologies, namely the number of disconnected components (β_0) and the number of loops (β_1). Now to make them distinguishable by high powers of the Laplacian moments we would like one of the graphs to have exponentially many more holes than the other. We know that the best candidate for this is the complete k-partite graph. Fortunately, the k m-clique graph complement is exactly the k-partite graph with m parts. So taking the complement of both graphs keeps them WL indistinguishable and makes them Laplacian-moment distinguishable, see Figure 4.11. The complement also has the advantage of making both graphs connected (so they no longer differ in Betti 0) and the topology differs exponentially in order k, which can be pushed to high dimension either by making k larger or by using the procedure called suspension. Suspension has the advantage of making low order Laplacian moments more indistinguishable. Alternative low order obfuscation techniques have also been investigated with success such as using a perturbed k-partite graph versus a k-partite graph.

It is very believable that real-world mechanisms could correspond to producing circulant versus clique structures, for example neighbourhood vs hemisphere collaboration preferences of participants arranged on a circle (see an example with 10 nodes in Figure 4.12). Depending on the domain, going straight to the complement picture may be more natural, where k-partite edges are swapped for inner-partite edges, (see the complement picture in Figure 4.13).

Even if real-world mechanisms do not always produce cleanly separable structures, the incorporation of the Laplacian moments in an ML pipeline allows the network to automatically learn the value of Laplacian moments on a class-by-class, approximative and weighted basis. Finally, this engineered example does illustrate that the question of Laplacian moment advantage and then quantum advantage is very domain and use-case specific. In this project, the search has just begun and we have

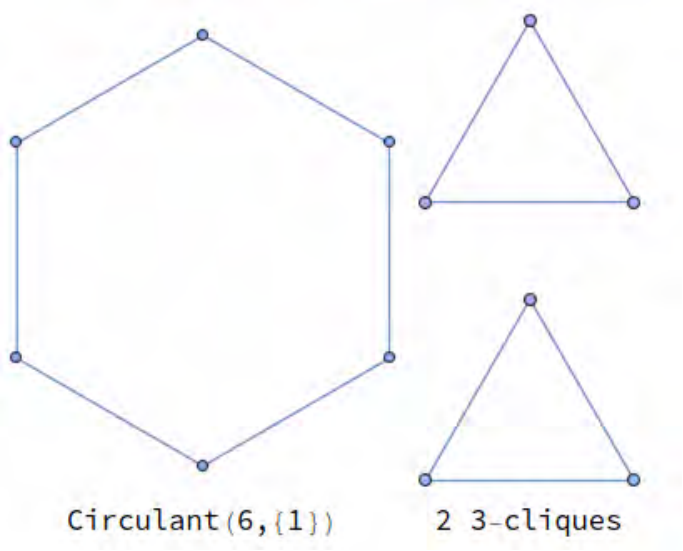


Figure 4.10: Standard example of WL indistinguishability

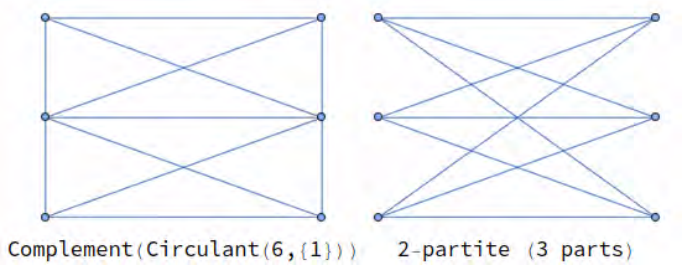


Figure 4.11: Modified example for Laplacian moment distinguishability

found our first case of Laplacian advantage (without quantum advantage). It is possible that there might even be use-cases whose Laplacian advantage only becomes evident in the quantum advantage regime. Nevertheless, this section makes a strong case for TDA data-scientists to embrace the Laplacian ML approach.

4.6 Role of Topological Data Analysis in Fraud Detection

In AI, performance is strongly influenced not only by the choice of models and learning algorithms but also by appropriate data analysis. Even without acquiring additional data, accuracy can sometimes be improved by analyzing the available data more deeply and transforming it into more informative representations before training. In real-world operations, data collection and labeling entail cost and constraints; therefore, the value of achieving improvements solely through analysis of existing data is substantial.

As a concrete example where this perspective is particularly important, this work focuses on the task of telecommunications fraud detection. Telecommunications fraud has been increasing year by year, and a CFCA report estimates that losses in 2023 reached $38.95 billion [48]. Given this scale, even a 1% improvement in detection performance could, by a simple calculation, correspond to suppressing losses by $0.3895 billion, implying a very large social and economic impact even for a modest gain [48]. At the same time, acquiring new data for telecommunications fraud tasks is not straightforward. Due to privacy and operational constraints, the difficulty of labeling, evolving fraud tactics, and the scarcity of positive examples, even a small accuracy improvement using the original data, without further expensive forensics, can lead to a meaningful reduction in losses.

In telecommunications fraud detection, interactions such as calls, messages, transfers, and ac-

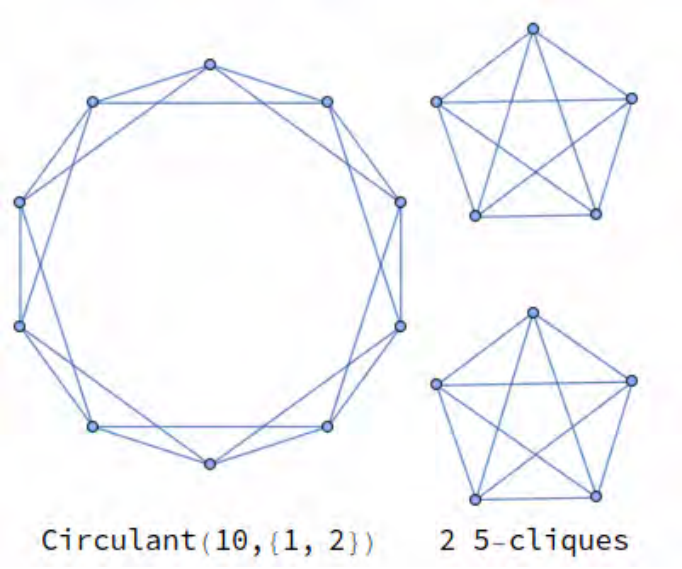


Figure 4.12: Extended example of WL indistinguishability

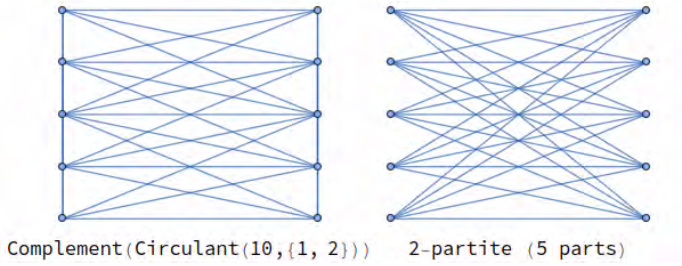


Figure 4.13: Modified extended example for Laplacian moment distinguishability

count relationships can be represented as a network, making graph neural network (GNN) approaches particularly promising [49–53]. GNNs respect graph structure and learn representations by aggregating information from neighborhoods, providing a natural framework for handling anomalies and fraud on telecommunications networks. In this work, we take GNNs as a central baseline and examine whether additional features obtained through data analysis can contribute to improving fraud detection performance.

However, it is not the case that adding arbitrary features will necessarily help. Because GNNs already incorporate basic graph features such as node degree through neighborhood aggregation, simply adding elementary local statistics is unlikely to yield substantial improvements. As a next step, one may compute more advanced structural features such as various centrality measures. However, as shown in Fig. 4.14, the distributions of centrality measures may exhibit little difference between normal and fraudulent nodes. This is consistent with the idea that fraudsters intentionally mimic normal behavior to evade detection, suggesting that discrimination is difficult at least through straightforward extensions of purely local structural features.

Nevertheless, fraudsters must not only evade detection rules but also accomplish their objectives, such as transferring funds or accumulating rewards. In other words, they must simultaneously satisfy both “remaining inconspicuous” and “achieving the intended outcome,” and this dual constraint can induce structural regularities in their behavior. As a result, patterns such as the following are often generated:

- distributing edges to avoid threshold-based detection,
- constructing detour paths to avoid direct connections,
- forming multi-stage structures by combining nodes with distinct roles (e.g., collection, relay, and

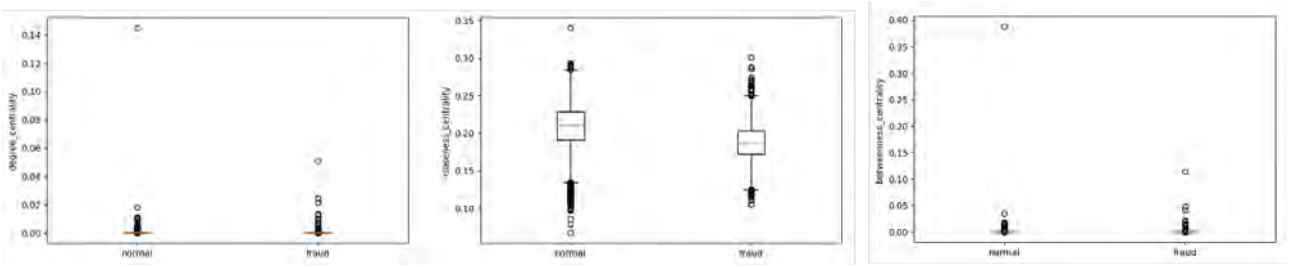


Figure 4.14: Comparison of centrality measures between normal and fraudulent nodes. From left to right: degree centrality, closeness centrality, and betweenness centrality for the BUPT dataset.

cash-out).

These patterns may not drastically change local quantities such as degree or neighborhood statistics, but they can manifest as more global geometric structure at the network level. Therefore, to extract differences that are difficult to capture using only local statistics or centrality measures, we require a framework capable of handling global shape information such as connectivity, loop structure, and cavity-like structures.

Therefore, in this work, we investigate whether applying QTDA – HPC – AI workflows to telecommunications network data can improve detection performance compared to conventional approaches.

4.7 Experimental Results on Fraud Detection

In this work, we use the BUPT dataset [49, 51], which is widely recognized as a benchmark dataset for fraud detection in telecommunications networks. The graph is relatively sparse, with low edge density, and therefore remains tractable for exact classical computation. Accordingly, in the analyses that follow, both Betti numbers and Laplacian moments are computed exactly rather than via approximation.

We first compute Betti numbers for ego-graphs in the communication network. We observe that

$$\beta_k = 0 \quad (k \geq 2).$$

For $k = 0$, there is no information because by definition the ego-graph has one exactly one connected component. For $2 \leq k \leq 4$ there are sometimes k-simplices present but there is no homology. For greater k there are simply no k-simplices. In the range, $2 \leq k \leq 4$, it is interesting that higher-dimensional homology is entirely absent; the ego-graphs contain no nontrivial higher-dimensional topological features when they could have. This observation has important implications. Pure topological invariants are extremely global quantities. When higher-dimensional structure is absent, they assign identical values (zero) to all nodes and therefore lack discriminative power. In other words, topological invariants can be “too global” and discard information essential for classification.

Here is where Laplacian moments become crucial. When finite-d Laplacian moments are computed for ego-graphs in the communication network, the resulting distributions exhibit nontrivial higher-dimensional contributions, even though the corresponding Betti numbers vanish (see Fig. 4.16). This indicates that Laplacian moments retain spectral information beyond zero modes and therefore capture structural differences invisible to pure topology. Moreover, for moderately large d, Laplacian moments incorporate increasingly global information that standard GNNs do not naturally encode through localized aggregation. Empirically, incorporating Laplacian moments as additional features in GNN-based fraud detection improves classification accuracy(see Table 4.1). We evaluate the statistical significance of this improvement using a two-sample t-test. Comparing the recall obtained with and without LM across 10 random seeds, we obtain a p-value of 0.026. This result indicates a statistically meaningful improvement and suggests a moderate level of advantage when LM features are incorporated.

By evaluating the model at multiple threshold values, we obtain the full Precision-Recall (PR)

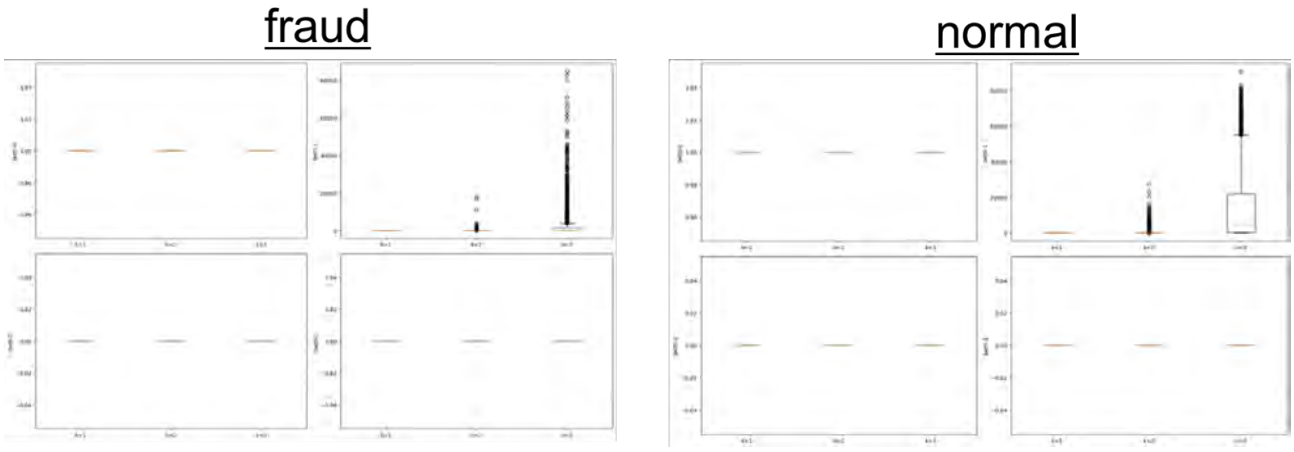


Figure 4.15: Comparison of Betti numbers between normal and fraudulent nodes.

	Macro AUC	Macro recall	Macro F1	G-mean
with LM	0.9243 ± 0.0169	0.8449 ± 0.0237	0.8415 ± 0.0122	0.8327 ± 0.0283
without LM	0.9204 ± 0.0169	0.8317 ± 0.0219	0.8315 ± 0.0101	0.8170 ± 0.0275

	With LM		Without LM	
	Fraud	Normal	Fraud	Normal
Fraud	242.1 ± 21.5	97.1 ± 14.2	230.5 ± 21.6	108.7 ± 14.8
Normal	106.8 ± 31.1	6972.0 ± 41.1	109.5 ± 30.2	6969.3 ± 37.0

Table 4.1: (Top) In threshold = 0.5, performance comparison on BUPT dataset. (Bottom) Confusion Matrices Comparison. We use 15% of the total nodes as the test set.

curve (see Left of Fig. 4.17). At a given operating point, if we fix precision and compare the corresponding recall values, the difference in recall directly reflects the improvement in fraud detection achieved by incorporating LM as additional features. Using this improvement in recall, we can estimate how much of the annual telecom fraud loss could potentially be mitigated. In this analysis, we focus on approximately 25% of the global annual loss, corresponding to about USD 10 billion, as the target loss volume for evaluation. Our analysis indicates that the gain translates into a reduction on the order of several hundred million dollars annually (see Right of Fig. 4.17). For example, at a fixed precision of 0.9, the improvement in recall corresponds to an estimated reduction of approximately USD 0.220 billion in annual fraud losses. In addition, we also evaluate the performance with and without LM at the precision level where the difference in recall is the largest (see Table 4.2). Importantly, finite-d Laplacian moments are computationally far simpler than exact topological invariants such as Betti numbers, yet in practice they contain strictly richer structural information for discrimination tasks.

We conclude the results discussion by noting that the fraud dataset turned out to be too sparse to contain higher-order cliques. There are at least two strategies we could employ to find more suitable graphs. Firstly, we could investigate pre-processing steps that increase the edge-density but in an actually useful way, for example, contracting nodes which meaningfully belong together. Secondly, we could take inspiration from the synthetic dataset constructions and look specifically for related real-world graph-generating mechanisms. For example, the synthetic graphs have hidden repeating dense regions which could correspond to waves of compression and rarefaction in high-dimensional embedding space. From a fraud detection point of view, this suggests we should look beyond simple anomalous behaviour in static datasets, perhaps Laplacian-based TDA is better suited for detecting *co-ordinated* and *dynamic* anomalous behaviour such as is, for example, found in Denial-of-Service (DoS) attacks.

Although the communication network considered contains little higher-dimensional structure and

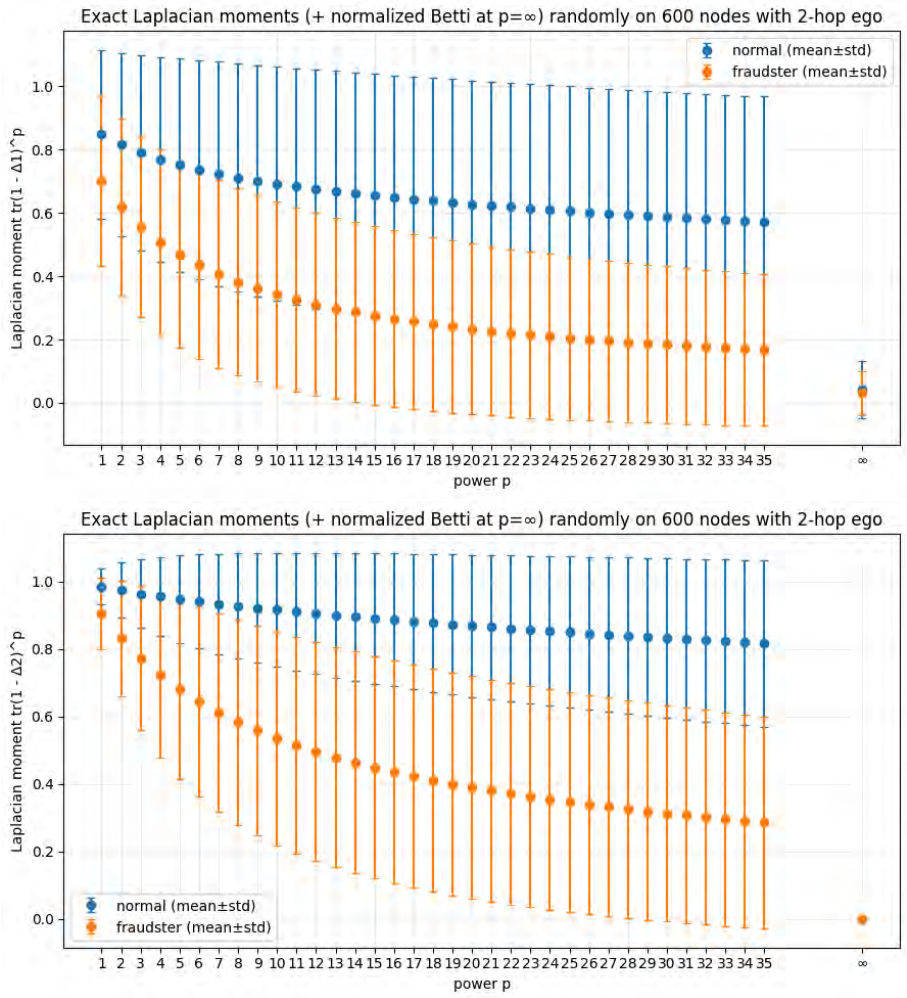


Figure 4.16: Comparison of Laplacian moments (LM) computed on 600 randomly selected fraud and normal nodes with 2 hop ego-graph. The upper panel shows the one-dimensional LM, while the lower panel shows the two-dimensional LM. Note that, in this case, the trace is computed by exact diagonalization rather than by Monte Carlo estimation.

therefore does not directly demonstrate quantum advantage, the observed improvement arises from quantities that are naturally implementable on quantum hardware. The advantage thus stems from a paradigm shift: rather than directly targeting pure topological invariants, we begin with quantities that are naturally compatible with quantum implementation. In doing so, new structural information and practical advantages emerge organically. In this sense, the benefit of this approach is not merely computational speedup, but a transformation in the framework of geometric information extraction itself.

4.8 Discussion and Outlook

This exploratory project into the usefulness of Laplacian moments for real-world data has yielded promising results. With a theoretically possible (under widely-held computational complexity assumptions) exponential quantum advantage on offer for high-order normalised Betti numbers the goal was to find evidence for real-world realisations of this advantage. We provided motivation for the claim that powers of the Laplacian may be valuable carriers of a mix of geometric and topological information beyond Betti numbers. We developed classical tools to calculate approximations to the normalised Laplacian moments to allow searching for *Laplacian advantage* in the regime before the power becomes too high for classical tractability. Indeed this suggested not only a path to discovering *quantum*

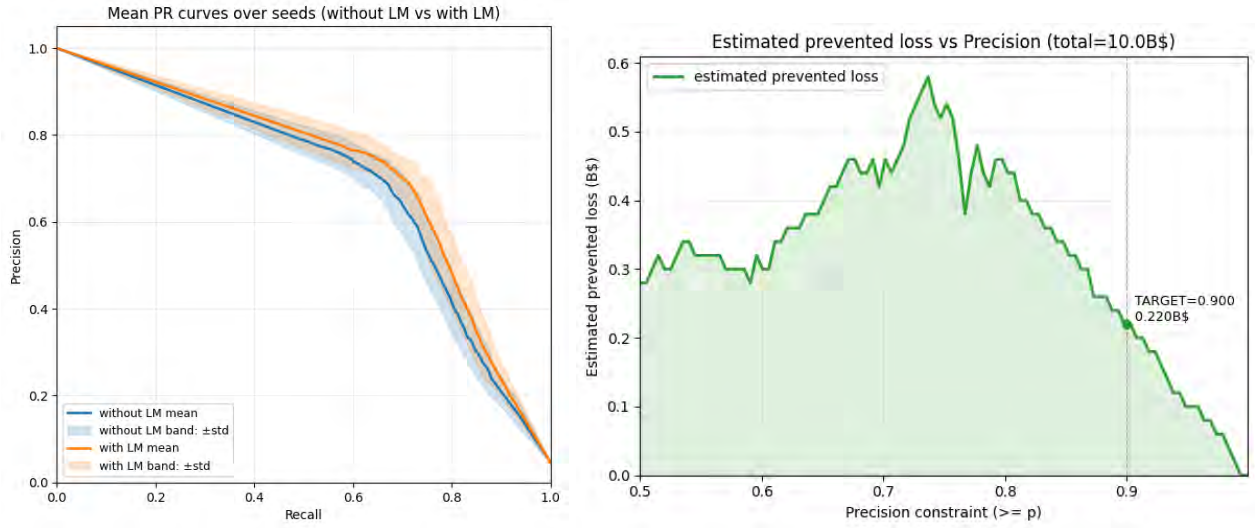


Figure 4.17: (Left) Comparison of PR curves with and without LM. (Right) Estimated reduction in the total fraud loss (USD 10 billion) achieved by incorporating LM, calculated based on the difference in recall at fixed precision levels between the two models.

	Macro recall	Macro F1	G-mean
with LM	0.8270 ± 0.0253	0.8409 ± 0.0150	0.8106 ± 0.0315
without LM	0.8119 ± 0.0205	0.8311 ± 0.0125	0.7920 ± 0.0259

	With LM		Without LM	
	Fraud	Normal	Fraud	Normal
Fraud	225.9 ± 20.6	113.3 ± 18.3	215.6 ± 18.7	123.6 ± 14.7
Normal	83.5 ± 9.2	6995.3 ± 21.8	81.7 ± 11.3	6997.1 ± 22.1

Table 4.2: (Top) In precision = 0.737, performance comparison on BUPT dataset, where 0.737 is the precision level at which the largest difference in recall is observed when precision is held constant. (Bottom) Confusion Matrices Comparison. We use 15% of the total nodes as the test set.

advantage in real-world datasets when large enough quantum computers become available, but also allows for a plausible case to be made with purely classical compute: in particular, if we could demonstrate that high-powers of *increasing* Laplacian order contain valuable information, then it is believable that this trend would continue into the regime that is classically unreachable. This more subtle goal was achieved in all the settings on the spectrum of synthetic to real-world data that we considered. We passed through three categories: firstly highly structured, engineered, examplary datasets; secondly marginally-structured statistical models of graphs and finally real-world datasets. In all three categories, we established the goal, namely that higher order can indeed provide extra independent predictive value. For the first category, we constructed an example with provable maximal advantage of arbitrarily high-order Laplacian moments (which can be pushed into the quantum advantage territory) over classical GNNs. In the second category (see Appendix B.2), using the rigourous statistical measure of mutual information, we empirically demonstrated that higher-order moments increase the mutual information with the predictive label. In the third category, for real-world fraud data, using simple statistical standard deviation from the mean, we showed that the second-order Laplacian moment has a larger separation than the first-order. Unfortunately, we could not go to even higher-order because of the dearth of higher order structure. Nevertheless, for the two orders meaningfully present we showed that both orders improved ML prediction scores, thereby validating the end-to-end ML pipeline including the idea of augmenting node features with Laplacian moments and training large classical ML models on powerful GPU/HPC architecture.

Having discovered classical Laplacian advantage for real-world problems and designed synthetic examples with quantum advantage, we can therefore confidently invite the TDA community to join us in researching this promising, fresh approach of normalised higher-order Laplacian moments that has benefit starting in the classical regime today and that seamlessly extends into a potentially exponentially powerful algorithm in the quantum future.

5 Roadmap and Use-Case Timeline

5.1 Aligning Hardware Evolution with Executable Use Cases

This chapter translates the technical analyses presented in the previous chapters into a structured, forward-looking roadmap. The goal is not to describe quantum computing capability in abstract terms, but to provide a disciplined assessment of how executable problem classes expand as hardware matures.

We anchor the analysis to four successive generations of quantum hardware on Quantinuum's development roadmap: Helios (current generation), Sol (anticipated for 2027), Apollo (anticipated for 2029), and it's large-scale fault-tolerant generation, Lumos, anticipated in the 2030s. These milestones correspond to projected increases in qubit counts, improvements in physical error rates, and growing feasibility of logical-layer execution with quantum error correction.

For each generation, we assess how projected device-level characteristics—available qubits, connectivity, error rates, and realistic error-correction overhead—interact with the resource requirements identified in Chapters 3 and 4.

In quantum chemistry, feasibility is expressed in terms of the number of spin orbitals, which directly reflects molecular complexity and electronic degrees of freedom. This provides a chemically meaningful scaling axis that connects hardware capability to industrially relevant molecular systems. Within this framework, we distinguish two execution paradigms. At the physical level, circuits are implemented directly on hardware, optionally with quantum error detection (QED), enabling larger spin-orbital counts but only for shallow circuits due to the exponential decay of success probability, $P_{\text{succ}} \sim e^{-pN}$. At the logical level, computations are performed on error-corrected qubits, where the available number of spin orbitals is reduced by encoding overhead but deep circuits become feasible, enabling algorithms such as quantum phase estimation (QPE) for systematically accurate simulations.

In TDA, feasibility is characterized by graph size and structural parameters, particularly the number of clusters and nodes per cluster in complete k-partite graphs. These parameters allow computational complexity to be controlled in a reproducible manner and serve as a bridge between abstract algorithmic scaling and real-world network analysis tasks.

We summarize these feasibility boundaries in Table 5.1 and Table 5.2, respectively. Importantly, the progression is not binary. Rather than a sudden transition from "impossible" to "possible," the hardware roadmap suggests a gradual widening of the executable envelope: from small proof-of-concept circuits, to pilot-scale workloads, and ultimately to application-relevant problem sizes.

Uncertainty naturally increases toward later milestones. Logical overhead, non-Clifford gate synthesis costs—particularly for T gates—and the realized performance of physical qubits will influence effective computational capacity. Nevertheless, the trajectory outlined here reflects a consistent expansion of viable problem classes rather than speculative discontinuities.

5.2 Domain Progression, Joint Vision, and Strategic Positioning

As outlined in Chapters 1 and 2, the purpose of this white paper is to construct a use-case timeline that connects technical feasibility to strategic action. The joint research between SoftBank and Quantinuum is conducted under a shared vision: the long-term realization of quantum AI data centers in which quantum processors operate as integrated computational accelerators alongside AI and HPC systems.

Hardware Generation	Year (Target)	Expected Chemistry (Spin Orbitals) – Physical/QED	Expected Chemistry (Spin Orbitals) – Logical (FTQC)	Expected Technical Characteristics	Potential Representative Use Cases	Potential Customer Segments	Expected Improvements
Helios	Current (2025)	· ~50 spin orbitals · $p_{phys} \sim 10^{-3}$ · Subspace methods, dynamics simulation	· ~10 spin orbitals · $p_L \sim 10^{-3}$ · $d = 3–5$ · Steane code early QPE/T-gate benchmarks	Proof-of-concept focused	Small-molecule excited-state PoC	Materials R&D divisions	Technical validation
Sol	2027	· ~100 spin orbitals · QED with high rejection rate · Subspace methods, dynamics simulation	· ~10 spin orbitals · $p_L \sim 10^{-4}$ · $d = 5–7$ · Early logical stabilization, toy QPE (small active space)	Initial scaling phase	Limited excited-state applications	Pharmaceutical and materials firms	R&D efficiency gains
Apollo	2029	· ~100 spin orbitals · QED or partial fault-tolerant · Dynamics simulation	· ~100 spin orbitals · $p_L \sim 10^{-7}$ · $d = 7–9$ · Excited-state QPE	Transition toward application-relevant regimes	Medium-scale materials modeling	Global manufacturers	Quantifiable R&D cost reduction
Lumos	2030+	N/A	· 100+ spin orbitals · $p_L \sim \leq 10^{-10}$ · Chemical accuracy · Scalable workflows	Full fault-tolerant regime	Integrated materials discovery workflows	Industrial infrastructure operators	Sustained competitive advantage

Table 5.1: Hardware roadmap for quantum chemistry applications across hardware generations.

Hardware Generation	Year (Target)	Expected TDA (Graph Size)	Expected Characteristics	Potential Representative UseCases	Potential Primary Customer Segments	Expected Improvements
Helios	Current (2025)	· Complete k-partite graphs $K(m,k)$ for $k \times m < 90$ · Error mitigation on noisy qubits	Fault-tolerant quantum computing bridge validation stage	Structured IRSF validation	Telecom innovation units	Detection accuracy demonstration
Sol	2027	· 100 nodes, 30 steps · Early advantage demonstrations on more general graph families · Integration with error correcting codes	Improved logical layer stability	Pilot fraud detection deployment	Tier-1telecom operators	Measurable fraud-loss reduction (few percentrange)
Apollo	2029	· 100–300 total nodes · Structured real-world approximations · Enhanced anomaly detection modeling	Transition toward application-relevant regimes	Commercial IRSF exploration	Telecom operators	Annual fraud-loss mitigation
Lumos	2030+	· Hundreds to thousands of nodes · Production-scale quantum-enhanced graph analytics	Infrastructure-grade capability	Large-scale security analytics platforms	Telecom and finance sectors	Infrastructure level service revenue

Table 5.2: Hardware road map for TDA applications across hardware generations.

Within this shared vision, quantum chemistry and TDA play complementary roles. Quantum chemistry represents a longer-term, high-impact domain in which large-scale fault-tolerant capability unlocks transformative industrial applications in materials discovery and molecular design. TDA—particularly in use cases such as IRSF (International Revenue Share Fraud) detection—offers earlier opportunities for commercialization, as even incremental performance improvements can produce measurable financial impact.

The staged progression across hardware generations therefore naturally maps onto a staged business strategy for deploying quantum AI data center services. Early milestones establish technical credibility and ecosystem positioning. Intermediate milestones enable pilot deployments and early customer engagement. Later milestones support scalable, infrastructure-level service offerings embedded within a quantum AI data center model. The use-case timeline therefore functions not only as a technical roadmap, but as an integrated strategic framework for potential future customers.

A Technical Details of Quantum Chemistry

A.1 Photochromic reactions

Photochromic reactions are reversible reactions that happen under light stimulation. The reactants undergo geometrical and electronic changes to yield isomeric products with different colors. Such photochromic systems are broadly used as switches with various applications. Common steps in photochromic reactions include *cis*/*trans* isomerization and ring opening/closure. Examples of photochromic molecules are stilbene, diarylethene, spiropyran, and rhodopsin. These conjugated systems can be used as ligands to construct more complex structures, e.g., transition-metal complexes, metal-organic frameworks, dendrimers, and polymers. More interestingly, the new structures retain the photochromic properties of the parent ligands, thereby expanding the spectrum of applications [54–57].

The elucidation of photochromic reaction mechanisms has been an active field of research for many years. It is common for these reactions to proceed through multiple local minima, which are difficult to predict computationally. More importantly, the existence of one or more conical intersections (CoIn) ensures the reversibility of the reactions, which is important for the above-mentioned applications. From a computational chemistry perspective, predicting geometry and electronic properties is a well-known challenging problem [58]. Typically, two potential energy surfaces (PESs) are degenerate at the CoIn point, facilitating reversible transitions between electronic states. For this reason, single referenced methods, e.g., Hartree-Fock (HF) and Density Functional Theory (DFT), fail to reproduce this degeneracy and often exhibit a cusp at the CoIn point [58–60].

Selecting the active space is a crucial step. In the context of the calculation of PES for a reaction, the most important orbitals are those that involve electronic density in the spatial region of the system that is changing for each step of the reaction, *e.g.* dissociating bond, formation of a new bond etc. Evidently, the number and type of orbitals that need to be considered may vary along the reaction coordinate. This creates the problem of consistency of the active space throughout the different structures considered along the reaction pathway. The development of tools for automated selection of orbital sets for chemical reactions is an active area of research.

Ethylene is usually used as a benchmark for excited-state methods. The system exhibits a CoIn when the HC=CH dihedral angle is set to 90°. Common TDDFT approaches exhibit a cusp at the CoIn region. In recent years, TD-DFT has been augmented with corrections that can capture the degeneracy of the excited states at the CoIn region [59]. Here, we present our CASCI results in Figure A.1(a). We used a minimal active space of two spatial orbitals, the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) and the STO-3G basis set for all the atoms. Evidently, the degeneracy of the ground state and the first excited state at 90°is well reproduced. In the middle and right panels of Figure A.1 we show the valence occupied (with red) and unoccupied (with blue) orbitals and energy levels at 0°and 90°respectively. As expected, the HOMO-LUMO gap becomes smaller at the CI level and the symmetry of the obritals is reduced at the CoIn level.

Now we proceed to the photoisomerization process of stilbene. Stilbene is a photochromic molecule that transforms from *trans* to *cis* configuration reversibly, by twisting the central C=C bond, similarly to ethylene. We focus on the *trans-cis* isomerization process, which involves a CoIn. We show in Figure A.2(a) the CASCI+NEVPT2 results for the ground state and the first singlet excited state. We employed 12 electrons in twelve spatial orbitals (12,12). The associated geometries were taken from Ref. [61] and the 6-31G basis set was used for all the atoms. The CASCI+NEVPT2 cal-

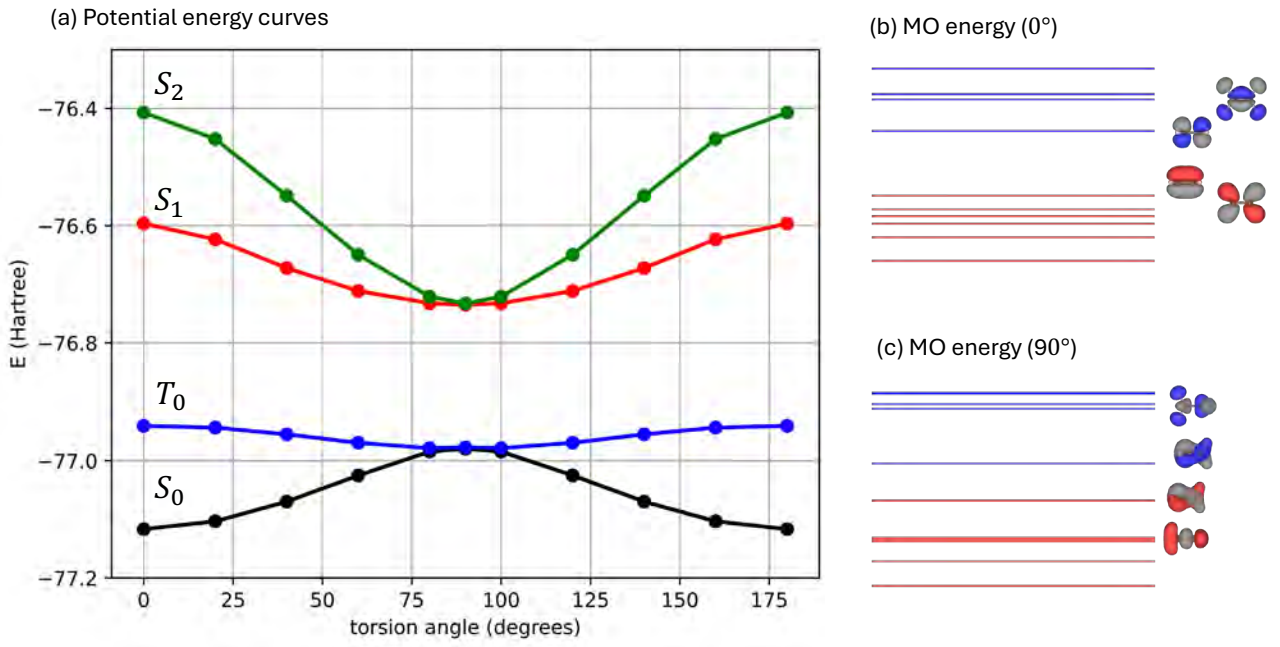


Figure A.1: a) PES of the ground and three first excited states. The calculations were performed with CASCI with an active space of two electrons in two spatial orbitals. (b) The HF eigenvalues and the plots of HOMO-1, HOMO, LUMO, LUMO+1 for the 0° configuration, and (c) the 90° configuration.

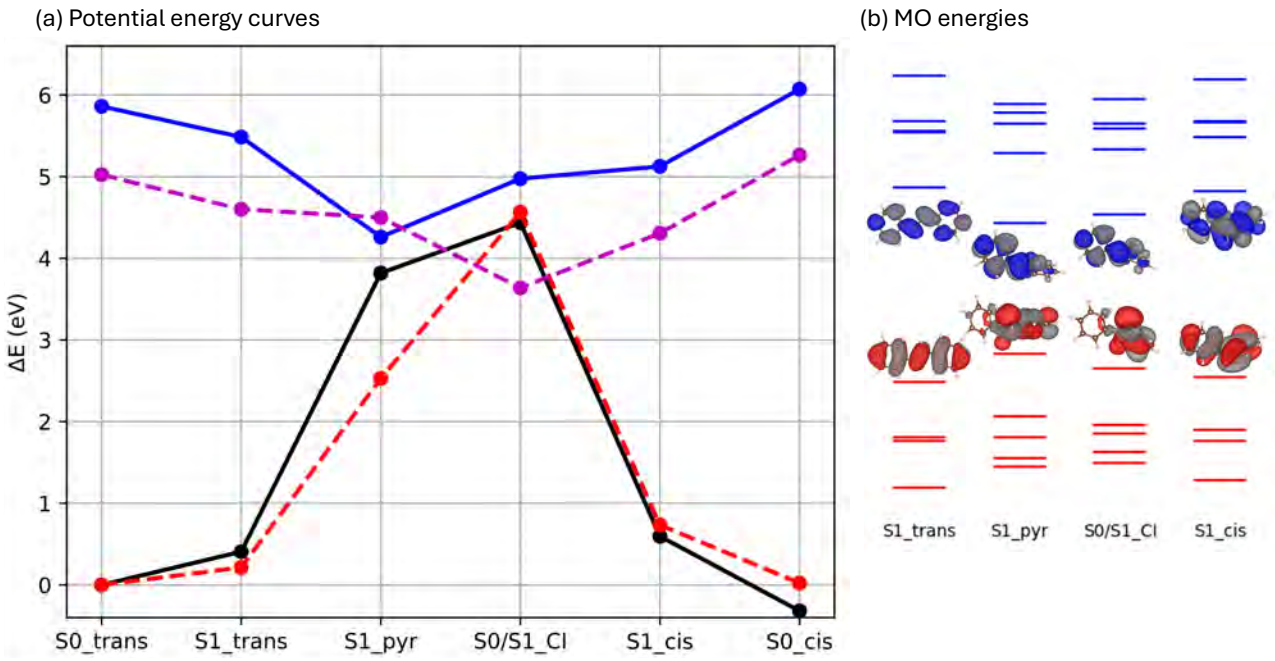


Figure A.2: a) PES of the ground and the excited states of a stilbene molecule. The calculations were performed with CASCI with an active space of (12,12). (b) The RHF eigenvalues and the plots of HOMO and LUMO.

culation qualitatively reproduces the behavior of CoIn, where the low-lying energy levels are closer to each other at the S0/S1_CI geometry. The spatial distributions of HOMO and LUMO are not continuously connected, and thus tracking experiments along the reaction coordinates are likely to break down. As such, the chemistry around the conical intersection is sensitive to the level of computation and the chemist's intuition. The canonical approach to overcoming this problem is to extend the active space to include all relevant orbitals, and a quantum computer may, in the long term, allow us to take this approach because the exponential scaling of resource requirements in classical computation does not appear.

As shown above, the excited electronic states are far from the single-reference picture even for a relatively simple stilbene molecule. The consistent description of photochemical reactions requires an impractically large active space for classical computers, highlighting the potential of quantum computers to push the boundaries of quantum chemistry.

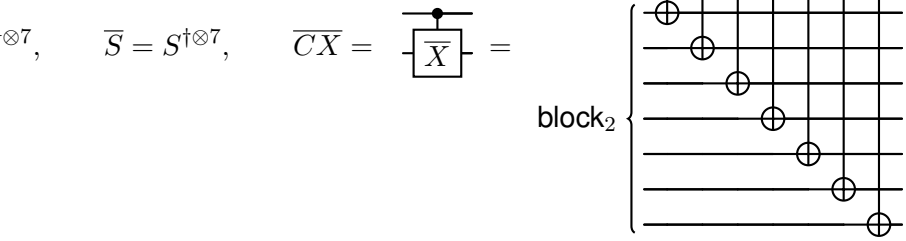
A.2 Logical components

We collect all the logical components in the encoded quantum circuits used in the main text.

A.2.1 Clifford operations

The logical Clifford gates $\{\overline{S}, \overline{H}, \overline{CX}\}$ admitted by the color code can be implemented transversally.

$$\overline{H} = H^{\otimes 7}, \qquad \overline{S} = S^{\dagger\otimes 7}, \qquad \overline{CX} =$$



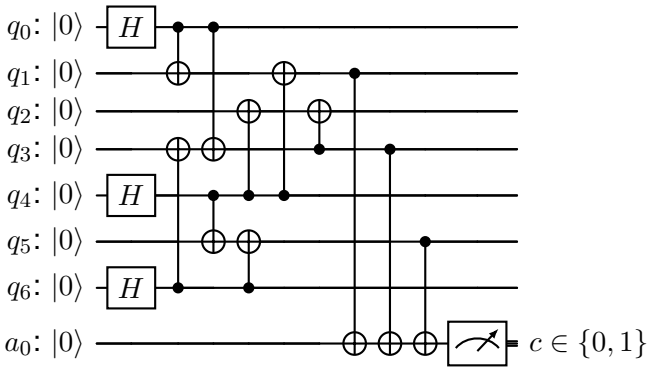
(A.1)

where $H^{\otimes 7}$ and $S^{\dagger\otimes 7}$ act on all the qubits transversally in a single code block, and block$_i$ in the $\overline{CX}$ circuit represents a code block.

A.2.2 Preparation of $|\overline{0}\rangle$

We summarize a circuit to fault-tolerantly encode the state $|\overline{0}\rangle$. A FT preparation of the state $|\overline{0}\rangle$ is accomplished by the circuit [62]:

$$|\overline{0}^{\mathrm{FT}}\rangle =$$



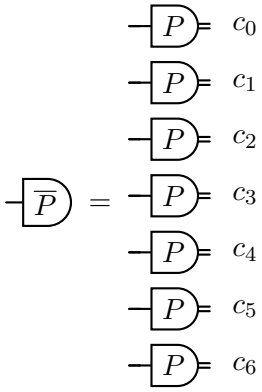
(A.2)

where a_0 represents the ancillary qubit. The circuit first encodes the state $|\overline{0}\rangle$ non-fault-tolerantly and then measures the bottom qubit to check if the state is correctly prepared up to a weight-1 X error. A

successful encoding is indicated by $c = 0$. The circuit is discarded when $c = 1$ is measured, and the protocol repeats the circuit executions until $c = 0$ is measured.

A.2.3 Measurement

The quantum circuit to destructively measure a logical Pauli $\overline{P} \in \{\overline{X}, \overline{Z}\}$ is given by



(A.3)

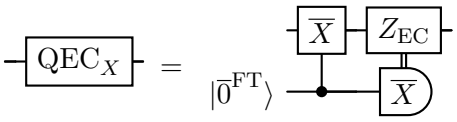
From the seven bits of measurement outcome $\{c_i\}$, we extract the syndromes s_1, s_2, s_3, associated to the stabilizers S_{X0}, S_{X1}, S_{X2} in Fig. 3.2, to identify and correct a weight-one error. Then, we evaluate the logical readout $\overline{c}$,

$$\overline{c} = c_0 \oplus c_1 \oplus c_4, \tag{A.4}$$

with the corrected bits.

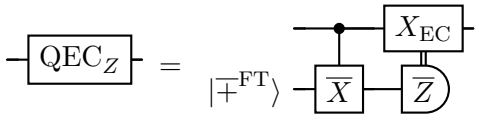
A.2.4 Steane QEC gadgets

The Steane QEC gadget using all the X stabilizers is given by,



(A.5)

The gate $\boxed{Z_{\mathrm{EC}}}$ represents the Z gates that correct the detected Z errors. Similarly, the Steane QEC gadget using the Z stabilizers is given by



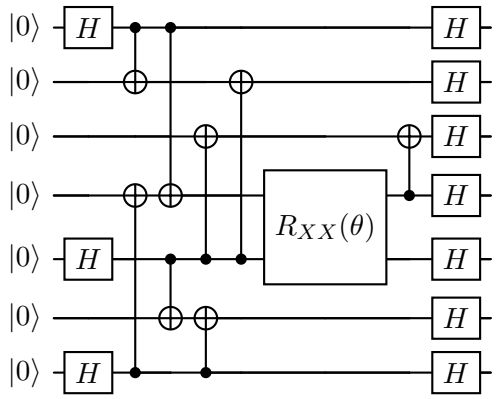
(A.6)

A.2.5 Preparation of $|\overline{\theta}\rangle$

We show quantum circuits to encode the state $|\overline{\theta}\rangle := \overline{R}_Z(\theta)\,|\overline{+}\rangle$.

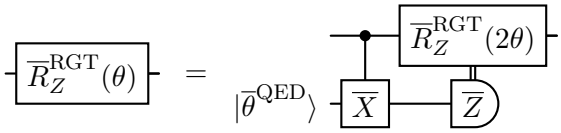
The non-FT $|\overline{\theta}\rangle$ preparation circuit is expressed as

$$|\overline{\theta}^{\mathrm{noQED}}\rangle =$$



(A.7)

This circuit is combined with the recursive gate teleportation gadget:



(A.8)

A.3 Default noise parameters

Tabs. A.1 and A.2 list the default noise parameters employed in the H2 and Helios emulators, respectively, for the simulations. The one and two-qubit gate faults are modeled by anisotropic Pauli channels. The initialization of a qubit suffers from a bit-flip with the probability p_{init}. The readout fault flips a result from "1" to "0" with the probability $p_{\mathrm{r},1\to0}$, and from "0" to "1" with the probability $p_{\mathrm{r},0\to1}$. The memory noise on idling or moving qubits is modeled by the channel Eq. (3.34). See [63, 64] for a complete list of the noise models and corresponding parameters.

Noise model		Parameter
1-qubit-gate fault (p_1)		7.3×10^{-5}
two-qubit-gate fault (p_2)		1.29×10^{-3}
Initialization fault (p_{init})		4×10^{-5}
Readout fault	"1"→"0" ($p_{\mathrm{r},1\to0}$)	1.8×10^{-3}
	"0"→"1" ($p_{\mathrm{r},0\to1}$)	0.9×10^{-3}
Memory noise rate	coherent (f)	4.3×10^{-2}rad $\cdot$ s^{-1}
	incoherent (g)	2.8×10^{-3}s^{-1}

Table A.1: The default noise parameters used in our numerical simulations with the H2 emulator. The memory noise rates f and g are defined in Eq. (3.34).

Noise model		Parameter
1-qubit-gate fault (p_1)		2.5×10^{-5}
two-qubit-gate fault (p_2)		8.0×10^{-4}
Initialization fault (p_{init})		5×10^{-4}
Readout fault	"1"→"0" ($p_{\text{r},1\to0}$)	1.0×10^{-6}
	"0"→"1" ($p_{\text{r},0\to1}$)	1.0×10^{-6}
Memory noise rate	coherent (f)	5.8×10^{-2}rad $\cdot$ s^{-1}
	incoherent (g)	7.9×10^{-3}s^{-1}

Table A.2: The default noise parameters used in our numerical simulations with the Helios emulator. The memory noise rates f and g are defined in Eq. (3.34).

B Quantum Algorithms for Topological Data Analysis

B.1 Quantum Circuits for Laplacian Moments

The algorithm of Akhalwaya et al.[45] for estimating Laplacian Moments $\widetilde{T}_k^{(d)}$ for a given input graph $G = (V, E)$ which we presented in Section 4 relies on the modular circuit construction given in Figure 4.4. We present some technical details for understanding this circuit construction and the resource estimates in Section 4.3.2.

B.1.1 Representing cliques on qubits

The circuit operates on two registers: the main register and an auxilliary register.

Main register qubits In the standard implementation of this algorithm, the main register contains n qubits, with the ith qubit corresponding to the ith vertex of the graph under some ordering. There is a one-to-one mapping between computational basis states $|\mathbf{x}\rangle = |x_1 x_2 \dots x_n\rangle$ on this register and subsets $V_{\mathbf{x}} = \{v_i \in V | x_i = 1\}$ of vertices in the graph G. At the beginning of a shot of this circuit estimating $\widetilde{T}_{v,k}^{(d)}$ for some randomly chosen clique v in G, we prepare the computational basis state $|x_1 x_2 \dots x_n\rangle$ with $x_i = 1$ iff $v_i \in v$, we write this state as $|v\rangle$.

Auxilliary qubits To perform projection operators (red boxes in Figure 4.4), we create clean ancilla qubits, entangle them with qubits from the main register and then measure and post-select on desired outcomes to perform the projections. We provide more details on these projections in the next section. In the qubit counts provided in Section 4.3.2 we count the max number number of qubits required simulataneously at any point in the algorithm. In reality, the ancilla qubits will be continuously measured and replaced throughout the computation.

B.1.2 Circuit components

There are three components of the circuit in Figure 4.4. These are a unitary B which depends only on the number of nodes in the graph G and two projections P_{Γ_G} (onto the clique states of G) and $I - P_{\Gamma_G}$ onto the non-clique states which depend on the edges in the complement graph G^c. We describe them briefly in this section.

Unitary Boundary Operator As described by Akhalwaya et al. [65], the (unrestricted) Hermitian boundary operator can be written in terms of even fermionic Majorana operators $B = \frac{1}{\sqrt{n}} \sum_i \gamma_{2i}$. This gives rise to a simple v-shaped circuit for the the unitary $\frac{1}{\sqrt{n}} B$ in the Jordan-Wigner encoding shown in Figure B.1. In this project, we have explored alternative constructions of this operator and these feature in the resource estimates given in Section 4.3.2.

Clique projection The clique projection is defined as $P_{\Gamma_G} = \sum_{v \in \Gamma_G} |v\rangle \langle v|$ where v ranges over subsets of the vertices V of G and Γ_G is the clique complex of G containing all subsets v s.t. $x, y \in v \implies (x, y) \in E$. This can be broken up as a product of simpler projections for each of the edges in the complement graphs as $P_{\Gamma_G} = \prod_{e \in E(G^c)} P_e$ where the edge projection is defined in Figure B.2.

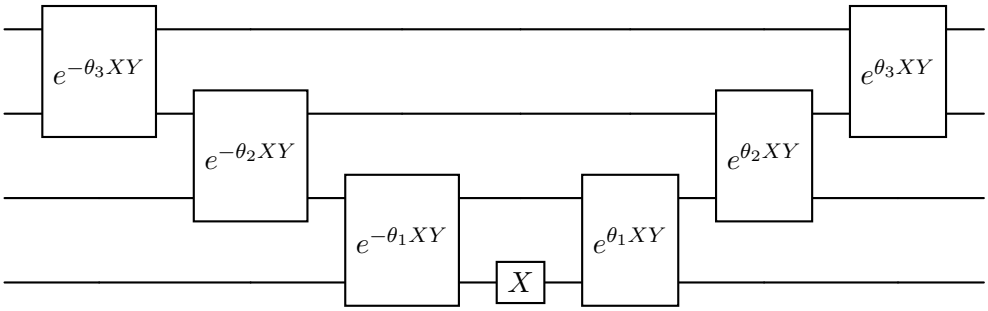


Figure B.1: V-shaped Unitary Boundary Operator under the Jordan-Wigner mapping from [65].

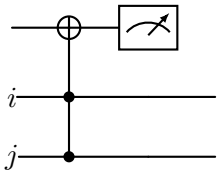


Figure B.2: The edge projection $P_{(i,j)}$ is performed by entangling the ith and jth qubits of the main register with a clean ancilla qubit and measuring. The projection is performed by post-selecting on the outcome 0.

Non-clique projection As noted in [45], to estimate the normalised Betti number we need both the clique projection and its complement $I - P_{\Gamma_G}$. In that work an unoptimised implementation of this circuit is given requiring $\mathcal{O}(n^2)$ ancilla qubits to perform the complement of P_{Γ_G} from the circuit for P_{Γ_G}. The same effect can be achieved by introducing a count register of size $2 \times \lceil \log(n) \rceil$ and replacing the Toffolis from P_{Γ_G} with controlled additions to this register, performing uncomputation as you go. Then project onto non-zero states of the count register and uncompute. We incorporate resources for an $O(n)$-ancilla qubit, optimised version of this in Section 4.3.2.

B.2 Laplacian features for graph-based tasks

Having introduced the Laplacian moments and the machine learning pipeline (see section 4.4) from ego-graphs to classification we now present evidence that the Laplacian moments make good features for mildy structure graph-generation models. The statistical framework we employ to investigate the usefulness of the Laplacian moments in a machine learning pipeline is Bayesian statistics and Mutual Information (see B.3.3 for definitions). This is not only a principled and fundamental approach to studying arbitrary correlations between input and output variables, it also gives us, in the form of Bayes' Law a method to extract predictive value and anticipate and explain the improvement in accuracies we might see from the ML pipeline.

B.2.1 Heterophily and the Stochastic Block Model (SBM)

We began by investigating how Betti numbers and more generally the vector of Laplacian moments correlate with SBM graph structure. In particular, we are interested in the correlations between Betti numbers and the degree of heterophily. Heterophily is the notion of the frequency with which dissimilar labelled nodes are connected to each [66]. There are many reasons to be interested in heterophily and the topic has been extensively studied in the field of Graph Neural Networks[66]. We use it as our starting point to investigate the usefulness of Betti numbers. Our setup is to assume that graphs are generated under a stochastic block model process with unknown parameters and our task is to learn those parameters [67]. The stochastic block model posits a partitioning of graph nodes with fixed probabilities assigned to the occurrence of edges within and between partitions. In

the case of the same inter- and intra- edge probabilities being used for all pairs of partitions, we get the so called Planted Clique model [68].

B.2.2 Predicting the SBM parameters

Using the Bayesian perspective, with the observed data being the Betti numbers of a drawn graph, we can write the probability of the model's parameters in terms of our prior and the probability of the Betti number given different model settings. The latter allows us to run experiments and empirically visualise how the Betti numbers change as a function of model parameters, the more 'invertible' and rapidly changing the landscape the more valuable the information (which could be formalised with the notion of Fisher information).

$$\arg\max_{\mathcal{M}} \Pr[\mathcal{M} \mid \overrightarrow{\beta}] = \arg\max_{\mathcal{M}} \frac{\Pr[\overrightarrow{\beta} \mid \mathcal{M}] \Pr[\mathcal{M}]}{\Pr[\overrightarrow{\beta}]} = \arg\max_{\mathcal{M}} \Pr[\overrightarrow{\beta} \mid \mathcal{M}] \Pr[\mathcal{M}]$$

In Figure B.4, we can visually see how β_1-β_5 correlates well with the generating process' parameters. Any one Betti number has regions of the parameter plane were the variability is poor, but taken together, a full set of observed betti numbers can 'place' ourselves on the SBM parameter plane well.

B.2.3 Predicting the number of clusters

While the SBM experiment begins to demonstrate the value of Betti numbers, we still need to build up a proper use-case. To this end, we can imagine the following two step process. First we have a random process where we choose the number of partitions/clusters and then we have an SBM process with *known* inter and intra- edge probabilities. Now the task is to estimate the number of clusters. This situation can be found in the real-world where the mechanism generating edges is understood (for example, biological assay statistics) but the number of operative clusters is the sought after piece of information (for example, the number of sets of interacting genes). Indeed clustering raw data is a well known, highly important primitive in unsupervised machine learning.

$$\text{argmax}_k P(k|G, \overrightarrow{\beta}) = \text{argmax}_k P(G, \overrightarrow{\beta}|k)P(k)$$

In Figure B.5, we can see how the Laplacian moments of random graphs generated by the above two-step process distribute. Each figure corresponds to a different power of the Laplacian moment. Each 3D datapoint in each figure corresponds to a triple of Laplacian moments of orders one through three of a randomly drawn graph from the two-step process. The colour of the datapoint corresponds to the setting of the number of clusters in the SBM that generated that graph. The main insight is that even though different clusters have overlapping Laplacian moments in any one dimension, taken together (across orders, for some high power), the clusters begin to separate.

B.2.4 Mutual Information Results

The separation of the marginally overlapping clusters constitutes a wonderful example of the usefulness of joint mutual information. If one were to naively focus on the mutual information of any one input variable with the number of clusters, one would get a mediocre result. However, the joint mutual information grows to a high value because each extra variable jointly brings additional discriminative power. In Figure B.3, we can see the mutual information increase as we add the laplacian moments from higher order. Going to higher order (and higher power) is the regime of classical intractable.

B.3 Laplacian features for graph-based tasks: Figures and Definitions

B.3.1 Predicting the SBM parameters

The figures B.4 that demonstrate the value of Betti numbers for predicting SBM parameters should be understood collectively. Taken together, a given graph's sequence of Betti numbers will make regions of the SBM parameter space more likely.

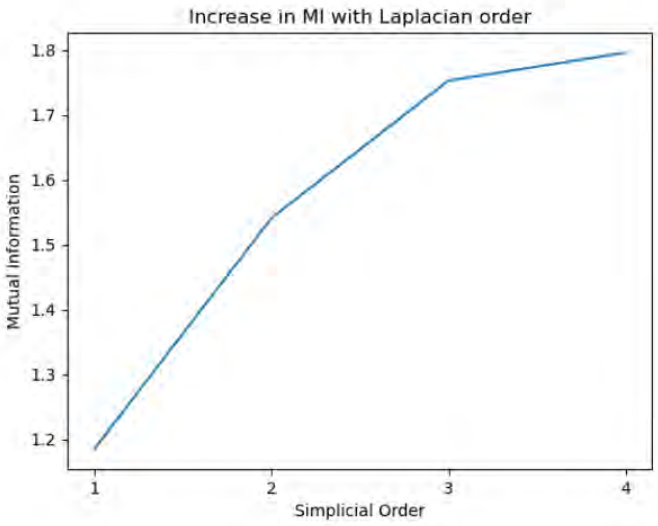


Figure B.3: Mutual Information between k and growing set of Laplacian moments

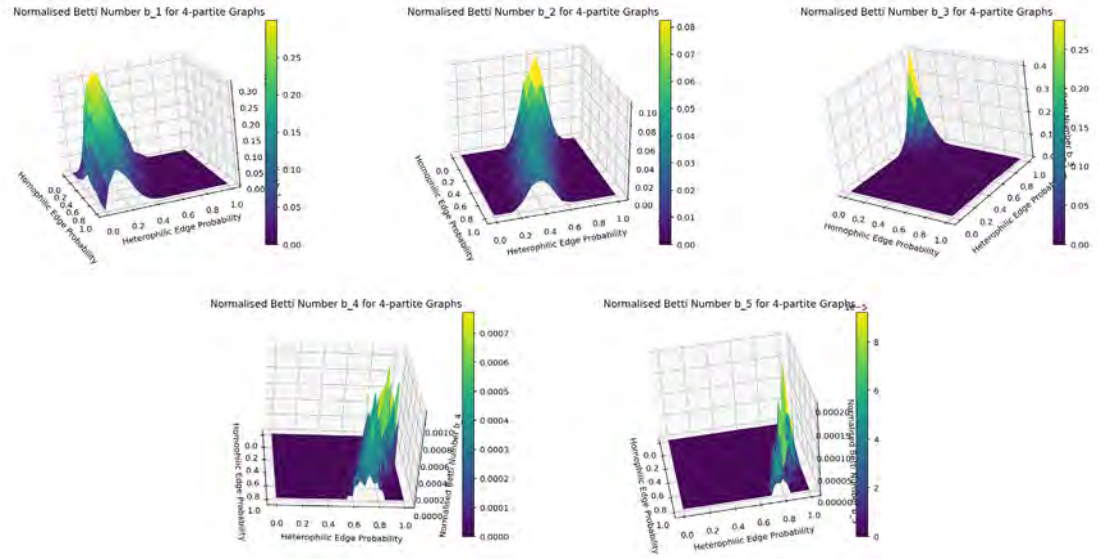


Figure B.4: Average Betti$_i$ vs SBM parameters, for $i = 1, \ldots, 5$.

B.3.2 Predicting the number of clusters Figures

The figures B.5 demonstrating how cluster sizes correlate with Laplacian moments should be read by noticing that the colours are separable in moment space. An extra observation is how even though there is some redundancy in the discriminatory power of moments of different powers, the pattern changes as the power increases and for very high power the points start to separate less well. This shows that for some applications, intermediate powers may be more valuable than Betti numbers.

B.3.3 Bayes' Law

Bayes' law from a Bayesian statistics point of view, is a simple theorem building on Kolmogorov's axioms of probability theory that allow a rational agent to update the parameters of their model of the world given new experimental data.

Bayes's law simply states that the probability of a model (called the posterior), with it's internal parameters set to specific values, given data from experiments, can be calculated from the predictions made by the model, multiplied by our initial probability of the model (called the prior), divided by the

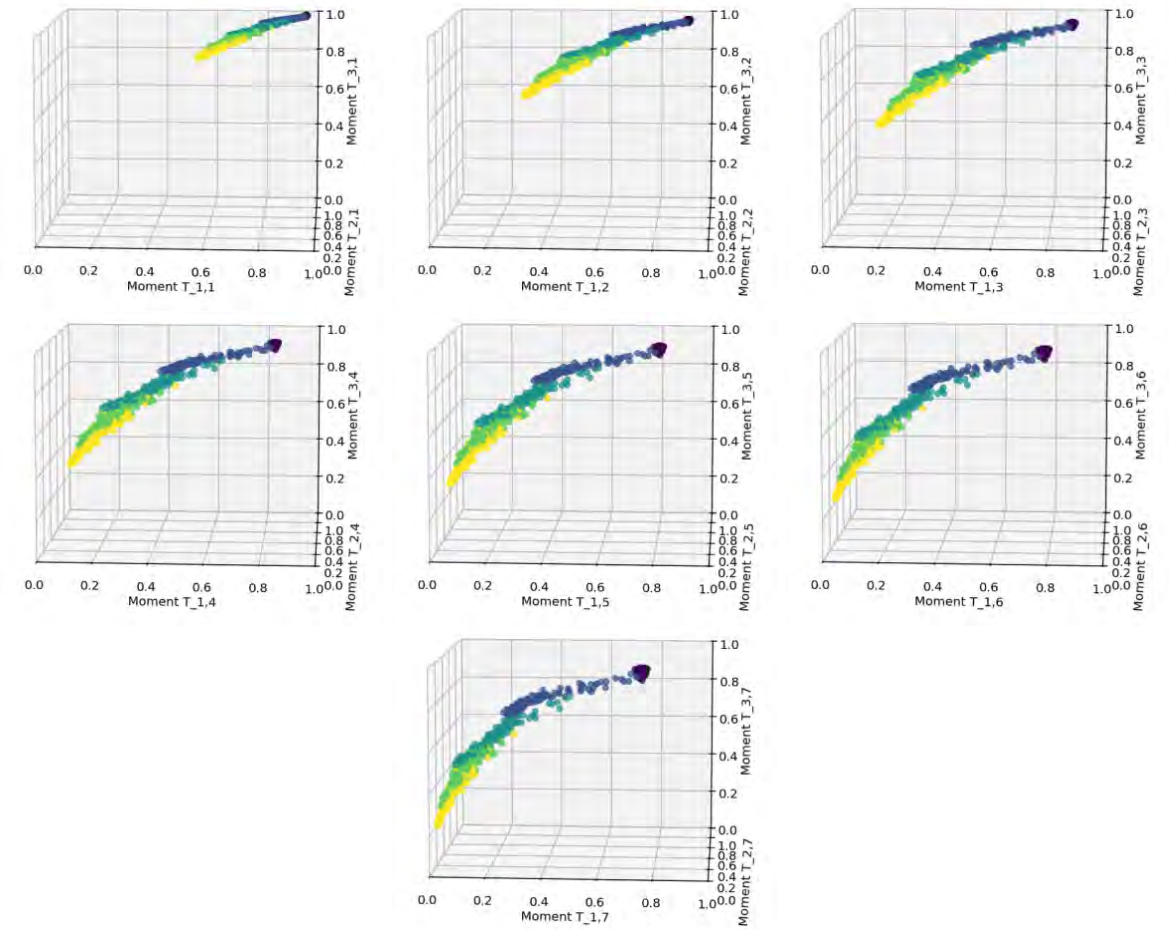


Figure B.5: k-colour-coded Laplacian Moment (order 1-3, power p) for $p = 1, 2, \ldots, 7$

probability of the data:

$$\Pr[\mathcal{M} \mid D] = \frac{\Pr[D \mid \mathcal{M}] \Pr[\mathcal{M}]}{\Pr[D]}$$

In our investigations of the usefulness of Laplacian moments, our model, $\mathcal{M}$, produces a probability distribution over graphs. Graphs drawn from an 'unknown' distribution and their properties constitute the data that we are trying to model.

B.3.4 Mutual Information

While Bayes' law allows us to think about modelling the correlations in our data, mutual information is a model-agnostic measure of the correlations present in the data waiting to be learnt. Given the feature space X and label space Y, the mutual information is given by

$$\mathrm{MI} = \sum_{x \in X, y \in Y} \Pr[x, y] \log_2 \frac{\Pr[x, y]}{\Pr[x] \Pr[y]}$$

Here for example X would be the laplacian moments and Y would be some label that we would like to predict. This quantity can be estimated by taking a subset of the data set as the "test dataset" and empirically approximating the joint probability with an advanced 'histogram' method. Difficulties with the last step is the reason why mutual information does not feature more in machine learning research. For example, in the space of natural images it would be intractable to approximate the mutual information between pixels and the label of the picture. For us, things are more manageable since we have a small set of features that we are interested in studying, namely the mutual information.

Bibliography

[1] Kentaro Yamamoto, Yuta Kikuchi, David Amaro, Ben Criger, Silas Dilkes, Ciarán Ryan-Anderson, Andrew Tranter, Joan M. Dreiling, Dan Gresh, Cameron Foltz, Michael Mills, Steven A. Moses, Peter E. Siegfried, Maxwell D. Urmey, Justin J. Burau, Aaron Hankin, Dominic Lucchetti, John P. Gaebler, Natalie C. Brown, Brian Neyenhuis, and David Muñoz Ramo. Quantum Error-Corrected Computation of Molecular Energies, September 2025.

[2] David R. Yarkony. Diabolical conical intersections. *Reviews of Modern Physics*, 68(4):985–1013, October 1996.

[3] Trygve Helgaker, Poul Jorgensen, and Jeppe Olsen. *Molecular electronic-structure theory*. John Wiley & Sons, 2013.

[4] Andrew Tranter, Ciaran Ward, Cono Di Paola, David Muñoz Ramo, David Zsolt Manrique, Duncan Gowland, Evgeny Plekhanov, Gabriel Greene-Diniz, Georgia Christopoulou, Georgia Prokopiou, Harry D J Keen, Iakov Polyak, Irfan T Khan, Jerzy Pilipczuk, Josh J M Kirsopp, Kentaro Yamamoto, Kripa Panchagnula, Ludmila Szulakowska, Maria Tudorovskaya, Michal Krompiec, Michelle Sze, Nathan Fitzpatrick, Oliver Backhouse, Regina Finsterhoelzl, Robert J Anderson, and Vardhini Bhasker. InQuanto: Quantum Computational Chemistry, 2022.

[5] Qiming Sun, Timothy C. Berkelbach, Nick S. Blunt, George H. Booth, Sheng Guo, Zhendong Li, Junzi Liu, James D. McClain, Elvira R. Sayfutyarova, Sandeep Sharma, Sebastian Wouters, and Garnet Kin-Lic Chan. PySCF: The Python-based simulations of chemistry framework. *WIREs Computational Molecular Science*, 8(1):e1340, 2018.

[6] Krysta M. Svore, Matthew B. Hastings, and Michael Freedman. Faster phase estimation. *Quantum Information and Computation*, 14(3&4):306–328, March 2014.

[7] Kenji Sugisaki, Chikako Sakai, Kazuo Toyota, Kazunobu Sato, Daisuke Shiomi, and Takeji Takui. Bayesian phase difference estimation: A general quantum algorithm for the direct calculation of energy gaps. *Physical Chemistry Chemical Physics*, 23(36):20152–20162, September 2021.

[8] Shu Kanno, Kenji Sugisaki, Hajime Nakamura, Hiroshi Yamauchi, Rei Sakuma, Takao Kobayashi, Qi Gao, and Naoki Yamamoto. Tensor-based quantum phase difference estimation for large-scale demonstration. *Proceedings of the National Academy of Sciences*, 122(30):e2425026122, July 2025.

[9] Kenji Sugisaki, V. S. Prasannaa, Satoshi Ohshima, Takahiro Katagiri, Yuji Mochizuki, B. K. Sahoo, and B. P. Das. Bayesian phase difference estimation algorithm for direct calculation of fine structure splitting: Accelerated simulation of relativistic and quantum many-body effects. *Electronic Structure*, 5(3):035006, September 2023.

[10] Andrew Steane. Multiple-particle interference and quantum error correction. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 452(1954):2551–2577, January 1997.

[11] C Ryan-Anderson, J G Bohnet, K Lee, D Gresh, A Hankin, J P Gaebler, D Francois, A Chernoguzov, D Lucchetti, N C Brown, T M Gatterman, S K Halit, K Gilmore, J A Gerber, B Neyenhuis, D Hayes, and R P Stutz. Realization of real-time fault-tolerant quantum error correction. *Phys. Rev. X*, 11(4):041058, December 2021.

[12] Riki Toshio, Yutaro Akahoshi, Jun Fujisaki, Hirotaka Oshima, Shintaro Sato, and Keisuke Fujii. Practical quantum advantage on partially fault-tolerant quantum computer, August 2024.

[13] S. A. Moses, C. H. Baldwin, M. S. Allman, R. Ancona, L. Ascarrunz, C. Barnes, J. Bartolotta, B. Bjork, P. Blanchard, M. Bohn, J. G. Bohnet, N. C. Brown, N. Q. Burdick, W. C. Burton, S. L. Campbell, J. P. Campora, C. Carron, J. Chambers, J. W. Chan, Y. H. Chen, A. Chernoguzov, E. Chertkov, J. Colina, J. P. Curtis, R. Daniel, M. DeCross, D. Deen, C. Delaney, J. M. Dreiling, C. T. Ertsgaard, J. Esposito, B. Estey, M. Fabrikant, C. Figgatt, C. Foltz, M. Foss-Feig, D. Francois, J. P. Gaebler, T. M. Gatterman, C. N. Gilbreth, J. Giles, E. Glynn, A. Hall, A. M. Hankin, A. Hansen, D. Hayes, B. Higashi, I. M. Hoffman, B. Horning, J. J. Hout, R. Jacobs, J. Johansen, L. Jones, J. Karcz, T. Klein, P. Lauria, P. Lee, D. Liefer, S. T. Lu, D. Lucchetti, C. Lytle, A. Malm, M. Matheny, B. Mathewson, K. Mayer, D. B. Miller, M. Mills, B. Neyenhuis, L. Nugent, S. Olson, J. Parks, G. N. Price, Z. Price, M. Pugh, A. Ransford, A. P. Reed, C. Roman, M. Rowe, C. Ryan-Anderson, S. Sanders, J. Sedlacek, P. Shevchuk, P. Siegfried, T. Skripka, B. Spaun, R. T. Sprenkle, R. P. Stutz, M. Swallows, R. I. Tobey, A. Tran, T. Tran, E. Vogt, C. Volin, J. Walker, A. M. Zolot, and J. M. Pino. A Race-Track Trapped-Ion Quantum Processor. *Physical Review X*, 13(4):041052, December 2023.

[14] Quantinuum H2-2, Helios-1E. https://www.quantinuum.com/. Jauary 12–March 27, 2026.

[15] Anthony Ransford, M. S. Allman, Jake Arkinstall, J. P. Campora III, Samuel F. Cooper, Robert D. Delaney, Joan M. Dreiling, Brian Estey, Caroline Figgatt, Alex Hall, Ali A. Husain, Akhil Isanaka, Colin J. Kennedy, Nikhil Kotibhaskar, Ivaylo S. Madjarov, Karl Mayer, Alistair R. Milne, Annie J. Park, Adam P. Reed, Riley Ancona, Molly P. Andersen, Pablo Andres-Martinez, Will Angenent, Liz Argueta, Benjamin Arkin, Leonardo Ascarrunz, William Baker, Corey Barnes, John Bartolotta, Jordan Berg, Ryan Besand, Bryce Bjork, Matt Blain, Paul Blanchard, Robin Blume-Kohout, Matt Bohn, Agustin Borgna, Daniel Y. Botamanenko, Robert Boutelle, Natalie Brown, Grant T. Buckingham, Nathaniel Q. Burdick, William Cody Burton, Varis Carey, Christopher J. Carron, Joe Chambers, John Children, Victor E. Colussi, Steven Crepinsek, Andrew Cureton, Joe Davies, Daniel Davis, Matthew DeCross, David Deen, Conor Delaney, Davide DelVento, B. J. DeSalvo, Jason Dominy, Ross Duncan, Vanya Eccles, Alec Edgington, Neal Erickson, Stephen Erickson, Christopher T. Ertsgaard, Bruce Evans, Tyler Evans, Maya I. Fabrikant, Andrew Fischer, Cameron Foltz, Michael Foss-Feig, David Francois, Brad Freyberg, Charles Gao, Robert Garay, Jane Garvin, David M. Gaudiosi, Christopher N. Gilbreth, Josh Giles, Erin Glynn, Jeff Graves, Azure Hansen, David Hayes, Lukas Heidemann, Bob Higashi, Tyler Hilbun, Jordan Hines, Ariana Hlavaty, Kyle Hoffman, Ian M. Hoffman, Craig Holliman, Isobel Hooper, Bob Horning, James Hostetter, Daniel Hothem, Jack Houlton, Jared Hout, Ross Hutson, Ryan T. Jacobs, Trent Jacobs, Melf Johannsen, Jacob Johansen, Loren Jones, Sydney Julian, Ryan Jung, Aidan Keay, Todd Klein, Mark Koch, Ryo Kondo, Chang Kong, Asa Kosto, Alan Lawrence, David Liefer, Michelle Lollie, Dominic Lucchetti, Nathan K. Lysne, Christian Lytle, Callum MacPherson, Andrew Malm, Spencer Mather, Brian Mathewson, Daniel Maxwell, Lauren McCaffrey, Hannah McDougall, Robin Mendoza, Michael Mills, Richard Morrison, Louis Narmour, Nhung Nguyen, Lora Nugent, Scott Olson, Daniel Ouellette, Jeremy Parks, Zach Peters, Jessie Petricka, Juan M. Pino, Frank Polito, Matthias Preidl, Gabriel Price, Timothy Proctor, McKinley Pugh, Noah Ratcliff, Daisy Raymondson, Peter Rhodes, Conrad Roman, Craig Roy, Ciaran Ryan-Anderson, Fernando Betanzo Sanchez, George Sangiolo, Tatiana Sawadski, Andrew Schaffer, Peter Schow, Jon Sedlacek, Henry Semenenko, Peter Shevchuk, Susan Shore, Peter Siegfried, Kartik Singhal, Seyon Sivarajah, Thomas Skripka, Lucas Sletten, Ben Spaun, R. Tucker Sprenkle, Paul Stoufer, Mariel Tader, Stephen F. Taylor, Travis H. Thompson, Raanan Tobey, Anh Tran, Tam Tran, Grahame Vittorini, Curtis Volin, Jim Walker, Sam White, Douglas Wilson, Quinn Wolf, Chester Wringe, Kevin Young, Jian Zheng, Kristen Zuraski, Charles H. Baldwin, Alex Chernoguzov, John P. Gaebler, Steven J. Sanders, Brian Neyenhuis, Russell Stutz, and Justin G. Bohnet. Helios: A 98-qubit trapped-ion quantum computer. https://arxiv.org/abs/2511.05465v1, November 2025.

[16] Lorenza Viola, Emanuel Knill, and Seth Lloyd. Dynamical Decoupling of Open Quantum Systems. *Physical Review Letters*, 82(12):2417–2421, March 1999.

[17] Quantinuum. Dynamical decoupling on quantinuum systems. https://docs.quantinuum.com/systems/trainings/h2/getting_started/dd.html, 2026. Accessed: 2026-03-26.

[18] Michael A Perlin, Zichang He, Anthony Alexiades Armenakas, Pablo Andres-Martinez, Tianyi Hao, Dylan Herman, Yuwei Jin, Karl Mayer, Chris Self, David Amaro, Ciaran Ryan-Anderson, and Ruslan Shaydulin. Fault-tolerant execution of error-corrected quantum algorithms. *arXiv [quant-ph]*, 2026.

[19] C. Ryan-Anderson, N. C. Brown, M. S. Allman, B. Arkin, G. Asa-Attuah, C. Baldwin, J. Berg, J. G. Bohnet, S. Braxton, N. Burdick, J. P. Campora, A. Chernoguzov, J. Esposito, B. Evans, D. Francois, J. P. Gaebler, T. M. Gatterman, J. Gerber, K. Gilmore, D. Gresh, A. Hall, A. Hankin, J. Hostetter, D. Lucchetti, K. Mayer, J. Myers, B. Neyenhuis, J. Santiago, J. Sedlacek, T. Skripka, A. Slattery, R. P. Stutz, J. Tait, R. Tobey, G. Vittorini, J. Walker, and D. Hayes. Implementing Fault-tolerant Entangling Gates on the Five-qubit Code and the Color Code, August 2022.

[20] A. Paetznick, M. P. da Silva, C. Ryan-Anderson, J. M. Bello-Rivas, J. P. Campora, A. Chernoguzov, J. M. Dreiling, C. Foltz, F. Frachon, J. P. Gaebler, T. M. Gatterman, L. Grans-Samuelsson, D. Gresh, D. Hayes, N. Hewitt, C. Holliman, C. V. Horst, J. Johansen, D. Lucchetti, Y. Matsuoka, M. Mills, S. A. Moses, B. Neyenhuis, A. Paz, J. Pino, P. Siegfried, A. Sundaram, D. Tom, S. J. Wernli, M. Zanner, R. P. Stutz, and K. M. Svore. Demonstration of logical qubits and repeated error correction with better-than-physical error rates, November 2024.

[21] Ben W. Reichardt, David Aasen, Rui Chao, Alex Chernoguzov, Wim van Dam, John P. Gaebler, Dan Gresh, Dominic Lucchetti, Michael Mills, Steven A. Moses, Brian Neyenhuis, Adam Paetznick, Andres Paz, Peter E. Siegfried, Marcus P. da Silva, Krysta M. Svore, Zhenghan Wang, and Matt Zanner. Demonstration of quantum computation and error correction with a tesseract code, December 2024.

[22] C. Ryan-Anderson, N. C. Brown, C. H. Baldwin, J. M. Dreiling, C. Foltz, J. P. Gaebler, T. M. Gatterman, N. Hewitt, C. Holliman, C. V. Horst, J. Johansen, D. Lucchetti, T. Mengle, M. Matheny, Y. Matsuoka, K. Mayer, M. Mills, S. A. Moses, B. Neyenhuis, J. Pino, P. Siegfried, R. P. Stutz, J. Walker, and D. Hayes. High-fidelity teleportation of a logical qubit using transversal gates and lattice surgery. *Science*, 385(6715):1327–1331, September 2024.

[23] Rajeev Acharya, Laleh Aghababaie-Beni, Igor Aleiner, Trond I. Andersen, Markus Ansmann, Frank Arute, Kunal Arya, Abraham Asfaw, Nikita Astrakhantsev, Juan Atalaya, Ryan Babbush, Dave Bacon, Brian Ballard, Joseph C. Bardin, Johannes Bausch, Andreas Bengtsson, Alexander Bilmes, Sam Blackwell, Sergio Boixo, Gina Bortoli, Alexandre Bourassa, Jenna Bovaird, Leon Brill, Michael Broughton, David A. Browne, Brett Buchea, Bob B. Buckley, David A. Buell, Tim Burger, Brian Burkett, Nicholas Bushnell, Anthony Cabrera, Juan Campero, Hung-Shen Chang, Yu Chen, Zijun Chen, Ben Chiaro, Desmond Chik, Charina Chou, Jahan Claes, Agnetta Y. Cleland, Josh Cogan, Roberto Collins, Paul Conner, William Courtney, Alexander L. Crook, Ben Curtin, Sayan Das, Alex Davies, Laura De Lorenzo, Dripto M. Debroy, Sean Demura, Michel Devoret, Agustin Di Paolo, Paul Donohoe, Ilya Drozdov, Andrew Dunsworth, Clint Earle, Thomas Edlich, Alec Eickbusch, Aviv Moshe Elbag, Mahmoud Elzouka, Catherine Erickson, Lara Faoro, Edward Farhi, Vinicius S. Ferreira, Leslie Flores Burgos, Ebrahim Forati, Austin G. Fowler, Brooks Foxen, Suhas Ganjam, Gonzalo Garcia, Robert Gasca, Élie Genois, William Giang, Craig Gidney, Dar Gilboa, Raja Gosula, Alejandro Grajales Dau, Dietrich Graumann, Alex Greene, Jonathan A. Gross, Steve Habegger, John Hall, Michael C. Hamilton, Monica Hansen, Matthew P. Harrigan, Sean D. Harrington, Francisco J. H. Heras, Stephen Heslin, Paula Heu, Oscar Higgott, Gordon Hill, Jeremy Hilton, George Holland, Sabrina Hong, Hsin-Yuan Huang, Ashley Huff, William J. Huggins, Lev B. Ioffe, Sergei V. Isakov, Justin Iveland, Evan Jeffrey, Zhang Jiang, Cody Jones, Stephen Jordan, Chaitali Joshi, Pavol Juhas, Dvir Kafri, Hui Kang,

Amir H. Karamlou, Kostyantyn Kechedzhi, Julian Kelly, Trupti Khaire, Tanuj Khattar, Mostafa Khezri, Seon Kim, Paul V. Klimov, Andrey R. Klots, Bryce Kobrin, Pushmeet Kohli, Alexander N. Korotkov, Fedor Kostritsa, Robin Kothari, Borislav Kozlovskii, John Mark Kreikebaum, Vladislav D. Kurilovich, Nathan Lacroix, David Landhuis, Tiano Lange-Dei, Brandon W. Langley, Pavel Laptev, Kim-Ming Lau, Loïck Le Guevel, Justin Ledford, Kenny Lee, Yuri D. Lensky, Shannon Leon, Brian J. Lester, Wing Yan Li, Yin Li, Alexander T. Lill, Wayne Liu, William P. Livingston, Aditya Locharla, Erik Lucero, Daniel Lundahl, Aaron Lunt, Sid Madhuk, Fionn D. Malone, Ashley Maloney, Salvatore Mandrá, Leigh S. Martin, Steven Martin, Orion Martin, Cameron Maxfield, Jarrod R. McClean, Matt McEwen, Seneca Meeks, Anthony Megrant, Xiao Mi, Kevin C. Miao, Amanda Mieszala, Reza Molavi, Sebastian Molina, Shirin Montazeri, Alexis Morvan, Ramis Movassagh, Wojciech Mruczkiewicz, Ofer Naaman, Matthew Neeley, Charles Neill, Ani Nersisyan, Hartmut Neven, Michael Newman, Jiun How Ng, Anthony Nguyen, Murray Nguyen, Chia-Hung Ni, Thomas E. O'Brien, William D. Oliver, Alex Opremcak, Kristoffer Ottosson, Andre Petukhov, Alex Pizzuto, John Platt, Rebecca Potter, Orion Pritchard, Leonid P. Pryadko, Chris Quintana, Ganesh Ramachandran, Matthew J. Reagor, David M. Rhodes, Gabrielle Roberts, Eliott Rosenberg, Emma Rosenfeld, Pedram Roushan, Nicholas C. Rubin, Negar Saei, Daniel Sank, Kannan Sankaragomathi, Kevin J. Satzinger, Henry F. Schurkus, Christopher Schuster, Andrew W. Senior, Michael J. Shearn, Aaron Shorter, Noah Shutty, Vladimir Shvarts, Shraddha Singh, Volodymyr Sivak, Jindra Skruzny, Spencer Small, Vadim Smelyanskiy, W. Clarke Smith, Rolando D. Somma, Sofia Springer, George Sterling, Doug Strain, Jordan Suchard, Aaron Szasz, Alex Sztein, Douglas Thor, Alfredo Torres, M. Mert Torunbalci, Abeer Vaishnav, Justin Vargas, Sergey Vdovichev, Guifre Vidal, Benjamin Villalonga, Catherine Vollgraff Heidweiller, Steven Waltman, Shannon X. Wang, Brayden Ware, Kate Weber, Theodore White, Kristi Wong, Bryan W. K. Woo, Cheng Xing, Z. Jamie Yao, Ping Yeh, Bicheng Ying, Juhwan Yoo, Noureldin Yosri, Grayson Young, Adam Zalcman, Yaxing Zhang, Ningfeng Zhu, and Nicholas Zobrist. Quantum error correction below the surface code threshold. *Nature*, 638(8052):920–926, February 2025.

[24] Dolev Bluvstein, Simon J. Evered, Alexandra A. Geim, Sophie H. Li, Hengyun Zhou, Tom Manovitz, Sepehr Ebadi, Madelyn Cain, Marcin Kalinowski, Dominik Hangleiter, J. Pablo Bonilla Ataides, Nishad Maskara, Iris Cong, Xun Gao, Pedro Sales Rodriguez, Thomas Karolyshyn, Giulia Semeghini, Michael J. Gullans, Markus Greiner, Vladan Vuletić, and Mikhail D. Lukin. Logical quantum processor based on reconfigurable atom arrays. *Nature*, 626(7997):58–65, February 2024.

[25] Quantinuum. Quantinuum System Model Noise Model. https://docs.quantinuum.com/systems/user_guide/emulator_user_guide/noise_model.html, 2026. Accessed: 2026-03-26.

[26] David Cohen-Steiner, Herbert Edelsbrunner, and John Harer. Stability of persistence diagrams. In *Proceedings of the twenty-first annual symposium on Computational geometry*, pages 263–271, 2005.

[27] Ippei Obayashi, Takenobu Nakamura, and Yasuaki Hiraoka. Persistent homology analysis for materials research and persistent homology software: Homcloud. *journal of the physical society of japan*, 91(9):091013, 2022.

[28] Takenobu Nakamura, Yasuaki Hiraoka, Akihiko Hirata, Emerson G Escolar, and Yasumasa Nishiura. Persistent homology and many-body atomic structure for medium-range order in the glass. *Nanotechnology*, 26(30):304001, 2015.

[29] Yasuaki Hiraoka, Takenobu Nakamura, Akihiko Hirata, Emerson G Escolar, Kaname Matsue, and Yasumasa Nishiura. Hierarchical structures of amorphous solids characterized by persistent homology. *Proceedings of the National Academy of Sciences*, 113(26):7035–7040, 2016.

[30] Edelsbrunner, Letscher, and Zomorodian. Topological persistence and simplification. *Discrete & computational geometry*, 28(4):511–533, 2002.

[31] Michael Moor, Max Horn, Bastian Rieck, and Karsten Borgwardt. Topological autoencoders. In *International conference on machine learning*, pages 7045–7054. PMLR, 2020.

[32] Robin Vandaele, Bo Kang, Jefrey Lijffijt, Tijl De Bie, and Yvan Saeys. Topologically regularized data embeddings. *arXiv preprint arXiv:2110.09193*, 2021.

[33] Rickard Brüel Gabrielsson, Bradley J Nelson, Anjan Dwaraknath, and Primoz Skraba. A topology layer for machine learning. In *International Conference on Artificial Intelligence and Statistics*, pages 1553–1563. PMLR, 2020.

[34] Christoph Hofer, Florian Graf, Bastian Rieck, Marc Niethammer, and Roland Kwitt. Graph filtration learning. In *International Conference on Machine Learning*, pages 4314–4323. PMLR, 2020.

[35] Henry Adams, Tegan Emerson, Michael Kirby, Rachel Neville, Chris Peterson, Patrick Shipman, Sofya Chepushtanova, Eric Hanson, Francis Motta, and Lori Ziegelmeier. Persistence images: A stable vector representation of persistent homology. *Journal of Machine Learning Research*, 18(8):1–35, 2017.

[36] Christoph Hofer, Roland Kwitt, Marc Niethammer, and Andreas Uhl. Deep learning with topological signatures. *Advances in neural information processing systems*, 30, 2017.

[37] Peter Bubenik. The persistence landscape and some of its properties. In *Topological Data Analysis: The Abel Symposium 2018*, pages 97–117. Springer, 2020.

[38] Frédéric Chazal, Brittany Terese Fasy, Fabrizio Lecci, Alessandro Rinaldo, and Larry Wasserman. Stochastic convergence of persistence landscapes and silhouettes. In *Proceedings of the thirtieth annual symposium on Computational geometry*, pages 474–483, 2014.

[39] Sara Kališnik. Tropical coordinates on the space of persistence barcodes. *Foundations of Computational Mathematics*, 19(1):101–129, 2019.

[40] Dominic W. Berry, Yuan Su, Casper Gyurik, Robbie King, Joao Basso, Alexander Del Toro Barba, Abhishek Rajput, Nathan Wiebe, Vedran Dunjko, and Ryan Babbush. Analyzing prospects for quantum advantage in topological data analysis. *PRX Quantum*, 5:010319, Feb 2024.

[41] Daniel Lütgehetmann, Dejan Govc, Jason P Smith, and Ran Levi. Computing persistent homology of directed flag complexes. *Algorithms*, 13(1):19, 2020.

[42] Alexander Schmidhuber and Seth Lloyd. Complexity-theoretic limitations on quantum algorithms for topological data analysis. *PRX Quantum*, 4(4):040349, 2023.

[43] Chris Cade and P Marcos Crichigno. Complexity of supersymmetric systems and the cohomology problem. *Quantum*, 8:1325, 2024.

[44] Marcos Crichigno and Tamara Kohler. Clique homology is qma 1-hard. *Nature Communications*, 15(1):9846, 2024.

[45] Ismail Yunus Akhalwaya, Ahmed Bhayat, Adam Connolly, Steven Herbert, Lior Horesh, Julien Sorci, and Shashanka Ubaru. Comparing quantum and classical Monte Carlo algorithms for estimating Betti numbers of clique complexes. *Quantum*, 9:1901, October 2025.

[46] Simon Apers, Sander Gribling, Sayantan Sen, and Dániel Szabó. A (simple) classical algorithm for estimating Betti numbers. *Quantum*, 7:1202, December 2023.

[47] Julien Sorci. Analyzing and improving a classical betti number estimation algorithm, 2025.

[48] Communications Fraud Control Association (CFCA). Telecommunications fraud increased 12% in 2023 equating to an estimated $38.95 billion lost to fraud, 2023. Accessed: 2026-03-23.

[49] Ming Liu, Jianxin Liao, Jingyu Wang, and Qi Qi. Agrm: Attention-based graph representation model for telecom fraud detection. In *ICC 2019-2019 IEEE International Conference on Communications (ICC)*, pages 1–6. IEEE, 2019.

[50] Xinxin Hu, Hongchang Chen, Shuxin Liu, Haocong Jiang, Guanghan Chu, and Ran Li. Btg: A bridge to graph machine learning in telecommunications fraud detection. *Future Gener. Comput. Syst.*, 137(C):274 – 287, December 2022.

[51] Xinxin Hu, Haotian Chen, Junjie Zhang, Hongchang Chen, Shuxin Liu, Xing Li, Yahui Wang, and Xiangyang Xue. Gat-cobo: Cost-sensitive graph neural network for telecom fraud detection. *IEEE Transactions on Big Data*, 10(4):528–542, 2024.

[52] Juncheng Yang, Shuxia Li, Zijun Huang, and Junhang Wu. An improve fraud detection framework via dynamic representations and adaptive frequency response filter. *Scientific Reports*, 15(1):19051, 2025.

[53] An Tong, Bochao Chen, Zhe Wang, Jiawei Gao, and Chi Kin Lam. Gdfgat: Graph attention network based on feature difference weight assignment for telecom fraud detection. *Plos one*, 20(5):e0322004, 2025.

[54] Satoshi Nakazato, Tsutomu Takizawa, and Tatsuo Arai. Photoisomerization and energy transfer in naphthalene-terminated stilbene dendrimers. *Photochem. Photobiol. Sci.*, 11:885–888, 2012.

[55] Hong-Bo Cheng, Shuchun Zhang, Enying Bai, Xiaoqiao Cao, Jiaqi Wang, Ji Qi, Jun Liu, Jing Zhao, Liqun Zhang, and Juyoung Yoon. Future-oriented advanced diarylethene photoswitches: From molecular design to spontaneous assembly systems. *Adv. Mat.*, 34(16):2108289, 2022.

[56] Samraj Mollick and Jin-Chong Tan. Organic solid-state photochromism using porous scaffolds. *Nat. Rev. Mat.*, 10:519–535, 2025.

[57] Eunji Jin, Volodymyr Bon, Shubhajit Das, A. D. Dinga Wonanke, Martin Etter, Martin A. Karlsen, Ankita De, Nadine Bönisch, Thomas Heine, and Stefan Kaskel. Engineering photoswitching dynamics in 3d photochromic metal – organic frameworks through a metal – organic polyhedron design. *J. Am. Chem. Soc.*, 147(10):8568–8577, 2025.

[58] Spiridoula Matsika. Electronic structure methods for the description of nonadiabatic effects and conical intersections. *Chem. Rev.*, 121(15):9407–9449, 2021.

[59] Nicolas Ferré, Michael Filatov, and Miquel Huix-Rotllant. *Density-Functional Methods for Excited States*. Springer Cham, 2016.

[60] John M. Herbert and Aniket Mandal. *Time-Dependent Density Functional Theory*, chapter Spin-Flip TDDFT for Photochemistry. Jenny Stanford Publishing, 2022.

[61] Noriyuki Minezawa and Mark S. Gordon. Photoisomerization of stilbene: A spin-flip density functional theory approach. *J. Phys. Chem. A*, 115(27):7901–7911, 2011.

[62] Hayato Goto. Minimizing resource overheads for fault-tolerant preparation of encoded states of the steane code. *Sci. Rep.*, 6:19578, January 2016.

[63] Quantinuum. System Model H2 Emulators. https://docs.quantinuum.com/systems/user_guide/emulator_user_guide/emulators/h2_emulators.html, 2025. Accessed: 2026-03-27.

[64] Quantinuum. Quantinuum Helios. https://docs.quantinuum.com/systems/user_guide/emulator_user_guide/emulators/helios_emulators.html, 2025. Accessed: 2026-03-27.

[65] Ismail Yunus Akhalwaya, Yang-Hui He, Lior Horesh, Vishnu Jejjala, William Kirby, Kugendran Naidoo, and Shashanka Ubaru. Representation of the fermionic boundary operator. *Phys. Rev. A*, 106:022407, Aug 2022.

[66] Xin Zheng, Yi Wang, Yixin Liu, Ming Li, Miao Zhang, Di Jin, Philip S. Yu, and Shirui Pan. Graph Neural Networks for Graphs with Heterophily: A Survey, February 2024. arXiv:2202.07082 [cs].

[67] Tiago P. Peixoto. Bayesian stochastic blockmodeling. pages 289–332. November 2019. arXiv:1705.10225 [stat].

[68] Lizhi Zhang and Tiago P. Peixoto. Statistical inference of assortative community structures. *Physical Review Research*, 2(4):043271, November 2020. Publisher: American Physical Society.

Contributors

Adam Connolly	Quantinuum
Ahmed Bhayat	Quantinuum
Cahit Kargı	Quantinuum
Colin Krawchuk	Quantinuum
David Muñoz Ramo	Quantinuum
Georgia Prokopiou	Quantinuum
Hiroki Goto	SoftBank Corp.
Ifan Williams	Quantinuum
Ismail Akhalwaya	Quantinuum
Kentaro Yamamoto	Quantinuum
Kevin Xie	Quantinuum
Nikhil Khatri	Quantinuum
Pranav Kalidindi	Quantinuum
Rei Nishimura	SoftBank Corp.
Satoshi Kanno	SoftBank Corp.
Shintaro Maekawa	Quantinuum
Simon McAdams	Quantinuum
Steve Clark	Quantinuum
Thomas M. Soini	Quantinuum
Yoshi-aki Shimada	SoftBank Corp.


SoftBank
QUANTINUUM