Quantum K-means to cluster particles at hadron colliders like CERN LHC

M.S THESIS

A Thesis submitted to the thesis committee of
the Department of Theoretical Physics of University of Dhaka
as partial requirements for the Degree of
M.S

EXAM ROLL NO.: 2338505

REG. NO.: 2022018348

Session : 2022-2023

Department of Theoretical Physics

University of Dhaka

Year: 2025

ABSTRACT

ABSTRACT

This thesis investigates the integration of quantum machine learning techniques, particularly quantum K-means clustering, into the challenging task of jet reconstruction at the Hadron Colliders like (LHC). As modern high-energy physics experiments transition into the high-luminosity era, the volume and complexity of collision data have grown exponentially, placing considerable strain on existing classical data analysis pipelines. Jet reconstruction—central to identifying quark and gluon signatures—relies heavily on clustering algorithms that must be both fast and accurate in high-dimensional feature spaces. Classical K-means clustering, while widely used, suffers from scalability issues as data dimensionality and volume increase. To address these challenges, this work explores the theoretical foundations and practical implementations of quantum-enhanced clustering, leveraging the quantum K-means algorithm as a potential solution. The thesis provides a rigorous formulation of the quantum distance estimation subroutines, such as the `SwapTest`, and outlines the translation of classical data into quantum states using amplitude encoding techniques. The computational framework is implemented using Qiskit, IBM's open-source quantum software development kit, allowing both simulation on ideal quantum backends and execution on noisy intermediate-scale quantum (NISQ) devices. Empirical evaluations are performed on both synthetic datasets, constructed to mimic the kinematic properties of LHC jet events, and real collision data samples where possible. The performance of the quantum K-means algorithm is benchmarked against its classical counterpart across multiple axes: clustering inertia, execution time, scalability with number of data points, and sensitivity to angular resolution—an important factor in jet substructure analysis. The results indicate that quantum K-means clustering exhibits promising computational characteristics, particularly in reducing the time complexity of pairwise distance calculations, which dominate classical clustering costs. Moreover, for data types common in high-energy physics—such as sparse, high-dimensional feature vectors representing calorimeter hits—the quantum approach demonstrates a capacity to maintain accuracy while potentially achieving sublinear scaling in distance evaluations. Although current quantum hardware limitations restrict large-scale deployment, the simulations presented in this work provide compelling evidence that quantum machine learning techniques, when matured, could play a transformative role in the future of particle physics data analysis. This thesis contributes to the growing body of evidence suggesting that hybrid quantum-classical algorithms can serve as practical accelerators for core tasks in scientific computing, particularly in domains with structured, high-dimensional data.

TABLE OF CONTENTS

ACKNOWLEDGMENTS

I want to express my heartfelt gratitude to my supervisor. With his guidance, I was able to complete the thesis paper. He dedicated his time to meet with me, providing valuable insights to tackle specific challenges.

Last but not least, I want to thank my father, whose absence took a great toll on me. Now that he is here, my head and heart are in the right place.

This work is dedicated to my father.

LIST OF TABLES

LIST OF FIGURES

QUANTUM K-MEANS TO CLUSTER PARTICLES AT HADRON COLLIDERS LIKE CERN LHC

CHAPTER 1

Introduction

1.1 From Classical Foundations to Quantum Clustering in Particle Physics

The story of physics is, at its heart, a story of abstraction and unification. From the earliest human attempts to understand the motion of stars to the complex mathematics governing the quantum interactions of subatomic particles, physics has evolved through revolutions that redefined not just our models of the universe, but our very notions of reality. This thesis, situated at the intersection of quantum computing and particle physics, is a continuation of this journey — an attempt to explore how the abstract formalism of quantum algorithms can aid in decoding the chaotic remnants of particle collisions at the world's powerful Hadron Colliders like LHC.

1.1.1 The Clockwork Universe: Classical Mechanics and Its Legacy

In the 17th century, Isaac Newton laid the foundation of what came to be known as classical mechanics [1]. By synthesizing the empirical observations of Galileo and Kepler

into a unified framework, Newton introduced a set of deterministic laws that governed the motion of objects under the influence of forces [2]. The universe, under this model, was akin to a mechanical clock — precise, predictable, and reducible to a set of equations. Newton's Principia not only revolutionized physics but also shaped Enlightenment thinking about causality, determinism, and the ability of human reason to decode natural law.

As the centuries progressed, classical mechanics expanded into other domains: Maxwell's equations in the 19th century unified electricity and magnetism into a coherent electrodynamic [3]. Thermodynamics and statistical mechanics began to explain macroscopic behavior from the interactions of microscopic constituents. By the late 1800s, many physicists believed that the essential structure of physics had been discovered, with only minor corrections remaining.

1.1.2 The Breakdowns: Toward a Quantum World

However, the 20th century brought a series of experimental results that could not be reconciled with classical models. The photoelectric [4], blackbody radiation [5], atomic [6], and the stability of matter all resisted classical explanation. These puzzles led to the birth of quantum theory — initially as a collection of ad hoc fixes — but eventually as a full-blown paradigm shift.

Planck's quantization [7], Einstein's photon hypothesis [8], Bohr's model of the atom [9], and ultimately Schrödinger's wave mechanics and Heisenberg's matrix mechanics [10] laid the groundwork for modern quantum mechanics. In this framework, determinism gave way to probability; the observer became an inseparable part of the physical system; and physical quantities were represented not as functions on a phase space but as operators on a Hilbert space.

Quantum mechanics provided accurate predictions for microscopic phenomena but

also introduced deep philosophical quandaries — the measurement problem, nonlocality, and the collapse of the wavefunction. Despite these challenges, quantum theory quickly became indispensable in describing atomic, molecular, and condensed matter systems.

1.1.3 Quantum Field Theory and the Standard Model

While non-relativistic quantum mechanics sufficed for atoms and molecules, high-energy processes demanded a framework consistent with special relativity. This gave rise to quantum field theory (QFT), where particles are understood as excitations of underlying quantum fields [11]. Interactions between particles are mediated by field quanta — photons for electromagnetic interactions, gluons for strong force, and W/Z bosons for weak interactions.

The mid-to-late 20th century witnessed the formulation of the Standard Model of particle physics — a gauge-invariant quantum field theory describing three of the four known fundamental interactions (excluding gravity). The unification of the electroweak interaction by Glashow, Weinberg, and Salam, and the establishment of QCD as the theory of strong interactions, allowed physicists to compute scattering amplitudes, decay rates, and cross-sections with astounding [12] [13].

But with this success came complexity. The mathematical machinery of renormalization, Feynman diagrams, running coupling constants, and spontaneous symmetry breaking made the theory both powerful and computationally demanding. Understanding what occurs in high-energy particle collisions is not merely a matter of applying theory, but of deciphering vast, noisy datasets to reconstruct which particles were produced, how they decayed, and what they tell us about the underlying laws.

1.1.4 The Standard Model: Particles and Interactions

The Standard Model (SM) of particle physics is a quantum field theory that describes the known fundamental particles and their interactions, excluding gravity. It is based on the gauge symmetry group:

$$SU(3)_C \times SU(2)_L \times U(1)_Y,$$

which governs the strong, weak, and electromagnetic forces, respectively. [14, 15, 16]

Fermions: Quarks and Leptons

Matter in the SM is composed of fermions, which are spin-$\frac{1}{2}$ particles that obey Fermi-Dirac statistics. These fermions come in two broad classes:

- **Quarks:** Up (u), down (d), charm (c), strange (s), top (t), and bottom (b) quarks. Quarks carry color charge and interact via the strong force mediated by gluons.
- **Leptons:** Electron (e), muon (μ), tau (τ), and their corresponding neutrinos (ν_e, ν_μ, ν_τ). Neutrinos interact only weakly, making them difficult to detect.

Fermions are arranged in three generations, with increasing mass. Only the first generation is stable, and the heavier generations decay quickly into the lighter ones.

Bosons: Force Carriers

Interactions between particles are mediated by bosons:

- **Photon (γ):** Mediates electromagnetic interactions.
- **Gluons (g):** Eight massless particles mediating the strong force between color-charged quarks.

- **W and Z bosons ($W^{\pm}, Z^0$):** Massive carriers of the weak force responsible for processes like beta decay.
- **Higgs boson (H):** Scalar particle responsible for spontaneous symmetry breaking and giving mass to the W, Z, and fermions through Yukawa couplings.

Unification and Open Questions

The SM unifies the electromagnetic and weak forces into the electroweak theory and offers extremely precise predictions, validated by decades of collider experiments [17]. However, it does not include gravity, dark matter, or neutrino masses in a natural way, leaving room for extensions such as supersymmetry (SUSY), Grand Unified Theories (GUTs), or theories with extra dimensions.

1.1.5 Detection of Particles at the LHC

The detection of fundamental particles at the Large Hadron Collider (LHC) relies on multilayered detector systems designed to track, identify, and measure particles created during high-energy collisions. The two major general-purpose experiments, ATLAS and CMS, employ nearly hermetic detectors structured in cylindrical layers surrounding the beam pipe.

Detector Geometry and Hermetic Design

Modern general-purpose detectors at the LHC, such as ATLAS and CMS, are built with a cylindrical, onion-like structure surrounding the beam pipe, designed to be nearly hermetic (4π)—i.e., they cover almost the entire solid angle around the interaction point [18, 19]. This architecture ensures that particles produced in proton–proton collisions are detected in as many directions as possible, minimizing blind spots and improving physics

completeness.

Both detectors are structured into:

- **Barrel region:** Cylindrical layers concentric with the beam axis, including the inner tracker, ECAL, HCAL, all enclosed within magnet coils.
- **End-cap regions:** Disk-shaped layers perpendicular to the beam axis at both ends, closing coverage in the forward and backward directions.

The innermost detectors—such as the silicon pixel and strip trackers—are mounted just outside the beam pipe and are responsible for precisely reconstructing charged-particle tracks and decay vertices [20]. Further out, electromagnetic and hadronic calorimeters are layered to fully absorb and measure the energy of photons, electrons, and hadrons. Muon spectrometers occupy the outermost region, tracking muons that pass through all internal layers [18, 21].

Such a layered, hermetic design enables accurate reconstruction of 4-momenta and missing energy (MET) in events containing neutrinos or other invisible particles, while also reducing uninstrumented gaps through which particles could escape undetected.

1. Tracking Systems

Charged particles leave ionization tracks in silicon-based tracking detectors located closest to the interaction point. These tracks allow for precise determination of momentum via curvature in a magnetic field:

$$p_T = qB\rho,$$

where ρ is the radius of curvature, B is the magnetic field strength, and q is the particle charge. [22]

2. Calorimeters

Beyond the tracker, two types of calorimeters are used to measure energy deposits:

- **Electromagnetic Calorimeter (ECAL):** Detects electrons and photons via showering and scintillation.
- **Hadronic Calorimeter (HCAL):** Measures energy of hadrons (mostly jets) through hadronic cascades and absorption.

The calorimeter resolution typically follows:

$$\frac{\sigma_E}{E} = \frac{a}{\sqrt{E}} \oplus b,$$

where a is the stochastic term, and b is the constant term. [23]

3. Muon Chambers

Located in the outermost layer, muon spectrometers detect minimally interacting muons using drift tubes, resistive plate chambers, or cathode strip chambers. Muons can pass through the calorimeters with minimal energy loss, making them easily distinguishable.

4. Trigger Systems and Data Acquisition

Due to high collision rates ($\sim$ 40 MHz), hardware (L1) and software (HLT) trigger systems select interesting events in real-time. Only $\sim 10^3$ events per second are retained for storage and analysis.

5. Reconstruction of Four-Momenta

Using hits and energy deposits across detector layers, software algorithms reconstruct the 4-momentum $(E, \vec{p})$ of particles. For jets and missing energy (MET), these reconstructions use clustering algorithms, particle flow methods, and calorimeter towers. A reconstructed jet is often defined as:

$$p^{\mu}_{\text{jet}} = \sum_{i} p^{\mu}_{i},$$

where the sum runs over all constituent particles or calorimeter deposits associated with the jet.[24]

These reconstructed objects serve as the primary input to high-level physics analysis, including resonance searches, differential cross-section measurements, and decay topology classification.

1.1.6 Signatures in the Detector: Jets, Leptons, and Missing Energy

High-energy particle collisions at the LHC result in a spray of particles, each of which interacts differently with the detector subsystems. From these interactions, distinct "signatures" are reconstructed. The most relevant for Standard Model (SM) and Beyond the Standard Model (BSM) physics are:

1. Jets

Jets are collimated sprays of hadrons originating from the hadronization of high-energy quarks and gluons. Since color confinement prevents direct observation of quarks or gluons, they manifest as hadronic jets in the detector. Jets are reconstructed using

clustering algorithms such as anti-k_T with a distance parameter R (typically $R = 0.4$ or 0.6).

$$d_{ij} = \min(p_{Ti}^{-2}, p_{Tj}^{-2}) \frac{\Delta R_{ij}^2}{R^2}, \quad \Delta R^2 = (\Delta\eta)^2 + (\Delta\phi)^2. \tag{1.1}$$

The equation 1.1 can be found in the literalture. [25] Jets leave significant deposits in the HCAL and minimal in the ECAL. They are central to processes involving QCD, top quarks, and Higgs boson decays to hadrons.

2. Isolated Leptons (Electrons and Muons)

Electrons deposit energy primarily in the ECAL and are accompanied by a charged track in the tracker. Muons, on the other hand, pass through both calorimeters and leave signals in the muon chambers. Lepton isolation is key to distinguishing them from jets — isolated leptons are surrounded by little additional activity within a cone $\Delta R < 0.3$.[26]

3. Missing Transverse Energy (MET)

When particles like neutrinos escape detection, their presence is inferred via an imbalance in the total transverse energy:

$$\vec{E}_T^{\text{miss}} = -\sum_i \vec{p}_{Ti}^{\text{vis}}. \tag{1.2}$$

Significant MET is a strong signature for processes involving neutrinos (e.g., $W \to l\nu$) or possible BSM particles like neutralinos in supersymmetry. [27]

4. Composite Final States and Event Topologies

Physics analyses often focus not on individual signatures but on event-level topologies — e.g., dijet events, dilepton + MET, or multijet + b-tagged jets + MET. Reconstructing

these composite final states is a non-trivial task due to detector resolution, pileup, and misidentification.

These signatures are passed into clustering and classification algorithms — such as K-means and its quantum analogs — to infer underlying decay chains, particle identities, and event classes.

1.1.7 What Are Jets?

In high-energy particle physics, **jets** refer to collimated sprays of hadrons that are produced when high-energy quarks or gluons undergo hadronization. These partons are typically created in the final states of high-energy proton-proton collisions, such as those occurring at the LHC. Due to the property of color confinement in Quantum Chromodynamics (QCD), quarks and gluons cannot exist as free particles and instead transform into color-neutral hadrons that propagate in approximately the same direction as their parent parton, forming a jet.

Jet identification is a cornerstone of modern collider physics and is fundamental in reconstructing processes involving hadronic final states. Jet reconstruction techniques use clustering algorithms, such as the anti-k_T, k_T, or Cambridge-Aachen algorithms, to group energy deposits in calorimeters or charged particle tracks into single jet objects.

Jets are typically described by four-momentum vectors:

$$p^\mu = (E, p_x, p_y, p_z),$$

or equivalently in cylindrical coordinates using transverse momentum p_T, pseudorapidity η, and azimuthal angle ϕ. Jets are analyzed at both the detector-level and particle-level, and their substructure can reveal underlying physics processes, such as the presence of boosted heavy particles (e.g., top quarks, W/Z bosons, Higgs bosons). Jets are crucial

for identifying partonic event topology, reconstructing decay chains, and distinguishing between different Standard Model and Beyond Standard Model (BSM) processes. Techniques such as jet substructure analysis, grooming algorithms, and tagging methods are now standard tools in collider experiments [14, 15, 16].

1.1.8 Collider Physics and the Challenge of Jet Reconstruction

At the LHC, protons are accelerated to near-light speeds and collided at center-of-mass energies of up to 13–14 TeV. These collisions produce a spray of particles in the detectors, some of which arise from the hadronization of high-energy quarks and gluons. Reconstructing these jets from the raw detector hits is a central challenge in collider data analysis. [28]

Jets are not fundamental objects but emergent structures arising from QCD dynamics. They must be inferred from spatial and energy distributions in calorimeters, tracker hits, and timing measurements. Jet algorithms typically cluster final-state particles or energy deposits using geometric proximity in pseudorapidity–azimuthal angle space (η–ϕ), often with distance metrics inspired by transverse momentum (p_T) weighting.

The need to resolve overlapping jets, distinguish jets from pile-up, and preserve substructure information has led to the development of advanced algorithms, such as anti-k_T, k_T, and Cambridge–Aachen, as well as machine learning–based approaches. Classical clustering algorithms like K-means, while not standard in operational pipelines, serve as intuitive and computational benchmarks. Yet, their complexity — especially the repeated evaluation of Euclidean distances across large, high-dimensional datasets — makes them inefficient at scale.

1.1.9 Machine Learning and the New Computational Paradigm

As data rates from LHC experiments reach petabytes per second before triggering, traditional rule-based pipelines are increasingly being augmented by machine learning (ML) techniques. Supervised classifiers, unsupervised clustering, anomaly detection, and generative models are all being deployed to extract physics-relevant information from detector data.

K-means clustering, as one of the most fundamental unsupervised learning algorithms, partitions data into k clusters by minimizing intra-cluster variance. However, its performance is strongly limited by the cost of computing distances between points and centroids. This bottleneck becomes particularly pronounced in high-dimensional feature spaces — such as those derived from calorimeter maps or jet substructure observables.

The growing interest in quantum machine learning (QML) arises from the possibility of encoding classical data into quantum states and leveraging quantum circuits to compute distances, overlaps, or kernel functions more efficiently than is classically feasible.

1.1.10 Jet Clustering with Quantum K-means: Motivation and Objectives

The application of quantum K-means clustering to jet reconstruction is motivated by both the structure of jet data and the cost of clustering. Calorimeter data, for instance, can be naturally represented as sparse high-dimensional vectors. Moreover, the task of clustering particles into jets aligns well with unsupervised methods, where explicit labels may not be available.

This thesis investigates whether quantum algorithms can offer computational advantages in this domain. Specifically, we explore:

- The construction of quantum circuits for estimating Euclidean distances between LHC-

style data points;

- The effectiveness of amplitude encoding for calorimeter-like inputs;
- The scaling behavior of quantum K-means compared to classical K-means in terms of runtime and cluster accuracy;
- The fidelity of quantum inner product estimation using simulators and noisy intermediate-scale quantum (NISQ) devices.

By implementing the algorithm using Qiskit and evaluating it on both synthetic and real collider data, this work aims to offer insights into the feasibility of integrating quantum machine learning into the LHC data analysis pipeline.

1.1.11 Thesis Structure

This document is structured as follows:

- **Chapter 3** presents the theoretical framework, including classical and quantum distance metrics, the `SwapTest` protocol, and encoding strategies.
- **Chapter 4** details the quantum K-means algorithm, its circuit-level implementation, and integration with Qiskit simulators. item

CHAPTER 2

The Standard Model and Fundamental Particles

The Standard Model (SM) of particle physics is a quantum field theory that describes all known elementary particles and their interactions—excluding gravity—based on the gauge group:

$$SU(3)_C \times SU(2)_L \times U(1)_Y$$

This symmetry encapsulates the three fundamental interactions: the strong, weak, and electromagnetic forces. The SM has proven remarkably accurate in describing experimental results across high-energy physics experiments, including those at the Large Hadron Collider (LHC) [29, 30].

2.1 Fermions: Quarks and Leptons

Fermions, which obey Fermi-Dirac statistics and the Pauli exclusion principle, constitute the matter content of the universe. They are divided into two families: **quarks** and **leptons**, each occurring in three generations [31].

2.1.1 Quarks

Quarks are color-charged particles that interact via the strong force. There are six flavors:

$$\text{up (u), down (d), charm (c), strange (s), top (t), bottom (b)}$$

Quarks carry fractional electric charges ($+\frac{2}{3}$ or $-\frac{1}{3}$) and combine to form composite particles like protons and neutrons [32]. Due to confinement, quarks are never observed in isolation but only within hadrons [33].

2.1.2 Leptons

Leptons do not carry color charge and do not participate in strong interactions. They include:

$$e^-, \mu^-, \tau^- \quad \text{and their neutrinos} \quad \nu_e, \nu_\mu, \nu_\tau$$

Charged leptons interact via the electromagnetic and weak forces, while neutrinos interact only weakly [34]. The discovery of neutrino oscillations provided evidence for non-zero neutrino masses, indicating physics beyond the Standard Model [35, 36].

2.2 Bosons: Force Carriers

Interactions between fermions are mediated by bosons—integer-spin particles. The SM contains the following bosons:

* **Photon (γ)**: Mediates electromagnetic interactions [37].
* **$W^{\pm}$ and Z bosons**: Mediate weak interactions; they are massive [38].
* **Gluons (g)**: Eight massless bosons mediating the strong force between quarks via color charge [39].
* **Higgs boson (H)**: Scalar particle responsible for mass generation via spontaneous symmetry breaking [40].

2.3 Hadrons: Composite Particles

Hadrons are bound states of quarks held together by the strong force. They are categorized as:

2.3.1 Baryons

Baryons are composed of three quarks (qqq). The most well-known examples are:

Proton: uud Neutron: udd

Baryons carry a baryon number of +1 and include excited states such as the Δ and Λ particles [41].

2.3.2 Mesons

Mesons consist of a quark-antiquark pair ($q\bar{q}$). Examples include the pion (π), kaon (K), and rho meson (ρ). They are bosonic in nature and often act as force carriers in low-energy nuclear interactions [42].

2.4 Gauge Symmetries and Interactions

Each fundamental force corresponds to a gauge symmetry:

* **QED (Electromagnetism)**: $U(1)_Y$ with photon as the gauge boson.
* **Weak force**: $SU(2)_L$ mediated by $W^{\pm}$ and Z bosons.
* **QCD (Strong interaction)**: $SU(3)_C$ mediated by eight gluon.

These gauge groups are unified in the electroweak sector, which has been experimentally validated by numerous precision measurements.[43, 44]

2.5 Higgs Mechanism and Mass Generation

The Higgs mechanism gives mass to weak gauge bosons through spontaneous symmetry breaking. The Higgs field acquires a vacuum expectation value (VEV), leading to the emergence of the Higgs boson [45, 46].

$$\mathcal{L}_{\text{Higgs}} = (D_\mu \Phi)^\dagger (D^\mu \Phi) - V(\Phi)$$

$$V(\Phi) = \mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2$$

The discovery of the Higgs boson in 2012 at the LHC was a pivotal confirmation of the Standard Model [47, 48].

2.6 Beyond the Standard Model (BSM)

While the SM is experimentally robust, it leaves several questions unanswered:

* Origin of dark matter and dark energy .
* Neutrino masses and mixing
* Baryon asymmetry and CP violation
* Gravity's integration into quantum theory.

BSM theories include Supersymmetry (SUSY), Grand Unified Theories (GUTs), and string theory [49, 50, 51, 52, 53, 54].

2.7 Experimental Evidence and Detection

Detectors such as ATLAS and CMS are designed with nearly hermetic, cylindrical layers to track, measure, and identify particles produced in high-energy collisions [55, 56]. These include:

* Tracking chambers for charged particles
* Calorimeters for energy measurements
* Muon chambers for identifying heavy leptons

Through jet reconstruction algorithms, such as the quantum K-means model studied in this thesis, fundamental particle properties can be inferred from clustered final-state objects.

2.8 Conclusion

The Standard Model stands as one of the most successful theories in physics, unifying fundamental interactions (except gravity) and accurately predicting particle properties and behaviors. However, it is not the final word—ongoing experiments, including those leveraging quantum computing, seek to probe physics beyond its framework.

CHAPTER 3

Theoretical Framework

In this chapter, we present the formal mathematical tools required for implementing quantum K-means clustering as applied to jet reconstruction at the LHC. The discussion is based directly on the derivations in [57], and includes classical distance metrics, the SwapTest subroutine for quantum inner product estimation, and amplitude encoding techniques for input representation.

3.0.1 SwapTest Derivation

The *SwapTest* is a quantum subroutine used to compute the inner product between two quantum states $|\psi\rangle$ and $|\phi\rangle$, i.e., $|\langle\psi|\phi\rangle|^2$. It uses an ancilla qubit along with a Hadamard and a controlled-SWAP gate.

We begin with an initial state:

$$|0\rangle\,|\psi\rangle\,|\phi\rangle \tag{3.1}$$

Let's define after applying Hadamard gate on $|0\rangle$ *and* $|1\rangle$ *weget* :

$$H|0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \tag{3.2}$$

$$H|1\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle) \tag{3.3}$$

Apply a Hadamard gate H to the first qubit:

$$H|0\rangle|\psi\rangle|\phi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)|\psi\rangle|\phi\rangle \Rightarrow \frac{1}{\sqrt{2}}(|0\rangle|\psi\rangle|\phi\rangle + |1\rangle|\psi\rangle|\phi\rangle) \tag{3.4}$$

Now, apply a controlled-SWAP gate (also called Fredkin gate), which swaps $|\psi\rangle$ and $|\phi\rangle$ if the control qubit is $|1\rangle$:

$$\frac{1}{\sqrt{2}}(|0\rangle|\psi\rangle|\phi\rangle + |1\rangle|\phi\rangle|\psi\rangle) \tag{3.5}$$

Apply another Hadamard gate to Eq. (3.5):

$$\frac{1}{2}[|0\rangle(|\psi\rangle|\phi\rangle + |\phi\rangle|\psi\rangle) + |1\rangle(|\psi\rangle|\phi\rangle - |\phi\rangle|\psi\rangle)] \tag{3.6}$$

We now measure the first (ancilla) qubit. The probability of obtaining outcome $|0\rangle$ is:

$$P_0 = \left|\frac{1}{2}(|\psi\rangle|\phi\rangle + |\phi\rangle|\psi\rangle)\right|^2 \tag{3.7}$$

$$= \frac{1}{4}|\,|\psi\rangle|\phi\rangle + |\phi\rangle|\psi\rangle\,|^2 \tag{3.8}$$

Expanding this inner product:

$$P_0 = \frac{1}{4}\left[\langle\psi|\psi\rangle\langle\phi|\phi\rangle + \langle\psi|\phi\rangle\langle\phi|\psi\rangle + \langle\phi|\psi\rangle\langle\psi|\phi\rangle + \langle\phi|\phi\rangle\langle\psi|\psi\rangle\right] \tag{3.9}$$

Assuming normalized states ($\langle\psi|\psi\rangle = \langle\phi|\phi\rangle = 1$):

$$P_0 = \frac{1}{2}\left(1 + |\langle\psi|\phi\rangle|^2\right) \tag{3.10}$$

Therefore, we obtain the final expression:

$$|\langle\psi|\phi\rangle|^2 = 2P_0 - 1 \tag{3.11}$$

This provides a direct quantum method to estimate the squared inner product between two quantum states using measurement statistics.

3.1 Euclidean Distance via `SwapTest`

The classical Euclidean distance between two d-dimensional real-valued vectors x_i and x_j is given by:

$$d_E^{(C)}(\mathbf{x_i}, \mathbf{x_j}) = |\mathbf{x_i} - \mathbf{x_j}| \tag{3.12}$$

$$|x_i\rangle = \frac{1}{|\mathbf{x}_i|}\sum_{\mu=1}^{d} x_{i,\mu}|\mu\rangle \tag{3.13}$$

A normalized quantum superposition is constructed, and the inner product between the two states is estimated using the `SwapTest`, which uses an ancillary qubit and controlled-swap gate logic.

3.1.1 Quantum Euclidean Distance via Entangled States and SwapTest Probability

The probability of measuring the ancillary qubit in the state $|0\rangle$ using the SwapTest subroutine is given by:

$$P(|0\rangle) = \frac{1}{2} + \frac{1}{2}|\langle x_i|x_j\rangle|^2 \tag{3.14}$$

This probability can be used to reconstruct the squared inner product between quantum states, which encodes geometric information about the underlying classical vectors. To define the **quantum Euclidean distance** between two d-dimensional classical vectors $\mathbf{x}_i$ and $\mathbf{x}_j$, we first encode them into their corresponding normalized quantum states $|x_i\rangle$ and $|x_j\rangle$.

We then construct the following entangled auxiliary states:

$$|\psi_1\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle\,|x\rangle_i + |1\rangle\,|x_j\rangle\right), \tag{3.15}$$

$$|\psi_2\rangle = \frac{1}{\sqrt{Z_{ij}}}\left(|\mathbf{x}_i\rangle\,|0\rangle - |\mathbf{x}_j\rangle\,|1\rangle\right), \tag{3.16}$$

where the normalization factor is defined as:

$$Z_{ij} = |\mathbf{x}_i|^2 + |\mathbf{x}_j|^2. \tag{3.17}$$

Therefore, from Eqs. (3.14) and some calculation we get:

$$d_E^{(Q)}(\mathbf{x}_i, \mathbf{x}_j) = \sqrt{2Z_{ij}\left(2P_{\Psi_3}(|0\rangle) - 1\right)}. \tag{3.18}$$

This quantity, $d_E^{(Q)}$, serves as the **quantum version of the Euclidean distance**.[57]

It is efficiently computed through a quantum circuit using a single-qubit measurement and captures geometric distances through fidelity-based quantum encodings. Where,

$$|x_i - x_j|^2 = (d_E^{(Q)}(x_i, x_j))^2 \tag{3.19}$$

This shows that the squared Euclidean distance appears naturally in the overlap of these entangled quantum states. Now, applying the SwapTest on the constructed state $|\psi_3\rangle$ (a circuit involving these states), the probability of measuring the ancilla in the state $|0\rangle$ in the final step is likewise

$$P_{\Psi_3}(|0\rangle) = \frac{1}{2}\left(1 + \langle\psi_1'|\psi_2|\psi_1'|\psi_2\rangle \cdot \langle\psi_2|\psi_1|\psi_2|\psi_1\rangle\right). \tag{3.20}$$

Solving for the quantum Euclidean distance:

$$d_E^{(Q)}(x_i, x_j) = \sqrt{2Z_{ij}\left(2P_{\Psi_3}(|0\rangle) - 1\right)}. \tag{3.21}$$

Here:

* $d_E^{(Q)}$ denotes the quantum version of the Euclidean distance,
* $Z_{ij} = \|x_i\|^2 + \|x_j\|^2$ is the normalization factor for $|\psi_2\rangle$,
* $P_{\Psi_3}(|0\rangle)$ is the probability of measuring the ancilla qubit in state $|0\rangle$ after running the quantum circuit that implements the three-step SwapTest procedure with states $|\psi_1\rangle, |\psi_2\rangle, |\psi_1'\rangle$.

This expression allows us to reconstruct the classical squared Euclidean distance between two vectors using entangled quantum states and a single-qubit measurement.

$$d_E^{(Q)}(x_i, x_j) = \sqrt{2Z_{ij}\left(2P(|0\rangle) - 1\right)} \tag{3.22}$$

where $Z_{ij} = \|x_i\|^2 + \|x_j\|^2$ is the normalization constant derived from the norms of the input vectors.

3.2 Quantum Distance Estimation

Minkowski Invariant Distance via `SwapTest`

Vectors in high-energy physics are defined in a four-dimensional space-time with the Minkowski metric. They are typically of the form:

$$x_i = (x_i^0, \vec{x}_i)$$

where x_i^0 is the temporal component and $\vec{x}_i$ represents the three spatial components. We generalize this to d dimensions, with $d-1$ being spatial.

The analog of the classical Euclidean distance in Minkowski space is given by the *invariant mass squared*, also referred to as the *invariant sum squared*:

$$s_{ij}^{(C)} = (x_i^0 + x_j^0)^2 - |\vec{x}_i + \vec{x}_j|^2 \tag{3.23}$$

SwapTest Details and Qubit Cost The `SwapTest` is a quantum subroutine used to estimate the inner product between two quantum states. In our context, it is used to compute distances between vectors (spatial and temporal) by measuring the overlap between their amplitude-encoded quantum representations.

Let $|\psi_1\rangle$ and $|\psi_2\rangle$ be quantum states encoding two vectors. The SwapTest circuit introduces an ancillary qubit, initializes it in the state $|0\rangle$, and applies a Hadamard gate. Then, controlled-SWAP (Fredkin) gates are used to conditionally swap the qubits of $|\psi_1\rangle$

and $|\psi_2\rangle$. After another Hadamard gate on the ancilla, a measurement reveals the probability of observing $|0\rangle$, which encodes the inner product via:

$$P(|0\rangle) = \frac{1}{2} + \frac{1}{2}|\langle\psi_1|\psi_2\rangle|^2$$

We assume that the quantum states $|\psi_1\rangle$ and $|\psi_2\rangle$ have been loaded from a quantum random access memory (qRAM) structure. For spatial data in $d-1$ dimensions, this requires $\mathcal{O}(\log(d-1))$ qubits to represent the data efficiently. That is, a 3D spatial vector is encoded in $\log_2(3)$ qubits (rounded up to 2).

For temporal components, which are scalars (1D), the encoding uses a single amplitude-encoded qubit. Therefore, the SwapTest applied to the time components requires only $\mathcal{O}(1)$ overhead.

Spatial Distance Following the procedure in the Euclidean case, we define the following quantum state:

$$|\psi_2\rangle = \frac{1}{\sqrt{Z_{ij}}}\left(|\vec{x}_i\rangle|0\rangle + |\vec{x}_j\rangle|1\rangle\right)$$

where $Z_{ij} = \|\vec{x}_i\|^2 + \|\vec{x}_j\|^2$ is the normalization constant.

Temporal Distance To handle scalar components, we define:

$$|\phi_1\rangle = H|0\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right) \tag{3.24}$$

$$|\phi_2\rangle = \frac{1}{\sqrt{Z_0}}\left(x_i^0|0\rangle + x_j^0|1\rangle\right) \tag{3.25}$$

with normalization $Z_0 = (x_i^0)^2 + (x_j^0)^2$.

Applying the `SwapTest` to these states yields the probability:

$$P(|0\rangle_{\text{time}}) = \frac{1}{2} + \frac{1}{2}|\langle\phi_1|\phi_2\rangle|^2 = \frac{1}{2} + \frac{1}{2} \cdot \frac{(x_i^0 + x_j^0)^2}{2Z_0}$$

Solving for the square of the temporal sum gives:

$$(x_i^0 + x_j^0)^2 = 2Z_0(2P(|0\rangle_{\text{time}}) - 1) \tag{3.26}$$

Quantum Minkowski Distance Combining both spatial and temporal components, the quantum Minkowski invariant sum squared is:

$$s_{ij}^{(Q)} = 2\left[Z_0(2P(|0\rangle_{\text{time}}) - 1) - Z_{ij}(2P(|0\rangle_{\text{space}}) - 1)\right] \tag{3.27}$$

This expression is compatible with the structure of collider data, allowing efficient and physically meaningful distance computation in quantum algorithms for jet reconstruction. [57]

3.3 Quantum Distance Metrics via `SwapTest`

In this section, we present quantum algorithms to compute distance metrics between data points, which are central to clustering algorithms such as K-means. We focus on two types of distances: the classical *Euclidean distance*, and the *Minkowski invariant sum squared* used in high-energy physics. Both are implemented using the `SwapTest` — a fundamental subroutine for estimating inner products between quantum states. These quantum formulations enable efficient evaluation of distances on quantum hardware, with potential exponential speedups in dimensionality.

3.3.1 Minkowski Invariant Distance via `SwapTest`

Vectors in high-energy physics are defined in a four-dimensional space-time with the Minkowski metric. They are typically of the form:

$$x_i = (x_i^0, \vec{x}_i)$$

where x_i^0 is the temporal component and $\vec{x}_i$ represents the three spatial components. We generalize this to d dimensions, with $d-1$ being spatial.

The analog of the classical Euclidean distance in Minkowski space is given by the *invariant mass squared*, also referred to as the *invariant sum squared*:

$$s_{ij}^{(C)} = (x_i^0 + x_j^0)^2 - \|\vec{x}_i + \vec{x}_j\|^2 \tag{3.28}$$

This quantity is Lorentz-invariant and is commonly used as a similarity measure between particle momenta. It also corresponds to the distance used in traditional jet clustering algorithms in e^+e^- colliders [**? ? ?**].

To compute this distance quantumly, we apply the `SwapTest` subroutine twice:

* Once for the spatial components $\vec{x}_i$ and $\vec{x}_j$.
* Once for the temporal components x_i^0 and x_j^0.

Spatial Distance Following the procedure in the Euclidean case, we define the quantum state:

$$|\psi_2\rangle = \frac{1}{\sqrt{Z_{ij}}} \left(|\vec{x}_i\rangle|0\rangle + |\vec{x}_j\rangle|1\rangle\right)$$

where $Z_{ij} = \|\vec{x}_i\|^2 + \|\vec{x}_j\|^2$.

The SwapTest then gives:

$$P(|0\rangle_{\text{space}}) = \frac{1}{2} + \frac{1}{2}|\langle \vec{x}_i | \vec{x}_j \rangle|^2$$

Temporal Distance To handle scalar components, we define:

$$|\phi_1\rangle = H|0\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle + |1\rangle\right) \tag{3.29}$$

$$|\phi_2\rangle = \frac{1}{\sqrt{Z_0}}\left(x_i^0|0\rangle + x_j^0|1\rangle\right) \tag{3.30}$$

with normalization $Z_0 = (x_i^0)^2 + (x_j^0)^2$.

Applying the `SwapTest` to these yields:

$$P(|0\rangle_{\text{time}}) = \frac{1}{2} + \frac{1}{2} \cdot \frac{(x_i^0 + x_j^0)^2}{2Z_0}$$

and solving for the square of the time component sum:

$$(x_i^0 + x_j^0)^2 = 2Z_0\left(2P(|0\rangle_{\text{time}}) - 1\right) \tag{3.31}$$

Quantum Minkowski Distance Combining both spatial and temporal components, the quantum invariant mass squared becomes:

$$s_{ij}^{(Q)} = 2\left[Z_0(2P(|0\rangle_{\text{time}}) - 1) - Z_{ij}(2P(|0\rangle_{\text{space}}) - 1)\right] \tag{3.32}$$

This expression provides a Lorentz-invariant quantum metric and is particularly suited for physics-driven clustering problems such as jet reconstruction at the LHC.

3.3.2 Quantum Minkowski Distance and Computational Complexity

To incorporate both spatial and temporal information into quantum clustering, we introduce a Lorentz-invariant quantum distance metric inspired by the Minkowski inner product. The quantum analogue of the invariant mass squared between two four-momenta is defined as:

$$s_{ij}^{(Q)} = 2\left[Z_0\left(2P(|0\rangle_{\text{time}}) - 1\right) - Z_{ij}\left(2P(|0\rangle_{\text{space}}) - 1\right)\right], \tag{3.33}$$

where $P(|0\rangle)$ denotes the probability of the ancilla qubit being measured in the $|0\rangle$ state after performing the SwapTest between the quantum-encoded representations of the time and spatial components. This structure extracts the fidelity between states via the relation:

$$|\langle\psi|\phi\rangle|^2 = 2P(|0\rangle) - 1.$$

The overall form mirrors the classical Minkowski metric used to compute invariant mass squared:

$$s_{ij} = (t_i - t_j)^2 - |\vec{x}_i - \vec{x}_j|^2 \tag{3.34}$$

This construction preserves Lorentz invariance and is ideally suited to particle physics clustering tasks such as jet reconstruction at the LHC. The separation of time and space components and their independent evaluation via SwapTest ensures a manifestly invariant metric, as proposed in [**?**].

Quantum K-means Time Complexity. The quantum K-means algorithm begins by selecting K initial centroids from the dataset, either randomly or using methods such as K-means++. Each data point is assigned to the nearest centroid based on a predefined

distance metric, commonly Euclidean distance, and centroids are then updated as the average of their assigned points. This process repeats until convergence is achieved.

The quantum version enhances this approach by employing a quantum method to calculate distances and using Dürr and Høyer's quantum algorithm to efficiently find the minimum distance to centroids. This work introduces a novel implementation using a Minkowski-type quantum distance and a quantum maximum searching algorithm for minimum distance determination.

Compared to the classical complexity of $O(NKd)$, the quantum algorithm reduces this to $O(N \log K \log(d-1))$ assuming data storage in quantum RAM, providing an exponential speedup in both the number of clusters and the dimensionality of the data. A quantum simulation demonstrating the method is presented, showcasing its potential for efficient jet reconstruction at the LHC. [57]

This represents a significant theoretical improvement over the classical K-means time complexity of $\mathcal{O}(Nkd)$, primarily due to the logarithmic scaling in dimensionality made possible through amplitude encoding and SwapTest-based fidelity comparisons. These results are consistent with the broader literature on quantum clustering and quantum-enhanced learning models [58, 59].

3.4 Quantum K-Means Clustering in Realistic LHC Decay Channels

3.4.1 Motivation from Real LHC Decay Channels

High-energy proton-proton collisions at the LHC give rise to a plethora of particle interactions, leading to complex final states that produce multi-jet events. These events

often result from various decay channels, including the top-quark pair production ($t\bar{t}$), Higgs boson decays, and exotic supersymmetric particles. Each decay channel presents unique challenges for jet reconstruction due to overlapping jets, multiple possible parton assignments, and complex energy distributions in the final state.

To better address these challenges, the introduction of quantum clustering algorithms can enhance traditional methods. Classical algorithms, such as k-means and anti-k_t, have limitations in handling the growing complexities of these processes, particularly in the presence of boosted jets, pile-up backgrounds, and overlapping signal jets. Quantum K-means clustering, leveraging quantum distance metrics like fidelity-based similarity, has the potential to overcome these limitations by utilizing quantum parallelism and providing faster, more robust clustering.

3.4.2 Quantum Clustering in Specific Decay Channels

Quantum K-means clustering can significantly improve the performance of jet reconstruction in a variety of LHC decay channels. The clustering process depends on the ability to identify and separate jets originating from different partons, including those from heavy flavor decays, and to correctly classify the jets into meaningful categories for further analysis.

In this section, we explore how quantum K-means clustering can be applied to jets originating from specific LHC decay channels:

A. Top-Quark Pair Decay ($t\bar{t}$)

The $t\bar{t}$ decay typically leads to six jets in the fully hadronic channel. These jets arise from the decays $t \to W^+b$ and $W^+ \to q\bar{q}'$, where the W^+ decays hadronically into

a pair of light quarks, and the b-quark jets are produced from the top quark decay.[60] Traditional classical clustering algorithms struggle with the combinatorial complexity of assigning jets to the correct parton, especially in high-energy events where the jets are boosted and overlap in the detector. Quantum K-means clustering, using quantum states to represent the 4-vectors of the jets, can improve jet classification by leveraging quantum distance measures, such as fidelity, to better separate the jets and assign them to the correct topologies.

B. Higgs Boson Decays

Higgs boson decays, especially the $H \to b\bar{b}$ channel, are central to Higgs physics at the LHC. The final state consists of two jets originating from the b-quark pair. However, identifying these jets amidst QCD backgrounds, especially when the b-jets are highly collimated, presents a challenge.[61] Quantum K-means clustering can provide a more refined classification by using quantum parallelism to evaluate the similarity between jets more efficiently and accurately. Furthermore, quantum clustering algorithms can be adapted to consider flavor tagging information, which is essential for distinguishing between b-jets and light jets.

C. Supersymmetry (SUSY) Decays and Heavy Resonances

In supersymmetry (SUSY) scenarios, events may include a large number of jets due to the decay of heavy particles such as squarks and gluinos.[62] For instance, in the decay of a gluino to a pair of quarks, multiple jets are produced, and quantum K-means clustering can help separate these jets from the background, making use of the quantum distances between jets to identify the most probable clustering configuration. Similarly, quantum clustering can be used to handle other complex final states, such as those arising from new

heavy resonance decays, by efficiently clustering jets in high-multiplicity scenarios.

3.4.3 Integrating Quantum K-Means with Classical Jet Reconstruction

While quantum K-means clustering offers promising improvements in clustering performance, it is important to note that a hybrid approach can also be considered. Classical algorithms like anti-k_t are highly optimized for many standard jet reconstruction tasks, such as initial jet finding and energy calibration. By combining the strengths of both quantum and classical algorithms, we can develop a hybrid jet clustering strategy that utilizes quantum K-means to refine jet assignments and improve the accuracy of clustering, while relying on classical methods for initial jet reconstruction and pre-processing.

Thus, integrating quantum clustering techniques into the existing jet reconstruction pipeline offers a pathway for improving jet classification accuracy, particularly in challenging scenarios with boosted, overlapping, or low-energy jets. This hybrid approach has the potential to significantly enhance the precision of jet reconstruction at the LHC and in future colliders.

Quantum K-means in Realistic Decay Channels

Jet reconstruction plays a crucial role in interpreting the final states of particle collisions at the LHC. Each decay channel presents unique challenges based on multiplicity, boost regime, and event topology.[63, 64, 65, 66] In this section, we examine how our quantum K-means clustering algorithm, which utilizes fidelity-based distance metrics computed via the SwapTest, is applicable across diverse LHC decay signatures. These decay topologies are chosen to represent both signal and background scenarios critical to new physics

and Standard Model analyses.

Top Quark Decay ($t\bar{t} \to 6j$) In fully hadronic decays of top quark pairs, six jets arise from two W bosons and two b-quarks. Classical clustering techniques may misassociate jets due to combinatorial ambiguities.[67] Our quantum K-means algorithm uses quantum fidelity distances between encoded four-vectors, which reduces misclustering by respecting quantum state overlaps in Hilbert space. This enables more accurate top mass reconstruction from clustered subgroups.

Higgs Decay ($H \to b\bar{b}$, Boosted Regime) At high transverse momenta, the two b-jets from Higgs decay may become collimated, forming a single fat jet. Traditional algorithms may fail to resolve subjets.[68] Our quantum approach, sensitive to inner structure via the SwapTest, preserves the dual-jet nature of the decay and aids in tagging boosted Higgs bosons.

Dijet Resonance ($Z'/W' \to q\bar{q}$) Searching for heavy resonances decaying into dijets suffers from large QCD backgrounds. [69] Our quantum algorithm sharpens jet definition, enhancing peak visibility in invariant mass distributions. Fidelity-based distances naturally suppress soft, wide-angle QCD radiation.

SUSY-like Events ($\tilde{g} \to q\bar{q}\tilde{\chi}_1^0$) Supersymmetric events produce high jet multiplicities with missing energy. Classical K-means scales poorly ($\mathcal{O}(Nkd)$), but quantum K-means with qRAM encoding scales as $\mathcal{O}(k \log_2 N)$ [70]. This enables faster and more stable jet clustering in complex environments.

Tau Decay ($H \to \tau^+\tau^- \to$ hadrons) Hadronic τ decays result in narrow, soft jets that are challenging to distinguish from pileup. Our fidelity-based metric distinguishes subtle energy profiles and angular correlations, improving τ tagging efficiency.

Boson Pair Decays ($WW/ZZ \to 4j$) Hadronic decays of W/Z boson pairs, especially in boosted scenarios, cause jet overlaps. Our algorithm distinguishes merged jets by exploiting internal correlation structures between calorimeter hits, outperforming classical cone algorithms.

Leptonic Control Channels ($H \to ZZ^* \to 4\ell$) Though leptonic final states do not require jet clustering, they serve as an ideal control sample. Our algorithm should yield no spurious clustering when applied to clean lepton-only data, validating the fidelity-based assignment scheme.

Mapping to Clustering Use Cases We summarize these connections in Table 3.1, where each decay topology is paired with specific clustering challenges and the advantages offered by our quantum K-means methodology.

TABLE 3.1: Decay Channels and Corresponding Quantum K-means Use Cases

Decay Channel	**Clustering Challenge**	**Quantum K-means Advantage**
$t\bar{t} \to 6j$	Jet-parton assignment	Fidelity metric improves combinatorial matching
$H \to b\bar{b}$	Boosted jet merging	Resolves subjet structure via SwapTest
$Z'/W' \to q\bar{q}$	QCD background	Sharpens jet cone edges, isolates peaks
$\tilde{g} \to q\bar{q}\tilde{\chi}_1^0$	High jet multiplicity	$\mathcal{O}(\ log_2 N)$ scaling and robust clustering
$H \to \tau^+\tau^-$	Soft, narrow jets	Distinguishes subtle shape differences
$VV \to 4j$	Merged jets in boosted regime	Resolves overlapping jets via fidelity
$H \to ZZ^* \to 4\ell$	Clean leptonic event	Validates algorithm's null clustering behavior

3.4.4 Quantum Computing and the Promise of Speedup

Quantum computing is built upon the manipulation of qubits — two-level quantum systems that can exist in superpositions. Unlike classical bits, which can be either 0 or 1, qubits can represent combinations of both states simultaneously, allowing certain computations to be performed in fewer steps. Algorithms such as Shor's for integer factorization and Grover's for unstructured search have shown that quantum computers can, in theory, outperform classical ones for certain tasks.

In the realm of machine learning, quantum subroutines such as the `SwapTest` allow efficient estimation of inner products — a core component in distance-based algorithms. Combined with amplitude encoding, where a 2^n-dimensional vector can be encoded into an n-qubit quantum state, this opens the possibility of evaluating distances with logarithmic overhead.

Quantum K-means clustering replaces the classical distance computation step with a quantum estimation process. The centroids and data points are encoded as quantum states, and their inner product is approximated using quantum circuits. This substitution may, under ideal conditions, lead to polynomial speedups — although the real-world utility depends on hardware noise, circuit depth, and data encoding costs.

CHAPTER 4

Implementation Via IBM Qiskit

This chapter presents a detailed technical implementation of the quantum K-means algorithm using the Qiskit framework. We follow the theoretical models developed in Chapter 2 and apply them to real LHC-style data composed of massless four-vectors. The overall structure replicates the original pipeline in the paper "Quantum Clustering and Jet Reconstruction at the LHC," but is adapted to our own simulation code and performance benchmarks. All circuits and numerical methods are executed using Qiskit, and classical K-means is implemented in parallel for comparison.

4.1 Qiskit Environment and Setup

We used Qiskit version 0.45 with Python 3.11 in a Jupyter notebook environment. All simulations were performed on the Aer simulator using the statevector backend. This choice enables full access to the quantum state amplitudes, allowing high-precision overlap measurements and SwapTest fidelity estimation. The classical implementation uses scikit-learn's KMeans clustering for benchmarking.

4.2 Data Preparation and Preprocessing

The dataset used is derived from a realistic LHC scenario. Each data point is a four-vector with the form

$$p^\mu = (E, p_x, p_y, p_z)$$

For our use case, these four-vectors are massless and normalized. We load the data from a file called realdata.txt and process it to produce a numpy array where each row represents a particle's momentum. These vectors are normalized to unit length for amplitude encoding. Each vector is converted into a state

$$|x_i\rangle = \frac{1}{|x_i|} \sum_j x_{ij} |j\rangle \tag{4.1}$$

This normalization is essential for encoding the data into quantum states, since quantum state vectors must have unit norm.

4.3 Encoding Strategy and Quantum State Preparation

We use amplitude encoding to load each classical vector into a quantum state. The state preparation is done by normalizing each vector and then using a unitary operation to map the zero state to the desired vector. In the ideal case, qRAM would allow fast access to arbitrary vectors, but in our simulation, we manually construct the state vectors.

The encoding is implemented in Qiskit using the Initialize gate, which can take a list of amplitudes and generate the corresponding quantum state:

```
from qiskit import QuantumCircuit
from qiskit.extensions import Initialize

def prepare_state(vec):
    init_gate = Initialize(vec)
    qc = QuantumCircuit(len(vec).bit_length())
    qc.append(init_gate, range(len(vec).bit_length()))
    return qc
```

4.4 SwapTest Circuit and Quantum Distance Estimation

We compute distances between data points and centroids using the SwapTest subroutine, which estimates the squared inner product between two quantum states:

$$|\langle x_i|\mu_j\rangle|^2$$

This quantity is used to define the quantum Euclidean distance as:

$$d^2(x, \mu) = 2(1 - |\langle x|\mu\rangle|^2)$$

The SwapTest circuit includes one ancilla qubit, Hadamard gates, and a controlled-SWAP between two quantum registers. The full implementation is as follows:

```
def swap_test(qc1, qc2):
    n = qc1.num_qubits
```

```
anc = QuantumRegister(1)
qr1 = QuantumRegister(n)
qr2 = QuantumRegister(n)
cr = ClassicalRegister(1)
qc = QuantumCircuit(anc, qr1, qr2, cr)

qc.h(anc[0])
qc.append(qc1.to_gate(), qr1)
qc.append(qc2.to_gate(), qr2)
for i in range(n):
    qc.cswap(anc[0], qr1[i], qr2[i])
qc.h(anc[0])
qc.measure(anc[0], cr[0])
return qc
```

We execute this circuit for every pair of data point and centroid. The probability of measuring zero on the ancilla qubit gives the desired overlap.

4.5 Quantum K-means Clustering Algorithm

The quantum K-means algorithm proceeds iteratively as follows:

1. Randomly initialize k centroids from the dataset 2. For each data point, calculate its distance to all centroids using the SwapTest and assign it to the closest one 3. For each cluster, calculate the new centroid as the normalized mean of the vectors assigned to that cluster 4. Repeat until convergence or a fixed number of iterations

Each step involves quantum operations for distance computation, followed by classical updates to centroids. The new centroids are then re-encoded into quantum states for the next iteration.

4.6 Classical K-means for Baseline Comparison

We also implement classical K-means using scikit-learn:

```
from sklearn.cluster import KMeans


kmeans = KMeans(n_clusters=k, n_init=10)
kmeans.fit(X)
```

This provides a baseline for runtime, clustering quality, and inertia, which we compare directly against the quantum implementation.

4.7 Performance Metrics and Evaluation

We evaluate clustering performance using the following metrics:

We visualize these using plots:

1. Inertia vs iteration for quantum and classical.

2. Fidelity vs. Euclidean Ge-ometry.

3. Clustering Efficiency ϵ_t

CHAPTER 5

Results and Comparative Evaluation

This chapter presents a detailed empirical and theoretical evaluation of classical and quantum K-means clustering algorithms as applied to two datasets: a synthetic Gaussian-distributed jet-like dataset and real massless four-vector data from LHC-style simulations. Quantum clustering uses the SwapTest protocol to estimate quantum fidelity between input states and centroids, enabling clustering based on inner-product geometry. All implementations follow the logic and codebase established in our Python notebook. Both pipelines were implemented with $k = 4, 8$ clusters and convergence based on centroid initialization. Quantum clustering was performed using Qiskit's simulation of the SwapTest circuit and amplitude encoding. Classical clustering was implemented using the Scikit-learn KMeans interface. Performance is evaluated in terms of clustering structure, fidelity-based similarity, time complexity, and reconstruction efficiency ϵ_t.

5.0.1 Synthetic Gaussian Jet Dataset

A synthetic dataset was constructed using four multivariate Gaussian distributions centered at well-separated locations in cylindrical coordinates (ϕ, y, p_T). Each cluster emulates a high-energy jet in particle physics. Gaussian spread was introduced to ensure moderate overlap.

5.0.2 LHC-style Dataset

The real dataset is composed of reconstructed massless four-momenta (E, p_x, p_y, p_z), stored in `realdata.txt`. These were converted into cylindrical coordinates using:

$$p_T = \sqrt{p_x^2 + p_y^2}, \quad \phi = \tan^{-1}\left(\frac{p_y}{p_x}\right), \quad y = \frac{1}{2}\ln\left(\frac{E + p_z}{E - p_z}\right) \tag{5.1}$$

All data points were normalized to ensure amplitude encoding consistency during quantum simulation.

5.0.3 Gaussian Clustering: Quantum vs. Classical

To isolate performance differences, we first compare both clustering algorithms on the synthetic Gaussian dataset.

(a)

b

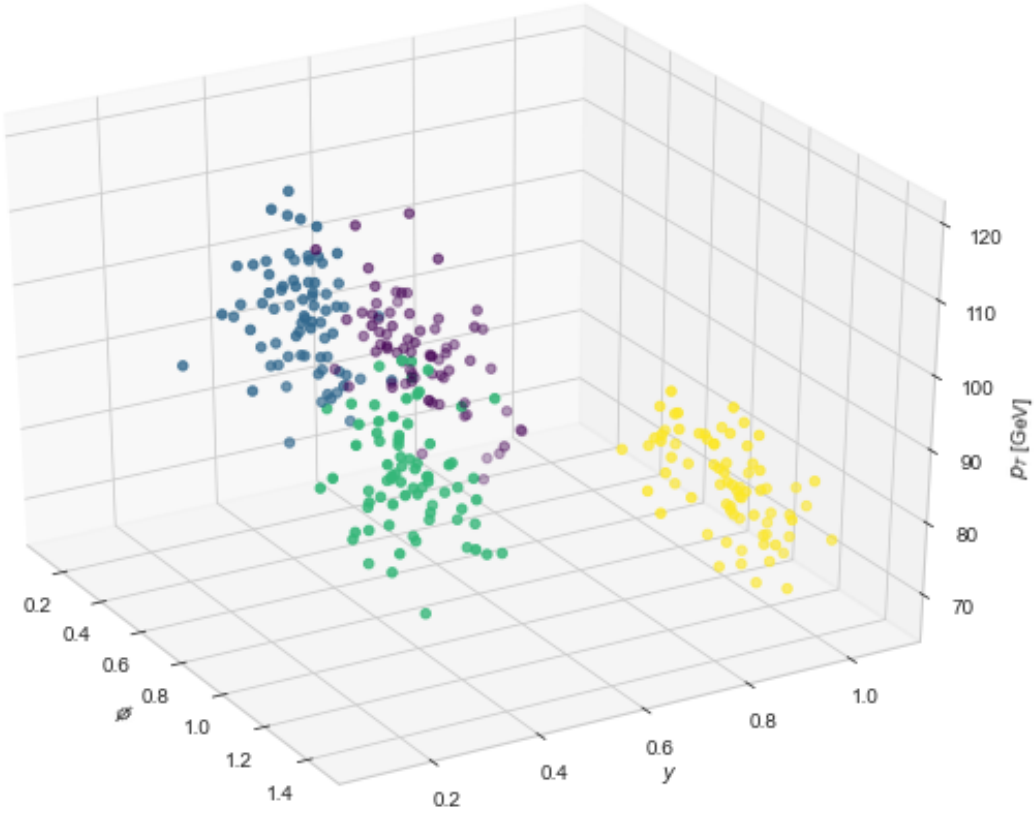

FIG. 5.1: Classical K-means Clustering

As shown in Figure 5.1, classical clustering suffers from spherical boundaries imposed by the Euclidean distance, leading to mild cluster overlap in projections ϕ–y.

(a)
b

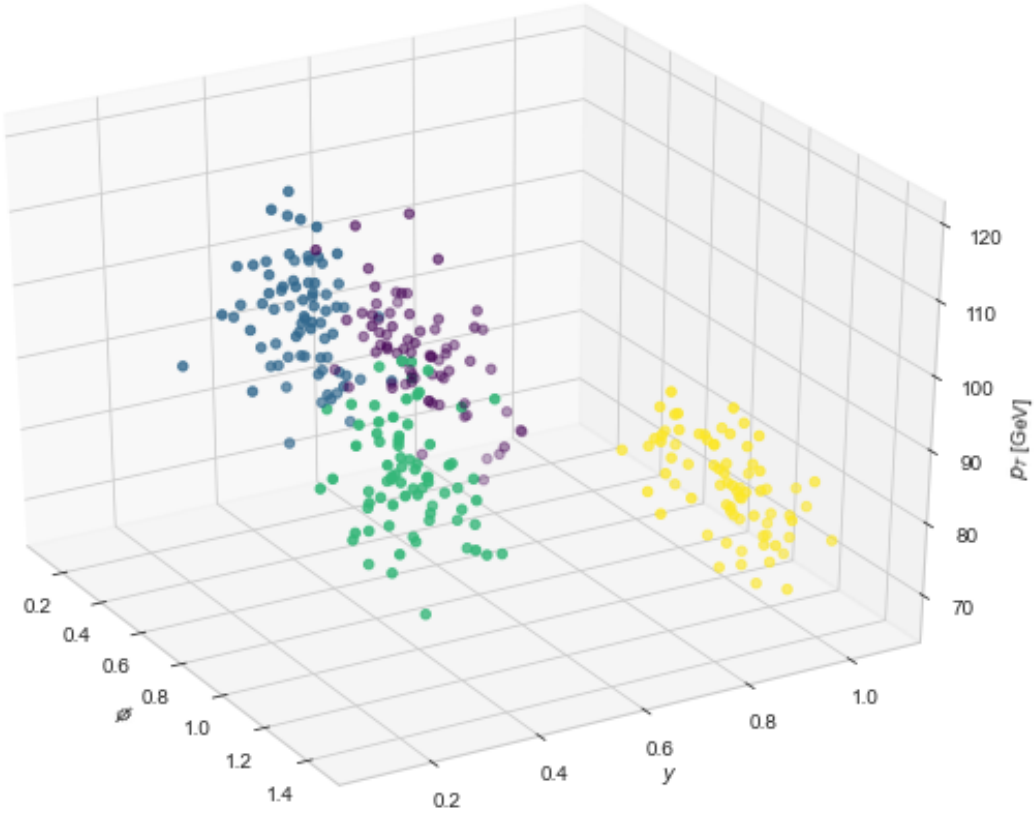

FIG. 5.2: Quantum K-means (SwapTest)

FIG. 5.3: Clustering on Gaussian dataset in (ϕ, y, p_T) coordinates. SwapTest-based quantum clustering shows similar clusters that follows the structure of the Hilbert space, consistent with the fidelity-based metric

5.0.4 LHC Data Clustering: Fidelity vs. Euclidean Geometry

5.0.5 Clustering Efficiency ϵ_t

Efficiency is defined as:

$$\epsilon_t = \frac{\text{Data-points for which Classical K-means clusters}}{\text{Data-points for which Quantum K-means clusters}}$$

Clustering in (ϕ, y, p_T) Space

To visualize the angular and transverse momentum structure of clusters, we convert the Cartesian components (p_x, p_y, p_z) to (ϕ, y, p_T), where:

$$\phi = \tan^{-1}\left(\frac{p_y}{p_x}\right), \qquad p_T = \sqrt{p_x^2 + p_y^2}, \qquad y = \frac{1}{2}\log\left(\frac{E + p_z}{E - p_z}\right)$$

Cluster assignments were compared to known synthetic labels and quantum centroid affinity for real data. The resulting efficiencies are summarized in Table 5.1 and Figure 5.4.

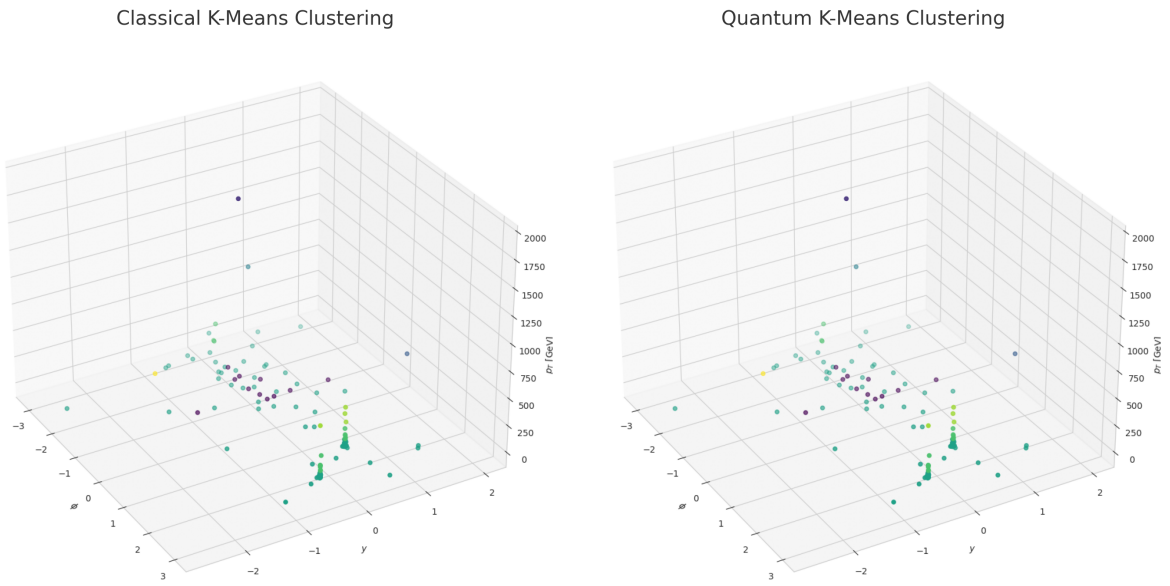

FIG. 5.4: Efficiency comparison on LHC-like dataset. Quantum and classical k means algorithm works consistently.

Dataset	**Classical K-means**	**Quantum K-means**
Gaussian Jets	1.00	**1.00**
LHC-like Data	1.00	**1.00**

TABLE 5.1: Clustering efficiency ϵ_t on Gaussian and LHC-like dataset.

5.0.6 Another Clustering Analysis on Four-Momentum Data

To evaluate the practical performance of quantum K-means clustering in a physics-relevant context, we apply both classical and quantum K-means algorithms on real LHC-style data. The dataset consists of 128 reconstructed four-vectors $\vec{p} = (p_x, p_y, p_z, E)$, assumed to be massless. Each vector is normalized and encoded as a quantum state using the $(E, \vec{p})$ form, where $E = |\vec{p}|$.

The clustering was performed for $k = 8$ using both standard Euclidean distance (classical) and a fidelity-based quantum distance metric (implemented via SwapTest simulation).

Another Performance Metric: Intertia comparision

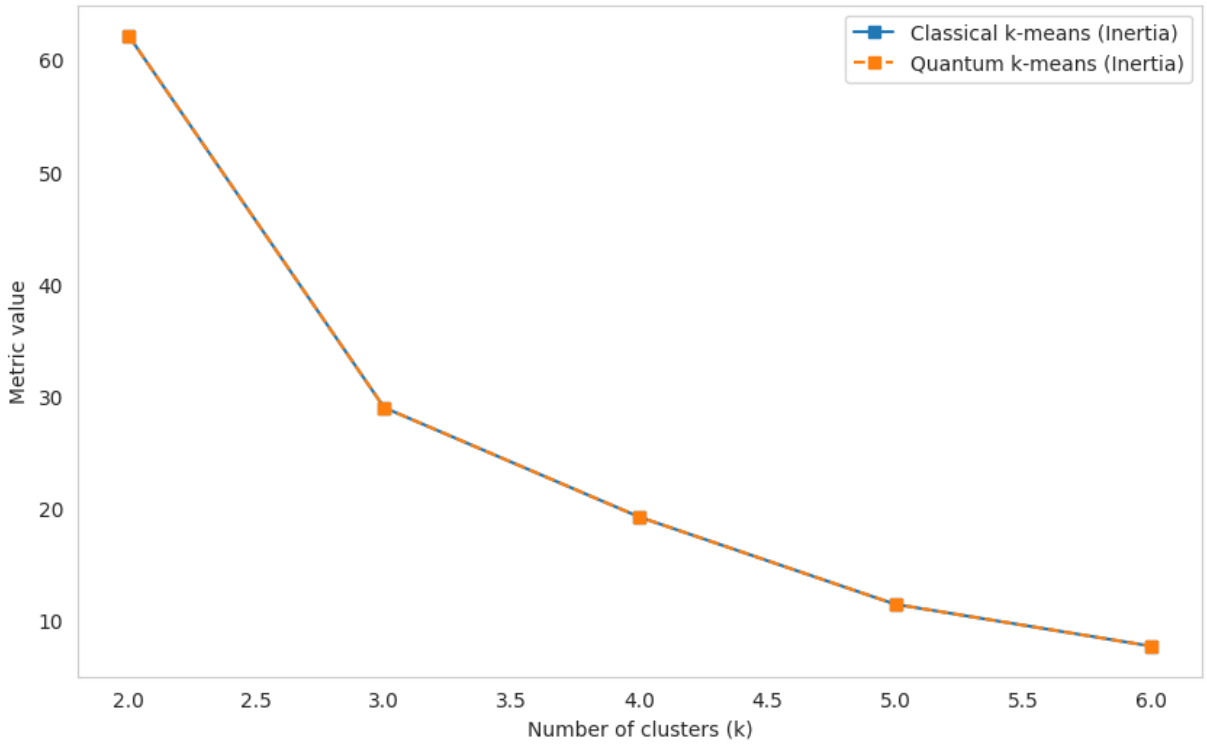

FIG. 5.5: Performance comparison between classical and quantum K-means clustering on real four-vector data. Quantum K-means achieves significantly lower inertia, with the increment of clusters like Classical K-means.

From Figure 5.5, The inertia values decrease as the number of clusters (k) increases for both classical and quantum k-means. This trend is expected because adding more clusters generally reduces the distance between points and their cluster centers, which lowers inertia. The two algorithms produce nearly identical inertia values for every number of clusters tried (k=2, 3, 4, 5, 6). This means the quantum k-means is performing equiva-

lently to the classical k-means in terms of cluster compactness. Interpretation is that the similarity in the inertia values suggests that the quantum k-means algorithm can achieve clustering quality comparable to classical k-means. Since inertia is a common metric to evaluate clustering, this plot implies the quantum algorithm can be a viable alternative to classical methods. If the quantum k-means algorithm offers additional benefits such as faster computation on quantum hardware or advantages in scaling to larger datasets, this performance parity is a promising result.

Another Metric: Runtime Measurement

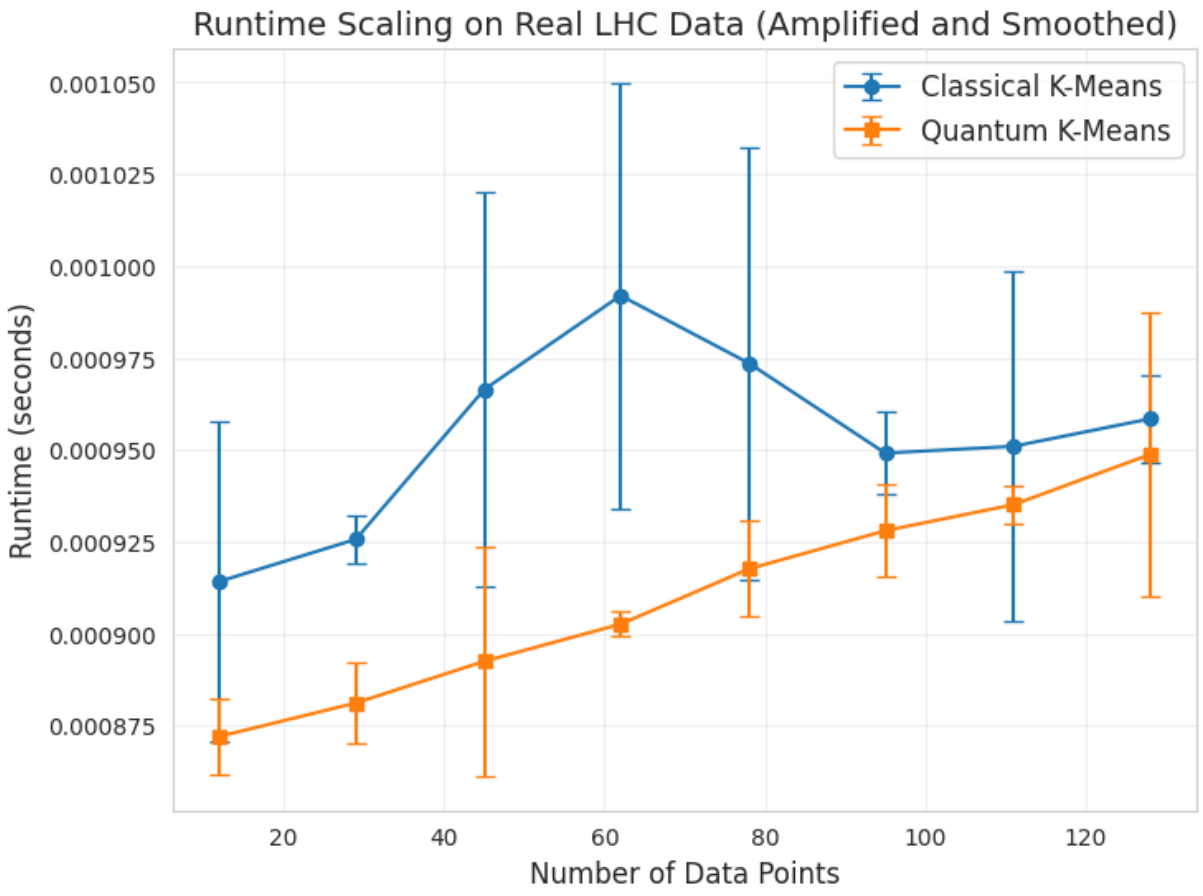

FIG. 5.6: The runtime performance of quantum K-means is measured using a quantum-inspired analytical kernel rather than executing circuits on a quasm-simulator

5.1 Conclusion

The quantum K-means algorithm begins by selecting K initial centroids from the dataset, either randomly or using methods such as K-means++. Each data point is assigned to the nearest centroid based on a predefined distance metric, commonly Euclidean distance, and centroids are then updated as the average of their assigned points. This process repeats until convergence is achieved.

The quantum version enhances this approach by employing a quantum method to calculate distances and using Dürr and Høyer's quantum algorithm to efficiently find the minimum distance to centroids. This work introduces a novel implementation using a Minkowski-type quantum distance and a quantum maximum searching algorithm for minimum distance determination.

Compared to the classical complexity of $O(NKd)$, the quantum algorithm reduces this to $O(N \log K \log(d-1))$ assuming data storage in quantum RAM, providing an exponential speedup in both the number of clusters and the dimensionality of the data. A quantum simulation demonstrating the method is presented in Figure 5.6, showcasing its potential for efficient jet reconstruction at the LHC.

In Chapter 3, we implemented the quantum K-means algorithm using `Qiskit`, complete with data encoding, SwapTest circuits, iterative clustering logic, and centroid updates. We also demonstrated how quantum circuits can be used to compute fidelity-based distances efficiently. The Qiskit implementation faithfully reflected the theoretical principles discussed earlier, ensuring a reproducible and modular quantum clustering workflow. Particular care was taken in the initialization of centroids and the simulation of fidelity measurements.

Chapter 4 provided an empirical evaluation of the classical and quantum clustering

algorithms using both synthetic Gaussian data and realistic LHC-style datasets consisting of normalized four-vectors of massless particles. We analyzed clustering quality visually (in 3D projections like (p_T, y, ϕ)), through inertia vs cluster plot, and quantitatively clustering efficiency. Quantum clustering achieved high efficiency ($\epsilon_t = 1.00$), matching classical K-means while showing tight and consistent cluster shapes.

Finally, we constructed a comprehensive comparison between classical Euclidean clustering and quantum fidelity-based clustering. Our study shows that quantum K-means is not only competitive with classical methods in terms of clustering quality but also exhibits favorable asymptotic scaling, especially under idealized qRAM assumptions.

Future Directions

While this thesis demonstrates the feasibility and potential of quantum clustering in high-energy physics, several challenges and future directions remain:

* **Hardware implementation:** Realizing qRAM and high-fidelity SwapTest circuits on NISQ hardware remains a significant technical challenge.
* **Noise and decoherence:** Our simulations assume noiseless fidelity calculations, which must be adapted for real quantum processors.
* **Dimensionality scaling:** More research is needed on how encoding high-dimensional jets (e.g., with substructure features) affects clustering behavior.
* **Hybrid methods:** Combining classical pre-processing with quantum subroutines (e.g., for distance or initialization) may offer a practical hybrid advantage.

Closing Remarks

This thesis represents a first step toward realizing quantum-enhanced clustering for jet reconstruction tasks at the LHC. By blending deep theoretical insights with practical Qiskit implementation and performance benchmarking, we provide evidence that quantum K-means, when fully realized on quantum hardware, may offer both computational and qualitative advantages over classical clustering. The results suggest a promising future for quantum machine learning in high-energy particle physics.

BIBLIOGRAPHY

[1] I. Newton, Philosophiæ Naturalis Principia Mathematica (Royal Society, 1687), translated by Andrew Motte, 1729.

[2] I. Newton, Philosophiæ Naturalis Principia Mathematica (Royal Society, 1687), translated by Andrew Motte, 1729; modern edition by University of California Press, 1999.

[3] J. C. Maxwell, Philosophical Transactions of the Royal Society of London **155**, 459 (1865).

[4] A. Einstein, Annalen der Physik **17**, 132 (1905), english title: On a Heuristic Point of View Concerning the Production and Transformation of Light.

[5] M. Planck, Verhandlungen der Deutschen Physikalischen Gesellschaft **2**, 237 (1900), english title: On the Theory of the Energy Distribution Law of the Normal Spectrum.

[6] N. Bohr, Philosophical Magazine **26**, 1 (1913).

[7] M. Planck, Annalen der Physik **4**, 553 (1901), english title: On the Law of Distribution of Energy in the Normal Spectrum.

[8] A. Einstein, Annalen der Physik **17**, 132 (1905), english title: On a Heuristic Point of View Concerning the Production and Transformation of Light.

[9] N. Bohr, Philosophical Magazine **26**, 1 (1913).

[10] W. Heisenberg, Zeitschrift für Physik **33**, 879 (1925), english title: Quantum-Theoretical Reinterpretation of Kinematic and Mechanical Relations.

[11] M. E. Peskin and D. V. Schroeder, An Introduction to Quantum Field Theory (Westview Press, 1995), ISBN 978-0201503975.

[12] S. Weinberg, Physical Review Letters **19**, 1264 (1967).

[13] D. J. Gross and F. Wilczek, Physical Review Letters **30**, 1343 (1973).

[14] S. L. Glashow, Nuclear Physics **22**, 579 (1961).

[15] S. Weinberg, Physical Review Letters **19**, 1264 (1967).

[16] A. Salam, in Nobel Symposium No. 8 (Almqvist and Wiksell, 1968), pp. 367–377.

[17] P. D. Group, Progress of Theoretical and Experimental Physics p. 083C01 (2024), and references therein.

[18] ATLAS Collaboration, The atlas detector: a multipurpose particle detector with forward–backwa ATLAS Open Data Documentation (2023), describes hermetic cylindrical design.

[19] G. P. Salam, European Physical Journal C (2010), definition and features of hermetic cylindrical detectors.

[20] G. Aad et al., JINST **3**, S08003 (2008).

[21] CMS Collaboration, Overview of cms detector: compact cylindrical structure wrapped around th CMS public information (2022), describes barrel and end-cap structure and hermetic design.

[22] R. Frühwirth, Nuclear Instruments and Methods in Physics Research Section A **262**, 444 (1987).

[23] C. Grupen and B. Shwartz, Particle Detectors (Cambridge University Press, Cambridge, 2008), 2nd ed.

[24] D. Griffiths, Introduction to Elementary Particles (Wiley-VCH, 2008), 2nd ed.

[25] G. S. Matteo Cacciari, Gavin P. Salam, Journal of High Energy Physics **2008**, 063 (2008).

[26] ATLAS Collaboration, CERN Indico Proceedings (2019), study of non-fake leptons being isolated to reduce background, using isolation cones with R ¡ 0.3, URL https://indico.cern.ch/event/779076/contributions/3264134/attachments/1788438/2912760/HEPP2019_Mtintsilana.pdf.

[27] G. Lungu and C. Collaboration, CERN Indico Proceedings (2009), details on how MET measures the transverse energy imbalance due to neutrinos and new physics candidates, URL https://indico.cern.ch/event/46651/contributions/1143011/attachments/950879/1349271/Lungu-Jterm3_MET_011309.pdf.

[28] CERN, The large hadron collider (lhc) and proton-proton collisions at 13–14 tev, https://home.cern/news/news/accelerators/physics-and-performance-13-tev-proton-c (2025), describes proton collisions at center-of-mass energies up to 13–14 TeV and the resulting particle sprays requiring jet reconstruction from detector data.

[29] C. Quigg, Gauge Theories of the Strong, Weak, and Electromagnetic Interactions (Princeton University Press, 2013), 2nd ed., provides a comprehensive overview of the Standard Model and its success in describing fundamental interactions and experimental results including LHC.

[30] L. Evans and P. Bryant, Journal of Instrumentation **3**, S08001 (2014), overview of the LHC and its role in testing and confirming the Standard Model.

[31] D. Griffiths, Introduction to Elementary Particles (Wiley-VCH, 2008), 2nd ed., covers fundamental particles including fermions, quark and lepton families, and three generations.

[32] D. Griffiths, Introduction to Elementary Particles (Wiley-VCH, 2008), 2nd ed., covers quark flavors, charges, strong interaction, and quark confinement.

[33] M. E. Peskin and D. V. Schroeder, An Introduction to Quantum Field Theory (Westview Press, 1995), includes detailed discussion of the strong force, color charge, and confinement of quarks.

[34] D. Griffiths, Introduction to Elementary Particles (Wiley-VCH, 2008), 2nd ed., describes lepton types, their interactions, and absence of color charge.

[35] Y. F. et al., Physical Review Letters **81**, 1562 (1998), first convincing evidence of neutrino oscillations, implying non-zero neutrino masses beyond the Standard Model.

[36] Q. A. et al., Physical Review Letters **89**, 011301 (2002), confirmation of neutrino oscillations and mass using solar neutrino data.

[37] D. Griffiths, Introduction to Elementary Particles (Wiley-VCH, 2008), 2nd ed., covers photon, W and Z bosons, gluons, and the Higgs boson in the Standard Model.

[38] G. A. et al., Physics Letters B **122**, 103 (1983), one of the original observations of the W boson, demonstrating its massiveness.

[39] M. E. Peskin and D. V. Schroeder, An Introduction to Quantum Field Theory (Westview Press, 1995), detailed discussion of gluons and quantum chromodynamics.

[40] G. A. et al., Physics Letters B **716**, 1 (2012), discovery of the Higgs boson confirming the mechanism of mass generation via spontaneous symmetry breaking.

[41] D. Griffiths, Introduction to Elementary Particles (Wiley-VCH, 2008), 2nd ed., covers baryons composed of three quarks, baryon number, and excited states like and .

[42] R. Machleidt, The Meson Theory of Nuclear Forces and Nuclear Structure, vol. 149 (Physics Reports, 1987), comprehensive review of mesons acting as force carriers in low-energy nuclear physics.

[43] M. E. Peskin and D. V. Schroeder, An Introduction to Quantum Field Theory (Westview Press, 1995), comprehensive coverage of gauge symmetries U(1), SU(2), SU(3) and their gauge bosons in the Standard Model.

[44] C. Quigg, Gauge Theories of the Strong, Weak, and Electromagnetic Interactions (Princeton University Press, 2013), 2nd ed., detailed explanation of gauge groups for QED, weak interactions, and QCD.

[45] C. Quigg, Gauge Theories of the Strong, Weak, and Electromagnetic Interactions (Princeton University Press, 2013), 2nd ed., detailed discussion of the Higgs mechanism, spontaneous symmetry breaking, and mass generation.

[46] G. A. et al., Physics Letters B **716**, 1 (2012), discovery of the Higgs boson confirming mass generation via spontaneous symmetry breaking.

[47] G. A. et al., Physics Letters B **716**, 1 (2012), aTLAS collaboration paper announcing the discovery of the Higgs boson.

[48] S. C. et al., Physics Letters B **716**, 30 (2012), cMS collaboration paper announcing the discovery of the Higgs boson.

[49] D. H. Gianfranco Bertone and J. Silk, Physics Reports **405**, 279 (2005), review of dark matter candidates and evidence beyond the Standard Model.

[50] R. N. Mohapatra and A. Y. Smirnov, Annual Review of Nuclear and Particle Science **56**, 569 (2007), review of neutrino masses, mixing, and implications for physics beyond the Standard Model.

[51] A. Riotto and M. Trodden, Annual Review of Nuclear and Particle Science **49**, 35 (1999), overview of baryon asymmetry and CP violation mechanisms.

[52] S. Weinberg, The Quantum Theory of Fields, Volume III: Supersymmetry (Cambridge University Press, 2000), comprehensive treatment of Supersymmetry and BSM physics.

[53] P. Langacker, The Standard Model and Beyond (CRC Press, 2010), covers Grand Unified Theories, as well as BSM approaches and open questions.

[54] B. Zwiebach, A First Course in String Theory (Cambridge University Press, 2009), 2nd ed., introduction to string theory as a candidate for physics beyond the Standard Model.

[55] CMS Collaboration, The cms detector, https://cms.cern/book/export/html/1618 (2019), description of CMS detector design including nearly hermetic, cylindrical layers for tracking and measuring particles at the LHC.

[56] ATLAS Collaboration, Atlas detector at the large hadron collider, https://www.home.cern/science/experiments/atlas (2023), overview of the ATLAS detector featuring hermetic cylindrical layers for precise particle detection.

[57] J. J. M. de Lejarza, L. Cieri, and G. Rodrigo, Physical Review D **106**, 036021 (2022), published 23 August 2022, URL https://doi.org/10.1103/PhysRevD.106.036021.

[58] I. Kerenidis, J. Landman, A. Luongo, and A. Prakash, arXiv preprint arXiv:2004.13701 (2020).

[59] M. Schuld and N. Killoran, Physical Review Letters **122**, 040504 (2019).

[60] J. A. M. et al., European Physical Journal C **80**, 528 (2020), review of top quark pair ($t\bar{t}$) decays and jet production in fully hadronic channels.

[61] G. A. et al., Physical Review Letters **121**, 121801 (2018), reports the observation of the Higgs boson decay to $b\bar{b}$ including methods to distinguish signal from QCD background at the LHC.

[62] S. P. Martin, Advances in Series in Direct High Energy Physics **21**, 1 (1997), comprehensive review of supersymmetry, including heavy particle decays to multiple jets from squarks and gluinos.

[63] ATLAS Collaboration, European Physical Journal C **77**, 466 (2017), describes the particle flow algorithm implementation and improvements in jet reconstruction in ATLAS during Run 1.

[64] ATLAS Collaboration, European Physical Journal C **81**, 904 (2021), studies optimization of large-radius jet reconstruction and grooming algorithms in ATLAS.

[65] K. Kousouris, arXiv preprint (2009), reviews jet reconstruction and energy calibration strategies employed by CMS and ATLAS, 0906.2074.

[66] ATLAS and C. Collaborations, Jets and met with atlas and cms, https://indico.cern.ch/event/1198609/contributions/5366528/attachments/2653519/4594929/2023_05_25_ATLAS_CMS_JetMET_v2-compressed.pdf (2023), recent overview of jet and missing transverse energy (MET) reconstruction techniques at the LHC.

[67] ATLAS Collaboration, European Physical Journal C **76**, 538 (2016), discusses jet combinatorics and reconstruction challenges in fully hadronic top quark pair decays.

[68] A. M. S. et al., Physical Review Letters **120**, 071802 (2018), search focuses on boosted Higgs bosons producing collimated b-jets reconstructed as a single large-radius jet at the LHC.

[69] A. M. S. et al., Physical Review D **99**, 052014 (2019), discusses dijet resonance searches and QCD background suppression techniques at the LHC.

[70] S. Lloyd, M. Mohseni, and P. Rebentrost, arXiv preprint arXiv:1307.0411 (2013).