Variational Quantum Algorithms as Hybrid Control Systems: Structured Optimization and Classical Regularization

Yury Chernyak

Contents

Chapter 1

Introduction

1.1 Variational Quantum Algorithms and the NISQ Regime

Variational Quantum Algorithms (VQAs) have emerged as a prominent approach for leveraging near-term quantum hardware to solve problems in quantum chemistry, condensed matter physics, and optimization [1], [2]. Unlike fault-tolerant quantum algorithms, VQAs are explicitly designed to operate within the constraints of the noisy intermediate-scale quantum (NISQ) regime, where circuit depth, qubit counts, and coherence times remain severely limited [3]. The defining feature of these algorithms is their hybrid structure: a parameterized quantum circuit is repeatedly evaluated on quantum hardware, while a classical optimization routine updates the circuit parameters based on measurement outcomes [2].

This hybrid formulation enables VQAs to bypass some of the stringent requirements of error correction by offloading learning and adaptation to classical computation. At the same time, it introduces a new class of challenges that are absent in fully classical or fully quantum algorithms. In practice, VQAs frequently exhibit unstable convergence, strong sensitivity to initialization, and large variability in performance across repeated runs, even when applied to relatively small systems and to idealized simulations with fixed noise models [1], [2], [4], [5]. These difficulties persist across different hardware platforms, ansatz families, and problem instances, indicating that they are not solely artifacts of experimental imperfections.

A substantial body of work has identified barren plateaus—regions of parameter space in which the cost-function gradient vanishes exponentially in the number of qubits, leaving optimizers without informative descent directions [4]—together with sampling noise—the statistical fluctuation of cost estimates due to the finite number of measurement shots used to reconstruct each Pauli expectation value [6]—and hardware errors as contributing factors to these difficulties. While these effects undoubtedly shape the structure of the variational landscape, they do not, on their own, fully explain the wide range of observed optimization behaviors. In particular, identical quantum circuits evaluated under identical noise models can exhibit dramatically different convergence properties depending on the choice of classical optimization strategy [2], [4], [7], [8].

These observations motivate a closer examination of the dynamical role of the classical component of VQAs. Rather than viewing optimization difficulties solely as consequences of quantum expressivity or noise, it is necessary to analyze how classical parameter-update dynamics interact with the quantum-generated objective landscape. Establishing this perspective is the first step to-

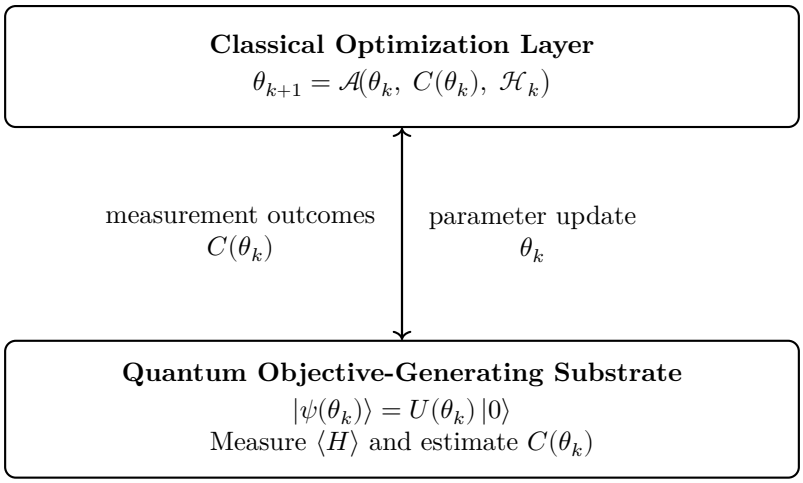

Figure 1.1: Variational quantum algorithms as a hybrid feedback system. The quantum objective-generating substrate evaluates the cost function for queried parameters via state preparation and measurement, while the classical optimization layer implements the update dynamics that produce the next parameter iterate.

ward a systematic understanding of stability, robustness, and failure modes in variational quantum algorithms.

1.2 Hybrid Quantum–Classical Structure of VQAs

At an operational level, a variational quantum algorithm consists of two tightly coupled but functionally distinct components: a quantum evaluation layer and a classical optimization layer. The quantum component prepares a parameterized quantum state and evaluates a cost function—(typically) an expectation value of a Hamiltonian—by repeated measurement. The classical component processes these measurement outcomes and determines how the parameters of the quantum circuit are updated between iterations. Crucially, this classical update process does not obey quantum mechanical principles such as superposition, entanglement, or unitary evolution, but follows classical dynamical rules.

More concretely, the quantum layer implements a mapping from a vector of real-valued parameters to a scalar objective value. Given a parameter vector $\theta \in \mathbb{R}^P$, a parameterized circuit prepares a quantum state $|\psi(\theta)\rangle$, from which an estimate of the cost function $C(\theta)$ is obtained through measurement statistics. This mapping may be highly nonlinear, noisy, and problem-dependent, but it is entirely specified by the circuit architecture, the observable being measured, and the underlying hardware or simulator. The stochasticity arises from measurement statistics and hardware imperfections rather than from any intrinsic temporal memory or adaptivity within the quantum subsystem. This interaction can be schematically represented as a closed-loop feedback system, as shown in Figure 1.1.

The classical layer operates on the output of this mapping. Based on one or more evaluations of the cost function, a classical algorithm produces an updated parameter vector according to an update rule of the form

$$\theta_{k+1} = \mathcal{A}(\theta_k, C(\theta_k), \mathcal{H}_k), \tag{1.1}$$

where $\mathcal{A}$ denotes the chosen optimization procedure and $\mathcal{H}_k$ represents any internal state maintained by the algorithm. This update rule may depend on factors including, but not limited to, gradient estimates, finite differences, stochastic sampling, population statistics, or internal memory accumulated over previous iterations. Importantly, all such update mechanisms are implemented purely in classical computation and obey classical dynamical rules.

The hybrid nature of VQAs lies in the feedback loop formed between these two layers. The quantum device supplies information about the objective landscape at queried points in parameter space, while the classical optimizer uses this information to guide subsequent queries. Although these components are executed sequentially within each iteration, their interaction defines the overall behavior of the algorithm.

This structural separation is central to understanding the behavior of variational algorithms. While the quantum component determines the structure of the objective landscape, the classical component governs the dynamical process by which that landscape is explored. As a result, properties such as convergence speed, stability, sensitivity to noise, and susceptibility to stagnation cannot be attributed to the quantum circuit alone, but emerge from the combined dynamics of both layers.

Importantly, the parameterized circuit architecture restricts the set of quantum states accessible during optimization to a structured subset of the full Hilbert space, thereby imposing a strong inductive bias on the resulting objective landscape [9].

1.3 Limitations of Existing Optimization Approaches

From the perspective established above, variational quantum algorithms may be viewed as constrained learning systems, in which optimization proceeds over a restricted family of quantum states determined a priori by circuit design (i.e., by the choice of variational ansatz implemented as a parameterized quantum circuit). The choice of variational ansatz imposes a strong inductive bias on the optimization process by restricting the set of accessible quantum states. While this restriction is necessary for trainability on NISQ devices, it also fundamentally alters the optimization landscape encountered by the classical optimizer. In particular, the objective function is evaluated under stochastic noise, may exhibit large flat regions, and is confined to a parameterized subset of Hilbert space. As a result, classical optimization strategies developed for smooth, deterministic objective functions may be poorly suited to this setting [2], [4], [9], [10].

The performance of variational quantum algorithms depends critically on the choice of classical optimization strategy. In practice, a wide range of optimizers have been applied to VQAs, including gradient-based methods, gradient-free local search algorithms, and heuristic global optimizers [1], [2], [4], [7]. Despite this diversity, no single approach has emerged as uniformly reliable across problem classes, ansatz structures, or hardware platforms.

Gradient-based optimizers are often favored due to their efficiency in classical machine learning under deterministic and high-precision conditions and optimization tasks. However, in the variational quantum setting they face several well-documented challenges. Gradient estimates are affected by finite-shot noise, hardware imperfections, and sampling fluctuations, which can severely degrade update accuracy. In addition, the presence of flat regions in the cost landscape can lead to vanishing gradients, rendering local descent ineffective even when the global optimum is well defined [2], [4], [7], [11].

Gradient-free methods, such as finite-difference schemes or local heuristic searches, are less sensi-

tive to gradient estimation noise but typically suffer from poor scaling with parameter dimension. As the number of variational parameters increases, these methods often exhibit slow convergence and strong dependence on initialization. Moreover, they may become trapped in local minima or oscillatory regimes without clear indicators of failure.

Importantly, many studies attribute these difficulties primarily to properties of the quantum landscape itself, such as expressibility, entanglement structure, or hardware-induced noise. While these factors undeniably shape the objective function, they do not fully account for the wide variability observed in optimization outcomes across repeated runs and optimization strategies. In particular, identical quantum circuits evaluated under identical noise models can display dramatically different convergence behavior when paired with different classical optimizers. This variability cannot be explained by static properties of the objective landscape alone, but reflects differences in the induced optimization dynamics [2], [7], [8].

This observation suggests that optimization failures in VQAs cannot be understood solely by analyzing the quantum component in isolation. Instead, they arise from the interaction between the objective landscape and the classical update dynamics. A framework that treats the classical optimizer as an active dynamical component—rather than a passive numerical subroutine—is therefore required to systematically analyze convergence, stability, and robustness in variational quantum algorithms.

While no classical optimization strategy can compensate for an inadequately expressive variational ansatz, the choice of optimizer critically shapes the dynamical behavior of optimization within a viable variational family.

1.4 Thesis Contributions and Structure

This thesis develops a system-level perspective on variational quantum algorithms by explicitly treating them as hybrid quantum–classical systems with distinct operational roles. These roles are operational rather than ontological, reflecting differences in function rather than in physical description. Rather than attributing optimization behavior solely to properties of quantum circuits or hardware noise, the thesis emphasizes the role of classical parameter-update dynamics in shaping convergence, stability, and robustness. This perspective enables a unified analysis of a wide range of observed behaviors in variational algorithms without introducing additional assumptions about quantum foundations or interpretational frameworks.

The core contributions of the thesis can be summarized as follows. First, it provides a clear operational decomposition of VQAs into a quantum objective-generating substrate and a classical optimization layer, highlighting their interaction through a feedback loop. Second, it analyzes population-based optimization strategies as explicit controller designs for variational algorithms, with particular emphasis on harmonic oscillator particle swarm optimization (HOPSO) as a non-gradient, stochastic update mechanism. Third, it examines the role of classical regularization techniques as optimizer-agnostic stabilization tools that modify effective optimization dynamics without altering the underlying search strategy.

The remainder of the thesis is organized as follows. Chapter 2 introduces the quantum substrate and formalizes the interface between circuit parameters and measurement statistics. Chapter 3 develops the central conceptual framework of variational quantum algorithms as hybrid control systems, establishing the separation between objective generation and parameter-update dynamics. Chapter 4 surveys population-based classical optimization methods in the context of variational

algorithms. Chapter 5 focuses specifically on harmonic oscillator particle swarm optimization and its application to variational quantum eigensolvers. Chapter 6 analyzes classical regularization techniques and their impact on robustness and convergence in variational optimization. Finally, Chapter 7 discusses implications, limitations, and directions for future work.

Chapter 2

Quantum Substrate and Objective Construction

2.1 Parameterized Quantum States

Variational quantum algorithms operate by preparing a family of quantum states parameterized by a set of real-valued variables. Let $\theta \in \mathbb{R}^P$ denote a vector of variational parameters. A parameterized quantum circuit implements a unitary transformation $U(\theta)$ acting on a fixed reference state, typically chosen as the computational basis state $|0\rangle^{\otimes n}$. The resulting quantum state is given by

$$|\psi(\theta)\rangle = U(\theta)|0\rangle^{\otimes n}. \tag{2.1}$$

The circuit architecture—defined by the choice of gate set, entangling operations, parameter placement, and circuit depth—fully specifies the mapping from the parameter vector θ to the prepared quantum state. Different ansatz constructions therefore generate distinct families of states, even when acting on the same number of qubits and optimizing the same cost function.

Crucially, the set of states reachable by varying θ is generally a strict subset of the full Hilbert space associated with the quantum system. This restriction arises from both practical and structural considerations: near-term quantum devices impose limits on circuit depth and connectivity, while variational algorithms deliberately employ finite parameterizations to maintain trainability. As a result, the variational ansatz defines a restricted family of quantum states that can be efficiently prepared and parameterized on quantum hardware. The dimensionality of this family is governed by the number of parameters and the circuit design, rather than by the exponential dimension of the underlying Hilbert space.

This restriction plays a central role in the behavior of variational algorithms. On the one hand, it enables optimization to be performed with a finite number of parameters on noisy hardware. On the other hand, it imposes a strong inductive bias on the optimization process by constraining which quantum states can be explored. The success of a variational algorithm therefore depends on whether the chosen ansatz is sufficiently expressive to approximate the target state of interest, while remaining sufficiently structured to allow efficient and stable optimization [2], [4], [9].

Throughout this thesis, parameterized quantum circuits are treated as state-preparation mechanisms that define the accessible search space for optimization. The classical optimizer does not

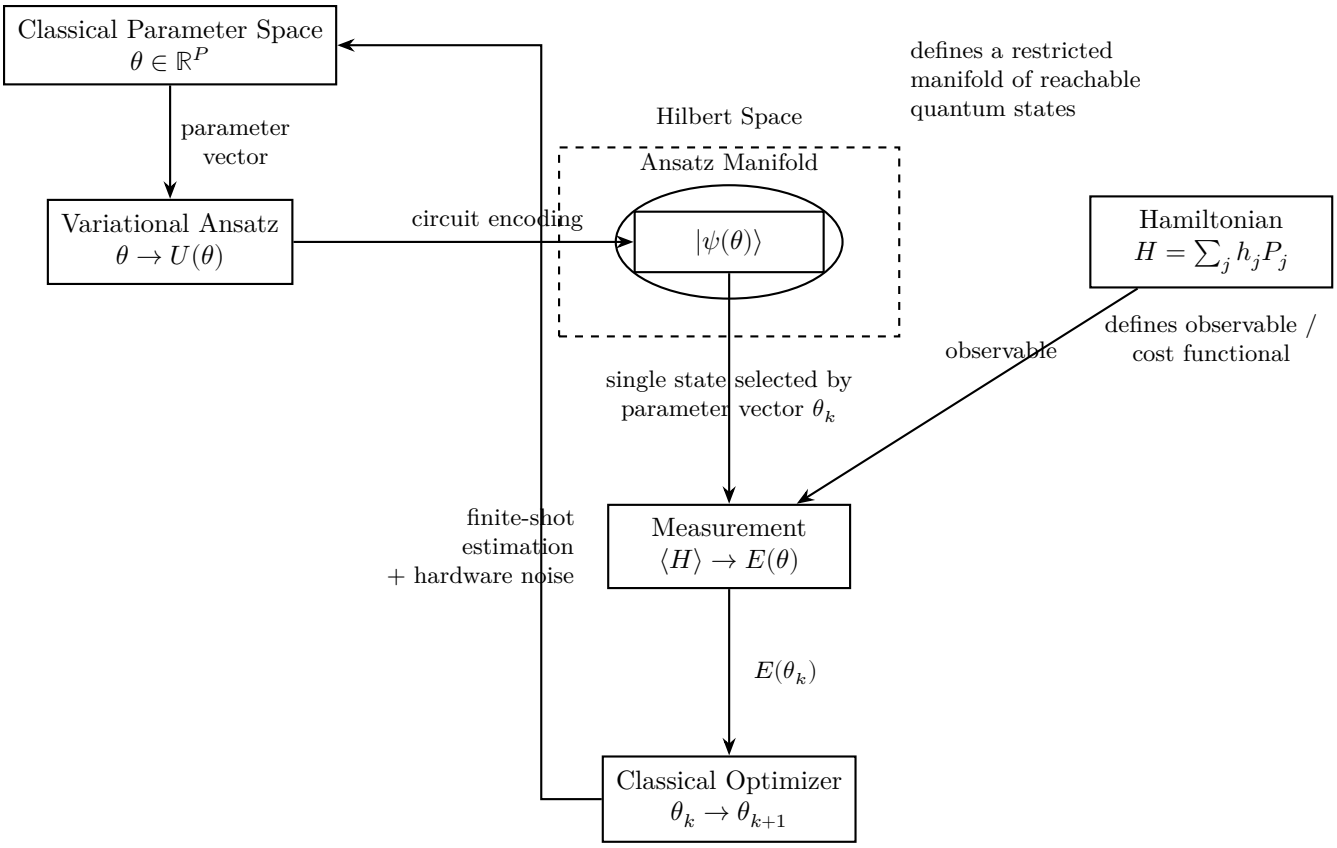

Figure 2.1: Variational quantum algorithm as a hybrid control system. The classical parameter space $\theta \in \mathbb{R}^P$ is mapped through the variational ansatz to quantum states $|\psi(\theta)\rangle$ that lie on a restricted manifold within Hilbert space. The Hamiltonian H defines the objective functional but does not parameterize the circuit. Measurement yields noisy energy estimates $E(\theta)$ subject to finite-shot statistics and hardware noise. The classical optimizer processes these estimates and updates parameters via feedback, forming a closed control loop.

interact directly with quantum states, but instead manipulates the parameter vector θ, thereby indirectly steering the prepared state within the restricted family defined by the ansatz. No feedback or adaptivity occurs within the quantum state preparation itself; all adaptive behavior is confined to the classical parameter-update process.

2.2 Hamiltonians and Cost Functions

In variational quantum algorithms, the objective function is derived from a Hamiltonian operator H acting on the Hilbert space of the quantum system. The Hamiltonian encodes the physical problem of interest—for example, the electronic structure of a molecule or the energy landscape of a spin model—and is represented as a Hermitian operator whose eigenvalues correspond to measurable energy levels [2].

For practical implementation on quantum hardware, the Hamiltonian is expressed as a weighted sum of Pauli operators,

$$H = \sum_j h_j P_j, \tag{2.2}$$

where each P_j is a tensor product of single-qubit Pauli matrices and the identity, and the coefficients $h_j \in \mathbb{R}$ are determined by the problem formulation. In quantum chemistry applications, this decomposition typically arises from a second-quantized fermionic Hamiltonian that has been mapped to qubits using transformations such as Jordan–Wigner or Bravyi–Kitaev [12], [13], [14]. Regardless of the mapping, the resulting Hamiltonian takes the form of a linear combination of Pauli strings, which are directly measurable on quantum hardware.

Given a parameterized quantum state $|\psi(\theta)\rangle$, the cost function is defined as the expectation value of the Hamiltonian,

$$C(\theta) = \langle\psi(\theta)|H|\psi(\theta)\rangle. \tag{2.3}$$

This scalar quantity represents the expected energy of the prepared state with respect to the problem Hamiltonian. Minimizing $C(\theta)$ over the parameter space corresponds to approximating the ground-state energy of H within the family of states accessible to the chosen variational ansatz.

The expectation value formulation is central to the variational approach. Rather than attempting to prepare an exact eigenstate of the Hamiltonian, which may be infeasible on near-term devices, the algorithm seeks parameters that minimize the energy expectation value. By the variational principle, this quantity provides an upper bound on the true ground-state energy, ensuring that improvements in the cost function correspond to physically meaningful progress toward the target solution.

From the perspective of the classical optimizer, the Hamiltonian enters only through the scalar-valued cost function $C(\theta)$. All operator-level structure is therefore collapsed into a single real-valued signal prior to classical processing. The optimizer does not access the Hamiltonian matrix, its eigenstates, or the quantum state itself. Instead, all information about the problem is mediated through repeated evaluations of the expectation value at selected points in parameter space. This separation between the operator-level problem definition and the scalar objective supplied to the optimizer is a defining feature of variational quantum algorithms and has significant implications for optimization strategy and performance [2].

2.3 Expectation Value Estimation and Stochasticity

On idealized simulators, the cost function $C(\theta) = \langle\psi(\theta)|H|\psi(\theta)\rangle$ can be evaluated exactly. On quantum hardware, however, this quantity must be estimated from finite measurement statistics. Each Pauli term P_j in the Hamiltonian decomposition is measured independently, and the corresponding expectation values are reconstructed from repeated projective measurements performed in appropriate measurement bases [2], [4].

As a consequence, evaluations of the cost function are inherently stochastic. For a fixed parameter vector θ, repeated executions of the same circuit yield different measurement outcomes due to finite-shot sampling. The resulting estimator of $C(\theta)$ therefore exhibits statistical fluctuations whose magnitude depends on the number of measurement shots, the variance of the individual Pauli terms, and the grouping strategy used during measurement. In practical settings, these fluctuations cannot be eliminated entirely and instead represent a fundamental source of noise in the optimization process [2], [4], [6].

In addition to sampling noise, hardware imperfections introduce further sources of variability. Gate errors, decoherence, readout noise, and calibration drift all perturb measurement outcomes, leading to systematic and stochastic deviations from ideal expectation values. These effects may vary over

time and across qubits, causing the effective objective function experienced by the optimizer to differ from the nominal cost function defined by the Hamiltonian [2], [3], [15].

From the perspective of the classical optimizer, the objective function is therefore not a deterministic scalar function of the parameters, but a noisy estimate obtained through repeated interaction with the quantum device. The stochasticity resides in the objective evaluations themselves rather than in any memory or state retained by the quantum subsystem. Importantly, the optimizer does not have direct access to the underlying noise-free objective or its exact gradients. Instead, it must infer update directions based solely on noisy scalar evaluations of $C(\theta)$ [2].

This stochasticity has significant implications for optimization dynamics. Classical algorithms that rely on precise gradient information or assume smooth, deterministic objective functions may exhibit degraded performance or instability when applied to variational quantum algorithms. Even in the absence of hardware noise, finite-shot sampling alone is sufficient to introduce non-negligible uncertainty into cost evaluations, particularly for large Hamiltonians or shallow measurement budgets.

Throughout this thesis, the expectation value estimation process is treated as an integral component of the variational algorithm rather than as a secondary implementation detail. The stochastic nature of cost evaluations is assumed to be intrinsic to the problem setting and is explicitly accounted for when analyzing optimization behavior, convergence properties, and stabilization techniques.

2.4 Interface to Classical Optimization

The quantum component of a variational quantum algorithm supplies information to the classical optimizer exclusively through scalar-valued estimates of the cost function $C(\theta)$. From the perspective of the classical optimization layer, the quantum device acts effectively as an oracle that maps parameter vectors $\theta \in \mathbb{R}^P$ to noisy objective evaluations obtained through measurement. No additional structural information about the quantum state, the Hamiltonian operator, or the underlying Hilbert space is directly accessible to the optimizer [2], [4].

This restricted interface has important consequences. The classical optimizer does not have direct access to the quantum state $|\psi(\theta)\rangle$, to the Hamiltonian matrix H, or to exact gradients of the cost function: any information about the objective landscape must be inferred indirectly from repeated, stochastic evaluations at selected points in parameter space. Assumptions commonly made in classical optimization—such as smoothness, determinism, or reliable gradient information—are therefore not generally satisfied in the variational quantum setting. The resulting picture is a feedback loop between two distinct dynamical systems: the quantum layer defines a stochastic objective whose structure is fixed by the circuit, Hamiltonian, and hardware characteristics, while the classical layer implements an update rule that selects new parameter vectors from previously observed objective values and any internal state maintained by the algorithm [2].

Crucially, the classical optimizer operates entirely within the parameter space $\mathbb{R}^P$, while the quantum system evolves within a restricted family of states embedded in Hilbert space. The optimizer influences the quantum state only indirectly, through changes in the circuit parameters, and observes the effects of these changes only through noisy scalar feedback. This separation of roles implies that optimization performance depends not only on the expressivity of the variational ansatz, but also on how well the classical update dynamics are matched to the stochastic, constrained nature of the objective landscape.

In the subsequent chapters, this interface perspective serves as the foundation for treating variational quantum algorithms as hybrid control systems. By analyzing the classical optimization layer as an active dynamical component—rather than a passive numerical routine—it becomes possible to systematically investigate stability, robustness, and convergence behavior, as well as to motivate the use of alternative optimization strategies tailored to the variational quantum setting.

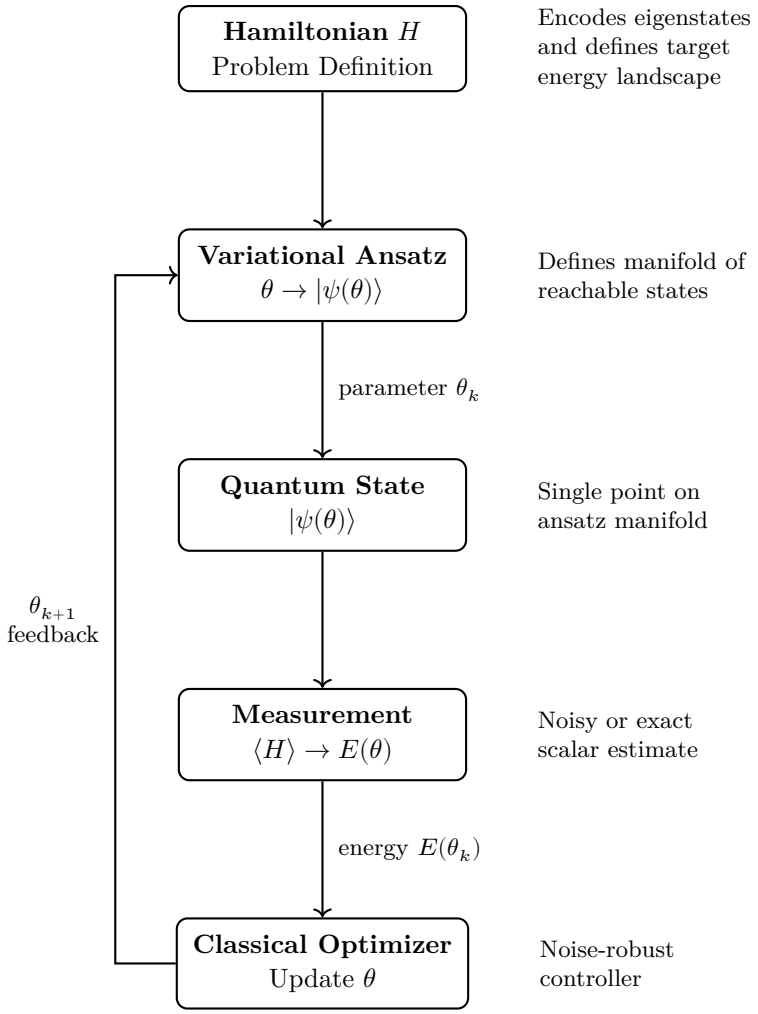

Figure 2.2: Complete variational quantum algorithm pipeline. The Hamiltonian defines the problem; the ansatz translates classical parameters into quantum states; measurement yields scalar objective values; and the classical optimizer updates parameters via feedback. Optimization corresponds to motion along the ansatz-defined manifold toward the ground state. Note: The Hamiltonian does not enter the variational ansatz or the state-preparation circuit; instead, it defines the objective functional through which prepared states are evaluated. H acts only during measurement, not during state preparation, and does not parameterize the circuit.

Chapter 3

Variational Quantum Algorithms as Hybrid Control Systems

3.1 Variational Optimization as a Feedback-Control Loop

Variational quantum algorithms may be naturally interpreted as discrete-time feedback-control systems [2], [16]. At each iteration, a classical optimization routine selects a parameter vector θ_k, which is passed to the quantum device for state preparation and measurement. The resulting measurement outcomes are processed to produce an estimate of the cost function $C(\theta_k)$, which is then fed back to the classical optimizer to determine the next parameter update. This process is repeated iteratively, forming a closed-loop interaction between the classical and quantum components.

From this perspective, the quantum subsystem functions, at the level of the variational interface, as a stochastic objective generator. Given an input θ_k, it produces an output $C(\theta_k)$ that reflects both the structure of the underlying Hamiltonian and the effects of finite sampling and hardware noise. The quantum device does not store memory across iterations within the variational loop, nor does it implement adaptive behavior; all temporal structure in the algorithm arises from the classical update dynamics.

The classical optimizer, by contrast, acts as a controller that governs the evolution of the parameter vector over successive iterations. Its update rule defines a discrete-time dynamical system whose state may include not only the current parameter vector, but also auxiliary variables such as momentum terms, population statistics, or historical evaluations. The choice of optimizer therefore determines how information obtained from the quantum device is accumulated, filtered, and transformed into subsequent parameter updates.

Crucially, the feedback loop couples two systems operating in different domains. The quantum subsystem evolves within a restricted family of states embedded in Hilbert space, while the classical subsystem evolves within a finite-dimensional parameter space. The optimizer does not observe the quantum state directly, but instead infers its behavior indirectly through noisy scalar feedback. As a result, the closed-loop dynamics cannot be reduced to properties of the quantum circuit or the classical algorithm in isolation, but must be understood as an emergent property of their interaction.

Interpreting variational quantum algorithms in this feedback-control framework provides a unifying lens through which stability, convergence, and failure modes can be analyzed. Rather than treating optimization difficulties as static features of an objective landscape, this viewpoint emphasizes the

role of update dynamics and feedback structure in shaping algorithmic behavior. This perspective forms the foundation for the system-level analysis of variational quantum algorithms as hybrid control systems developed in the remainder of this chapter.

3.2 Separation of Objective Generation and Update Dynamics

A central feature of variational quantum algorithms is the separation between objective generation and parameter-update dynamics. Although these components are coupled through a feedback loop, they play fundamentally different roles within the overall system. The quantum subsystem defines the objective function through state preparation and measurement, while the classical subsystem determines how that objective is explored over time.

The quantum component specifies the cost landscape implicitly. For a given circuit architecture, Hamiltonian, and hardware configuration, the mapping from parameters θ to expectation values $C(\theta)$ is fixed. This mapping determines which regions of parameter space correspond to low or high energy, the presence of flat regions or sharp features, and the degree of stochastic variability arising from sampling and noise.

By contrast, the classical optimizer defines the dynamics of exploration. Given access only to noisy scalar evaluations of the cost function, the optimizer selects parameter updates according to its internal update rules and state. Two optimizers interacting with the same quantum objective may therefore generate qualitatively different trajectories in parameter space, leading to markedly different convergence behavior even when all quantum resources are held fixed [2], [4].

This separation clarifies a common source of confusion in the analysis of variational algorithms. Properties such as expressivity, entanglement structure, or Hamiltonian locality influence the form of the objective landscape, but they do not uniquely determine optimization performance. Convergence speed, robustness to noise, sensitivity to initialization, and susceptibility to stagnation depend critically on the update dynamics imposed by the classical optimizer. From a system-level perspective, as a result, failures observed in practice cannot be attributed solely to deficiencies of the quantum circuit or to features of the Hamiltonian.

Treating objective generation and update dynamics as distinct components also enables a more precise classification of optimization behavior. For example, flat regions in the cost landscape may arise from properties of the ansatz or Hamiltonian, while prolonged stagnation within such regions is a dynamical phenomenon determined by the optimizer's response to noisy or uninformative feedback. Similarly, oscillatory or unstable behavior may reflect predominantly properties of the update rule rather than intrinsic features of the quantum objective.

This conceptual separation motivates a system-level approach to variational quantum algorithms. Rather than analyzing the quantum and classical components in isolation, it becomes possible to study how specific choices of update dynamics interact with the structure and noise characteristics of the quantum-generated objective. In the following sections, classical optimizers are therefore examined explicitly as discrete-time dynamical systems whose behavior shapes the overall performance of variational algorithms.

3.3 Classical Optimizers as Discrete-Time Dynamical Systems

The classical optimization component of a variational quantum algorithm can be formally viewed as a discrete-time dynamical system evolving in parameter space. At each iteration k, the optimizer maintains an internal state that includes the current parameter vector θ_k and, depending on the algorithm, additional auxiliary variables such as momentum terms, population members, velocity vectors, or historical cost evaluations. The update rule defines a state transition of the form

$$(\theta_{k+1}, \mathcal{H}_{k+1}) = \mathcal{F}(\theta_k, \mathcal{H}_k, C(\theta_k)), \tag{3.1}$$

where $\mathcal{H}_k$ denotes the internal memory of the optimizer and $\mathcal{F}$ specifies the update dynamics.

This formulation emphasizes that optimization in VQAs is not a static process of minimizing a fixed function, but a time-evolving trajectory governed by algorithm-specific dynamics. Even when the underlying objective landscape remains unchanged, different update rules induce different trajectories through parameter space. As a result, the behavior of the optimizer must be understood in terms of its dynamical properties rather than solely through its asymptotic convergence guarantees.

The stochastic nature of cost evaluations introduces an additional forcing term into these dynamics. Noise arising from finite-shot sampling and hardware imperfections acts as a source of random perturbations, effectively driving the optimizer with a noisy input signal. In this setting, update rules that rely on precise local information may amplify noise, while others may damp or average it over time. The response of an optimizer to stochastic forcing is therefore a critical determinant of its stability and robustness.

Classical optimization algorithms commonly used in variational quantum algorithms can be broadly distinguished by the structure of their internal dynamics. Gradient-based methods typically implement local descent dynamics that assume smoothness and reliable directional information. Momentum-based variants introduce additional state variables that couple past updates to current motion, modifying convergence behavior. Population-based methods, by contrast, evolve a collection of candidate solutions simultaneously, introducing interaction terms between population members and enabling collective exploration of parameter space.

Viewing these methods through a dynamical systems lens reveals that differences in optimization performance often stem from differences in update dynamics rather than from properties of the objective landscape itself. For instance, stagnation may arise when update rules fail to respond effectively to weak or noisy signals, while oscillatory behavior may result from overly aggressive updates or delayed feedback. These phenomena can occur even when the global structure of the objective landscape is favorable.

By framing classical optimizers as discrete-time dynamical systems driven by stochastic feedback, it becomes possible to analyze their behavior using concepts such as stability, attractors, and sensitivity to perturbations. This perspective provides a foundation for systematically comparing optimization strategies in the variational quantum setting and motivates the exploration of update rules designed explicitly to operate under noise, limited information, and constrained search spaces.

3.4 Stability, Convergence, and Failure Modes in Hybrid Variational Systems

Within the feedback-control interpretation of variational quantum algorithms, concepts such as stability and convergence must be understood in operational rather than purely analytical terms.

The objective function supplied by the quantum subsystem is stochastic, potentially non-smooth, and accessible only through finite sampling. As a result, classical notions of convergence based on exact gradients or deterministic descent trajectories are generally inapplicable.

In this context, convergence is most naturally defined in this work as the emergence of stable parameter dynamics over repeated iterations. An optimization process may be considered convergent if the parameter updates settle into a bounded region of parameter space and produce cost values that fluctuate within a narrow range around a minimum. Importantly, such convergence need not imply monotonic descent of the objective function, nor does it guarantee that the global optimum has been reached. Instead, it reflects the stabilization of the closed-loop dynamics formed by the interaction of the optimizer and the quantum substrate.

Operationally, stability refers to the system's response to perturbations, including stochastic fluctuations in cost evaluations and variations in initialization. A stable optimization process exhibits resilience to these perturbations, maintaining bounded parameter trajectories and consistent behavior across repeated runs. In contrast, unstable dynamics may manifest as unbounded parameter growth, persistent oscillations, or chaotic trajectories that prevent meaningful progress toward low-energy states.

Several characteristic failure modes are observed in practice. Stagnation occurs when parameter updates become negligibly small or ineffective, often due to flat regions in the objective landscape or insufficient sensitivity to noisy feedback. Oscillatory behavior may arise when update rules overreact to stochastic fluctuations, leading to repeated overshooting without net improvement. Divergence can occur when accumulated noise or aggressive update steps drive parameters outside regions where the objective landscape is informative or well-behaved.

Crucially, these failure modes cannot be attributed solely to properties of the quantum objective landscape. The same landscape may support stable convergence under one optimization strategy and exhibit persistent failure under another. This observation underscores the importance of treating optimization behavior as an emergent property of the hybrid system rather than as a static feature of the cost function.

By classifying optimization outcomes in terms of stability, convergence, and failure modes, it becomes possible to evaluate classical optimizers based on their dynamical compatibility with the variational quantum setting. This classification provides a framework for comparing optimization strategies beyond asymptotic performance metrics and highlights the need for update dynamics that are robust to noise, insensitive to precise gradient information, and capable of sustaining progress under constrained and stochastic feedback.

3.5 Implications for Optimizer Design in Variational Algorithms

The system-level perspective developed in the preceding sections has direct implications for the design and selection of classical optimizers in variational quantum algorithms. Because optimization behavior emerges from the interaction between stochastic objective generation and classical update dynamics, effective optimizers must be evaluated not only in terms of asymptotic convergence guarantees, but also in terms of their dynamical compatibility with noisy, constrained, and indirectly observed objective landscapes.

First, robustness to stochastic feedback is a fundamental requirement. Since cost evaluations are subject to finite-shot noise and hardware imperfections, optimizers must tolerate significant variabil-

ity in objective values without destabilizing parameter updates [3], [15]. Update rules that rely on precise local information or small differences between successive evaluations may perform poorly when noise levels are comparable to the signal being optimized. In contrast, strategies that aggregate information across iterations or across multiple candidate solutions may better suppress stochastic fluctuations.

Second, optimizers must be capable of operating effectively in landscapes that contain extended flat regions or weak gradients. Such regions may arise from properties of the variational ansatz, the Hamiltonian, or the measurement process itself [4]. In these settings, local descent methods may fail to generate meaningful update directions, leading to stagnation. Optimization strategies that incorporate exploration mechanisms or population-level dynamics can mitigate this risk by maintaining diversity and enabling movement even when local information is uninformative.

Third, sensitivity to initialization plays a critical role in practical performance. Because the variational landscape is often non-convex and constrained by circuit structure, different initial parameter choices can lead to qualitatively different trajectories and outcomes. Optimizers that exhibit strong dependence on initialization may produce highly variable results across repeated runs, complicating both empirical evaluation and practical deployment. Designs that reduce initialization sensitivity or enable recovery from unfavorable starting points are therefore advantageous.

Finally, optimizer design must account for the restricted interface between the classical and quantum components: only scalar-valued objective estimates are accessible at the classical–quantum boundary, and no direct access to gradients or state information is provided. Optimizers should therefore be formulated to operate under limited observability, and assumptions of smoothness, determinism, or exact gradient access should be avoided unless they can be justified within the variational quantum setting.

Taken together, these considerations suggest that optimization strategies developed specifically for noisy, partially observed, and constrained dynamical systems may be better suited to variational quantum algorithms than methods optimized for deterministic, high-precision classical objectives. In the following chapter, population-based optimization methods are examined from this perspective, with particular emphasis on their dynamical properties and their potential advantages in the variational quantum regime.

Chapter 4

Population-Based Optimization Methods in Variational Quantum Algorithms

4.1 Motivation for Population-Based Optimization

The design criteria identified in Chapter 3 from a system-level perspective highlight several challenges that classical optimizers must address in the variational quantum setting. Cost evaluations are stochastic, gradients may be unreliable or unavailable, and the objective landscape is constrained by the structure of the variational ansatz. Under these conditions, optimization strategies based on a single evolving parameter trajectory can be dynamically fragile, as they rely heavily on local information that may be weak, noisy, or misleading.

Population-based optimization methods offer a system-level alternative approach by evolving an ensemble of candidate solutions in parallel. Rather than maintaining a single parameter vector, these methods operate on a collection of parameter vectors that collectively explore the search space. Information about the objective landscape is obtained not only from individual evaluations, but also from interactions among population members, such as shared best-known solutions or aggregate statistics.

This ensemble-based structure provides several advantages in the variational quantum context. First, population-level dynamics naturally introduce redundancy, which can mitigate the impact of stochastic fluctuations in cost evaluations. Noise affecting individual candidates may be averaged out across the population, allowing more stable identification of promising regions in parameter space. Second, the simultaneous exploration of multiple trajectories reduces sensitivity to initialization, as poor starting points may be compensated by others that encounter lower-energy regions early in the optimization process.

Population-based methods are also well suited to landscapes containing extended flat regions or weak gradients. When local descent information is unavailable or unreliable, maintaining a diverse set of candidate solutions enables continued exploration without relying on precise directional cues. This property is particularly relevant for variational quantum algorithms, where barren plateaus and sampling noise can obscure local structure in the objective function [4].

From a dynamical systems perspective, population-based optimizers replace single-trajectory dynamics with coupled dynamics among multiple agents. Each agent evolves according to its own update rule while interacting with others through shared information channels. The resulting collective behavior can exhibit qualitatively richer dynamics than those of single-point optimizers, including adaptive exploration, self-correction, and resilience to perturbations.

These considerations motivate the examination of population-based optimization methods as candidates for variational quantum algorithms. By leveraging ensemble dynamics and information aggregation, such methods address several of the limitations identified in earlier chapters. The following sections formalize the structure of population-based optimizers and analyze their behavior in the context of noisy, constrained variational objectives.

4.2 General Structure of Population-Based Optimizers

Population-based optimization methods operate by evolving a set of candidate solutions simultaneously rather than a single parameter vector. Let $\{\theta_k^{(i)}\}_{i=1}^N \subset \mathbb{R}^P$ denote a population of N parameter vectors at iteration k. Each population member represents a candidate solution whose quality is evaluated through the same objective function accessible only through stochastic measurement $C(\theta)$ generated by the quantum subsystem.

At each iteration, the optimizer applies update rules that determine how each candidate evolves. In general form, the update dynamics can be written as

$$\theta_{k+1}^{(i)} = \mathcal{G}(\theta_k^{(i)}, \mathcal{I}_k^{(i)}, C(\theta_k^{(i)})), \tag{4.1}$$

where $\mathcal{G}$ denotes the population update rule and $\mathcal{I}_k^{(i)}$ represents information exchanged between population members. This information may include global or local best-known solutions, neighborhood statistics, or other aggregate measures derived from the population's collective performance.

A defining feature of population-based optimizers is the presence of interaction mechanisms that couple the dynamics of individual candidates. Unlike independent multi-start local search strategies, population methods allow information discovered by one candidate to influence the evolution of others. These interaction channels introduce collective behavior that cannot be reduced to a set of independent trajectories and give rise to emergent exploration–exploitation dynamics.

From a system-level perspective, population-based optimizers can be viewed as collections of coupled discrete-time dynamical systems driven by a shared, stochastic objective signal. Each candidate experiences its own noisy evaluations of the cost function, while the population as a whole aggregates information across multiple evaluations and iterations. The structure of the interaction rules determines how rapidly information propagates through the population and how diversity is maintained or suppressed over time.

Importantly, population-based methods do not require access to gradient information or detailed structural properties of the objective landscape. All updates are based on scalar objective evaluations and population-level statistics, making these methods naturally compatible with the restricted interface provided by variational quantum algorithms. The ability to operate under limited observability and stochastic feedback is therefore intrinsic to their design rather than an added feature.

This abstract formulation encompasses a wide range of algorithms, including particle swarm optimization, evolutionary strategies, and genetic algorithms [17], [18]. While these methods differ in their specific update rules and interaction mechanisms, they share a common structural foundation:

parallel exploration of parameter space coupled through information exchange. In the following sections, this structure is examined in greater detail, with particular emphasis on how population dynamics influence noise tolerance and optimization behavior in variational quantum settings.

4.3 Noise Tolerance and Information Aggregation

A key advantage of population-based optimization methods in the variational quantum setting lies in their inherent tolerance to noisy objective evaluations. As established in Chapter 2, cost function estimates obtained from quantum hardware are subject to finite-shot sampling noise and device imperfections [3], [15]. From the perspective of the optimizer, this noise acts as a stochastic perturbation on the objective signal and cannot be eliminated entirely without prohibitive measurement overhead.

Population-based methods mitigate the impact of such noise through information aggregation across multiple candidates and iterations. Because each population member evaluates the objective independently, stochastic fluctuations affecting individual evaluations are partially decorrelated across the population. Aggregating performance information—such as identifying population-wide best solutions or computing summary statistics—effectively averages out noise contributions and enhances the signal associated with consistently low-energy regions of parameter space.

This averaging effect is particularly important when the variance of the objective estimator is comparable to the magnitude of meaningful cost differences. In such regimes, single-trajectory optimizers may respond erratically to noisy evaluations, mistaking random fluctuations for genuine descent directions. Population-based methods, by contrast, rely on collective trends rather than isolated measurements, reducing sensitivity to outliers and transient noise-induced artifacts.

Information aggregation also occurs temporally. Many population-based algorithms retain memory of previously encountered low-cost solutions, allowing favorable information to persist even when subsequent evaluations are degraded by noise. This temporal smoothing stabilizes update dynamics and reduces the likelihood of catastrophic reversals caused by single noisy measurements. In effect, the population acts as a distributed memory system that filters stochastic feedback over time.

Importantly, noise tolerance in population-based methods does not require explicit noise modeling or adaptive measurement strategies. In many cases, robustness emerges naturally from the structure of the optimization dynamics. By maintaining multiple candidates and coupling their evolution through shared performance information, these methods transform stochastic objective evaluations into more reliable guidance signals.

In the context of variational quantum algorithms, where noise is intrinsic and often dominant, this property represents a significant practical advantage. Population-based optimizers can continue to make progress under noise levels that would severely degrade gradient-based or single-trajectory methods. The following section examines how these same population dynamics support a balance between global exploration and local refinement, further contributing to robust optimization behavior.

4.4 Exploration–Exploitation Tradeoffs in Population Dynamics

Effective optimization in variational quantum algorithms requires a balance between exploration of the parameter space and exploitation of regions associated with low cost values [19]. Exploration

enables the discovery of promising areas in complex, non-convex landscapes, while exploitation refines candidate solutions once such areas have been identified. Achieving this balance is particularly challenging in the presence of stochastic objective evaluations and constrained search spaces.

Population-based optimization methods address this tradeoff through their collective dynamics. By maintaining multiple candidate solutions simultaneously, these methods support parallel exploration of distinct regions of parameter space. Diversity among population members allows the optimizer to sample broadly, reducing the risk of premature commitment to suboptimal regions driven by noise or local structure.

At the same time, interaction mechanisms within the population promote exploitation. Information about high-performing candidates is shared across the population, guiding other members toward regions associated with lower cost values. This coupling introduces an implicit bias toward exploitation without requiring explicit gradient information or deterministic descent rules. The relative strength of exploratory versus exploitative behavior is governed by the update dynamics and interaction structure of the algorithm.

Importantly, in contrast to many deterministic schemes, population-based methods do not enforce a fixed exploration–exploitation schedule. Instead, the balance emerges dynamically as a result of population interactions and objective feedback. Early in optimization, diversity is typically high, and exploration dominates. As candidate solutions begin to cluster around favorable regions, exploitative dynamics become more pronounced. This adaptive behavior is well suited to variational quantum settings, where the reliability of local information may change over the course of optimization.

However, this same flexibility introduces potential challenges. Excessive exploitation may lead to premature convergence, reducing diversity before the population has adequately explored the landscape [20]. Conversely, excessive exploration may prevent refinement, resulting in slow convergence or persistent variability. Understanding how population dynamics regulate this tradeoff is therefore essential for designing optimizers that remain effective under noisy and constrained conditions.

In the variational quantum context, the ability to sustain exploration in the absence of reliable local gradients is particularly valuable. Population-based methods can continue to probe the parameter space even when objective evaluations provide weak or ambiguous signals, while still exploiting aggregate information when consistent trends emerge. The following section examines limitations of standard population-based approaches and motivates the introduction of structured dynamics to address these challenges.

4.5 Limitations of Standard Population-Based Methods

Despite their advantages in noisy and constrained optimization settings, standard population-based methods in their unstructured forms exhibit limitations that are particularly relevant in variational quantum algorithms. While ensemble dynamics can improve robustness and exploration, they do not, by themselves, guarantee stable or efficient convergence under stochastic feedback.

A common issue is premature convergence [20]. Population members may rapidly cluster around a candidate solution that appears favorable due to stochastic fluctuations rather than genuine structure in the objective landscape. Once diversity is lost, the population's ability to explore alternative regions is severely diminished, increasing the likelihood of stagnation in suboptimal regions. In variational quantum settings, where noise can obscure true cost differences, this effect

may be exacerbated.

Population-based methods can also be sensitive to hyperparameter choices that govern interaction strength, exploration rates, or update magnitudes [21]. Poorly tuned parameters may lead to overly aggressive attraction dynamics, causing oscillatory or unstable behavior, or to excessively diffusive dynamics that prevent meaningful refinement. Because the objective signal is noisy and indirectly observed, identifying appropriate hyperparameter regimes can be challenging and problem-dependent.

Another limitation concerns the lack of explicit dynamical structure in many standard formulations. In algorithms such as conventional particle swarm optimization, update rules are often motivated heuristically rather than derived from a principled dynamical model [17]. As a result, stability properties under stochastic forcing are not always well understood, and behavior may degrade unpredictably as noise levels increase or objective feedback weakens.

Computational overhead also merits consideration. Population-based methods require multiple objective evaluations per iteration, which may increase measurement cost on quantum hardware. While parallelism can mitigate wall-clock time, total sampling effort may still be significant, particularly for Hamiltonians with many measurement terms. Balancing population size against measurement budget is therefore an important practical concern.

These limitations do not negate the suitability of population-based optimization for variational quantum algorithms, but they do indicate that naive or unstructured population dynamics may be insufficient. To fully exploit the benefits of ensemble-based approaches, it is desirable to introduce update mechanisms with well-defined dynamical properties that promote stability, preserve diversity, and regulate exploration under stochastic feedback. The following chapter introduces harmonic oscillator particle swarm optimization as one such structured approach, designed to address these challenges within the variational quantum setting.

4.6 Summary and Transition

This chapter examined population-based optimization methods as candidates for variational quantum algorithms, motivated by the stochastic, constrained, and indirectly observed nature of variational objectives. By evolving multiple candidate solutions in parallel and coupling their dynamics through information exchange, population-based methods address several limitations of single-trajectory optimizers, including sensitivity to noise, weak local information, and strong dependence on initialization.

The analysis highlighted how ensemble dynamics support noise tolerance through information aggregation, enable adaptive exploration in the absence of reliable gradients, and promote robustness under stochastic feedback. At the same time, it identified important limitations of standard population-based approaches, such as premature convergence, sensitivity to hyperparameters, and the absence of explicitly structured dynamics capable of regulating behavior under noisy conditions.

These observations suggest that while population-based optimization provides a promising foundation for variational quantum algorithms, additional structure is required to ensure stable and reliable performance. In particular, update mechanisms that incorporate well-defined dynamical properties may offer improved control over exploration–exploitation tradeoffs and enhanced robustness to stochastic forcing.

The following chapter introduces harmonic oscillator particle swarm optimization as a structured

population-based method designed to address these challenges. By embedding particle dynamics within a harmonic oscillator framework, this approach provides explicit mechanisms for stability and regulation, making it a natural candidate for optimization in variational quantum algorithms.

Chapter 5

Harmonic Oscillator Particle Swarm Optimization for Variational Quantum Algorithms

5.1 Motivation and Context

The performance of variational quantum eigensolvers depends critically on the behavior of the classical optimizer responsible for updating circuit parameters [1], [2], [7]. Quantum objective functions are high-dimensional, non-convex, periodic, and inherently stochastic due to finite-shot measurement noise [4], [6]. These features impose structural demands on optimization dynamics that differ substantially from those encountered in deterministic classical benchmarks.

This chapter develops Harmonic Oscillator Particle Swarm Optimization (HOPSO) and its adaptation to variational quantum eigensolvers as the central technical contribution of the thesis. The algorithm was introduced in [22] as a structured alternative to heuristic swarm dynamics on classical benchmarks, and the VQE adaptation, periodic-boundary handling, and shot-noise-aware analysis reported here were developed in joint work by the author with Mohammad Aamir Sohail, Martin Plesch, and Ijaz Ahmad [23]. HOPSO replaces the heuristic velocity update of standard PSO with trajectories derived from the analytical solution of a damped harmonic oscillator, introducing explicit dynamical structure into particle motion and providing direct control over convergence through physically interpretable parameters.

We first present the HOPSO formulation, then describe its adaptation to the periodic, stochastic parameter landscapes that arise in VQE. The chapter concludes with an empirical evaluation on molecular Hamiltonians (H_2 and LiH), comparing HOPSO against standard PSO, COBYLA, and Differential Evolution under noiseless and shot-noise-limited estimation.

5.2 Harmonic Oscillator Particle Swarm Optimization

Let $x_{j,d}$ denote the position of particle j in dimension d. Following the standard particle-swarm convention, the *personal best* $p_{j,d}$ is the coordinate at which particle j has so far observed the lowest cost-function value, and the *global best* g_d is the corresponding coordinate of the lowest-cost position seen by any particle in the swarm to date. For each particle and dimension, an attractor

is defined as a weighted combination of the personal best and the global best,

$$a_{j,d} = \frac{c_1 p_{j,d} + c_2 g_d}{c_1 + c_2}, \tag{5.1}$$

where $c_1, c_2 > 0$ are attraction coefficients.

Particle motion is modeled using the closed-form solution of a damped harmonic oscillator centered at a. Rather than discretizing a differential equation, HOPSO directly evaluates the analytical trajectory

$$x(t) = A_0 e^{-\lambda t} \cos(\omega t + \theta) + a, \tag{5.2}$$

where

- $\lambda > 0$ is the damping parameter,
- $\omega > 0$ is the oscillation frequency,
- A_0 is the initial amplitude,
- θ is the initial phase.

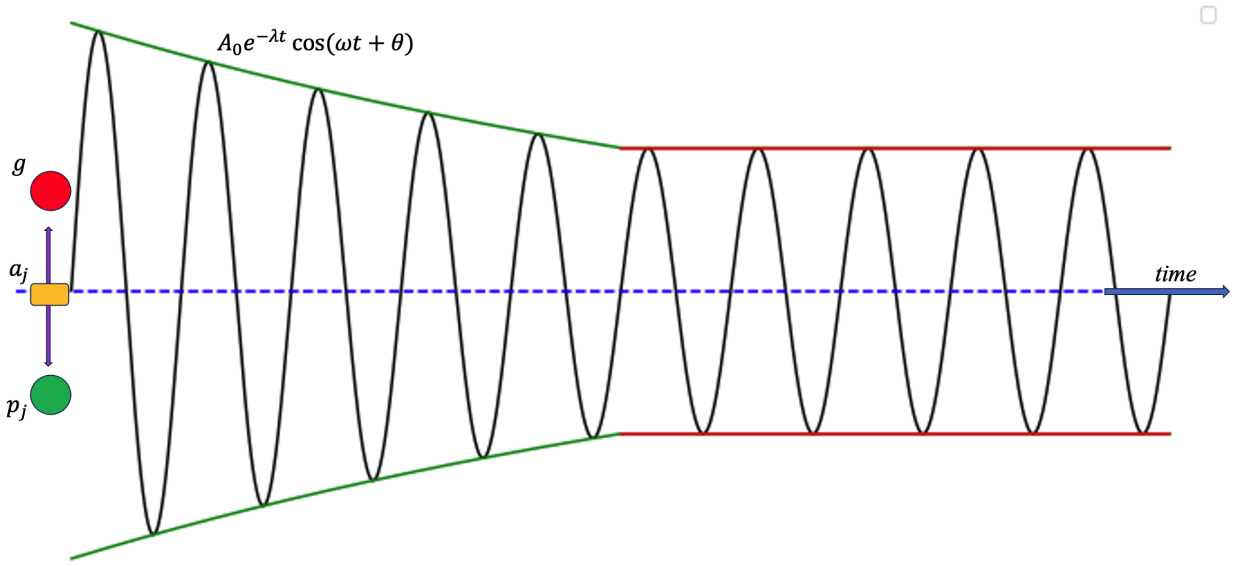

Figure 5.1: Damped-oscillator trajectory of an HOPSO particle. The particle position $x(t)$ evolves analytically toward the attractor a_j, weighted between the personal best p_j and the global best g, with amplitude bounded by the envelope $A_0 e^{-\lambda t}$.

Velocity is obtained by differentiation,

$$v(t) = -\omega A_0 e^{-\lambda t} \sin(\omega t + \theta) - \lambda(x(t) - a). \tag{5.3}$$

Given initial position $x(0)$ and velocity $v(0)$, the initial amplitude is

$$A_0 = \sqrt{(x(0) - a)^2 + \frac{(v(0) + \lambda(x(0) - a))^2}{\omega^2}}, \tag{5.4}$$

and the phase is determined by

$$\theta = \arccos\left(\frac{x(0) - a}{A_0}\right). \tag{5.5}$$

Instead of updating velocity via recursion, HOPSO samples the analytical trajectory at random time increments,

$$t_{k+1} = t_k + \text{rand}(0, t_{ul}), \tag{5.6}$$

where t_{ul} is typically chosen as one oscillation period. This stochastic time sampling provides exploration without introducing additive noise into the dynamics.

To prevent premature collapse of oscillatory motion, a minimum amplitude threshold is enforced,

$$A_{th} = \frac{|p_{j,d} - g_d|}{2}\, m, \tag{5.7}$$

where m is a tunable parameter. The effective amplitude is updated according to

$$A_0^{new} = \max(A_{current}, A_{computed}, A_{th}), \tag{5.8}$$

ensuring persistent exploration between personal and global best positions.

The damping parameter is linked to the total evaluation budget B and population size N through

$$\lambda = s\left(\frac{B}{N}\right)^{-1}, \tag{5.9}$$

where s is a scaling factor. This relation allows convergence rate to scale systematically with available computational resources.

This formulation replaces heuristic velocity recursion with analytically derived damped trajectories. Stability arises from exponential decay governed by λ, while exploration is maintained through oscillatory motion and amplitude control.

5.3 Adapting HOPSO to Variational Quantum Eigensolvers

When applied to VQE, HOPSO must account for two defining features of quantum parameter landscapes: periodicity and stochastic evaluation noise.

Circuit parameters corresponding to rotation gates are periodic with period 2π. To avoid discontinuities in attractor computation, personal best positions are confined to a dynamically selected angular window of length 2π, defined by an offset $r \in [0, 2\pi)$ sampled at initialization. When a personal best leaves this window, it is mapped back using modulo arithmetic, while particle positions themselves remain unrestricted.

Attractor computation respects circular topology. If the angular separation between personal and global best exceeds π, the shorter arc across the periodic boundary is used. Coordinate shifting and modular correction ensure averaging occurs along the minimal arc.

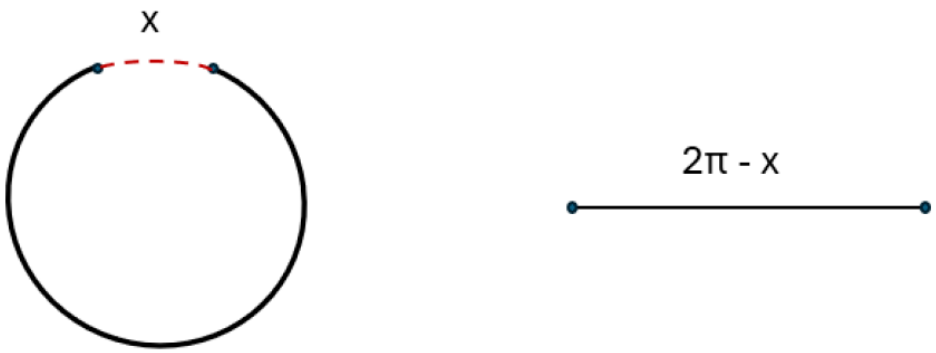

Figure 5.2: Wrapped angular distance on a periodic parameter axis. The two candidate paths between two points on the circle have lengths x and $2\pi - x$; HOPSO selects the shorter arc when computing attractors and amplitude thresholds.

The amplitude threshold is similarly modified using wrapped angular distance instead of linear distance. This guarantees meaningful oscillatory exploration near periodic boundaries.

For numerical stability, if phase computation becomes undefined due to floating-point inaccuracies, the corresponding particle is marked inactive and removed from further updates. This prevents discontinuities in the harmonic dynamics.

These modifications preserve the oscillator structure while aligning the algorithm with the topology of quantum parameter spaces.

5.4 Interaction with Shot Noise

Expectation values in VQE are estimated from finite measurement samples. Shot noise introduces stochastic fluctuations into cost evaluations.

Because HOPSO evolves particles according to structured harmonic motion around attractors, instantaneous fluctuations in cost do not directly translate into abrupt parameter updates. Population-level aggregation of personal and global best positions further reduces sensitivity to isolated noisy evaluations.

Exploration arises from oscillatory phase evolution and stochastic time sampling rather than from direct reaction to cost gradients. Damping suppresses large deviations, while amplitude thresholds maintain exploration. This combination yields stable convergence behavior under moderate shot noise.

5.5 Empirical Evaluation: Hydrogen Molecule

The algorithm was evaluated on the hydrogen molecule encoded as a four-qubit Hamiltonian via the Jordan–Wigner transformation. A hardware-efficient ansatz with three layers of parameterized

single-qubit rotations and linear entanglement was employed.

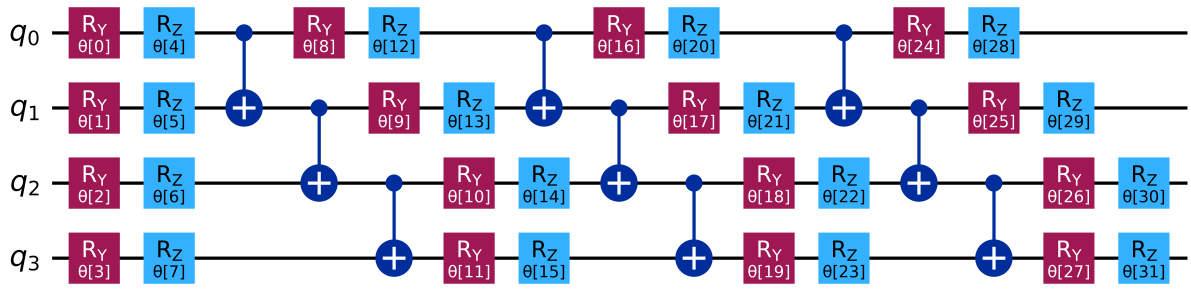

Figure 5.3: Hardware-efficient ansatz used for the hydrogen molecule benchmark in HOPSO-VQE. Four qubits, alternating R_Y–R_Z rotation blocks with a linear-chain CNOT entangler, repeated to yield a thirty-two-parameter circuit.

In the noiseless statevector setting, both HOPSO and COBYLA reached the chemical-accuracy threshold across all runs, with HOPSO producing the tightest dispersion of final energies. Standard PSO recovered the ground state for most runs but exhibited substantially wider variance, and Differential Evolution failed to converge below approximately -1.4 Ha.

When shot noise corresponding to 1000 measurement shots per evaluation was introduced, the relative ordering changed sharply. HOPSO retained the tightest dispersion and the lowest median energy, while COBYLA degraded substantially, exhibiting both a higher median and the largest interquartile range of any optimizer. Standard PSO degraded mildly, and Differential Evolution remained the worst performer.

These results indicate that while local optimizers may perform strongly in deterministic settings, structured population-based harmonic dynamics provide a clear advantage when objective evaluations are stochastic.

5.6 Empirical Evaluation: Lithium Hydride

The lithium hydride molecule was studied using an eight-qubit Hamiltonian obtained through symmetry tapering. The ansatz comprised four layers of parameterized rotations and entangling gates, yielding an 80-dimensional parameter space.

Large-scale parallel execution enabled extensive sampling across independent runs. In the noiseless case, both HOPSO and standard PSO achieved energies within ten times chemical accuracy of the exact ground-state value of ≈ -8.91 Ha, with HOPSO exhibiting marginally tighter spread (visible in the inset). Differential Evolution displayed substantially larger dispersion, with final energies between -6.0 Ha and -8.2 Ha.

Under shot noise, HOPSO retained strong performance and produced tightly clustered energy estimates near the ground state, with median energy below -8.7 Ha. Standard PSO exhibited markedly increased variability with median around -8.1 Ha, and Differential Evolution degraded further to a median above -6.5 Ha.

These observations demonstrate that harmonic oscillator–based dynamics scale effectively to higher-dimensional, noisy variational landscapes.

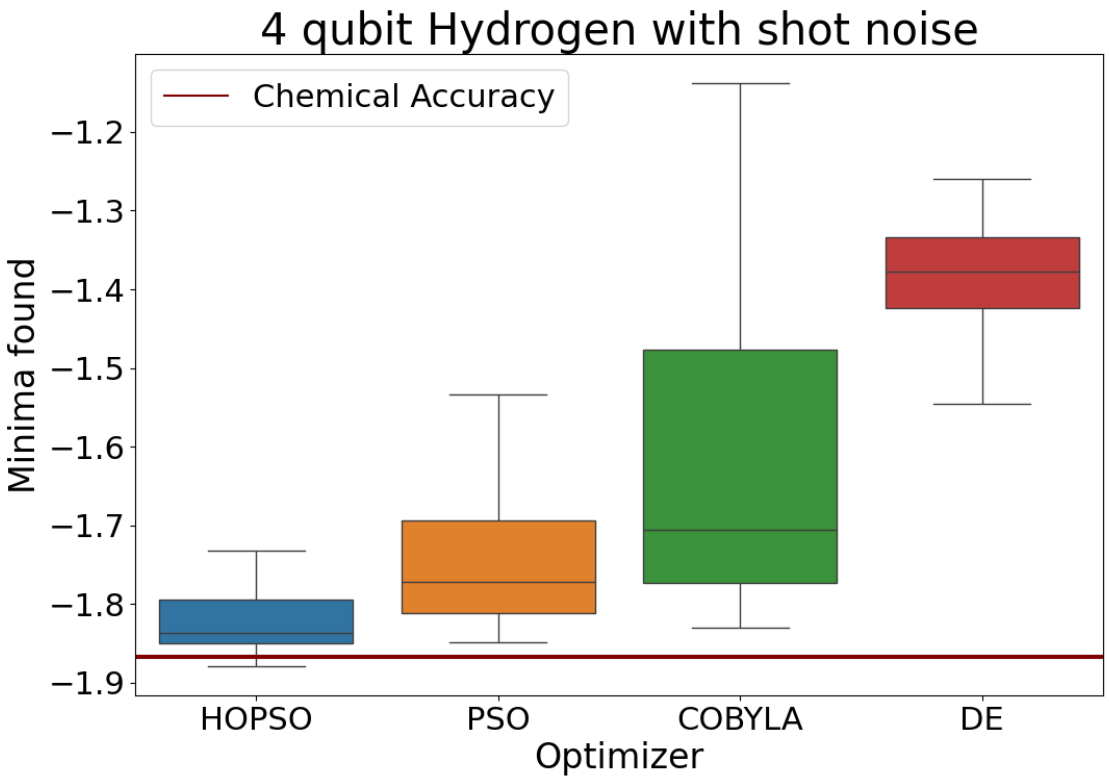

Figure 5.4: Final minima found by HOPSO, PSO, COBYLA, and Differential Evolution on the four-qubit hydrogen molecule under noiseless estimation. HOPSO and COBYLA both reach chemical accuracy (red line) with negligible spread; PSO recovers the minimum on most runs but with wider dispersion; DE fails to converge below −1.4 Ha.

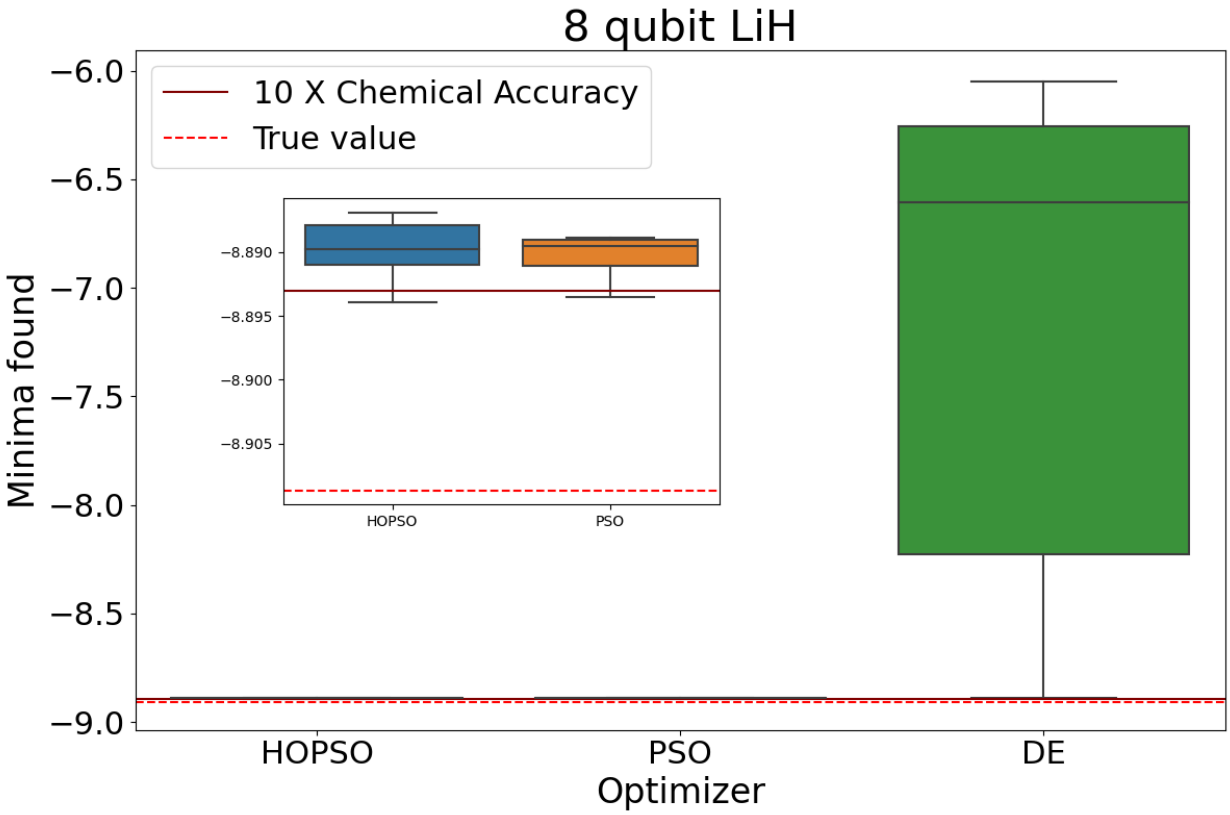

Figure 5.5: The same comparison under shot noise (1000 shots per evaluation). HOPSO retains a tight, low-energy distribution; COBYLA, the noiseless winner, degrades dramatically once the cost-function estimates become stochastic; PSO degrades mildly; DE remains the worst performer.

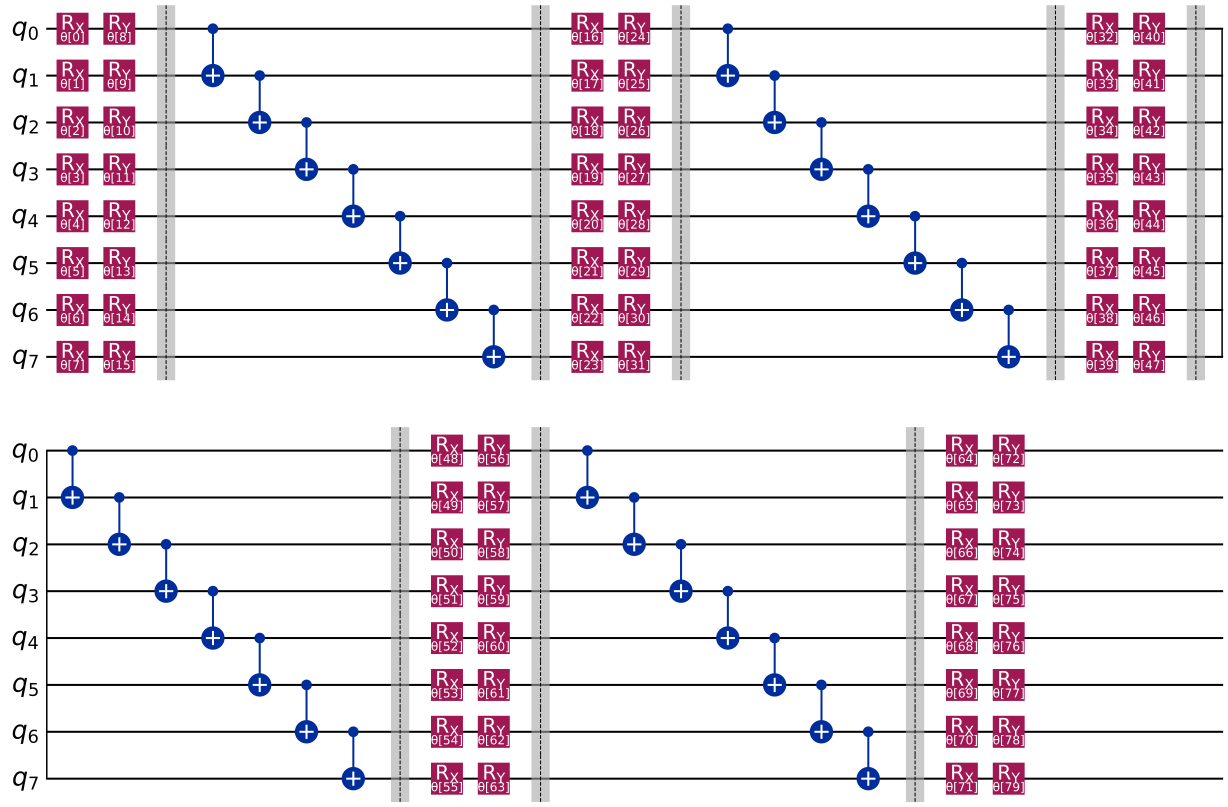

Figure 5.6: Eight-qubit ansatz used for the lithium hydride benchmark. Two repetition blocks of R_Y rotations interleaved with linear-chain CNOTs, yielding the eighty-dimensional parameter space referenced above.

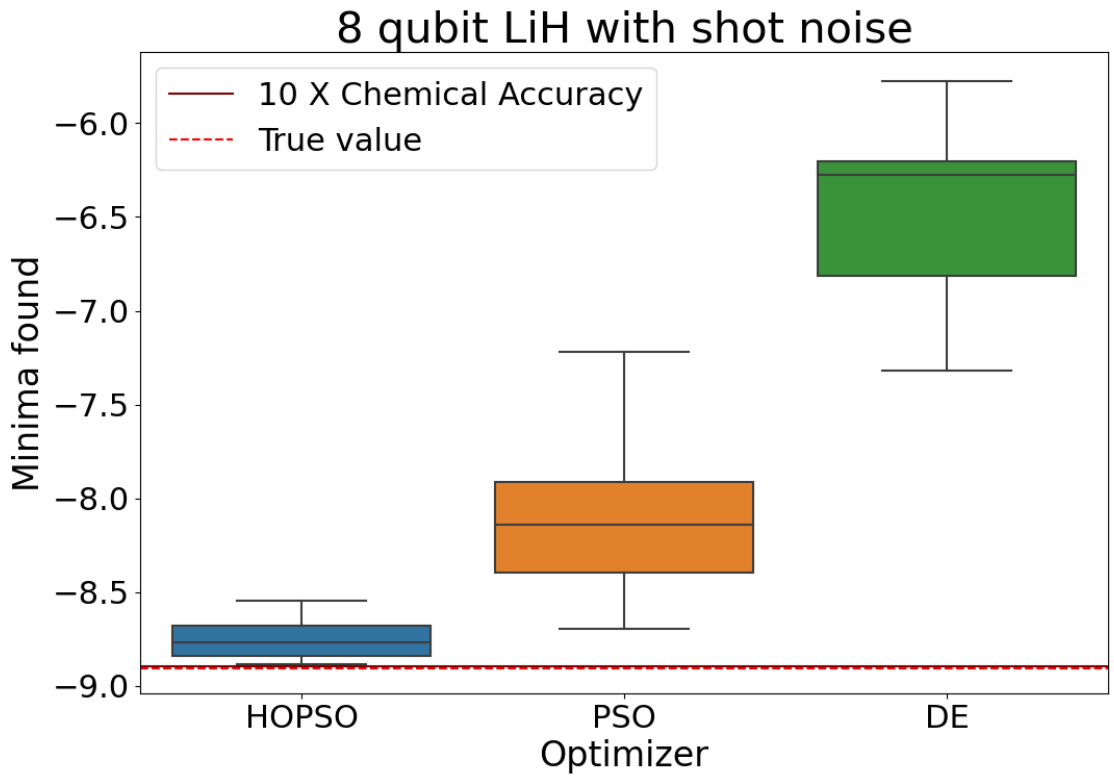

Figure 5.7: Final minima found by HOPSO, PSO, and Differential Evolution on the eight-qubit lithium hydride Hamiltonian under noiseless estimation. The inset zooms onto the HOPSO and PSO distributions; both cluster near the true ground-state value (dashed red line) within 10× chemical accuracy. DE fails to reach the same regime.

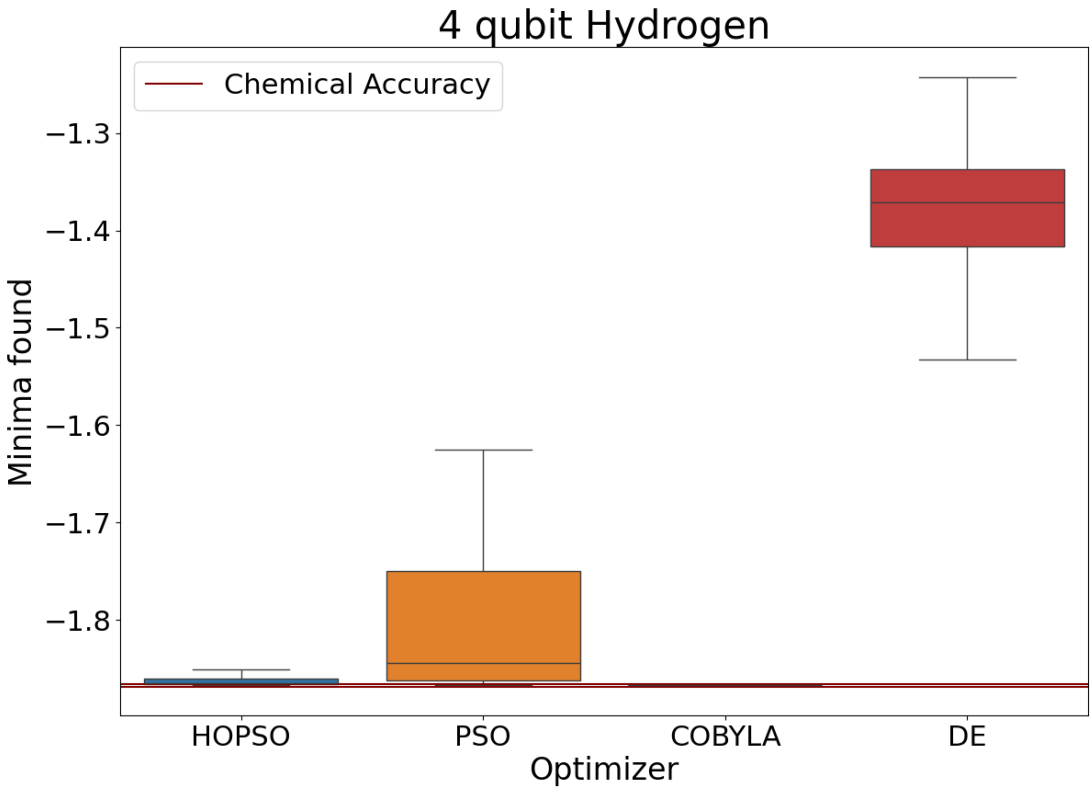

Figure 5.8: Same comparison under shot noise. HOPSO retains a narrow distribution near the true ground state; PSO median drifts upward and its dispersion widens; DE degrades further. The relative robustness of HOPSO observed for H_2 persists at the larger eight-qubit scale.

5.7 Discussion

HOPSO replaces heuristic swarm updates with analytically derived damped harmonic trajectories. Convergence behavior is governed explicitly by damping and amplitude parameters rather than by empirically tuned inertia factors.

The algorithm does not universally outperform local methods in noiseless settings. However, its structured oscillatory dynamics provide clear advantages when objective evaluations are noisy and parameter landscapes exhibit periodic structure.

The results support the view that explicitly structured dynamical models can enhance robustness in hybrid quantum classical optimization without modifying the underlying quantum circuit.

Chapter 6

Classical Regularization in Variational Quantum Eigensolvers

6.1 Motivation

Variational Quantum Eigensolvers often display unstable convergence, pronounced sensitivity to initialization, and significant dispersion across repeated optimization runs. These behaviors persist even under noiseless statevector simulation, indicating that they are not solely consequences of hardware noise or finite sampling effects. Rather, they reflect structural properties of the classical optimization landscape induced by expressive parameterized ansätze.

When a variational circuit contains many tunable parameters, the mapping from parameter space to quantum states is frequently redundant. Distinct parameter vectors may generate identical or nearly identical physical states while differing substantially in Euclidean norm. When the energy expectation value is regarded purely as a function of the classical parameter vector, such redundancies manifest as ill-conditioned curvature, shallow valleys, and narrow ridges. Optimization trajectories may therefore become erratic, exhibit high run-to-run variability, or stagnate in regions that are geometrically unstable rather than physically suboptimal.

To address these effects, we investigate quadratic L_2 regularization applied directly to the classical objective function. This approach modifies only the classical landscape supplied to the optimizer and does not alter the quantum circuit, Hamiltonian, or measurement procedure.

6.2 Regularized Objective

In standard VQE, the objective function is defined as

$$E(\theta) = \langle \psi(\theta) | \widehat{H} | \psi(\theta) \rangle, \tag{6.1}$$

where $\theta \in \mathbb{R}^P$ parameterizes the variational circuit.

We introduce a quadratic penalty on the parameter vector and define the regularized objective

$$\tilde{E}(\theta) = E(\theta) + \lambda \|\theta\|_2^2, \tag{6.2}$$

with regularization strength $\lambda \geq 0$.

Unlike its role in supervised learning, where L_2 regularization is typically associated with statistical generalization, its function here is purely geometric. The penalty term reshapes the curvature of the classical objective landscape by suppressing directions in parameter space that inflate the norm without meaningfully reducing the energy. The physical ground state of the Hamiltonian remains unchanged; only the geometry through which it is approached is modified.

6.3 Two-Stage Regularization Protocol

To ensure that regularization enhances conditioning without biasing the final reported energies, optimization is performed using a two-stage protocol.

In the first stage, regularization is active and its strength follows a cosine decay schedule,

$$\lambda(t) = \frac{\lambda_0}{2}\left[1 + \cos\left(\frac{\pi t}{T_A}\right)\right], \qquad 0 \leq t \leq T_A, \tag{6.3}$$

where λ_0 denotes the initial regularization strength and T_A is the number of iterations in this stage. This schedule ensures a gradual reduction of the penalty and avoids abrupt changes in the effective objective.

In the second stage, regularization is removed by setting $\lambda = 0$. Optimization resumes from the best parameter vector identified during Stage A and proceeds using the original objective function $E(\theta)$. This separation guarantees that the final energy values correspond to the physical Hamiltonian and are not biased by the penalty term.

6.4 Numerical Evaluation

The regularization framework was evaluated on three representative systems: the hydrogen molecule encoded on four qubits with a forty-parameter TwoLocal ansatz, lithium hydride encoded on eight qubits with an eighty-parameter ansatz of similar structure, and a twelve-qubit Random Field Ising Model with a single-layer ansatz. All simulations were performed under noiseless statevector estimation to isolate the geometric effects of regularization from sampling noise.

For the hydrogen molecule and the Random Field Ising Model, optimization was performed using the Conjugate Gradient method. For lithium hydride, L-BFGS-B was employed to accommodate the higher dimensional parameter space.

The initial regularization strength λ_0 was scanned across a discrete set of values spanning zero to moderate penalty strengths. For each value, thousands of independent random initializations were evaluated in order to obtain statistically reliable success rates. Confidence intervals were computed using Wilson interval estimates.

6.5 Success Rates and Stabilization Behavior

For a chosen chemical-accuracy threshold ΔE, a run is considered successful if

$$|E_{\text{final}} - E_{\text{exact}}| \leq \Delta E. \tag{6.4}$$

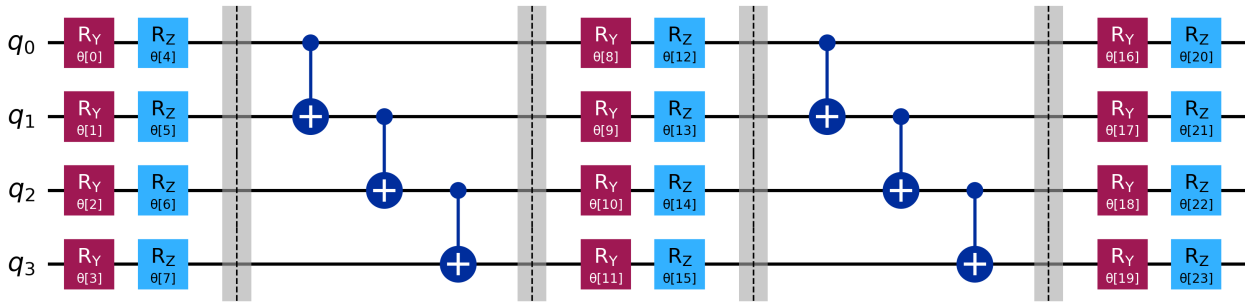

Figure 6.1: Hardware-efficient TwoLocal ansatz used for the four-qubit hydrogen Hamiltonian. Each layer applies single-qubit R_Y–R_Z rotations followed by a linear chain of CNOT gates; three such layers yield twenty-four trainable parameters per repetition.

Across all three systems, the dependence of success rate on λ_0 exhibits a consistent non-monotonic pattern. As the regularization strength increases from zero, performance initially improves. A broad interval of moderate regularization strengths produces substantially higher success rates and reduced dispersion across random initializations. For sufficiently large λ_0, performance deteriorates due to over-constraining of the parameter space.

The optimal regularization window λ_{opt} is defined as the contiguous range of λ_0 values for which the mean success rate remains at least ninety percent of its global maximum under the selected accuracy threshold. This definition captures a plateau of robust performance rather than a sharply tuned optimum.

For the hydrogen molecule, stabilization occurs approximately for λ between 0.10 and 0.18. For lithium hydride, the window shifts to smaller values, approximately between 0.005 and 0.025. For the Random Field Ising Model, stabilization appears roughly between 0.05 and 0.12. A shallow dip in success probability is observed at very small but nonzero regularization strengths, indicating that insufficient penalties can perturb the landscape without providing meaningful conditioning. As stricter accuracy thresholds are imposed, the preferred regularization strength shifts toward larger values.

Within the stabilization window, the interquartile range of final energies contracts significantly, parameter norms decrease systematically, and convergence trajectories become smoother. Importantly, the mean final energies obtained after the unregularized second stage remain statistically indistinguishable from the baseline case, confirming that moderate regularization improves conditioning without shifting the physical optimum.

6.6 Scaling Considerations

The optimal regularization strength decreases with increasing parameter dimension. This behavior is consistent with the heuristic scaling

$$\lambda_{\text{scale}} \sim \frac{\sum_i |c_i|}{P}, \tag{6.5}$$

where c_i are Hamiltonian coefficients and P is the number of variational parameters. Despite differences in system size and Hamiltonian structure, the presence of a broad stabilization window is

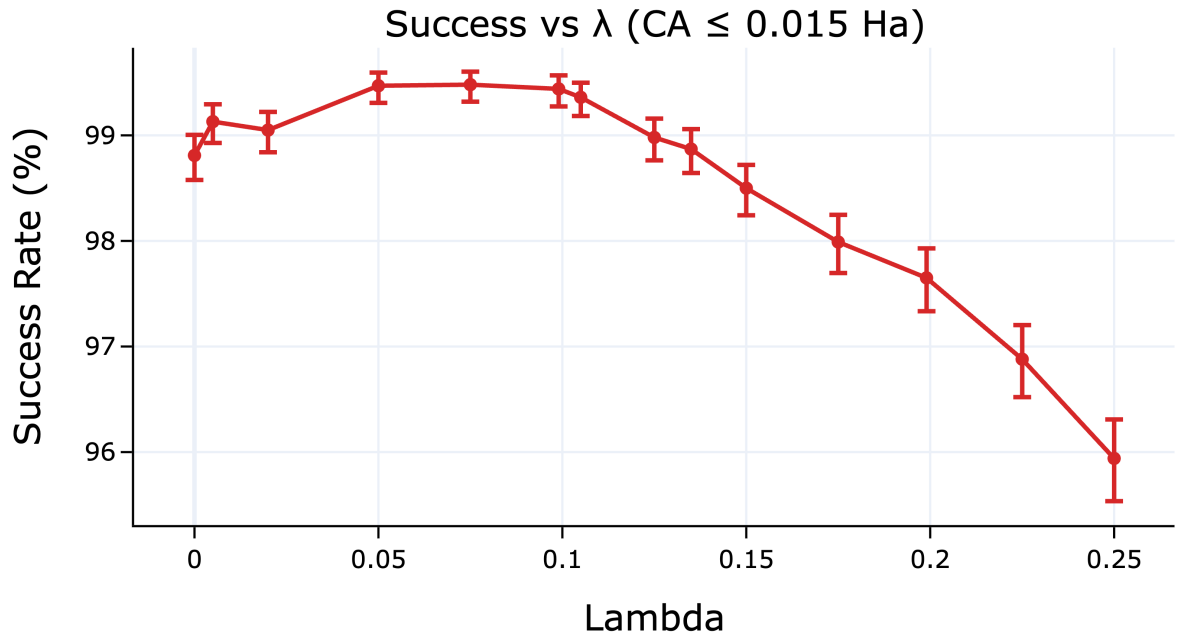

Figure 6.2: Hydrogen molecule success rates as a function of λ_0, evaluated at $\Delta E \leq 0.015$ Ha. A broad plateau of high success rates is visible across moderate regularization strengths.

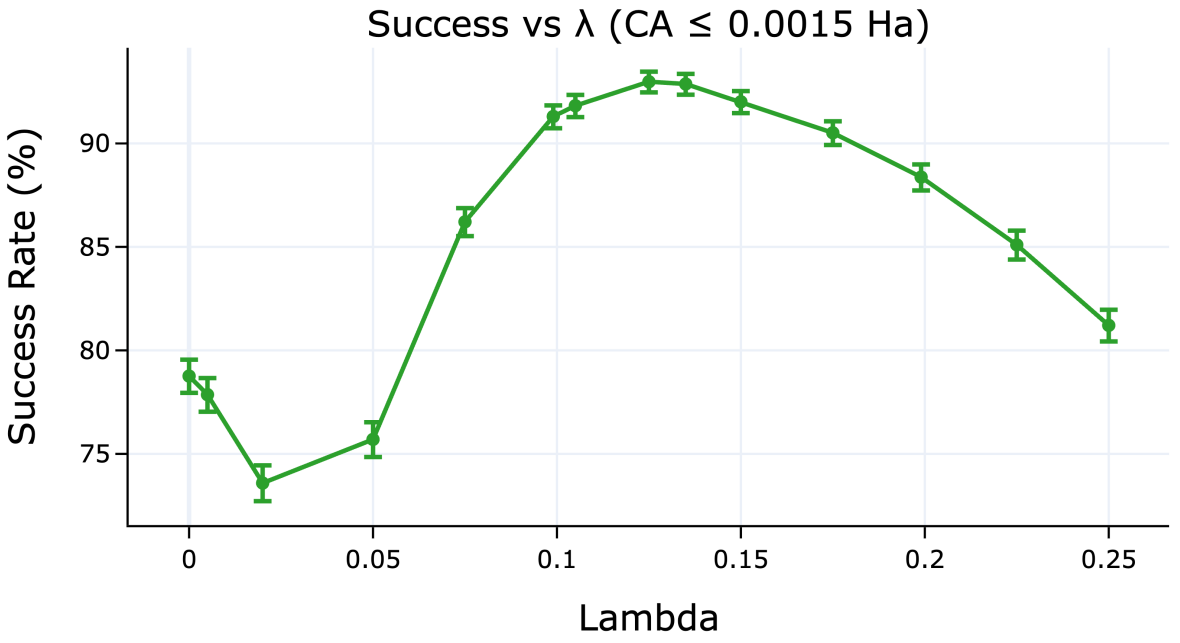

Figure 6.3: Hydrogen molecule, $\Delta E \leq 0.0015$ Ha (chemical accuracy). The non-monotonic dependence on λ_0 persists, with the optimal window remaining in the same neighbourhood.

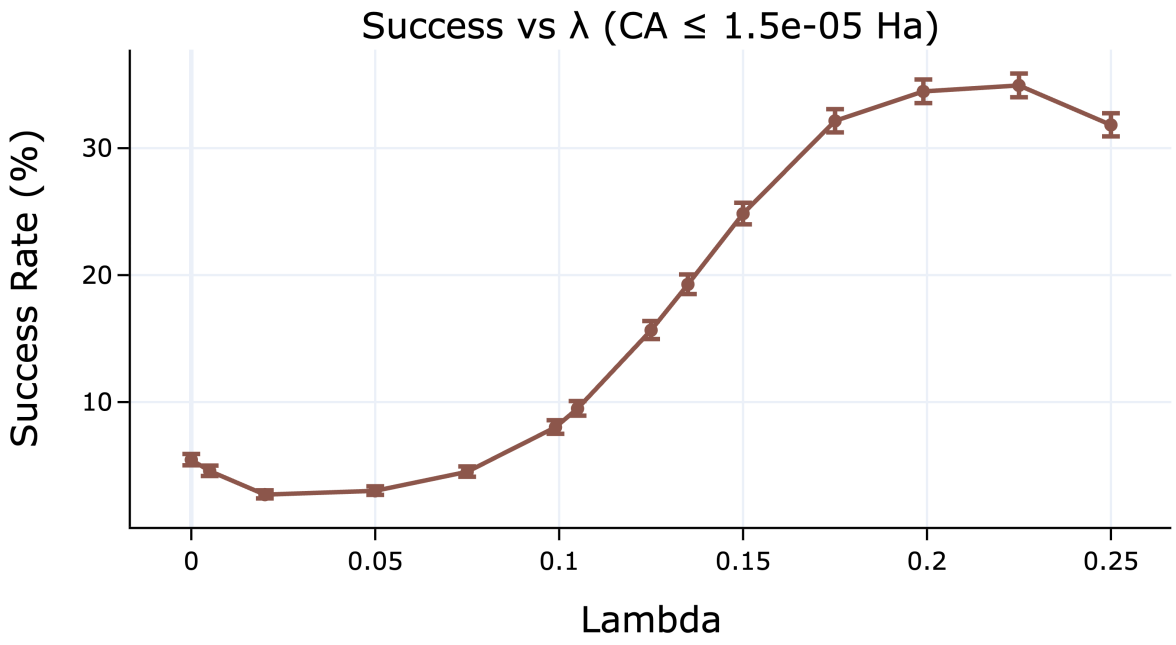

Figure 6.4: Hydrogen molecule, $\Delta E \leq 1.5 \times 10^{-5}$ Ha. As the threshold tightens, the preferred regularization strength shifts toward larger values of λ_0.

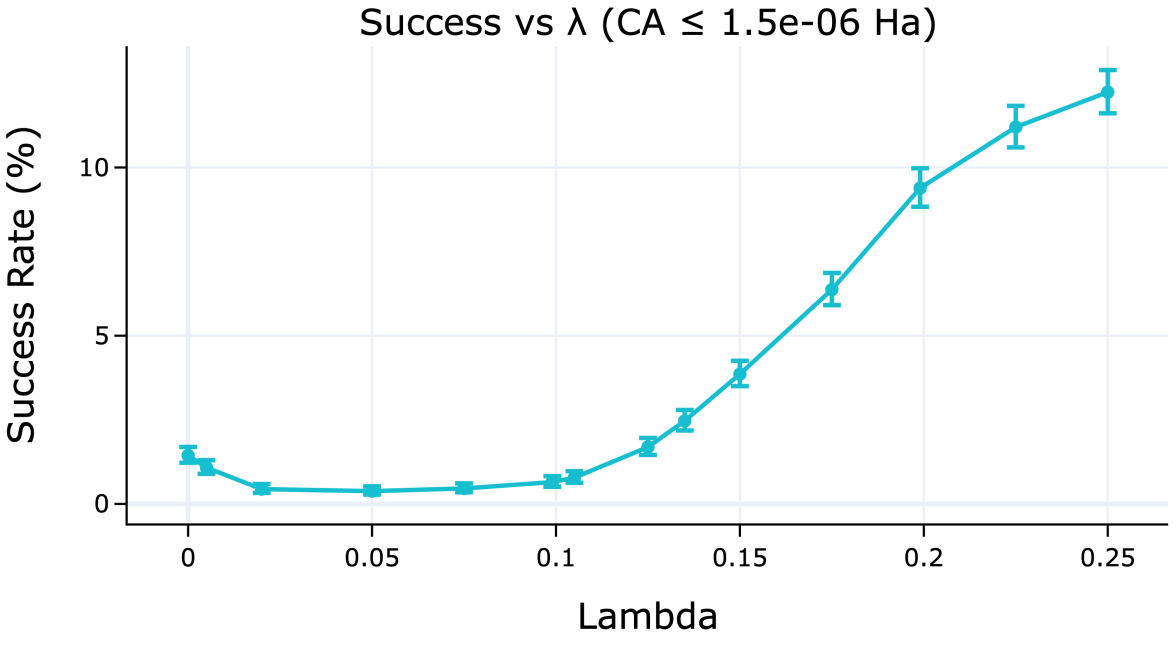

Figure 6.5: Hydrogen molecule, $\Delta E \leq 1.5 \times 10^{-6}$ Ha. The shallow dip near zero regularization becomes clearly visible.

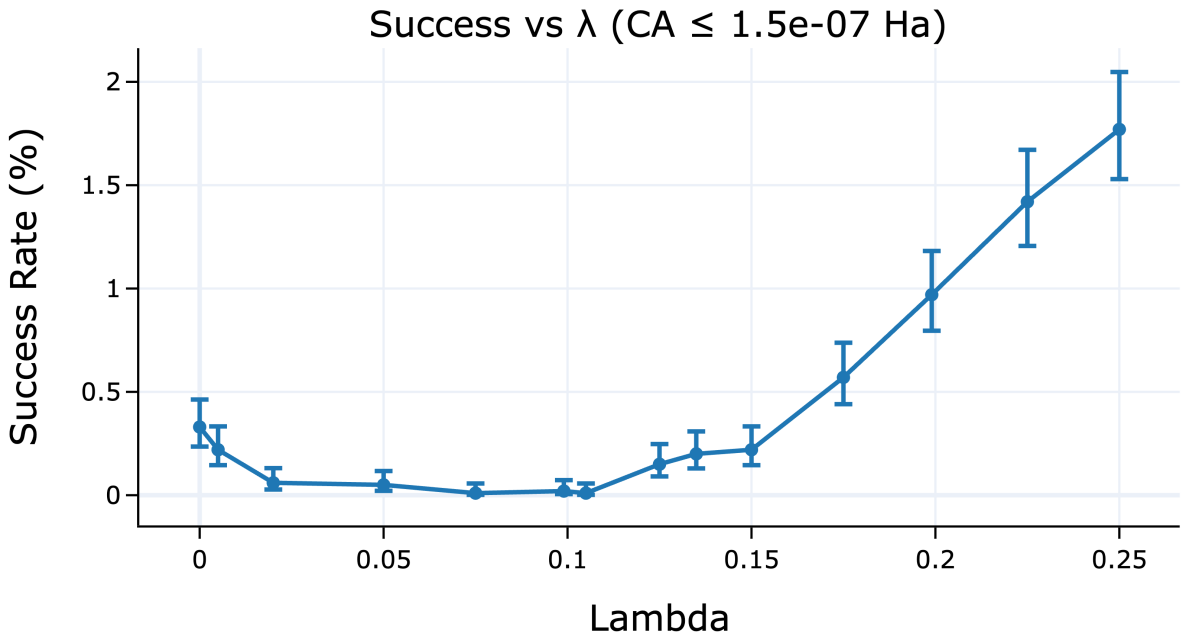

Figure 6.6: Hydrogen molecule, $\Delta E \leq 1.5 \times 10^{-7}$ Ha (sub-microhartree). At the most demanding threshold the curve is monotone-increasing across the explored range, confirming that strict accuracy benefits from stronger regularization.

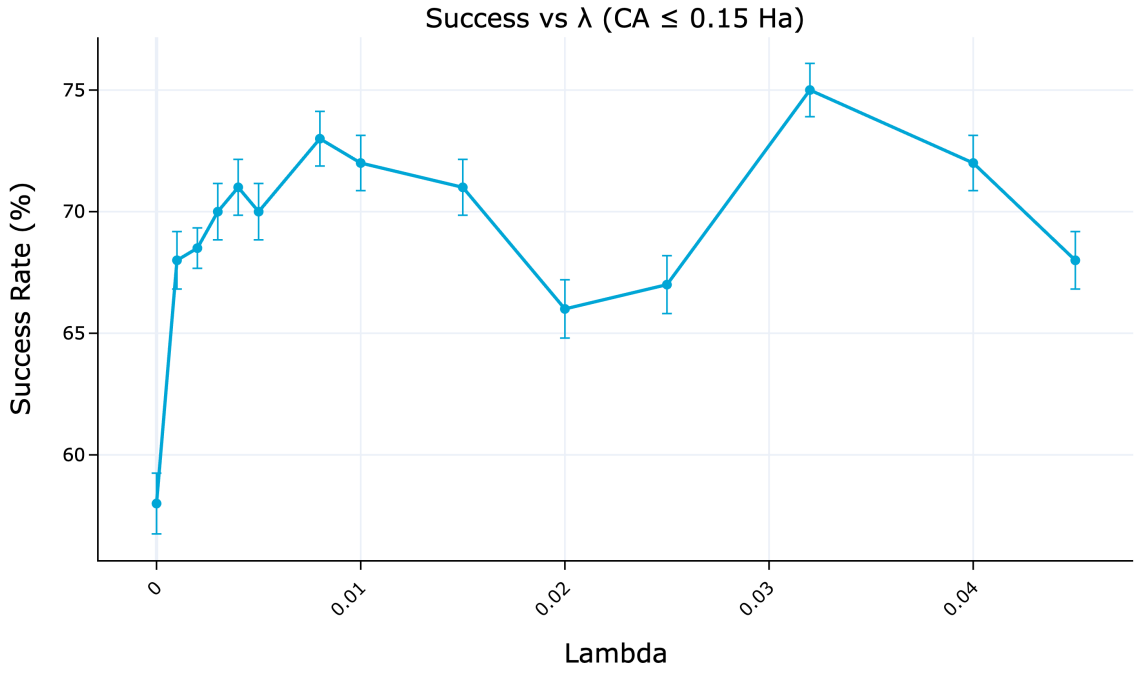

Figure 6.7: Lithium hydride success rates, $\Delta E \leq 0.15$ Ha. The stabilization window shifts to smaller λ_0 as the parameter dimension grows from 40 to 80.

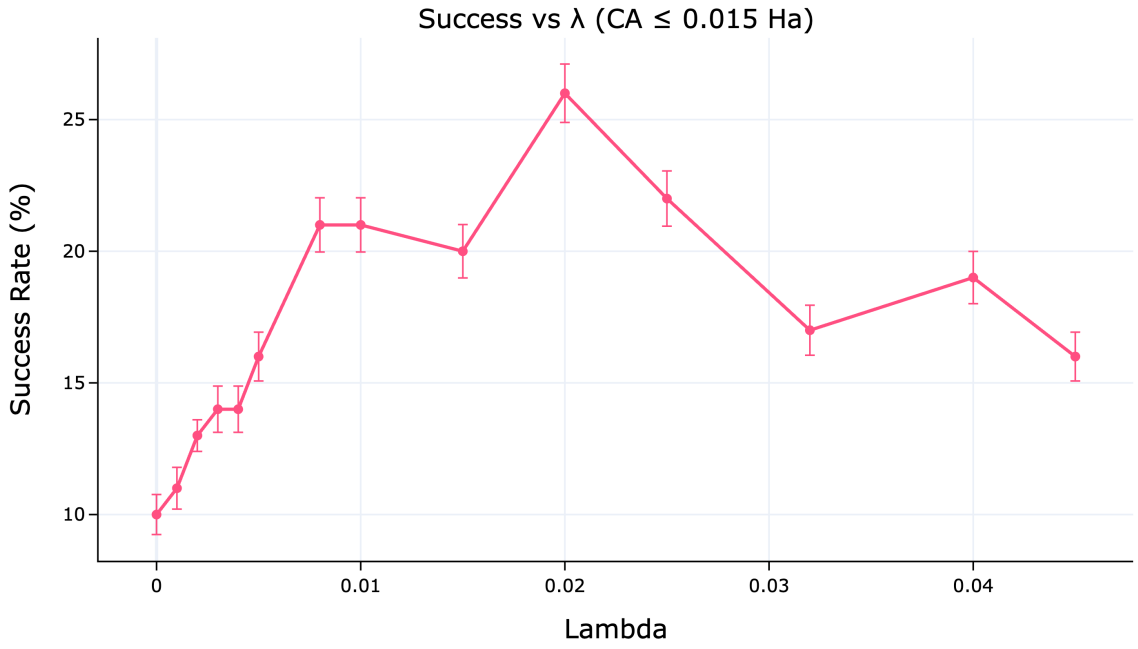

Figure 6.8: Lithium hydride, $\Delta E \leq 0.015$ Ha. The peak success rate near $\lambda_0 \approx 0.02$ confirms the dimension-dependent rescaling discussed in §6.6.

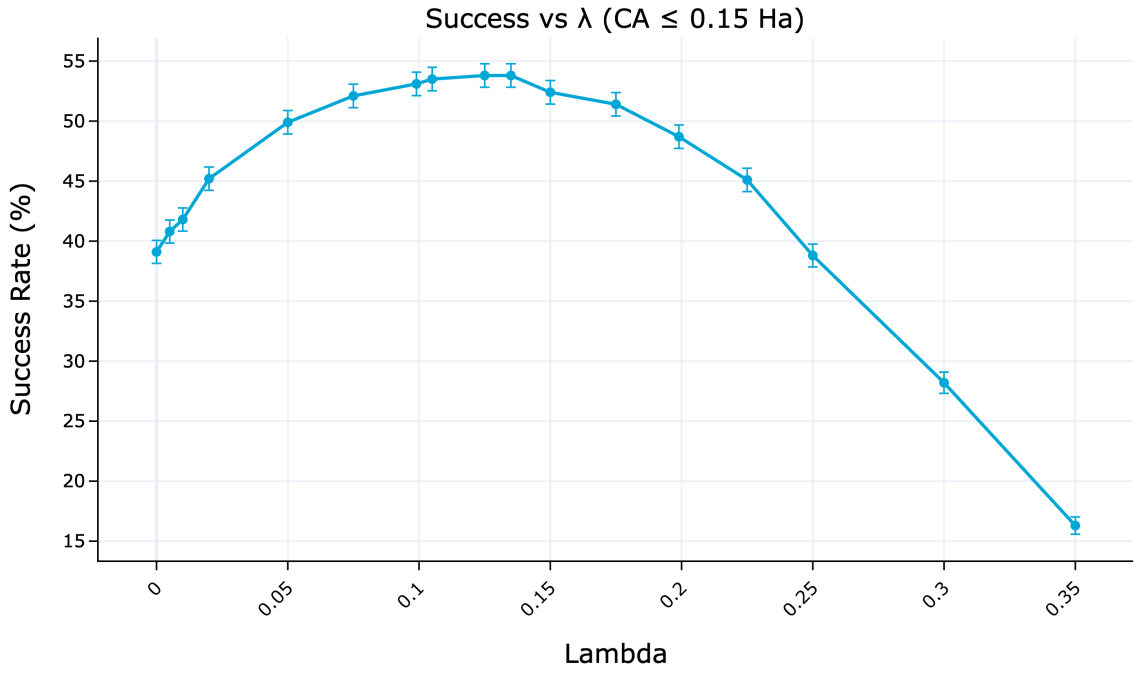

Figure 6.9: Random Field Ising Model success rates, $\Delta E \leq 0.15$ Ha. A broad stabilization plateau persists for the twelve-qubit, single-layer ansatz.

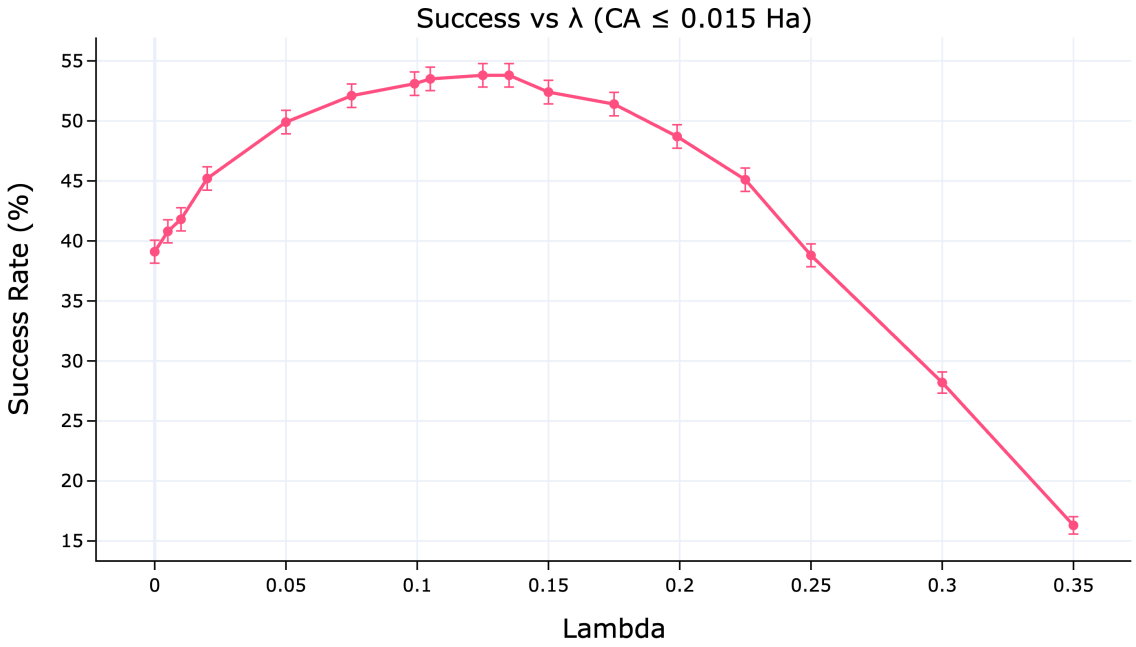

Figure 6.10: Random Field Ising Model, $\Delta E \leq 0.015$ Ha. The optimal regularization window lies between $\lambda_0 = 0.05$ and $\lambda_0 = 0.12$.

consistently observed.

6.7 Discussion and Summary

Quadratic L_2 regularization operates purely at the classical level by reshaping the curvature of the optimization landscape. Its stabilizing effect arises from improved geometric conditioning rather than from statistical considerations or circuit modification. The observed non-monotonic dependence on λ_0 indicates that an intermediate regime provides maximal robustness, balancing sufficient exploration with suppression of unstable directions.

Across molecular and spin systems, classical regularization systematically reduces variability and improves reproducibility while preserving unbiased final energy estimates. These findings demonstrate that landscape conditioning at the classical level can significantly enhance the reliability of variational quantum optimization.

Chapter 7

Conclusions, Implications, and Future Directions

7.1 Summary of Contributions

This thesis developed a system-level understanding of variational quantum algorithms by treating them explicitly as hybrid quantum–classical dynamical systems. Rather than attributing optimization behavior solely to properties of quantum circuits or hardware noise in isolation, the work emphasized the interaction between quantum objective generation and classical parameter-update dynamics as the dominant determinant of convergence, stability, and robustness.

The first contribution of the thesis is a clear operational decomposition of variational quantum algorithms into two functionally distinct components: a quantum substrate that generates stochastic objective evaluations, and a classical optimization layer that governs parameter evolution through discrete-time update dynamics. By making explicit this separation, the thesis reframes VQAs as feedback systems in which optimization behavior emerges from coupled dynamics rather than from either component in isolation.

Building on this perspective, the thesis analyzed classical optimization strategies as controller designs acting on noisy and weakly informative objective signals. Population-based optimization methods were examined as a natural response to limited observability and stochastic feedback, highlighting their ability to aggregate information, tolerate noise, and sustain exploration in the absence of reliable gradients. This analysis established population dynamics as a principled alternative to single-trajectory optimization in the NISQ regime [3].

The central technical contribution of the thesis is the formulation and evaluation (theoretical and empirical) of harmonic oscillator particle swarm optimization [22], [23] as a structured population-based optimizer for variational quantum eigensolvers. By embedding particle motion within a damped harmonic oscillator framework, HOPSO introduces explicit dynamical regulation that promotes bounded trajectories, controlled oscillatory exploration, and stability under stochastic forcing. Both theoretical analysis and empirical results suggest that these structured dynamics improve robustness, reduce sensitivity to initialization, and enhance repeatability across optimization runs.

A further contribution is the systematic treatment of classical regularization techniques as dynamical modifications rather than heuristic add-ons. By categorizing regularization methods according

to whether they act on parameter evolution or objective evaluation, the thesis clarified how regularization reshapes optimization dynamics without altering the underlying search objective in the intended small-regularization regime.

Collectively, these contributions advance a unified framework for understanding and designing variational quantum algorithms under realistic constraints. By emphasizing dynamical structure, stability, and feedback interactions, the thesis provides both conceptual clarity and practical guidance for developing robust hybrid quantum–classical optimization strategies on near-term quantum hardware.

7.2 Implications for Variational Quantum Algorithm Design

The results of this thesis have several important implications for the design and evaluation of variational quantum algorithms on near-term quantum hardware. Chief among these is the recognition that optimization behavior in VQAs is not determined solely by the expressivity of the variational ansatz or the characteristics of the quantum device, but emerges from the interaction between stochastic objective generation and classical update dynamics.

One immediate implication is that the choice of classical optimizer should be treated as a design decision of comparable importance to ansatz selection. In many studies, optimization is treated as a secondary concern, with emphasis placed on circuit architecture or noise mitigation techniques. The system-level perspective developed here suggests that even highly expressive ansätze may perform poorly if paired with update dynamics that are unstable or poorly matched to stochastic objective feedback.

The analysis further indicates that optimization strategies developed for deterministic or gradient-rich settings may be ill-suited to variational quantum algorithms [4], [24]. In the presence of finite-shot noise, flat regions, and weak local information, aggressive descent-based methods may stagnate or behave erratically. Population-based methods, particularly those with structured dynamics, offer an alternative that aligns more naturally with the constraints of the NISQ regime by emphasizing robustness, information aggregation, and bounded exploration.

Another implication concerns the role of benchmarking and performance evaluation. Because optimization outcomes in VQAs can exhibit substantial variability across runs, comparing algorithms based on single trajectories or best-case performance may be misleading under stochastic evaluation and limited measurement budgets. Stability, repeatability, and robustness to initialization and noise should be treated as primary performance metrics alongside convergence speed or final energy values, particularly in stochastic and measurement-limited regimes [25]. This shift in evaluation criteria has consequences for both algorithm development and experimental validation.

The results also suggest a modular approach to optimizer design. Structured update dynamics and classical regularization can be combined to address different sources of instability without entangling these mechanisms with quantum circuit design. Such modularity facilitates fair comparison across optimizers and supports portability across hardware platforms and problem classes.

Finally, the system-level framing underscores the importance of viewing variational quantum algorithms as feedback-controlled processes rather than static pipelines. Design choices at the classical layer influence not only efficiency but also the qualitative behavior of optimization trajectories under uncertainty. Incorporating principles from dynamical systems and control theory into VQA design

therefore represents a promising pathway toward more reliable and predictable hybrid quantum–classical algorithms.

7.3 Limitations of the Present Work

While this thesis develops a coherent framework for analyzing optimization dynamics in variational quantum algorithms, several limitations should be acknowledged. These limitations delineate the scope of the results and clarify the conditions under which the conclusions apply.

First, the analysis focuses primarily on fixed variational ansätze with finite-dimensional parameter spaces. Although this setting encompasses many practical VQE implementations, it does not address adaptive or dynamically growing ansatz strategies [26], [27], where the structure of the parameter space evolves during optimization. In such cases, the interaction between optimization dynamics and ansatz adaptation may introduce additional complexity not captured by the present framework.

Second, the empirical evaluations presented in this work are restricted to representative benchmark problems, such as small molecular Hamiltonians [25]. While these systems are standard in the VQE literature and sufficient to illustrate key dynamical behaviors, they do not exhaust the range of problem instances encountered in practice. Extending the analysis to larger systems, different Hamiltonian classes, or alternative variational algorithms would provide a more comprehensive assessment of generality.

Third, the practical cost profile of population-based methods deserves a more careful framing than is sometimes asserted. At equal *measurement* budgets, the population-based methods studied here amortize their candidate evaluations across fewer outer iterations, so the per-evaluation classical bookkeeping is broadly comparable to single-trajectory methods rather than uniformly larger. Where population-based methods do incur additional cost is in memory footprint (maintaining N parameter vectors plus personal-best records) and in synchronization overhead when personal- and global-best information must be aggregated across the swarm. These costs may become limiting in tightly memory-constrained settings or when scaling to very high-dimensional parameter spaces, but they do not, on their own, undermine the measurement-budget comparison on which the empirical results in this thesis are based.

Fourth, the work does not explicitly incorporate hardware-aware optimization strategies or adaptive noise mitigation techniques. While the framework is compatible with such extensions, the present analysis treats noise as an exogenous stochastic influence rather than as a controllable or learnable feature of the system. Integrating hardware-specific feedback or adaptive measurement strategies remains an open area for future research.

Finally, the theoretical analysis emphasizes qualitative dynamical behavior rather than formal convergence guarantees. While the harmonic oscillator formulation provides intuitive and empirical evidence of improved stability, rigorous convergence proofs under stochastic objective evaluations are beyond the scope of this thesis [28]. Developing such guarantees would require additional assumptions about objective structure and noise statistics.

These limitations do not diminish the contributions of the work but rather highlight opportunities for refinement and extension. The following section outlines several directions for future research that build naturally on the system-level perspective developed here.

7.4 Future Research Directions

The system-level perspective developed in this thesis opens several non-exhaustive avenues for future research in the design and analysis of variational quantum algorithms. These directions extend the core ideas of structured dynamics, robustness under uncertainty, and hybrid feedback control beyond the specific settings examined here.

One promising direction is the integration of adaptive or evolving ansatz strategies with structured optimization dynamics. In such approaches, the parameter space itself changes over time as circuit elements are added, removed, or reparameterized. Understanding how classical controllers can remain stable and effective in the presence of a non-stationary search space represents an important challenge, particularly under stochastic objective feedback.

Another avenue concerns the application of control-theoretic tools to hybrid quantum–classical optimization loops. Viewing variational algorithms as feedback systems [29] suggests the possibility of formal stability analysis, controller synthesis, and robustness guarantees. Techniques from nonlinear control, stochastic stability, and system identification may provide deeper insight under suitable modeling assumptions into the conditions under which reliable convergence can be expected.

Extensions to other classes of variational algorithms also merit investigation. While this thesis focused primarily on variational quantum eigensolvers, the framework applies equally to algorithms such as the quantum approximate optimization algorithm, variational quantum classifiers, and hybrid quantum machine learning models [30], [31]. Each of these settings presents distinct objective landscapes and noise characteristics that may benefit from structured optimization dynamics.

The co-design of measurement strategies and classical optimization dynamics represents another important research direction. In current implementations, measurement budgets and optimization updates are often treated independently. Jointly optimizing how measurement resources are allocated and how optimization dynamics respond to noisy feedback could lead to more efficient use of limited quantum resources.

Hardware-aware and noise-adaptive optimization strategies also offer significant potential. Rather than treating noise as a static nuisance, future algorithms could adapt their dynamics in response to observed noise characteristics, hardware drift, or calibration changes. Incorporating such feedback would further strengthen the robustness of hybrid algorithms in realistic operating conditions.

Finally, broader theoretical work is needed to formalize convergence and stability properties of variational optimization under stochastic objective evaluations. Establishing rigorous bounds or guarantees, even under simplified assumptions, would strengthen the theoretical foundations of hybrid quantum–classical algorithms and complement the empirical insights developed in this thesis.

Together, these directions point toward a future in which variational quantum algorithms are designed as integrated dynamical systems, with classical control playing a central role in managing uncertainty and extracting reliable performance from near-term quantum hardware.

7.5 Broader Perspective

The analysis presented in this thesis suggests a broader way of thinking about variational quantum algorithms and, more generally, about hybrid quantum–classical computation. Rather than viewing these algorithms as pipelines that alternate between quantum state preparation and classical post-processing, it is more accurate to regard them as coupled dynamical systems operating under

uncertainty. Optimization behavior, success, and failure emerge from feedback interactions between components with fundamentally different informational and physical constraints.

This perspective challenges the tendency to localize difficulties in variational algorithms solely within the quantum layer. Phenomena such as barren plateaus, noise, or limited expressivity are often cited as primary obstacles, yet the results of this work support the view that classical update dynamics play an equally critical role. In many cases, instability, stagnation, or variability arise not because useful information is absent, but because available information is processed by update rules ill-suited to stochastic and weakly informative signals.

Framing variational algorithms as feedback-controlled systems highlights the relevance of classical control and dynamical systems theory to near-term quantum computing. Concepts such as stability, damping, boundedness, and response to stochastic forcing provide a language for reasoning about optimization behavior that complements traditional complexity-theoretic or circuit-centric analyses. This shift in viewpoint emphasizes reliability and predictability over idealized asymptotic performance.

More broadly, the work underscores that optimization in variational quantum algorithms differs fundamentally from optimization in classical machine learning. In classical settings, gradients are often abundant, objective evaluations are relatively inexpensive, and noise can be reduced through averaging. In contrast, variational quantum optimization operates with sparse, noisy feedback and strict resource constraints. These differences necessitate optimization strategies that prioritize robustness and controlled exploration over rapid descent.

The system-level framing also suggests that progress in near-term quantum algorithms will depend as much on advances in classical algorithm design as on improvements in quantum hardware. Carefully structured classical controllers can extract meaningful performance from imperfect quantum devices, extending the practical reach of NISQ-era algorithms. In this sense, hybrid quantum–classical computation should be viewed not as a temporary compromise, but as a domain where classical and quantum methods co-evolve.

By emphasizing dynamical structure and feedback interactions, this thesis contributes to a growing body of work that seeks to ground variational quantum algorithms in principled system design. This perspective provides a foundation for developing hybrid algorithms that are not only expressive in principle, but reliable in practice under realistic conditions.

7.6 Concluding Remarks

Where Section 7.5 situated this work within the broader research landscape, the present remarks return to the specific claims that the thesis itself substantiates. Three findings stand out.

First, optimizer choice is a primary, not secondary, design dimension for variational quantum algorithms: identical circuits and Hamiltonians yielded materially different success rates and dispersions purely as a function of the classical update rule. Second, structured dynamical models—HOPSO and two-stage classical regularization—delivered measurably tighter convergence distributions than their unstructured counterparts under shot noise, at no cost to the unregularized final energy. Third, these advantages survived a $40 \rightarrow 80 \rightarrow 12\,000$-step scaling across H_2, LiH, and the Random Field Ising Model, indicating that the underlying mechanism is not specific to a single problem instance.

Together these results support a single recommendation: classical update dynamics in NISQ-era variational algorithms should be designed and reported with the same care given to ansatz choice

and measurement strategy. The methods studied here do not exhaust the design space, but they demonstrate that explicitly structured controllers can extract reliable performance from imperfect hardware where unstructured local optimizers cannot.

Appendix A

Hamiltonian Encodings

This appendix provides technical background on the fermion-to-qubit mappings used to express molecular Hamiltonians in a form suitable for evaluation on quantum hardware. These encodings determine the structure of the Pauli operator decomposition referenced in Chapter 2 and directly influence the number of measurement terms, qubit requirements, and the form of the cost landscape encountered by the classical optimizer.

A.1 Jordan–Wigner and Bravyi–Kitaev Mappings

In second-quantized quantum chemistry, the electronic Hamiltonian is expressed in terms of fermionic creation and annihilation operators $a_j^\dagger$ and a_j, which obey canonical anticommutation relations. Since qubits are not fermions, a mapping is required to represent fermionic operators as qubit operators while preserving the correct anticommutation structure [12], [13].

A.1.1 Jordan–Wigner Transformation

The Jordan–Wigner (JW) transformation maps each fermionic mode to a single qubit and encodes the occupation number directly in the computational basis [12]. The creation operator for mode j is represented as

$$a_j^\dagger \mapsto \frac{1}{2}(X_j - iY_j) \otimes Z_{j-1} \otimes Z_{j-2} \otimes \cdots \otimes Z_0, \tag{A.1}$$

where the string of Z operators enforces the correct fermionic sign structure. The key properties of the JW mapping are:

- **Locality of occupation**: The occupation of mode j is stored in qubit j, making number operators local single-qubit measurements.
- **Non-locality of hopping**: Hopping terms between modes j and k produce Pauli strings of length $|k - j| + 1$, which can become long for distant modes.
- **Qubit count**: Requires exactly M qubits for M spin-orbitals.

The non-local Pauli strings arising from the JW mapping increase the number of distinct measurement terms in the Hamiltonian decomposition, which in turn affects the total measurement cost per iteration of the variational algorithm.

A.1.2 Bravyi–Kitaev Transformation

The Bravyi–Kitaev (BK) transformation provides an alternative encoding that balances locality of occupation and parity information [13], [14]. Rather than storing occupation numbers directly, the BK mapping encodes partial sums of occupation numbers in a tree-like structure. The resulting Pauli strings have length $O(\log M)$ rather than $O(M)$, reducing the weight of individual terms in the Hamiltonian decomposition.

The key properties of the BK mapping are:

- **Logarithmic Pauli weight**: Both occupation and parity queries produce Pauli strings of length at most $\lceil \log_2 M \rceil$.
- **Qubit count**: Also requires M qubits for M spin-orbitals.
- **Reduced measurement complexity**: Shorter Pauli strings may allow more efficient grouping of commuting terms for simultaneous measurement.

In practice, the choice between JW and BK mappings involves tradeoffs between implementation simplicity and measurement efficiency. The JW mapping is more intuitive and widely used, while the BK mapping can reduce the total number of required measurement circuits for large systems.

A.2 Pauli Operator Decomposition

Regardless of the fermion-to-qubit mapping chosen, the resulting qubit Hamiltonian takes the form of a weighted sum of Pauli strings,

$$H = \sum_{j=1}^{L} h_j P_j, \tag{A.2}$$

where each P_j is a tensor product of single-qubit Pauli operators $\{I, X, Y, Z\}$ and the coefficients $h_j \in \mathbb{R}$ are determined by the molecular integrals and the chosen encoding. The number of terms L depends on the molecular system, the basis set, and the encoding.

For a molecular system with M spin-orbitals, the number of Pauli terms scales as $O(M^4)$ in the worst case, arising from the two-electron integrals in the electronic Hamiltonian. In practice, symmetry considerations and sparsity in the molecular integrals often reduce this count significantly.

Each Pauli string P_j can be measured independently by preparing the appropriate measurement basis on each qubit. The expectation value of the full Hamiltonian is then reconstructed as

$$\langle H \rangle = \sum_{j=1}^{L} h_j \langle P_j \rangle, \tag{A.3}$$

where each $\langle P_j \rangle$ is estimated from finite measurement statistics. The variance of the total energy estimator depends on the individual variances of the Pauli term estimators and their coefficients, as discussed in Appendix B.

A.3 Implications for Hardware Implementation

The choice of Hamiltonian encoding has several practical consequences for variational quantum algorithm implementations:

Qubit connectivity requirements. Long Pauli strings require multi-qubit operations that may not be natively supported by the hardware topology. On devices with limited connectivity, additional SWAP gates are needed to implement non-local operations, increasing circuit depth and noise.

Circuit depth. The encoding affects the depth of circuits required to implement variational ansätze that respect the symmetries of the encoded Hamiltonian. Hardware-efficient ansätze bypass this concern by ignoring fermionic structure, but chemically inspired ansätze (such as UCCSD) inherit the locality properties of the chosen encoding.

Measurement overhead. The number of distinct Pauli strings determines the minimum number of measurement circuits required per energy evaluation. Encodings that produce fewer or shorter Pauli strings can reduce this overhead, directly affecting the measurement budget consumed per iteration of the optimization loop.

A.4 Encoding Choice and Optimization Landscape

The Hamiltonian encoding does not change the eigenvalues of the encoded operator—the ground-state energy is invariant under the choice of mapping. However, the encoding does affect the structure of the cost landscape as a function of circuit parameters, because different encodings produce different Pauli decompositions that interact differently with the variational ansatz.

In particular, the encoding influences:

- **The number and structure of local minima** in the variational landscape, since different Pauli decompositions may create different interference patterns in the cost function.
- **The variance of the energy estimator**, since encodings with more or higher-weight Pauli terms generally produce higher-variance estimates for a fixed shot budget.
- **The sensitivity of the cost function to individual parameters**, which affects gradient magnitudes and the susceptibility to barren plateaus.

From the perspective of the classical optimizer, these effects manifest as differences in the noise level, smoothness, and informational content of the objective signal. While the optimizer does not directly observe the encoding, its performance is indirectly shaped by the encoding through the properties of the cost landscape it navigates.

A.5 Practical Notes on Hardware Connectivity

The connectivity of quantum hardware plays a crucial role in the implementation of Hamiltonian encodings. The choice of encoding affects the length and structure of Pauli strings, which in turn determines the qubit connectivity requirements. Devices with limited connectivity may require additional SWAP gates to implement non-local operations, increasing circuit depth and noise.

In practice, the choice of encoding should be guided by the hardware topology and the specific requirements of the variational algorithm. By carefully considering the tradeoffs between implementation simplicity, measurement efficiency, and hardware connectivity, researchers can optimize the performance of their quantum algorithms and push the boundaries of what is possible with current hardware.

A.6 Summary

Fermion-to-qubit mappings such as the Jordan–Wigner and Bravyi–Kitaev transformations convert molecular Hamiltonians into weighted sums of Pauli operators suitable for evaluation on quantum hardware. The choice of encoding affects the number and weight of Pauli terms, the measurement overhead per energy evaluation, and the structure of the variational cost landscape. While the ground-state energy is invariant under the encoding, the practical difficulty of the optimization problem—including noise levels, measurement costs, and landscape structure—depends on the specific mapping used. These considerations are relevant to the system-level analysis of variational quantum algorithms developed in the main text, where the properties of the quantum-generated objective signal directly influence the behavior of the classical optimization layer.

Appendix B

Measurement Grouping and Shot Allocation

This appendix details the measurement process by which expectation values of qubit Hamiltonians are estimated on quantum hardware. The statistical properties of this estimation process determine the noise characteristics of the objective signal supplied to the classical optimizer, as discussed in Chapters 2 and 3.

B.1 Pauli Decomposition and Term-wise Measurement

As described in Appendix A, the qubit Hamiltonian is expressed as a weighted sum of Pauli strings,

$$H = \sum_{j=1}^{L} h_j P_j. \tag{B.1}$$

The expectation value $\langle H \rangle = \sum_j h_j \langle P_j \rangle$ is estimated by measuring each Pauli term separately. Each P_j is a tensor product of single-qubit Pauli operators, and its expectation value is obtained by preparing the variational state, rotating into the appropriate measurement basis, and performing projective measurements in the computational basis.

Because different Pauli strings may require different measurement bases, they cannot in general be measured simultaneously from a single circuit execution. The total number of distinct measurement circuits required per energy evaluation therefore depends on how many groups of simultaneously measurable terms can be formed.

B.2 Commuting Groups and Measurement Bases

Two Pauli strings P_j and P_k can be measured simultaneously if and only if they commute: $[P_j, P_k] = 0$. Commuting Pauli operators share a common eigenbasis, so a single measurement basis suffices for all terms in a commuting group.

Grouping Pauli terms into commuting sets reduces the number of distinct measurement circuits required per energy evaluation. Several grouping strategies have been studied [6]:

- **Qubit-wise commutativity (QWC)**: Two Pauli strings commute qubit-wise if, on every qubit, the corresponding single-qubit Pauli operators either match or at least one is the identity. QWC grouping is simple to implement and produces groups that share a tensor-product measurement basis.
- **General commutativity (GC)**: Two Pauli strings may commute even if they do not commute qubit-wise. GC grouping produces fewer groups but requires entangling rotations to diagonalize the group into a common basis, increasing circuit depth.
- **Graph-coloring approaches**: The grouping problem can be formulated as a graph coloring problem, where vertices represent Pauli terms and edges connect non-commuting pairs. The chromatic number of this graph gives the minimum number of measurement circuits.

In practice, QWC grouping is most commonly used due to its simplicity and compatibility with near-term hardware constraints. For a typical molecular Hamiltonian with L Pauli terms, QWC grouping reduces the number of measurement circuits from L to a smaller number $G \leq L$, with the reduction depending on the structure of the Hamiltonian.

B.3 Shot Noise and Estimation Variance

For a single Pauli string P_j, the expectation value $\langle P_j \rangle$ is estimated from S_j repeated measurements (shots). Each measurement yields an outcome ± 1, and the sample mean provides an unbiased estimator with variance

$$\mathrm{Var}[\hat{P}_j] = \frac{1 - \langle P_j \rangle^2}{S_j}. \tag{B.2}$$

The variance of the total energy estimator is then

$$\mathrm{Var}[\hat{E}] = \sum_{j=1}^{L} h_j^2 \, \mathrm{Var}[\hat{P}_j] = \sum_{j=1}^{L} \frac{h_j^2 (1 - \langle P_j \rangle^2)}{S_j}, \tag{B.3}$$

assuming independent measurements across groups. This expression reveals that the noise in the energy estimate depends on three factors: the Hamiltonian coefficients h_j, the true expectation values $\langle P_j \rangle$, and the number of shots allocated to each term.

For a uniform shot allocation $S_j = S/L$ (where S is the total shot budget), the variance scales as

$$\mathrm{Var}[\hat{E}] \propto \frac{L}{S} \sum_j h_j^2. \tag{B.4}$$

This shows that Hamiltonians with many terms or large coefficients produce noisier energy estimates for a fixed total shot budget, directly affecting the quality of the objective signal received by the classical optimizer.

B.4 Shot Allocation Strategies

The total shot budget S can be distributed across Pauli terms or commuting groups in several ways:

Uniform allocation. Each term or group receives an equal share of the total budget. This is simple to implement but suboptimal when Hamiltonian coefficients vary widely in magnitude.

Coefficient-weighted allocation. Shots are allocated proportionally to $|h_j|$, concentrating measurement effort on terms that contribute most to the energy estimate. This reduces the overall variance of the estimator compared to uniform allocation.

Variance-minimizing allocation. The optimal allocation that minimizes $\mathrm{Var}[\hat{E}]$ for a fixed total budget assigns shots proportionally to $|h_j|\sqrt{1-\langle P_j\rangle^2}$. Since the true expectation values are unknown during optimization, this allocation must be estimated adaptively or approximated using prior information.

In the context of variational quantum algorithms, the shot allocation strategy interacts with the optimization dynamics. Coefficient-weighted allocation reduces the variance of the energy estimate, providing a cleaner signal to the classical optimizer. However, the improvement is most significant when the Hamiltonian has a small number of dominant terms, which is common in molecular systems.

B.5 Practical Implications for Optimization Dynamics

The measurement process introduces a fundamental source of stochasticity into the variational optimization loop. From the perspective of the classical optimizer, the key consequences are:

Noise floor. For any fixed shot budget, there is a minimum variance below which the energy estimate cannot be refined without additional measurements. This noise floor sets a limit on the precision of cost function evaluations and determines the regime in which the optimizer must operate.

Signal-to-noise ratio. The ability of the optimizer to distinguish between parameter vectors with similar energies depends on the ratio of the true energy difference to the standard deviation of the estimator. When this ratio is small, the optimizer effectively receives uninformative feedback, which can lead to stagnation or erratic updates.

Budget-iteration tradeoff. Increasing the number of shots per evaluation improves the quality of each cost estimate but reduces the total number of evaluations (and hence optimizer iterations) available within a fixed budget. Population-based methods such as HOPSO face an additional dimension of this tradeoff, as discussed in Chapter 5, since each iteration requires evaluations for all population members.

Interaction with population methods. Population-based optimizers partially mitigate measurement noise through information aggregation across multiple candidates. Because each particle receives an independent noise realization, population-level statistics (such as the identification of the best-performing candidate) are less sensitive to individual noisy measurements than single-trajectory methods. This property allows population methods to operate effectively with lower per-evaluation shot counts, potentially reallocating measurement budget toward more iterations or larger populations.

These considerations underscore the importance of treating measurement noise as an integral feature of the optimization problem rather than as a secondary implementation detail. The system-level analysis developed in the main text explicitly accounts for these effects when analyzing optimizer stability, convergence behavior, and the benefits of structured dynamics under stochastic feedback.

Appendix C

Noise Models and Simulation Modes

This appendix describes the noise sources and simulation modes relevant to the numerical experiments discussed in Chapters 5 and 6. Understanding how noise enters the variational optimization loop is essential for interpreting empirical results and for appreciating why structured dynamics and regularization are needed in the NISQ regime.

C.1 Statevector vs Shot-Based Simulation

Numerical studies of variational quantum algorithms can be conducted at different levels of physical fidelity, each introducing distinct sources of variability into the cost function.

Statevector simulation. In statevector (or exact) simulation, the quantum state $|\psi(\theta)\rangle$ is represented as a full complex vector in $\mathbb{C}^{2^n}$, and expectation values are computed analytically as $C(\theta) = \langle\psi(\theta)|H|\psi(\theta)\rangle$. This mode introduces no measurement noise: the cost function is a deterministic, smooth function of the parameters. Statevector simulation is useful for isolating the effects of landscape structure and optimizer dynamics from measurement-induced stochasticity, but it does not reflect the conditions encountered on real quantum hardware.

Shot-based simulation. In shot-based simulation, expectation values are estimated from a finite number of simulated measurement outcomes, mimicking the statistical process on quantum hardware. For each Pauli term P_j, the simulator samples S_j binary outcomes from the appropriate probability distribution and computes a sample mean. The resulting energy estimate exhibits shot noise with variance determined by the shot count and the Hamiltonian structure, as detailed in Appendix B.

Shot-based simulation introduces the same type of stochastic variability that affects real quantum devices, making it the appropriate mode for evaluating optimizer robustness under realistic conditions. The experiments in this thesis primarily use shot-based simulation to ensure that conclusions about optimizer behavior apply to the noisy regime.

Hybrid modes. Some simulation frameworks allow selective introduction of noise sources. For example, shot noise can be included without hardware noise models, or specific noise channels can be added to an otherwise ideal simulation. These hybrid modes enable controlled studies of how individual noise sources affect optimization dynamics.

C.2 Noise Channels

In addition to finite-shot sampling noise, quantum hardware introduces errors through imperfect gate operations, qubit decoherence, and measurement imperfections. These errors are commonly modeled using quantum noise channels, which describe the transformation of an ideal quantum state into a noisy mixed state.

Depolarizing noise. A depolarizing channel replaces the quantum state with the maximally mixed state with probability p, modeling uniform random errors. For a single qubit,

$$\mathcal{E}_{\text{dep}}(\rho) = (1-p)\rho + \frac{p}{3}(X\rho X + Y\rho Y + Z\rho Z). \tag{C.1}$$

Depolarizing noise is commonly applied after each gate to model generic gate errors. It reduces the purity of the quantum state and biases expectation values toward zero, effectively compressing the range of the cost function.

Amplitude damping. Amplitude damping models energy relaxation (T_1 decay), in which an excited qubit spontaneously transitions to the ground state. This channel is characterized by the decay probability γ, related to the relaxation time T_1 and the gate duration. Amplitude damping introduces an asymmetric bias, preferentially driving qubits toward the $|0\rangle$ state.

Dephasing (phase damping). Dephasing models the loss of quantum coherence (T_2 decay) without energy exchange. Off-diagonal elements of the density matrix decay at a rate determined by the dephasing time T_2. Dephasing reduces the visibility of interference effects and degrades the sensitivity of the cost function to parameter changes.

Readout noise. Measurement errors occur when the classical outcome recorded by the detector does not match the true quantum state. Readout noise is typically modeled as a classical bit-flip channel applied to measurement outcomes, with error probabilities that may differ between $|0\rangle \to |1\rangle$ and $|1\rangle \to |0\rangle$ transitions. Readout errors directly corrupt the measurement statistics used to estimate expectation values.

In practice, multiple noise channels act simultaneously, and their combined effect depends on the circuit depth, gate count, and hardware-specific error rates. Noise models used in simulation are calibrated to match the characteristics of target hardware platforms.

C.3 Coherent vs Stochastic Noise

Noise in quantum devices can be broadly classified into two categories with distinct implications for optimization:

Stochastic noise arises from random processes such as thermal fluctuations, photon loss, and measurement randomness. Its effects vary between circuit executions and average out over many repetitions, contributing to the variance of the energy estimator. Stochastic noise is the primary concern for optimizer robustness, as it determines the signal-to-noise ratio of the objective feedback.

Coherent noise arises from systematic errors such as miscalibrated gate rotations, crosstalk between qubits, or persistent control errors. Unlike stochastic noise, coherent errors do not average out with repeated measurements. Instead, they introduce a systematic bias in the cost function, shifting the effective landscape experienced by the optimizer.

The distinction matters for optimization dynamics because:

- **Stochastic noise** increases the variance of cost evaluations, requiring the optimizer to tolerate fluctuations and extract trends from noisy signals. Population-based methods and temporal smoothing are effective countermeasures.
- **Coherent noise** shifts the location of minima in the effective landscape, potentially causing the optimizer to converge to a biased solution. Error mitigation techniques (such as zero-noise extrapolation or probabilistic error cancellation) are needed to address coherent errors, as they cannot be suppressed by optimizer design alone.

In the experiments presented in this thesis, the primary focus is on stochastic noise from finite-shot sampling, which represents the dominant and irreducible source of variability in variational quantum algorithms. Hardware-specific coherent errors are not modeled explicitly but are acknowledged as an additional source of bias in real-device implementations.

C.4 How Noise Enters the Energy Estimate

The energy estimate received by the classical optimizer is affected by noise at multiple stages of the evaluation process:

1. **State preparation errors.** Gate noise during circuit execution causes the prepared state to deviate from the intended $|\psi(\theta)\rangle$. The actual state is a mixed state $\rho(\theta)$ whose expectation values differ from the ideal values.
2. **Measurement basis rotation errors.** Noise in the single-qubit rotations used to change measurement bases introduces additional errors in the estimation of individual Pauli terms.
3. **Finite-shot sampling.** Even for a perfectly prepared state, the expectation value must be estimated from a finite number of measurement outcomes, introducing statistical fluctuations as described in Appendix B.
4. **Readout errors.** Classical bit-flip errors in the measurement apparatus corrupt the recorded outcomes, biasing the sample statistics.

The total noise in the energy estimate is a combination of all these sources. From the optimizer's perspective, the aggregate effect is a stochastic perturbation of the true cost function,

$$\tilde{C}(\theta) = C(\theta) + \eta(\theta), \tag{C.2}$$

where $\eta(\theta)$ represents the combined noise contribution. The variance and bias of η depend on the circuit depth, shot count, noise model, and Hamiltonian structure. Importantly, the optimizer does not have direct access to $C(\theta)$ or $\eta(\theta)$ separately; it observes only the noisy estimate $\tilde{C}(\theta)$.

C.5 Implications for Stability and Robustness

The noise characteristics described above have direct consequences for the stability and robustness of classical optimization in variational quantum algorithms:

Noise-dependent landscape. The effective cost landscape experienced by the optimizer is not the ideal landscape $C(\theta)$ but a noisy version whose features may be obscured, shifted, or distorted. Flat regions may appear flatter due to noise, local minima may be masked, and spurious features may emerge from stochastic fluctuations.

Noise-induced instability. Optimizers that rely on precise local information (such as exact gradients or small cost differences) may become unstable when noise magnitudes are comparable to the signal. This instability can manifest as oscillatory parameter trajectories, divergence, or convergence to noise-dominated artifacts rather than genuine minima [5].

Robustness through structured dynamics. The harmonic oscillator dynamics of HOPSO provide a natural mechanism for filtering stochastic noise, as discussed in Chapter 5. Damping suppresses noise-induced excursions, while the restoring force maintains bounded trajectories. These properties are particularly valuable when noise levels are high or when the objective landscape provides weak directional guidance.

Robustness through regularization. Classical regularization techniques, as discussed in Chapter 6, provide additional stabilization by constraining parameter updates or smoothing the objective signal. These mechanisms complement the intrinsic noise tolerance of structured dynamics and are most beneficial in high-noise or low-shot regimes.

Simulation mode selection. The choice of simulation mode determines which noise sources are present and therefore which aspects of optimizer behavior are being tested. Statevector simulation isolates landscape and optimizer effects; shot-based simulation adds measurement noise; full noise-model simulation includes hardware errors. Conclusions about optimizer robustness should specify the simulation mode used, as performance may differ substantially across modes.

The noise models and simulation modes described in this appendix provide the technical foundation for interpreting the empirical results presented in the main text. By explicitly accounting for the sources and characteristics of noise in variational quantum algorithms, the thesis ensures that conclusions about optimizer stability and robustness are grounded in realistic assumptions about the operating conditions of near-term quantum hardware.

References

[1] A. Peruzzo *et al.*, "A variational eigenvalue solver on a photonic quantum processor," *Nature Communications*, vol. 5, p. 4213, 2014.

[2] M. Cerezo *et al.*, "Variational quantum algorithms," *Nature Reviews Physics*, vol. 3, pp. 625–644, 2021.

[3] J. Preskill, "Quantum computing in the NISQ era and beyond," *Quantum*, vol. 2, p. 79, 2018.

[4] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, "Barren plateaus in quantum neural network training landscapes," *Nature Communications*, vol. 9, no. 1, p. 4812, 2018.

[5] S. Wang *et al.*, "Noise-induced barren plateaus in variational quantum algorithms," *Nature Communications*, vol. 12, no. 1, p. 6961, 2021.

[6] P. Gokhale *et al.*, "Minimizing state preparations in variational quantum eigensolver by partitioning into commuting families," *arXiv preprint arXiv:1907.13623*, 2019.

[7] A. Arrasmith, M. Cerezo, P. Czarnik, L. Cincio, and P. J. Coles, "Effect of barren plateaus on gradient-free optimization," *Quantum*, vol. 5, p. 558, 2021.

[8] D. Wierichs, J. Izaac, C. Wang, and C. Y.-Y. Lin, "Avoiding local minima in variational quantum algorithms with random restarts," *Physical Review Research*, vol. 2, no. 4, p. 043246, 2020.

[9] S. Sim, P. D. Johnson, and A. Aspuru-Guzik, "Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms," *Advanced Quantum Technologies*, vol. 2, no. 12, p. 1900070, 2019.

[10] L. Bittel and M. Kliesch, "Training variational quantum algorithms is hard," *Physical Review Letters*, vol. 127, no. 12, p. 120502, 2021.

[11] M. Schuld, V. Bergholm, C. Gogolin, J. Izaac, and N. Killoran, "Evaluating analytic gradients on quantum hardware," *Physical Review A*, vol. 99, no. 3, p. 032331, 2019.

[12] P. Jordan and E. Wigner, "Über das paulische Äquivalenzverbot," *Zeitschrift für Physik*, vol. 47, pp. 631–651, 1928.

[13] S. B. Bravyi and A. Yu. Kitaev, "Fermionic quantum computation," *Annals of Physics*, vol. 298, no. 1, pp. 210–226, 2002.

[14] J. T. Seeley, M. J. Richard, and P. J. Love, "The bravyi-kitaev transformation for quantum computation of electronic structure," *The Journal of Chemical Physics*, vol. 137, no. 22, p. 224109, 2012.

[15] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, J. M. Chow, and J. M. Gambetta, "Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets," *Nature*, vol. 549, pp. 242–246, 2017.

[16] M. Benedetti, E. Lloyd, S. Sack, and M. Fiorentini, "Parameterized quantum circuits as machine learning models," *Quantum Science and Technology*, vol. 4, no. 4, p. 043001, 2019.

[17] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95 – international conference on neural networks*, IEEE, 1995, pp. 1942–1948.

[18] H.-G. Beyer, *The theory of evolution strategies*. Berlin: Springer, 2001.

[19] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*, 2nd ed. Cambridge, MA: MIT Press, 2018.

[20] A. E. Eiben and J. E. Smith, *Introduction to evolutionary computing*. Berlin: Springer, 2003.

[21] M. Clerc and J. Kennedy, "The particle swarm – explosion, stability, and convergence in a multidimensional complex space," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 1, pp. 58–73, 2002.

[22] I. A. Mohammad, Y. Chernyak, and M. Plesch, "HOPSO: A robust classical optimizer for VQE." 2025. Available: https://arxiv.org/abs/2508.13651

[23] Y. Chernyak, I. A. Mohammad, N. Masnicak, M. Pivoluska, and M. Plesch, "Harmonic oscillator based particle swarm optimization," *PLOS One*, vol. 20, no. 6, p. e0326173, Jun. 2025, doi: 10.1371/journal.pone.0326173.

[24] K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii, "Quantum circuit learning," *Physical Review A*, vol. 98, no. 3, p. 032309, 2018, doi: 10.1103/PhysRevA.98.032309.

[25] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, "The theory of variational hybrid quantum-classical algorithms," *New Journal of Physics*, vol. 18, no. 2, p. 023023, 2016.

[26] H. R. Grimsley, S. E. Economou, E. Barnes, and N. J. Mayhall, "An adaptive variational algorithm for exact molecular simulations on a quantum computer," *Nature Communications*, vol. 10, no. 1, p. 3007, 2019.

[27] H. L. Tang *et al.*, "Qubit-ADAPT-VQE: An adaptive algorithm for constructing hardware-efficient ansätze on a quantum processor," *PRX Quantum*, vol. 2, no. 2, p. 020310, 2021.

[28] A. Nemirovski, A. Juditsky, G. Lan, and A. Shapiro, *Robust stochastic approximation approach to stochastic programming*, vol. 19. SIAM, 2009, pp. 1574–1609.

[29] K. J. Åström and R. M. Murray, *Feedback systems: An introduction for scientists and engineers*. Princeton University Press, 2010.

[30] E. Farhi, J. Goldstone, and S. Gutmann, "A quantum approximate optimization algorithm," *arXiv preprint arXiv:1411.4028*, 2014.

[31] M. Schuld and N. Killoran, "Quantum machine learning in feature hilbert spaces," *Physical Review Letters*, vol. 122, no. 4, p. 040504, 2019.