Quantum correlations and exceptional points in superconducting processors: collective measurements, generative learning, and dynamical control

Patrycja Tulewicz

Supervisor: dr hab. Karol Bartkiewicz, prof. UAM

Department of Quantum Information
Institute of Spintronics and Quantum Information
Faculty of Physics and Astronomy
ADAM MICKIEWICZ UNIVERSITY IN POZNAŃ, POLAND

Poznań 2026

Acknowledgements

I would like to thank my supervisor, prof. UAM dr hab. Karol Bartkiewicz, for his support, kindness, and excellent scientific guidance.
I would also like to thank my colleagues from the Institute of Spintronics and Quantum Information for many fruitful discussions, cups of coffee, and a fantastic atmosphere.
I am deeply grateful to my family for their unwavering support and understanding during all these years.

The dissertation was supported by the National Science Centre with funds awarded under the Maestro grant *Fundamental problems and implementations of dissipative quantum engineering* No. DEC-2019/34/A/ST2/00081.

NATIONAL SCIENCE CENTRE
POLAND

Abstract

Quantum systems are inherently linear at the level of their evolution equations, yet the properties that make them useful for computation, communication, and sensing are profoundly nonlinear. Entanglement cannot be captured by any single linear observable; generating a state with prescribed quantum correlations requires navigating a highly non-convex optimization landscape; and the dynamics of open quantum systems can develop singularities in parameter space that have no counterpart in ordinary Hermitian physics. This dissertation addresses all three aspects of this nonlinearity on current superconducting quantum hardware: how to measure it efficiently, how to generate it deliberately, and how to exploit it as a resource rather than treat it as an obstacle.

The characterization problem is addressed in Chapter 2 through a family of multicopy measurement protocols that operate on pairs or small sets of identically prepared state copies rather than on a single copy at a time. This approach reduces the number of measurement settings required to estimate entanglement from the exponential cost of full quantum state tomography, reducing the exponent base from 4^n to 2^n in the number of qubits, with a verified 67% reduction in overhead for two-qubit systems on IBM Quantum hardware. A central result of this chapter is the identification of partial-transpose moments as chirality-chirality correlators: the difference between partial-transpose and purity moments encodes the scalar spin chirality, a quantity familiar from condensed-matter physics as the order parameter of chiral spin liquids. This connection provides a physical interpretation for an otherwise abstract algebraic operation, and it extends naturally to bound entanglement detection through the realignment non-Hermiticity gap. Using the same controlled-SWAP circuit infrastructure, a multi-channel spectral classifier detects bound entanglement across all three known families of 3×3 positive-partial-transpose entangled states, including states invisible to any single realignment criterion, achieving 99.6% cross-validated recall at zero false positives in simulation and 94% recall on *ibm_fez* hardware. This is the first family-agnostic demonstration of bound entanglement detection on a gate-based superconducting processor. Gröbner basis decision trees provide an additional route to negativity estimation, converting spectral moments into entanglement values adaptively and achieving a five- to sevenfold reduction in measurement settings relative to full tomography.

The generation problem is addressed in Chapter 3 through the Synergic Quantum Generative Network (SQGEN), a quantum machine learning architecture that replaces the adversarial training paradigm of quantum generative adversarial networks with a cooperative one. Generator and discriminator are trained simultaneously by minimizing a single cost function, evaluated in a single circuit execution per training step by exploiting the unitarity of the discriminator. This eliminates the principal source of hyperparameter sensitivity in adversarial training and reduces the qubit requirement from $3n + 2$ to $n + 1$ for an n-qubit target state. On a single-qubit target, SQGEN requires an average of 33 circuit evaluations per training epoch compared to 150 for a standard quantum generative adversarial network, achieving generation fidelity $F \approx 0.92$ on *ibmq_manila* hardware. The architecture connects naturally to quantum kernel theory: training SQGEN is equivalent to learning a quantum kernel in 2^n-dimensional Hilbert space, a perspective that suggests analyzing its convergence within kernel learning theory rather than game theory.

The control problem is addressed in Chapter 4, which presents the first experimental demonstration that a quantum system can exhibit Liouvillian exceptional points in the complete absence of Hamiltonian exceptional points. Exceptional points are parameter values at which two or more eigenvalues and their associated eigenvectors simultaneously coalesce, rendering the governing operator non-diagonalizable. In semiclassical descriptions based on effective non-Hermitian Hamiltonians, exceptional points arise only when the Hamiltonian itself becomes defective. The full Lindblad superoperator, however, contains quantum jump contributions that have no semiclassical counterpart, and these can create spectral degeneracies in systems where the effective Hamiltonian remains diagonal and no Hamiltonian exceptional point is possible. This distinction is demonstrated experimentally on *ibm_nairobi* using quantum process tomography, which reconstructs the full Liouvillian superoperator and provides direct access to the eigenmatrix overlap that distinguishes a true exceptional point from an ordinary degeneracy. Near each of the two Liouvillian exceptional points identified, the spectral gap exhibits a threefold sensitivity enhancement, demonstrating that Liouvillian singularities are not merely theoretical curiosities but observable and potentially exploitable features of realistic quantum hardware.

Taken together, the three contributions form a modular toolkit for working with nonlinear quantum systems: characterizing their correlations without prior knowledge of the state, generating states with prescribed properties within a reduced hardware budget, and engineering the singular dynamics of open systems to enhance measurement sensitivity. The results demonstrate that the nonlinear properties of quantum systems are accessible, controllable, and practically useful on current noisy intermediate-scale quantum devices, and they outline a research agenda that extends naturally toward fault-tolerant quantum computing.

Table of contents

Chapter 1

Introduction: the relevance of nonlinear properties of states in quantum computing

1.1 The powerhouse of quantum computing

Quantum mechanics is fundamentally a linear theory because the evolution of quantum states, as described by Schrödinger's equation, adheres to the principle of superposition while observables act linearly on Hilbert space. However, it is the nonlinear functions of observables in quantum systems that describe the foundational quantities in quantum information processing and provide quantum advantage over classical computing.

To discuss nonlinear functions in the context of quantum systems, the term must first be properly defined. Quantum systems whose properties cannot be described as tensor products of subsystem states are entangled. The entanglement is not in general described by a single observable.

1. Entangled states are a key example of quantum resources. A two-qubit system is in a separable state if it can be written as a product of the state of qubit A and B:

$$|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle. \tag{1.1}$$

A state is considered entangled if such a decomposition is impossible. Some entangled states exhibit correlations that cannot be explained by any hidden variable theory, as demonstrated by the violation of Bell's inequality [1]. These nonlinear correlations are fundamental to quantum computing algorithms, quantum cryptography, and quantum teleportation.

2. Measurement processes introduce nonlinearity into the dynamics of quantum systems. According to the postulates of quantum mechanics, a measurement implements a nonlinear transformation of the quantum state, which is described by a density matrix ρ. This transformation results in a probability distribution:

$$\rho \xrightarrow{M} \rho' = \mathrm{Tr}_M(M \otimes \mathbb{I}\rho), \tag{1.2}$$

where M_m are elements of a POVM (positive operator-valued measure), and they span the complete space $\sum_m M_m^\dagger M_m = \mathbb{I}$. This nonlinearity is irreversible and poses a significant challenge to the characterization of quantum systems. It also motivates the measurement methods developed in Chapter 2 of this dissertation.

3. Unlike unitary evolution described by the Hermitian Hamiltonian, real quantum systems are open and interact with their environment, leading to nonlinear dynamics described by the Liouville superoperator. These dynamics create singularities in the parameter space where the spectral properties of the system's evolution operator change qualitatively, a phenomenon known as an *exceptional point*.

These three sources of nonlinearity underpin quantum advantage across the main application domains. Quantum algorithms such as Shor's factoring [2] and Grover's search [3] rely on entangled multi-qubit states; without nonclassical correlations, the computation could be efficiently simulated classically. Quantum key distribution protocols (BB84 [4], E91 [5]) use the nonlinear structure of entangled states to guarantee information-theoretic security, with Bell-inequality violations forming the basis of device-independent cryptography. Quantum error correction encodes logical information into highly entangled states of many physical qubits, using nonclassical correlations to detect and correct errors without directly measuring the protected state. In quantum sensing, entangled probe states and the singular response near exceptional points of open systems are expected to push measurement sensitivity beyond the standard quantum limit toward the Heisenberg limit.

The current era of noisy intermediate-scale quantum (NISQ) devices [6] makes the practical exploitation of these resources especially demanding. Uncontrolled dissipation degrades entanglement on timescales comparable to gate durations, while the limited qubit count and circuit depth restrict the complexity of states that can be reliably prepared and characterized. These constraints motivate the central questions of this dissertation:

1. How can the nonclassical properties of quantum systems be efficiently characterized (Chapter 2)? Traditional quantum state tomography scales as $\mathcal{O}(4^n)$ measurements for an n-qubit system, quickly becoming impractical on NISQ hardware. Efficient

methods for estimating entanglement and nonlocality without full state reconstruction are therefore needed.

2. How can quantum states with prescribed nonlinear properties be generated or controlled (Chapters 3 and 4)? On the generative side, machine learning offers a route to optimizing variational circuits without analytical design of gate sequences. On the dynamical side, open-system dissipation need not be treated purely as an obstacle: Liouvillian dynamics can be engineered to produce nontrivial mixed states, and the singular response near Liouvillian exceptional points can be harnessed for enhanced sensing.

The chapters that follow develop an integrated set of methods addressing these questions, progressing from characterization through generation to the constructive use of open-system singularities, with experimental validation on real quantum processors throughout.

1.2 Fundamental challenges

Working with nonlinear properties of quantum systems in the NISQ era raises three interconnected challenges that define the structure of this dissertation: how to measure nonclassical properties efficiently, how to generate states that possess them, and how to exploit rather than merely tolerate the open-system dynamics that governs realistic hardware.

1.2.1 How can the nonlinear properties of quantum systems be measured?

Entanglement and nonlocality are not directly observable in the sense of linear quantum mechanics. Measures such as negativity, based on the PPT criterion introduced by Peres [7] and Horodecki [8], are nonlinear functionals of the density matrix. They require knowledge of the eigenvalues of the partially transposed state rather than the expectation values of fixed observables.

The standard route to obtaining this information is *quantum state tomography (QST)*, which reconstructs the full density matrix from a complete set of measurements. The required number of measurement settings scales as:

$$N_{\text{QST}} = \mathcal{O}(4^n) \tag{1.3}$$

for an n-qubit system, rendering full tomography impractical even for moderate system sizes on hardware with limited coherence and non-negligible gate errors.

Chapter 2 addresses this challenge through three complementary approaches. Multicopy estimation (MCE) reduces the measurement overhead from $\mathcal{O}(4^n)$ to $\mathcal{O}(2^n)$ by operating on pairs of identically prepared states. Neural networks trained on MCE data can approximate nonlinear entanglement functionals without solving eigenvalue equations. A third approach constructs Gröbner basis decision trees that convert partial-transpose spectral moments directly into negativity values, achieving a further five- to sevenfold reduction in the number of required measurement settings.

1.2.2 How can nonclassical states be produced efficiently?

Preparing a quantum state with prescribed entanglement properties requires finding a gate sequence whose output has the desired nonlinear characteristics. In principle, any state can be reached by a suitable unitary circuit, but in practice two obstacles arise. The circuit depth needed to entangle n qubits grows with system size, accumulating errors on NISQ hardware before the target state is reached. The variational landscape of parameterized circuits is highly non-convex, with barren plateaus where gradients vanish exponentially and local minima that trap naive optimization.

Generative machine learning offers a systematic approach to this problem. Classical generative adversarial networks (GANs) learn complex probability distributions through a competition between a generator and a discriminator; their quantum counterparts, QGANs [9, 10], extend this idea to quantum state distributions. The standard QGAN architecture, however, demands independent generator and discriminator circuits together with ancilla qubits for state comparison, requiring:

$$N_{\text{qubits}}^{\text{QGAN}} = 3n + 2 \tag{1.4}$$

qubits to generate an n-qubit target state, eleven qubits for a three-qubit target, for instance. Chapter 3 replaces the adversarial paradigm with a cooperative one (SQGEN), where a single cost function drives generator and discriminator jointly, reducing the qubit overhead to:

$$N_{\text{qubits}}^{\text{SQGEN}} = n + 1, \tag{1.5}$$

and eliminating the need for manual hyperparameter tuning.

A complementary strategy for producing nontrivial mixed states is to engineer the open-system dynamics itself. Chapter 4 takes this perspective by considering Liouvillian exceptional points (LEPs): singularities in the spectrum of the Liouvillian superoperator where two eigenvalues and their associated eigenmatrices coalesce. At a second-order LEP the system's response to a small perturbation $\delta\kappa$ of a control parameter scales as

$\Delta\lambda \propto \sqrt{|\delta\kappa|}$ rather than linearly, a nonlinear sensitivity that can be exploited for enhanced quantum sensing. Locating and characterizing these singularities experimentally is nontrivial, requiring quantum process tomography to reconstruct the full Liouvillian from measurement data, which Chapter 4 carries out on a superconducting processor.

1.3 Theoretical foundations

1.3.1 The postulates of quantum mechanics

The mathematical structure of quantum mechanics is based on the concept of a Hilbert space, denoted by $\mathcal{H}$, which provides a suitable framework for describing quantum systems. As a complete vector space with a scalar product, Hilbert space allows us to define geometric concepts, such as angles and vector distances [11].

Postulate I: The state of an isolated quantum system is fully characterized by a unit vector $|\psi\rangle$ in the complex Hilbert space $\mathcal{H}$ assigned to the system. For a d-level system, the state vector can be expressed in any orthonormal basis $\{|i\rangle\}_{i=0}^{d-1}$ as:

$$|\psi\rangle = \sum_{i=0}^{d-1} c_i |i\rangle, \tag{1.6}$$

where the complex amplitudes satisfy the normalization condition $\langle\psi|\psi\rangle = \sum_{i=0}^{d-1} |c_i|^2 = 1$. The value of $|c_i|^2$ represents the probability of finding the system in state $|i\rangle$ upon measurement.

Postulate II: The time evolution of a closed quantum system is described by a unitary transformation $|\psi_2\rangle = U|\psi_1\rangle$, where U depends only on the time interval. Schrödinger's equation,

$$i\hbar \frac{d}{dt}|\psi(t)\rangle = H(t)|\psi(t)\rangle \tag{1.7}$$

regulates this evolution through the Hamiltonian operator $H(t)$. For time-independent Hamiltonians, the evolution operator is $U(t) = e^{-iHt/\hbar}$. Unitarity ensures preservation of the quantum state norm and the probabilistic interpretation.

Postulate III: Quantum measurements are described by a set of measurement operators $\{M_m\}$ satisfying the completeness condition $\sum_m M_m^\dagger M_m = \mathbb{I}$. The probability of obtaining outcome m is:

$$p(m) = \langle\psi|M_m^\dagger M_m|\psi\rangle, \tag{1.8}$$

and the post-measurement state is

$$|\psi_m\rangle = M_m|\psi\rangle/\sqrt{p(m)}. \tag{1.9}$$

For projective measurements of an observable A with eigenvalues a_m and eigenvectors $|a_m\rangle$, the measurement operators are projectors:

$$P_m = |a_m\rangle\langle a_m|. \tag{1.10}$$

Postulate IV: The state space of a composite physical system is the tensor product of its subsystems' state spaces. For systems in states $|\psi_i\rangle$, the joint state is

$$|\psi_1\rangle \otimes |\psi_2\rangle \otimes \cdots \otimes |\psi_n\rangle, \tag{1.11}$$

and the combined Hilbert space is

$$\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B. \tag{1.12}$$

These postulates have fundamental consequences for quantum information theory. Quantum states can exist in coherent superpositions, and the tensor product structure permits entangled states that cannot be decomposed into product states. The measurement process is essentially irreversible, introducing randomness, while the uncertainty principle limits simultaneous precise measurements of non-commutative observables. These properties will be crucial in later chapters when discussing methods for characterizing quantum entanglement and machine learning applications in quantum systems.

1.3.2 Formalism of density matrices

The formalism of state vectors, as presented in the previous section, is useful for describing pure states, that is, systems about which we have complete quantum information. However, pure states are an idealization. In practical situations, we often deal with mixed states, which result from incomplete knowledge of a system's state or interaction with its environment.

A *pure state* is a system that can be described by a single state vector, $|\psi\rangle$. All of the system's physical properties are fully determined by this vector. A *mixed state* arises when a system can exist in several possible pure states, $|\psi_i\rangle$, with specific classical probabilities, p_i. Unlike quantum superposition of pure states, which allows for interference between states, a statistical mixture of different states involves completely independent realizations.

The *density matrix* (*density operator*), ρ, generalizes the formalism of state vectors, allowing for a uniform description of pure and mixed states. For a mixed state, in which the

system is in state $|\psi_i\rangle$ with probability p_i, the density matrix is defined as follows:

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|, \tag{1.13}$$

where $p_i \geq 0$ for all i, $\sum_i p_i = 1$ (normalization of probabilities), $|\psi_i\rangle$ are pure states.

The density matrix has the following fundamental properties:

- Hermitian: $\rho^\dagger = (\sum_i p_i |\psi_i\rangle\langle\psi_i|)^\dagger = \sum_i p_i |\psi_i\rangle\langle\psi_i| = \rho$,
- unit trace: $\text{Tr}(\rho) = \sum_i p_i \langle\psi_i|\psi_i\rangle = \sum_i p_i = 1$,
- positive semidefiniteness: $\langle\phi|\rho|\phi\rangle = \sum_i p_i |\langle\phi|\psi_i\rangle|^2 \geq 0$,
- purity property: a state is pure if and only if: $\text{Tr}(\rho^2) = \begin{cases} 1 & \text{for a pure state,} \\ < 1 & \text{for a mixed state,} \end{cases}$ and more generally, for a system of dimension d, the purity satisfies:

$$\frac{1}{d} \leq \text{Tr}(\rho^2) \leq 1, \tag{1.14}$$

  where the lower bound $\text{Tr}(\rho^2) = 1/d$ corresponds to the maximally mixed state $\rho = \mathbb{I}/d$.

The expected value of any observable A in a given state, described by ρ, can be expressed as follows:

$$\langle A \rangle = \text{Tr}(\rho A). \tag{1.15}$$

When a measurement is performed using the operator M_m, the probability of obtaining the result m is given by:

$$p(m) = \text{Tr}(M_m^\dagger M_m \rho). \tag{1.16}$$

In composite systems, we are often interested in the state of a particular subsystem. A *partial trace* enables us to obtain the density matrix of a subsystem from that of the entire system.

For a system AB with a density matrix ρ^{AB}, the density matrix of subsystem A can be obtained by taking the partial trace with respect to subsystem B:

$$\rho^A = \text{Tr}_B(\rho^{AB}). \tag{1.17}$$

Constructing an explicit partial trace, we can write that if $\{|i\rangle^A\}$ and $\{|j\rangle^B\}$ are orthonormal bases in $\mathcal{H}^A$ and $\mathcal{H}^B$ respectively, then:

$$\rho^A = \sum_j \langle j|^B \rho^{AB} |j\rangle^B. \tag{1.18}$$

In matrix representation, the elements ρ^A are given by:

$$\langle i|\rho^A|k\rangle = \sum_j \langle ij|\rho^{AB}|kj\rangle. \tag{1.19}$$

A partial trace has the following properties:

- linearity: $\mathrm{Tr}_B(\alpha\rho_1 + \beta\rho_2) = \alpha\mathrm{Tr}_B(\rho_1) + \beta\mathrm{Tr}_B(\rho_2)$,
- trace preservation: $\mathrm{Tr}[\mathrm{Tr}_B(\rho^{AB})] = \mathrm{Tr}(\rho^{AB})$,
- product states: for separable product states $\rho^{AB} = \rho^A \otimes \rho^B$, the trace factorizes:

$$\mathrm{Tr}(\rho^A \otimes \rho^B) = \mathrm{Tr}_A(\rho^A) \cdot \mathrm{Tr}_B(\rho^B), \tag{1.20}$$

- irreversibility: ρ^{AB} cannot be reconstructed from ρ^A (loss of information).

Pauli representation for qubits

For single-qubit systems, the density matrix can be expressed in terms of the Pauli matrices σ_x, σ_y, σ_z and the identity $\mathbb{I}$:

$$\rho = \frac{1}{2}(\mathbb{I} + \vec{s}\cdot\vec{\sigma}), \tag{1.21}$$

where $\vec{s} = (s_x, s_y, s_z)$ is the Bloch vector and $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$. The Pauli matrices are defined as:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.22}$$

The Bloch vector $\vec{s}$ provides a complete geometric picture of single-qubit states. Every *pure* state $|\psi\rangle$ corresponds to a unique point on the surface of the unit sphere, the *Bloch sphere*, according to the parametrization

$$|\psi\rangle = \cos\tfrac{\theta}{2}\,|0\rangle + e^{i\varphi}\sin\tfrac{\theta}{2}\,|1\rangle, \tag{1.23}$$

where $\theta \in [0, \pi]$ is the polar angle from the z-axis and $\varphi \in [0, 2\pi)$ is the azimuthal angle, giving $\vec{s} = (\sin\theta\cos\varphi,\, \sin\theta\sin\varphi,\, \cos\theta)$. The six cardinal points of the sphere are the eigenstates of the Pauli operators: the north and south poles $|0\rangle$, $|1\rangle$ are the eigenstates of σ_z; the equatorial points $|\pm\rangle = (|0\rangle \pm |1\rangle)/\sqrt{2}$ are the eigenstates of σ_x; and $|\pm i\rangle = (|0\rangle \pm i|1\rangle)/\sqrt{2}$ those of σ_y (Fig. 1.1). Mixed states, described by density matrices with $\mathrm{Tr}(\rho^2) < 1$, correspond to points strictly inside the sphere, with the maximally mixed state $\rho = \mathbb{I}/2$ at the origin where $|\vec{s}| = 0$.

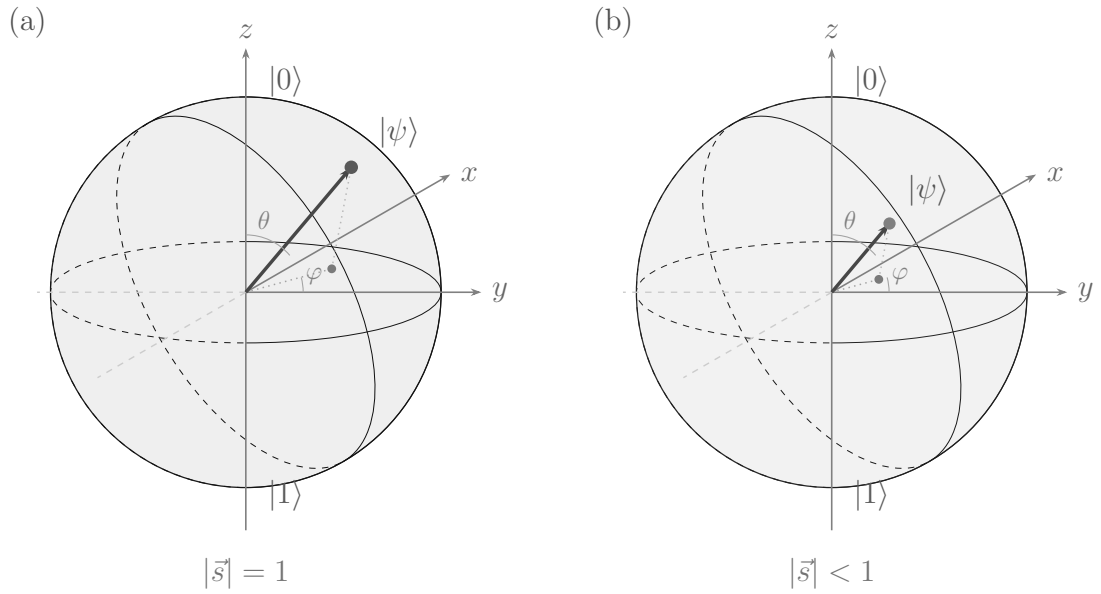


Fig. 1.1 **Bloch sphere representation of a single-qubit state.** (a) Pure state: the Bloch vector $\vec{s} = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta)$ reaches the unit sphere surface ($|\vec{s}| = 1$), corresponding to $|\psi\rangle = \cos(\theta/2)\,|0\rangle + e^{i\varphi}\sin(\theta/2)\,|1\rangle$. (b) Mixed state $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$: the Bloch vector lies strictly inside the sphere ($|\vec{s}| < 1$), with the maximally mixed state $\rho = \mathbb{I}/2$ at the origin.

The Bloch vector satisfies $|\vec{s}| \leq 1$, where $|\vec{s}| = 1$ corresponds to pure states and $|\vec{s}| < 1$ to mixed states.

For multi-qubit systems, the density matrix can be expanded in the basis of tensor products of Pauli matrices. For an n-qubit system, this expansion can be written as:

$$\rho = \frac{1}{2^n}\left(\mathbb{I}^{\otimes n} + \sum_{k=1}^{n} \vec{s}^{(k)} \cdot \vec{\sigma}^{(k)} + \sum_{k<l}\sum_{ij} \beta_{ij}^{(kl)}\, \sigma_i^{(k)} \otimes \sigma_j^{(l)} + \dots\right), \tag{1.24}$$

where $\vec{s}^{(k)} = (s_x^{(k)}, s_y^{(k)}, s_z^{(k)})$ are the local Bloch vectors for each qubit k, describing single-qubit properties, $\beta_{ij}^{(kl)}$ are the elements of correlation matrices describing two-qubit correlations between qubits k and l, and the ellipsis represents higher-order (three-body, four-body, etc.) correlation terms.

For the two-qubit case ($n = 2$), this expansion simplifies to:

$$\rho = \frac{1}{4}\left(\mathbb{I}\otimes\mathbb{I} + \vec{s}\cdot\vec{\sigma}\otimes\mathbb{I} + \mathbb{I}\otimes\vec{p}\cdot\vec{\sigma} + \sum_{i,j=x,y,z} \beta_{ij}\,\sigma_i\otimes\sigma_j\right), \tag{1.25}$$

where $\vec{s} = (s_x, s_y, s_z)$ and $\vec{p} = (p_x, p_y, p_z)$ are the Bloch vectors describing local properties of qubits A and B, respectively, observed individually. The correlations between the subsystems are described by the 3×3 correlation matrix $\hat{\beta}$ with elements β_{ij} (where $i, j \in \{x, y, z\}$). The normalization condition $\mathrm{Tr}(\rho) = 1$ is automatically satisfied by this parametrization.

Partial transposition

The *partial transpose* is a fundamental operation for detecting entanglement in bipartite systems. For a bipartite state ρ^{AB}, the partial transpose with respect to subsystem A is defined as:

$$\langle ij|\rho^{\Gamma_B}|kl\rangle = \langle il|\rho^{AB}|kj\rangle \tag{1.26}$$

where $|ij\rangle \equiv |i\rangle^A \otimes |j\rangle^B$. In operator form, if $\rho^{AB} = \sum_{ijkl} \rho_{ijkl} \, |i\rangle \, \langle k|^A \otimes |j\rangle \, \langle l|^B$, then:

$$\rho^{\Gamma_B} = \sum_{ijkl} \rho_{ijkl} |i\rangle\langle k|^A \otimes |l\rangle\langle j|^B. \tag{1.27}$$

The partial transpose preserves the trace and Hermiticity of the density matrix. However, unlike physical operations, it does not necessarily preserve positive semidefiniteness.

Separability criteria

The *separability criterion* indicates that the state ρ^{AB} is separable (unentangled) if and only if it can be written as follows:

$$\rho^{AB} = \sum_i p_i \rho_i^A \otimes \rho_i^B, \tag{1.28}$$

where ρ_i^A and ρ_i^B are the density matrices of subsystems A and B, respectively, and $p_i \geq 0$, and $\sum_i p_i = 1$. States that cannot be written in this form are called *entangled*.

The *positive partial transpose (PPT)* criterion, introduced by Peres [7] and generalized by the Horodecki family [8], is one of the most important and practical tools for detecting entanglement.

According to Peres's theorem [7], if state ρ^{AB} is separable, then its partial transpose is positive semidefinite:

$$\rho^{AB} \text{ separable} \Rightarrow \rho^{\Gamma_B} \geq 0. \tag{1.29}$$

The contraposition defines the entanglement criterion:

$$\rho^{\Gamma_B} \ngeq 0 \Rightarrow \rho^{AB} \text{ is entangled.} \tag{1.30}$$

A state with a negative partial transpose (NPT) is always entangled.

The Horodecki theorem [8] generalizes Peres's condition by stating that, for $2 \otimes 2$ and $2 \otimes 3$ systems, the PPT criterion is both necessary and sufficient for separability:

$$\text{for } 2 \otimes 2 \text{ and } 2 \otimes 3: \quad \rho^{AB} \text{ separable} \Leftrightarrow \rho^{\Gamma_B} \geq 0. \tag{1.31}$$

In higher dimensions, there exist bound entangled states with positive partial transposition, for which the PPT criterion is necessary but not sufficient.

Realignment criterion

The *realignment* (also called *reshuffling*) is another matrix operation useful for detecting entanglement, particularly in cases where the PPT criterion is inconclusive. For a bipartite state ρ^{AB} with matrix elements $\rho_{ij,kl} = \langle ij|\rho^{AB}|kl\rangle$ in the computational basis, the realigned matrix ρ^R is defined by rearranging the indices:

$$\langle ik|\rho^R|jl\rangle = \langle ij|\rho^{AB}|kl\rangle. \tag{1.32}$$

Equivalently, if we write the density matrix in block form as:

$$\rho^{AB} = \sum_{i,j,k,l} \rho_{ij,kl} |i\rangle\langle k|^A \otimes |j\rangle\langle l|^B, \tag{1.33}$$

then the realignment operation transforms it to:

$$\rho^R = \sum_{i,j,k,l} \rho_{ij,kl} |i\rangle\langle j|^A \otimes |k\rangle\langle l|^B. \tag{1.34}$$

The realignment criterion states that for separable states, $\|\rho^R\|_1 \leq 1$, where $\|\cdot\|_1$ denotes the trace norm. Violations of this bound, $\|\rho^R\|_1 > 1$, indicate entanglement. The realignment operation is particularly useful because it can detect certain entangled states that the partial transpose criterion fails to identify, including some bound entangled states with positive partial transpose, making it a complementary tool for entanglement detection [12, 13].

The density matrix formalism, along with operations such as partial transposition and realignment, links the formalism of state vectors to practical applications in quantum technologies, which inevitably involve mixed states and open systems. These tools will be essential in later chapters for characterizing and quantifying quantum correlations.

Makhlin invariants

Two-qubit states can be described using local unitary invariants, which are quantities that remain unchanged under local unitary transformations performed on each subsystem independently. Makhlin introduced eighteen independent locally unitary invariants for a two-qubit state described by the correlation matrix $\hat{\beta}$ and local Bloch vectors $\mathbf{s}$ and $\mathbf{p}$ for the first and second subsystems, respectively [14]. To express negativity and nonlocality, a subset of these invariants is required. This can be written as follows:

$$\begin{aligned} I_1 = \det\hat{\beta}, \quad I_2 = \mathrm{Tr}\left(\hat{\beta}^T\hat{\beta}\right), \quad I_3 = \mathrm{Tr}\left(\hat{\beta}^T\hat{\beta}\right)^2, \quad I_4 = \mathbf{s}^2, \quad I_5 = [\mathbf{s}\hat{\beta}]^2, \\ I_6 = \left[\mathbf{s}\hat{\beta}\hat{\beta}^T\right]^2, \quad I_7 = \mathbf{p}^2, \quad I_8 = [\hat{\beta}\mathbf{p}]^2, \quad I_9 = \left[\hat{\beta}^T\hat{\beta}\mathbf{p}\right]^2 \\ I_{12} = \mathbf{s}\hat{\beta}\mathbf{p}, \quad I_{13} = \mathbf{s}\hat{\beta}\hat{\beta}^T\hat{\beta}\mathbf{p}, \quad I_{14} = e_{ijk}e_{lmn}s_i p_l \beta_{jm}\beta_{kn}, \end{aligned} \tag{1.35}$$

where e_{ijk} is the Levi-Civita symbol. These invariants enable a full characterization of a two-qubit state with precision up to local unitary transformations. In other words, two states are considered locally equivalent if and only if they have the same values for all invariants.

1.3.3 Open quantum systems and Liouvillian formalism

The postulates presented in the previous section describe the evolution of isolated systems through unitary transformations. Realistic quantum systems, however, are never truly isolated: they inevitably exchange energy and information with their surrounding environment, giving rise to dissipation, decoherence, and loss of quantum coherence. These effects cannot be captured by unitary evolution alone. The framework for handling such scenarios is the theory of open quantum systems, which partitions the universe into a small *system* of interest S and a large *environment* (reservoir) E, and traces out the environmental degrees of freedom to obtain an effective equation of motion for the system alone.

The Lindblad master equation

When the system-environment coupling is weak and the environmental correlation time τ_{env} is much shorter than the system's characteristic timescale τ_{sys}, the *Markovian approximation* holds: the environment has no memory, and the reduced dynamics of $\rho_S(t) = \mathrm{Tr}_E(\rho_{SE}(t))$ takes the form of the Gorini–Kossakowski–Sudarshan–Lindblad (GKSL) master equation [15, 16]:

$$\frac{d\rho}{dt} = \mathcal{L}\rho = -i[H,\rho] + \sum_{\mu}\mathcal{D}[\Gamma_\mu]\rho, \tag{1.36}$$

where $\mathcal{L}$ is the Liouvillian superoperator, H is the system Hamiltonian, and the Lindblad dissipators read

$$\mathcal{D}[\Gamma_\mu]\rho = \Gamma_\mu \rho \Gamma_\mu^\dagger - \tfrac{1}{2}\{\Gamma_\mu^\dagger \Gamma_\mu, \rho\}. \tag{1.37}$$

The operators Γ_μ are the quantum jump operators, each representing a distinct dissipative channel such as spontaneous emission, dephasing, or inelastic scattering. The first term of Eq. (1.36) governs coherent evolution; the second accounts for irreversible information loss to the environment.

The master equation has a transparent physical interpretation in the quantum trajectory (Monte Carlo wave-function) picture [17, 18]. Between jumps, the state evolves under the effective non-Hermitian Hamiltonian

$$H_{\text{eff}} = H - \frac{i}{2}\sum_\mu \Gamma_\mu^\dagger \Gamma_\mu, \tag{1.38}$$

while a jump Γ_μ occurs stochastically at a rate proportional to $\langle \Gamma_\mu^\dagger \Gamma_\mu \rangle$. In terms of H_{eff}, the master equation takes the equivalent form:

$$\mathcal{L}\rho = -i\left[H_{\text{eff}}\rho - \rho H_{\text{eff}}^\dagger\right] + \sum_\mu \Gamma_\mu \rho \Gamma_\mu^\dagger, \tag{1.39}$$

which makes explicit the two contributions: continuous non-unitary decay through H_{eff}, and discrete quantum jumps Γ_μ. Averaging over many stochastic trajectories recovers the smooth Lindblad evolution.

Superoperator formalism and the Liouvillian spectrum

For spectral analysis and quantum process tomography it is convenient to represent the density matrix as a vector. A density matrix ρ can be flattened into a vector $|\tilde{\rho}\rangle$ using the vectorization operation. For a single-qubit density matrix:

$$\rho = \begin{pmatrix} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{pmatrix} \quad \rightarrow \quad |\tilde{\rho}\rangle = \begin{pmatrix} \rho_{00} \\ \rho_{10} \\ \rho_{01} \\ \rho_{11} \end{pmatrix}. \tag{1.40}$$

In this representation, superoperators acting on density matrices become matrices acting on vectors. The Liouvillian can be expressed as:

$$\tilde{\mathcal{L}} = -i\left(H \otimes \mathbb{I} - \mathbb{I} \otimes H^T\right) + \sum_\mu \Gamma_\mu \otimes \Gamma_\mu^* - \frac{1}{2}\sum_\mu \left(\Gamma_\mu^\dagger \Gamma_\mu \otimes \mathbb{I} + \mathbb{I} \otimes \Gamma_\mu^T \Gamma_\mu^*\right), \tag{1.41}$$

or equivalently, using the effective Hamiltonian:

$$\tilde{\mathcal{L}} = -i\Big(H_{\text{eff}} \otimes \mathbb{I} - \mathbb{I} \otimes H_{\text{eff}}^{T}\Big) + \sum_{\mu} \Gamma_{\mu} \otimes \Gamma_{\mu}^{*}. \tag{1.42}$$

The dynamics of the open quantum system is determined by the eigenspectrum of the Liouvillian. The right and left eigenproblems read:

$$\mathcal{L}\rho_n = \lambda_n \rho_n, \tag{1.43}$$

$$\mathcal{L}^{\dagger}\sigma_n = \lambda_n^{*} \sigma_n, \tag{1.44}$$

where ρ_n and σ_n are the right and left eigenmatrices, respectively. The eigenvalues $\lambda_n \in \mathbb{C}$ satisfy $\text{Re}(\lambda_n) \leq 0$; the unique eigenvalue with $\text{Re}(\lambda_0) = 0$ corresponds to the steady state ρ_{ss}, and the magnitudes of the remaining real parts set the relaxation rates toward equilibrium.

For a short time step dt, the density matrix evolution can be approximated as:

$$\rho(t + dt) = \mathcal{S}\,\rho(t), \qquad \mathcal{S} = \mathbb{I} + \mathcal{L}\,dt, \tag{1.45}$$

where $\mathcal{S}$ is the effective quantum operation. This short-time approximation forms the basis of quantum process tomography: by probing the system with multiple input states and measuring the outputs, one can reconstruct $\tilde{\mathcal{L}}$ from the data, as exploited in Chapter 4.

In the quantum trajectory picture, the system evolves under H_{eff} between jumps, and when a jump Γ_{μ} occurs the state transforms as:

$$\rho \to \frac{\Gamma_{\mu}\rho\Gamma_{\mu}^{\dagger}}{\text{Tr}(\Gamma_{\mu}\rho\Gamma_{\mu}^{\dagger})}. \tag{1.46}$$

The master equation recovers the smooth ensemble average over many such stochastic trajectories.

Exceptional points of the Liouvillian

Because $\tilde{\mathcal{L}}$ is generically non-Hermitian, its eigenvalues can coalesce together with the corresponding eigenmatrices at isolated points in parameter space. These *exceptional points* (EPs) are not merely degeneracies of the spectrum: at an EP of order N, both the eigenvalue and the eigenvector of the N coalescing modes become identical, so the Jordan normal form of $\tilde{\mathcal{L}}$ develops a non-trivial block structure. As a result, the response to a small perturbation

$\delta\kappa$ of a control parameter scales as

$$\Delta\lambda \propto |\delta\kappa|^{1/N}, \tag{1.47}$$

a sublinear sensitivity that has attracted interest for enhanced sensing [19].

A crucial distinction separates *Hamiltonian exceptional points* (HEPs), which arise from the non-Hermitian term H_{eff} alone, from *Liouvillian exceptional points* (LEPs), which are properties of the full superoperator $\tilde{\mathcal{L}}$ including the quantum-jump terms. Discarding the jump operators and retaining only H_{eff} can produce, cancel, or shift EPs relative to those of the complete Lindblad dynamics [20]. LEPs therefore represent a genuinely quantum generalization of classical EPs, since they incorporate the discrete stochastic nature of system-environment interactions that has no classical counterpart. The experimental realization and characterization of LEPs without accompanying HEPs is one of the central results of Chapter 4.

Markovian approximation and its validity

The Lindblad master equation relies on the Markovian approximation, which requires that the system-environment coupling be weak, the environment have a continuous spectrum with many degrees of freedom, and the environmental correlation time τ_{env} be much shorter than the system's characteristic time τ_{sys}: $\tau_{\text{env}} \ll \tau_{\text{sys}}$.

When these conditions are not satisfied, non-Markovian effects become important, requiring more sophisticated approaches such as hierarchical equations of motion or the pseudomode method. These methods can be formulated in terms of extended Liouvillian superoperators acting on enlarged state spaces that include auxiliary degrees of freedom representing environmental memory. The Markovian regime covers the superconducting qubit experiments of this dissertation, where environmental correlation times are orders of magnitude shorter than gate durations.

1.3.4 No-go theorems

Alongside the postulates that specify what quantum mechanics permits, a set of no-go theorems delineates what it forbids. These results are not merely formal curiosities: they impose hard constraints on the design of quantum algorithms, communication protocols, and measurement strategies, and several of them bear directly on the methods developed in this dissertation.

The no-cloning theorem

The no-cloning theorem is one of the most fundamental limitations of quantum mechanics. According to this principle, it is impossible to create an identical copy of any unknown quantum state. This concept is an extension of James L. Park's no-go theorem, which shows that there is no simple, perfect measurement scheme that does not cause interference. This result was presented independently by William Wootters and Wojciech H. Żurek [21], as well as by Dennis Dieks.

Theorem: There is no unitary operator U that can clone an arbitrary unknown quantum state, that is:

$$U(|\psi\rangle \otimes |0\rangle) = |\psi\rangle \otimes |\psi\rangle \tag{1.48}$$

for all possible states $|\psi\rangle$.

Proof: Assume such a unitary U exists. For two arbitrary states $|\psi\rangle$ and $|\phi\rangle$:

$$U(|\psi\rangle \otimes |0\rangle) = |\psi\rangle \otimes |\psi\rangle\,, \tag{1.49}$$

$$U(|\phi\rangle \otimes |0\rangle) = |\phi\rangle \otimes |\phi\rangle\,. \tag{1.50}$$

Taking the inner product:

$$\langle\psi|\phi\rangle = (\langle\psi| \otimes \langle 0|)U^{\dagger}U(|\phi\rangle \otimes |0\rangle) = \langle\psi|\phi\rangle\,\langle\psi|\phi\rangle = \langle\psi|\phi\rangle^2\,. \tag{1.51}$$

This equality holds only if $\langle\psi|\phi\rangle = 0$ or $\langle\psi|\phi\rangle = 1$, meaning the states must be either orthogonal or identical. Therefore, arbitrary quantum states cannot be cloned.

The no-cloning theorem has direct consequences for the methods developed here. Since quantum states cannot be copied, characterizing a system requires either multiple copies that are identically prepared or collective measurements on several copies at once. The multicopy approach discussed in Chapter 2 is a direct response to this constraint. The same theorem enforces the use of entanglement-based redundancy in quantum error correction rather than classical copying, and guarantees the security of QKD protocols because any eavesdropping attempt on transmitted quantum states necessarily disturbs them in a detectable way. The SQGEN architecture of Chapter 3 is explicitly designed around this limitation, avoiding the need to duplicate states during training [22].

The no-deleting theorem

The no-deleting theorem is the time-reversal counterpart of the no-cloning theorem. It states that quantum information cannot be completely removed from a composite system.

Theorem: There exists no unitary operation that can delete quantum information from a system while preserving it elsewhere:

$$U(|\psi\rangle \otimes |\psi\rangle \otimes |A\rangle) \neq |\psi\rangle \otimes |0\rangle \otimes |A_\psi\rangle \tag{1.52}$$

for all states $|\psi\rangle$, where $|A\rangle$ is an ancilla state.

This theorem demonstrates that quantum information exhibits a fundamental conservation property: while it can be moved or transformed, it cannot be created from nothing (no-cloning) or destroyed completely (no-deleting) by unitary evolution.

The no-broadcasting theorem

The no-broadcasting theorem generalizes the no-cloning theorem to mixed states: since quantum states cannot be copied, they cannot be broadcast, even approximately, to multiple independent receivers.

Theorem: For non-orthogonal pure states $\{|\psi_i\rangle\}$ with associated density matrices $\rho_i = |\psi_i\rangle \langle\psi_i|$, there exists no quantum operation $\mathcal{E}$ such that:

$$\mathcal{E}(\rho_i \otimes \rho_0) = \rho_i \otimes \rho_i \tag{1.53}$$

for all i, where ρ_0 is a fixed initial state.

The distinction between cloning and broadcasting is subtle: cloning requires perfect copying of the quantum state, while broadcasting allows the output systems to be in any state as long as the reduced density matrices are correct. The no-broadcasting theorem shows that even this weaker requirement cannot be satisfied for arbitrary quantum states.

The no-signaling principle

The no-signaling principle, also known as the no-communication theorem, states that it is impossible to transmit information faster than light during the measurement of an entangled state. This principle preserves the structure of causality in quantum mechanics, guaranteeing that the transmission of information does not violate the special theory of relativity.

Theorem: For a bipartite entangled state ρ^{AB}, local measurements performed by Alice on subsystem A cannot instantaneously affect the measurement statistics observed by Bob on subsystem B:

$$\rho^B = \mathrm{Tr}_A(\rho^{AB}) = \mathrm{Tr}_A[(M_A \otimes \mathbb{I}_B)\rho^{AB}] \tag{1.54}$$

for any measurement operator M_A acting on subsystem A.

Proof: Consider a general bipartite state ρ^{AB} and a measurement on subsystem A with operators $\{M_k\}$ satisfying $\sum_k M_k^\dagger M_k = \mathbb{I}$. The state of subsystem B after Alice's measurement, averaged over all outcomes, is:

$$\rho^B_{\text{avg}} = \sum_k \text{Tr}_A[(M_k \otimes \mathbb{I}_B)\rho^{AB}(M_k^\dagger \otimes \mathbb{I}_B)] = \text{Tr}_A\left[\left(\sum_k M_k^\dagger M_k\right) \otimes \mathbb{I}_B \rho^{AB}\right] = \text{Tr}_A(\rho^{AB}) = \rho^B. \tag{1.55}$$

Thus, Bob's reduced density matrix remains unchanged regardless of Alice's measurement choice.

The no-signaling principle has a precise implication: while entangled states violate Bell inequalities and exhibit nonlocal correlations, those correlations cannot be used to transmit information faster than light. For instance, quantum teleportation requires a classical communication channel in addition to a shared entangled state. The protocol is consistent with relativity because the classical channel is the bottleneck. Local measurements on one part of an entangled system leave the reduced state of the other part unchanged, though they do alter the joint correlations between subsystems.

Consequences for quantum information processing

Taken together, the no-cloning, no-deleting, no-broadcasting, and no-signaling results define the operational boundaries of quantum information. The impossibility of copying arbitrary states eliminates the most basic classical strategies for amplifying or re-transmitting quantum signals. However, it also enforces structural features, such as entanglement-based redundancy, collective measurements, and shared randomness over a classical channel, that distinguish quantum protocols. Each of the central chapters of this dissertation navigates these constraints directly: Chapter 2 through collective multicopy measurements, Chapter 3 through a generative architecture that never needs to clone its training states, and Chapter 4 by using the dissipative dynamics that ordinarily degrades quantum information as a resource for enhanced sensing.

Chapter 2

Characterization of nonlinear systems: collective measurement methods via nonlinear quantum computing

Chapter 1 introduced the theoretical framework for quantum systems, including density matrices, partial transposition, and the Liouvillian formalism. It also presented the central challenge that nonlinear properties, such as entanglement, generally cannot be extracted from a single measurement setting. This chapter addresses that challenge directly, developing three complementary methods for characterizing quantum correlations on near-term hardware without full state reconstruction. Section 2.2 presents a multicopy estimation approach that combines singlet-projection measurements with neural networks to reduce measurement overhead from $\mathcal{O}(4^n)$ to $\mathcal{O}(2^n)$. Section 2.3 introduces Gröbner basis decision trees that convert partial-transpose moments into negativity values adaptively, with a 5–7$\times$ gain in measurement efficiency. Section 2.4 identifies the physical content of those moments as a chirality-chirality correlator, extending the framework to bound entanglement detection via realignment spectral features on the same circuit infrastructure. Together these methods form a coherent measurement toolbox that is experimentally validated on IBM Quantum processors throughout.

2.1 Quantum entanglement: measures, detection, and applications

2.1.1 Entanglement definition and characteristics

Distinguishing between states that can be described as independent, locally correlated systems and states that exhibit quantum correlations that cannot be reproduced within classical physics is of fundamental importance.

Separable and entangled states

This discrimination is the basis of the theory of quantum entanglement and directly impacts quantum applications, including quantum computing, secure communication, and precision metrology. In the case of quantum bipartite states, we can distinguish between *separable states*, which can be prepared through local operations and classical communication, and *entangled states*, whose correlations exceed the capabilities of classical theories with hidden variables [23].

The formal definitions of separable and entangled states, the PPT criterion, and the realignment criterion were introduced in Chapter 1 (Section 1.3.2) and are assumed known here. For reference: a state ρ^{AB} is *separable* if it can be written as a convex combination $\sum_i p_i \rho_i^A \otimes \rho_i^B$, and *entangled* otherwise. The present section focuses on measures and detection strategies rather than definitions.

From EPR to Bell tests: nonlinearity of quantum correlations

The development of quantum entanglement theory has its origins in a fundamental discussion concerning the completeness of quantum mechanics. This discussion was initiated by Einstein, Podolsky, and Rosen in 1935 [24].

Einstein, Podolsky, and Rosen designed a thought experiment involving two particles with perfectly correlated positions and momenta. When one particle in such a system is at position x, the other particle is always at position $x + x_0$, where x_0 is the constant separation between the two particles. This state can be written as follows:

$$|\Psi\rangle_{\text{EPR}} = \int dx |x\rangle_A \otimes |x + x_0\rangle_B. \tag{2.1}$$

According to quantum mechanics, if we measure the position of the first particle and obtain the result x_1, then we immediately know the position of the second particle is $x_0 + x_1$, even without direct interaction. Similarly, measuring the momentum of the first particle

immediately determines the momentum of the second particle due to the conservation of momentum in the entire system.

The EPR paradox arises from the idea that, since particles can be separated by any distance, there cannot be a physical mechanism that allows a measurement of the first particle to "instantly" affect the state of the second particle. This contradicts the theory of relativity, which prohibits transmitting information at a speed greater than that of light. This implies that the properties of both particles (position and momentum) were determined at the moment of their creation. This would mean the existence of hidden local variables that quantum mechanics does not take into account. Therefore, quantum mechanics is an incomplete theory.

The publication of the EPR paper was followed by a debate between Einstein, who argued against the completeness of quantum mechanics, and Bohr, who supported it. Bohr claimed that the EPR authors were mistaken in adopting a classical view of physics, which assumes that physical properties exist independently of observation. According to quantum mechanics, quantum properties, such as position or momentum, only become objective reality after a measurement is made. Thus, measuring the first particle does not "reveal" a previously existing value for the second particle; rather, it defines the experimental context for the entire system. Bohr emphasized that the EPR authors' concept of "elements of reality" contradicts Heisenberg's uncertainty principle and necessitates rejecting the classical notion of an objective reality independent of the observer.

In 1964, John Bell demonstrated that the EPR-Bohr controversy could be resolved through experimentation [1]. He introduced Bell's inequalities, which all theories with local hidden variables must satisfy.

The practical implementation of Bell's idea was presented by Clauser, Horne, Shimony, and Holt (CHSH) [25] in the form of an inequality that can be tested directly in an experiment. In these experiments, spin correlations (or photon polarizations) between two spatially separated detectors are measured. Each detector can be set to one of two possible measurement directions. For spin measurements in the directions $\hat{a}, \hat{a}'$ on particle A and $\hat{b}, \hat{b}'$ on particle B *CHSH inequality* [25] can be written as:

$$|E(a,b) + E(a,b') + E(a',b) - E(a',b')| \leq 2, \tag{2.2}$$

where $E(a,b) = \langle A_a \otimes B_b \rangle$ is the correlation of measurements, calculated as the average of the products of the results ($+1$ or -1) for multiple repetitions of the experiment. All theories based on local hidden variables must satisfy this inequality. If reality is both local and realistic, then no combination of measurement directions can produce a result greater than 2.

According to Tsirelson's theorem [26], quantum mechanics allows Bell's inequality to be violated to the value:

$$\max_{QM} |E(a,b) + E(a,b') + E(a',b) - E(a',b')| = 2\sqrt{2}, \tag{2.3}$$

which was experimentally confirmed [27, 28] and led to the Nobel Prize in 2022.

Quantum entanglement and the violation of Bell's inequality are closely related but not equivalent concepts. All quantum states that violate Bell's inequality must be entangled. This follows directly from the fact that separable states can be described using local hidden variables and therefore cannot violate local boundaries. However, the reverse is not true: there are entangled states that do not violate Bell's inequality. This hierarchy of quantum correlations is significantly important for practical applications [27, 28].

2.1.2 Measures of entanglement

Hierarchy of quantum correlations

A practical characterization of quantum entanglement requires precise methods for detecting and quantifying the degree of correlation. In quantum systems, we can define a hierarchy of increasingly stronger types of correlation. As shown in Fig. 2.1, the quantum states of composite systems can be classified according to the degree of nonlocality exhibited by the correlations.

Quantum entanglement is the weakest level in the hierarchy and a necessary but insufficient condition for stronger quantum correlations. As presented in the previous chapter, a state is described as entangled if it cannot be expressed as a combination of product states. However, this definition does not specify all of this state's possible applications in quantum protocols.

An intermediate level in the hierarchy is *quantum steering*, which was introduced by Schrödinger. This phenomenon enables one side to "steer" the state of the other by selecting the proper measurement settings. In the case of a two-qubit state, the state is considered steerable when, by performing measurements on her qubit, Alice can prepare different states on Bob's qubit in a manner that cannot be explained by a local hidden variable model. The key difference between entanglement and steering is the asymmetry of the latter. A state can be unidirectionally steered from Alice to Bob, but not in the opposite direction. Steering can be experimentally verified using methods that are based on a measure independent device (MDI), which is an important practical advantage because it eliminates the need for complete trust in the measuring hardware on the part of the observer who is being steered [29].

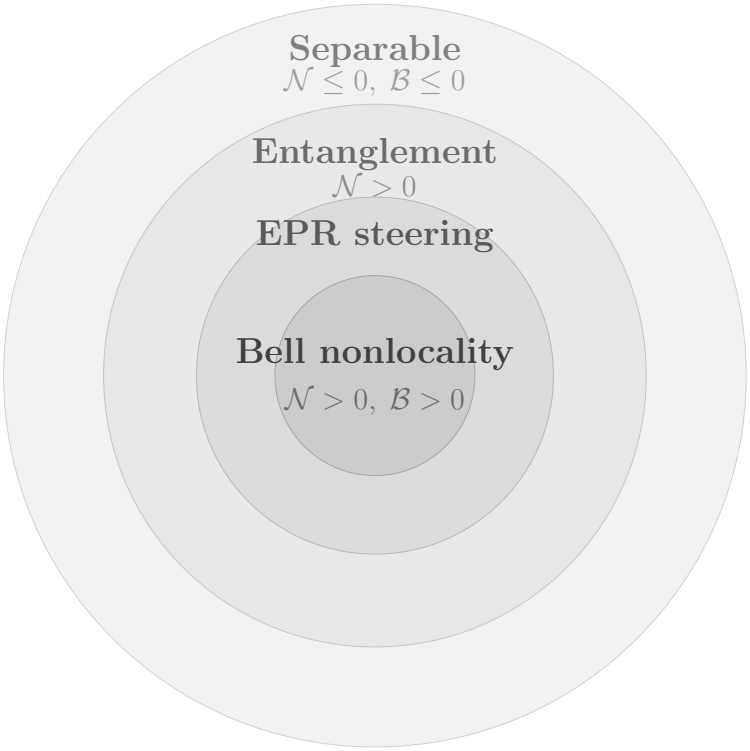


Fig. 2.1 **Hierarchy of quantum correlations in bipartite systems.** Bell nonlocality implies EPR steering, which implies entanglement ($\mathcal{N} > 0$), which implies non-separability. The containments are strict: there exist entangled states that do not exhibit steering, and steerable states that do not violate any Bell inequality.

Bell nonlocality, which manifests itself through the violation of Bell's inequality, is the strongest form of quantum correlation. Any state that violates Bell's inequality is entangled and exhibits steering. However, there are entangled states that do not violate Bell's inequality. These states are referred to as "weakly entangled." Similarly, some states exhibit steering but do not violate Bell's inequality.

This hierarchy is important for practical applications. Nonlocal states enable the implementation of device-independent quantum key distribution protocols. States that exhibit only steering can be used in one-sided device-independent protocols. States that are only entangled are suitable for standard quantum protocols that require trust in the measuring devices.

Measures of entanglement

Negativity ($\mathcal{N}$) is one of the basic measures used to calculate entanglement. It is based on the Peres-Horodecki separability criterion [7, 8]. The negativity of a bipartite quantum state $\hat{\rho}$ can be defined as:

$$\mathcal{N}(\rho) = \frac{\|\rho^{\Gamma}\|_1 - 1}{2} \tag{2.4}$$

where ρ^{Γ} denotes a partially transposed density matrix and $\|X\|_1 = \mathrm{Tr}|X| = \mathrm{Tr}\sqrt{X^{\dagger}X}$ is the trace norm or the sum of the singular values of the operator X. This is equivalent to the sum of the absolute values of the negative eigenvalues of the partial transpose:

$$\mathcal{N}(\rho) = \sum_{\lambda_i<0} |\lambda_i|, \tag{2.5}$$

where λ_i are the eigenvalues of $\hat{\rho}^{\Gamma}$.

For practical calculations, it is often more convenient to express negativity through the negative eigenvalues of the partially transposed density matrix. According to the PPT criterion, all eigenvalues of the partial transpose are non-negative for separable states. Therefore, any negative eigenvalue signals entanglement. This leads to the following equivalent definition:

$$\mathcal{N} = \max\{0, -2\sum_{\lambda_i<0} \lambda_i\} = 2\mu, \tag{2.6}$$

where μ is the absolute value of the negative eigenvalue of $\hat{\rho}^{\Gamma}$. The negativity increases with the degree of entanglement, reaching $\mathcal{N} = 1/2$ for maximally entangled two-qubit states (and $(d-1)/2$ for maximally entangled $d \times d$ states).

Negativity properties:

- $\mathcal{N}(\hat{\rho}) = 0$ for all separable states,
- $\mathcal{N}(\hat{\rho}) = 1/2$ for maximally entangled two-qubit states (general bound: $\mathcal{N} \leq (d-1)/2$ for $d \times d$ systems),
- does not increase under LOCC (Local Operations and Classical Communication),
- is a convex measure: $\mathcal{N}\left(\sum_i p_i\hat{\rho}_i\right) \leq \sum_i p_i\mathcal{N}(\hat{\rho}_i)$.

The key advantage of negativity is that it is easy to calculate for low-dimensional states, and it is directly related to the PPT criterion. For $2 \otimes 2$ and $2 \otimes 3$ systems, the PPT criterion is necessary and sufficient for separability. In these cases, negativity enables the explicit classification of entanglement. Throughout this dissertation, the negativity $\mathcal{N}$ is

used as defined in Eq. (2.4), with the convention $\mathcal{N} = \sum_{\lambda_i<0} |\lambda_i(\hat{\rho}^\Gamma)|$, yielding $\mathcal{N} = 1/2$ for two-qubit Bell states and $\mathcal{N} \in [0, 1/2]$ for all two-qubit states.

EPR steering, unlike entanglement (symmetric property) and Bell nonlocality (requires violation of inequalities), is inherently asymmetric and can be quantified through local unitary invariants.

For two-qubit systems, steering can be characterized using the realignment matrix ρ^R and its invariants. The realignment operation rearranges density matrix elements as it was shown in Eq. (1.32). The degree of steering can be quantified through realignment moments:

$$R_k = \text{Tr}[(\rho^R)^k], \tag{2.7}$$

where k denotes the moment order. Note that this R_k notation is used exclusively in this subsection to describe realignment moments in the context of steering; in Section 2.4.5 the realignment spectral features are denoted Σ_k (singular-value moments) and G_k (eigenvalue moments) to avoid ambiguity. Specifically, $R_2 = \text{Tr}[(\hat{\rho}^R)^2]$ provides a measure of correlations that is sensitive to steering.

The realignment criterion states that for separable states:

$$\|\rho^R\|_1 \leq \sqrt{I_2^2 - I_2 + \frac{1}{d}}, \tag{2.8}$$

where $I_2 = \text{Tr}[\hat{\rho}^2]$ is the purity of the state, and d is the dimension of the total Hilbert space ($d = 4$ for two-qubit systems). When this inequality is violated, the state exhibits quantum correlations beyond those captured by entanglement measures alone [12, 13].

Bell's nonlocality measure ($\mathcal{B}$) provides more information than just whether or not a system violates the CHSH inequality; it quantifies the degree to which the inequality is violated. For a two-qubit state ρ, the $\mathcal{B}$ measure determines the maximum possible violation of the CHSH inequality by optimizing across all possible measurement settings. Mathematically, the CHSH value for specific measurement settings can be written as follows:

$$S = |\langle A_1 \otimes B_1\rangle + \langle A_1 \otimes B_2\rangle + \langle A_2 \otimes B_1\rangle - \langle A_2 \otimes B_2\rangle|, \tag{2.9}$$

where A_1, A_2 and B_1, B_2 are observables with eigenvalues ± 1 measured by Alice and Bob, respectively. The nonlocality measure $\mathcal{B}$ is defined as:

$$\mathcal{B} = \frac{\max S - 2}{2\sqrt{2} - 2}, \tag{2.10}$$

where the maximum is taken after all possible choices of observables. This normalization ensures that $\mathcal{B} = 0$ for states that do not violate the CHSH inequality ($S \leq 2$) and $\mathcal{B} = 1$ for states that achieve maximum quantum violation consistent with the Tsirelson bound ($S = 2\sqrt{2}$, see Eq. (2.3)).

Not all entangled states (with $\mathcal{N} > 0$) violate Bell's inequality ($\mathcal{B} > 0$). There is a region of *weak entanglement* where $\mathcal{N} > 0$ but $\mathcal{B} = 0$. The inequality $\mathcal{B} \leq \mathcal{N}$ is useful for selecting appropriate cryptographic protocols: it captures the limitations of device-independent protocols and characterises the relative strength of correlations as a resource in quantum information theory.

2.1.3 Traditional measurement methods and their limitations

The standard approach to this problem is quantum state tomography, which reconstructs the full density matrix from measurements in a tomographically complete set of bases [31–33]. For an n-qubit system, QST requires:

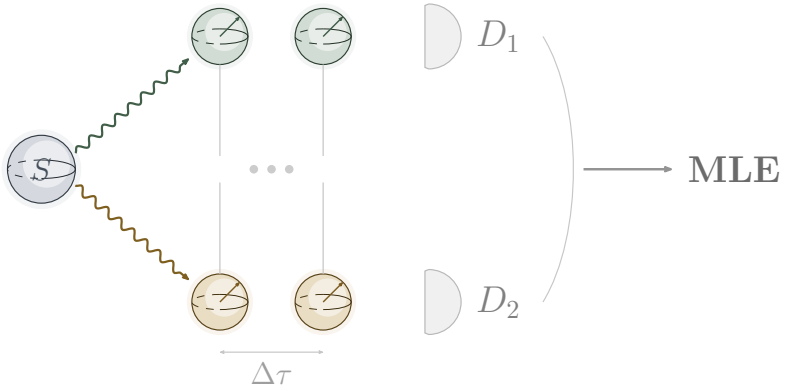


Fig. 2.2 **Schematic representation of the QST approach.** The symbol $\Delta\tau$ denotes the delay time between the generation of each entangled pair or a set of copies of entangled pairs. Detectors D_1 and D_2 measure products of eigenstates of Pauli's matrices. The detection rate is then processed using maximum likelihood estimation (MLE) to obtain the most probable physical combination of the measured experimental settings.

$$N^{\text{QST}}_{\text{measurements}} = \mathcal{O}(4^n), \tag{2.11}$$

independent measurement settings. This exponential complexity makes QST impractical even for systems with a few qubits. This is especially true in the context of NISQ devices [6],

which are characterized by high quantum gate noise, limited coherence time, and finite circuit depth, all of which contribute to error accumulation.

2.2 Resource-efficient quantum correlation measurement

This section presents a multicopy estimation (MCE) approach combined with artificial neural networks that reduces the measurement complexity from $\mathcal{O}(4^n)$, which is characteristic of full tomography, to $\mathcal{O}(2^n)$, while maintaining the ability to quantify both entanglement (via negativity) and nonlocality (via Bell inequality violation) [33].

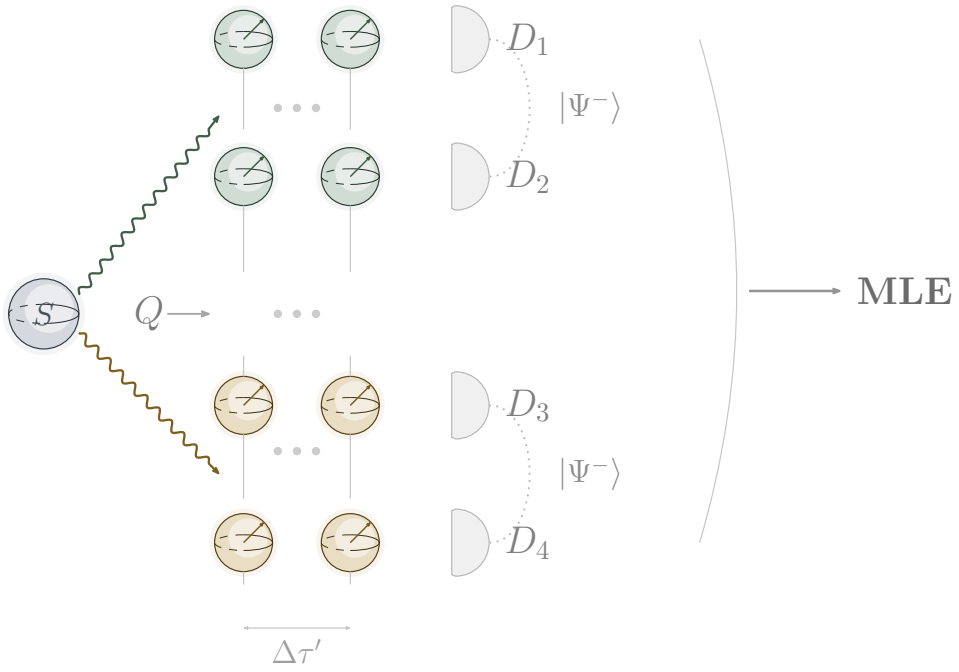


Fig. 2.3 **Schematic representation of the multicopy estimation (MCE) approach.** The symbol $\Delta\tau'$ denotes the delay time between the generation of each set of copies of entangled pairs. The multicopy state generation uses a quantum memory Q to store and release the collected pairs. The detection rates are then processed using maximum likelihood estimation (MLE) to obtain the most probable physical combination of the measured experimental settings.

2.2.1 Theoretical framework for multicopy estimation (MCE)

Makhlin invariants

An important observation in [34, 35] is that Makhlin invariants can be expressed as projections onto the singlet state, which are performed on multi-copies of a two-qubit system. Thus, we can express each invariant as a combination of measurements with results of ± 1. The singlet projection for two qubits in the k and l modes is defined as follows [35]:

$$\hat{P}^{-}_{k,l} = \frac{1}{4}(2\hat{\sigma}_0 \otimes \hat{\sigma}_0 - \hat{\sigma}_i \otimes \hat{\sigma}_i) = |\Psi^-\rangle\langle\Psi^-|, \tag{2.12}$$

where $|\Psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle)$ is a singlet state.

The multicopy measurement protocol uses three categories of projection configurations. Each configuration provides a coefficient that represents the probability of detecting a singlet state given a specific arrangement of qubit pairs.

(i) Local chain projections: **(**$c_1, ..., c_5$**)** - projections onto the singlet state are performed on each subsystem to provide information about local quantum properties. These projections have a chain-like structure, as shown in Fig. 2.4(d).

(ii) Local loop projections: **(**l_1, l_2**)** - these are similar to the chain projections but cover all pairs of qubits in a closed loop. This configuration preserves correlations between different copies of the same subsystem (see Fig. 2.4(a)).

(iii) Intersubsystem projections: **(**$\bar{c}_1, \bar{c}_2, \bar{c}_3, l_0, \bar{l}_1, \bar{l}_2$**)**, these are performed on qubits belonging to different subsystems (Alice and Bob) and reveal the nonlocal quantum properties of the studied state, as illustrated in Fig. 2.4(b) and 2.4(c).

As shown in [34], Makhlin invariants can be expressed as expectation values of measurements taken on multiple copies of a two-qubit system. Thus, we can express each invariant as

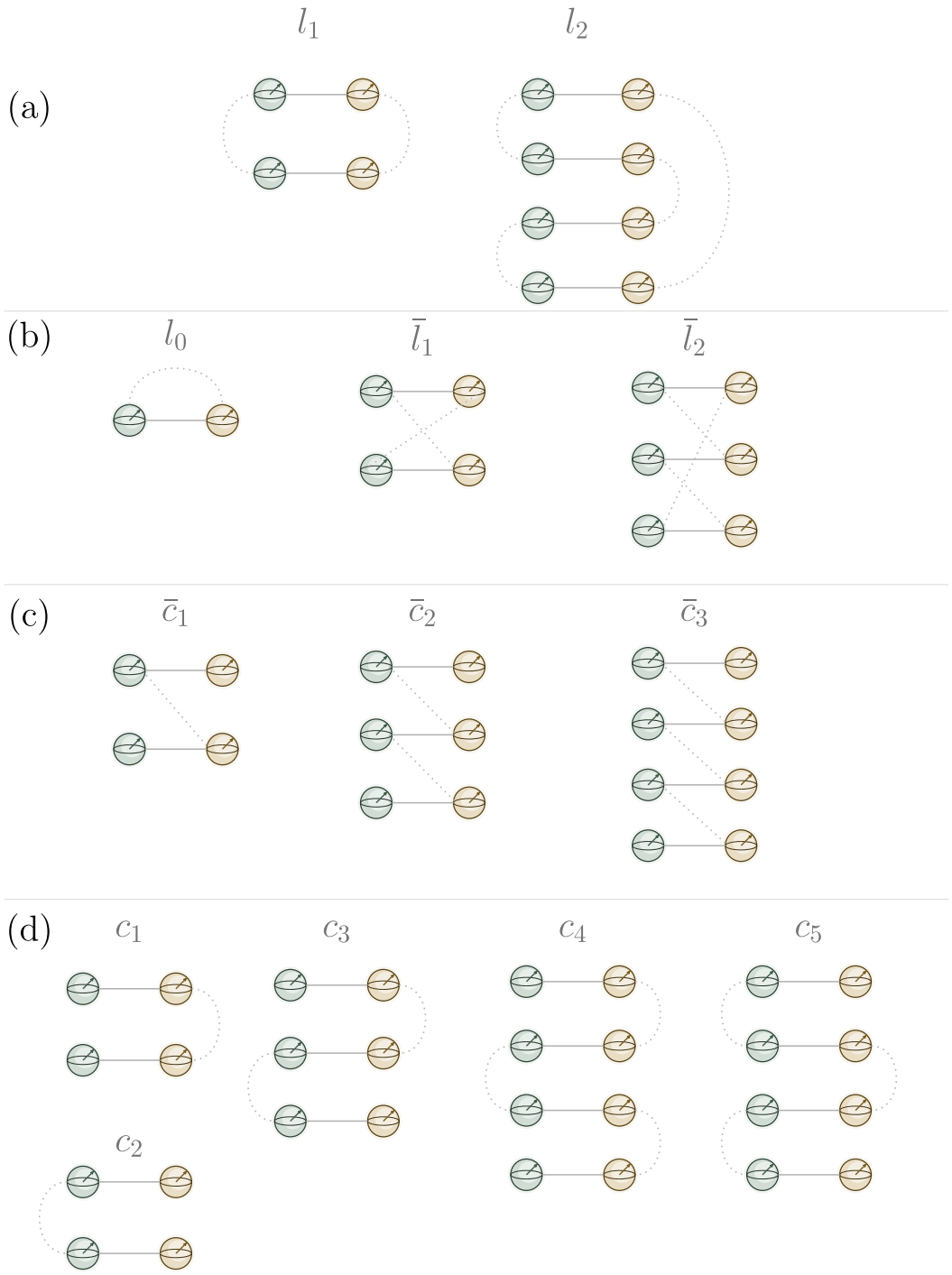


Fig. 2.4 **The graphs showing joint measurements of multi copies:** (a) the paired single–subsystem singlet projections l_1, l_2, (b) the paired cross–subsystem singlet projections $l_0, \bar{l}_1, \bar{l}_2$, (c) the chained cross–subsystem singlet projections $\bar{c}_1, \bar{c}_2, \bar{c}_3$. (d) The chained single–subsystem singlet projections $(c_1, ..., c_5)$. Black lines connect subsystems (green and yellow circles) of the same copy $\hat{\rho}$, while dotted lines correspond to projections of the multicopy system onto the singlet state. These graphical illustrations help to describe the concept of the different projection settings used in our approach to multicopy measurements.

a combination of singlet projection measurements:

$$
\begin{gathered}
I_1 = -\frac{8}{3}\left\{l_0\left[l_0\left(4l_0-3\right)+6\left(\bar{c}_1-2\bar{l}_1\right)\right]+3\bar{l}_1-6\bar{c}_2+8\bar{l}_2\right\}, \\
I_2 = 1+16l_1-4\left(c_2+c_1\right), \\
I_3 = 1+256\left(c_2^2+4c_3+c_1^2+l_2\right)-8\left(c_2+c_1\right), \\
I_4 = 1-4c_2, \\
I_5 = -4c_1+32c_3-64c_5+\left(1-4c_2\right)^2, \\
I_6 = 1-1024c_8-4\left(3c_2+2c_1\right)+16\left(3c_2^2+4c_3+2c_2c_1+c_1^2\right) \\
-64\left(c_2^3+4c_2c_3+2c_5+2c_3c_1+c_4\right), \\
I_7 = 1-4c_1, \\
I_8 = -4c_2+32c_3-64c_4+\left(1-4c_1\right)^2, \\
I_9 = 1-1024c_7-4\left(2c_2+3c_1\right)+16\left(c_2^2+4c_3+2c_2c_1+3c_1^2\right) \\
-64\left(2c_2c_3+c_5+4c_3c_1+c_1^3+2c_4\right)+256\left(c_3^2+2c_6+2c_1c_4\right), \\
I_{12} = 1+16c_3-4\left(c_2+c_1\right), \\
I_{13} = 1+256c_6-8\left(c_2+c_1\right)+16\left(c_2^2+3c_3+c_2c_1+c_1^2\right)-64\left[c_5+c_3\left(c_2+c_1\right)+c_4\right], \\
I_{14} = 16\left[l_0^2\left(1-4\bar{c}_1\right)+2l_0\left(4\bar{c}_2-\bar{c}_1\right)-\bar{l}_1+4\bar{c}_1\bar{l}_1+2\bar{c}_2-8\bar{c}_3\right]+256\left(c_3^2+2c_2c_5+2c_6\right).
\end{gathered}
\tag{2.13}
$$

Negativity and nonlocality expressed by Makhlin's invariants

We can derive alternative expressions for negativity and nonlocality measures using Makhlin invariants that are particularly useful in the context of multicopy measurements. For a two-qubit system, negativity is the positive solution to a fourth-degree polynomial equation:

$$
a_4N^4+a_3N^3+a_2N^2+a_1N+a_0=0, \tag{2.14}
$$

where the coefficients a_i are determined by combinations of singlet projections:

$$
\begin{gathered}
a_0 = -16[l_0^3+2\bar{l}_2+3(l_1^2-l_0^2\bar{c}_1-l_0\bar{l}_1+\bar{c}_1\bar{l}_1)-6(l_2-l_0\bar{c}_2+\bar{c}_3)], \\
a_1 = 24[l_0^2-\bar{l}_1-l_1+2(c_3-l_0\bar{c}_1+\bar{c}_2)]-32[l_0^3-3l_0\bar{l}_1+2\bar{l}_2], \\
a_2 = 12(c_2-2l_1+c_1), \\
a_3 = 6(1-\Pi_2), \\
a_4 = 3,
\end{gathered}
\tag{2.15}
$$

where $\Pi_n = \text{Tr}[(\rho^\Gamma)^n]$ is the n-th moment of the partially transposed density matrix. Specifically, $\Pi_2 = \text{Tr}[(\rho^\Gamma)^2]$ can be expressed in terms of singlet projections as:

$$\Pi_2 = 1 - 4(c_1 + c_2 - 2l_1) + 4[l_0^2 - \bar{l}_1 - l_1 + 2(c_3 - l_0\bar{c}_1 + \bar{c}_2)]. \tag{2.16}$$

The Bell nonlocality measure $\mathcal{B}$ can be expressed analogously through the Makhlin invariants I_2 and I_3 defined in Eq. (2.13) (not to be confused with the purity moments $\mathcal{I}_k = \text{Tr}[\rho^k]$ introduced in Section 2.3.1). It is determined by:

$$\mathcal{B} = I_2 - \min(r) - 1, \tag{2.17}$$

where r represents the roots of the characteristic equation:

$$r^3 - I_2 r^2 + \frac{1}{2}(I_2^2 - I_3)r + \frac{1}{6}[I_2^3 + (6I_1^2 - I_2^3)] = 0. \tag{2.18}$$

This representation enables a direct transition from measurable quantities (singlet projections) to nonlinear measures of entanglement and nonlocality. This reduces the computational and measurement complexity associated with traditional methods.

2.2.2 Multicopy state preparation and quantum memory considerations

A key aspect of implementing our protocol is precisely preparing multiple copies of the same quantum state. In real quantum processors, this process must take into account the device topology, the characteristic parameters of individual qubits, and gate error rates. We address this issue by averaging across different qubit mappings. The experimental implementation of multicopy measurements on currently available quantum platforms requires careful consideration of various implementation aspects. Ideally, we would prepare multiple identical copies of the quantum state and perform joint measurements covering all copies simultaneously. Unfortunately, current quantum processors lack sufficient quantum memory resources to enable such parallel manipulation without increasing the level of measurement errors.

To overcome this limitation, we employ a sequential approach, preparing each copy separately and using controlled operations to implement the equivalents of joint measurements. Specifically, we transform multicopy measurement operators into equivalent circuits that can be executed on the available hardware architecture. These circuits usually include preparing the target state, applying controlled operations that implement a particular projection configuration (like those shown in Fig. 2.4), and performing a final measurement in the proper

basis to extract the desired information. We then process the obtained measurement results to calculate the values of unitarily local invariants described by Eq. (2.13).

The fundamental circuit building blocks are shown in Fig. 2.5. Gate $U(k_a k_b)$ acts on the k-th entangled qubit pair: the upper qubit undergoes X, Z, a controlled-NOT, and H before measurement, while the lower qubit receives the CNOT target and is then measured; both outcomes are combined classically (cc) to yield the coalescence signal. Gate $U(k_a l_b | l_a k_b)$ acts on qubit pairs from two independent copies $\hat{\rho}_{ab}^{(k)}$ and $\hat{\rho}_{ab}^{(l)}$, applying the same Bell-basis rotation to each subsystem independently before joint classical postprocessing. The full six-copy

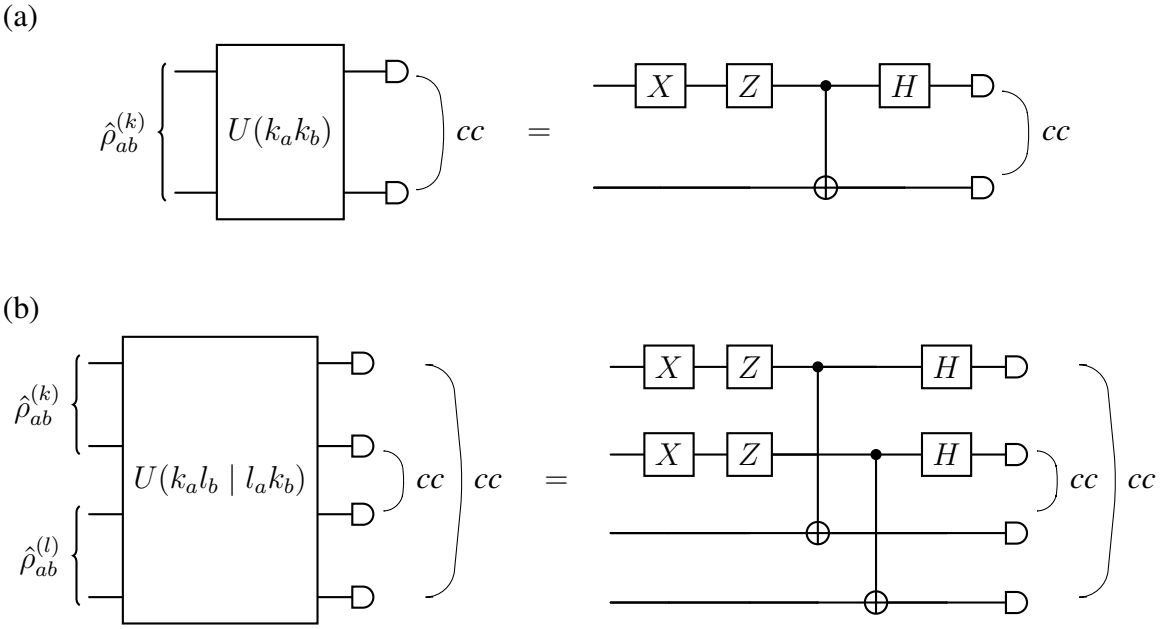
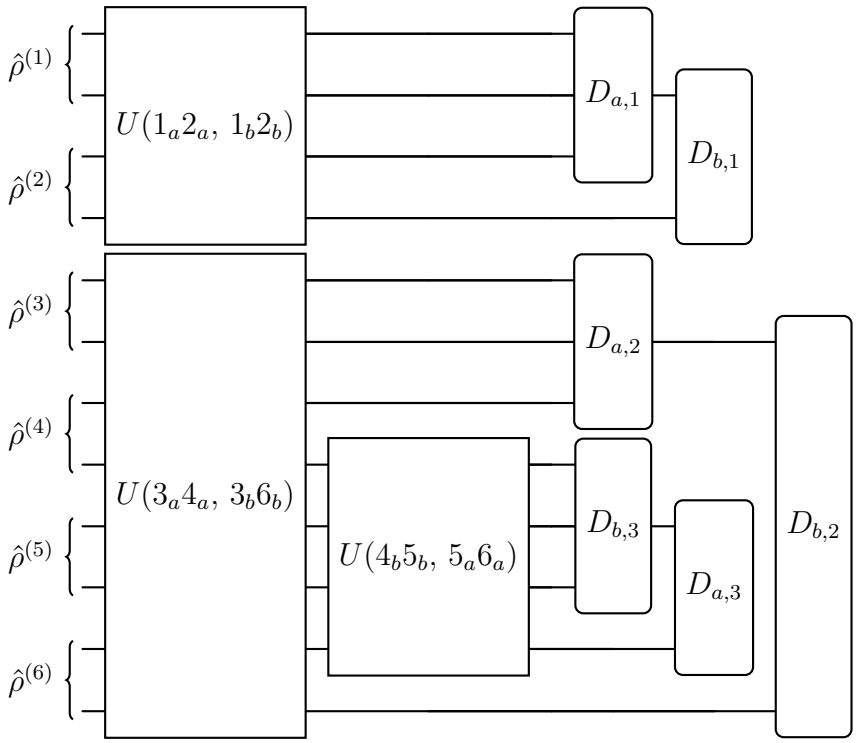


Fig. 2.5 **Fundamental circuit blocks for singlet projections.** (a) Gate $U(k_a k_b)$ acting on the k-th entangled qubit pair. (b) Gate $U(k_a l_b \mid l_a k_b)$ acting on two copies of qubit pairs from subsystems a and b. These blocks form the building elements for all projection measurements.

projection circuits assembled from these blocks are shown in Fig. 2.6: panel (a) implements the single-subsystem and chained projections (l_1, l_2, c_1–c_5, $\bar{c}_3$) using three sequential unitary stages, while panel (b) implements the intersubsystem projections (l_0, $\bar{c}_1$, $\bar{c}_2$, $\bar{l}_1$, $\bar{l}_2$) through a different coupling topology. In both cases, detector pairs $D_{a,n}$ and $D_{b,n}$ record coalescence and anticoalescence counts whose ratio yields the projection coefficient.

Table 2.1 shows the various detector settings used to measure the different correlation terms in Fig. 2.6. The rows correspond to different combinations of coalescence (a) and summed signal (s) across detector pairs $D_{a,n}$ and $D_{b,n}$. Projecting these results onto l_i and c_i, or their barred counterparts, efficiently acquires properties such as singlet projections

(a)



(b)

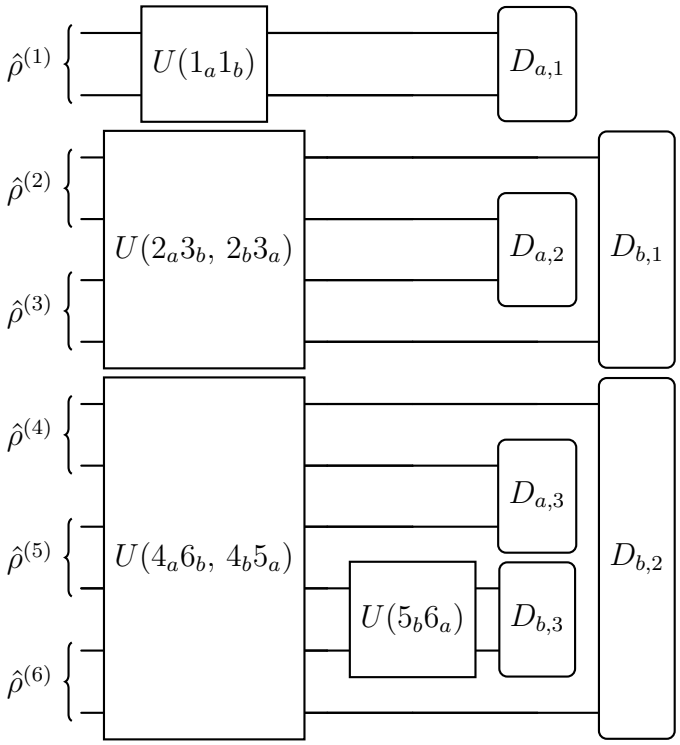


Fig. 2.6 **Implementation of singlet projection measurements.** (a) Circuit configuration for measuring l_1, l_2, c_1, c_2, c_3, $\bar{c}_3$, c_4, c_5. (b) Circuit configuration for measuring l_0, $\bar{c}_1$, $\bar{c}_2$, $\bar{l}_1$, $\bar{l}_2$. Detector pairs $D_{a,n}$ and $D_{b,n}$ ($n = 1, 2, 3$) measure coalescence and anticoalescence; gates $U(k_ak_b)$ and $U(k_al_b|l_ak_b)$ are defined in Fig. 2.5.

and correlation coefficients. This systematic labeling makes results more reproducible and transparent.

Table 2.1 The detector configurations and corresponding measurements are shown for the circuits in Fig. 2.6. Here a represents anticoalescence and s represents the sum signal.

$D_{a,1}$	$D_{a,2}$	$D_{a,3}$	$D_{b,1}$	$D_{b,2}$	$D_{b,3}$	(a)	(b)
a	s	s	a	s	s	l_1	–
s	a	a	s	a	a	l_2	–
s	s	s	s	s	a	c_1	–
s	s	a	s	s	s	c_2	–
s	a	s	s	a	s	c_3	–
s	s	a	s	a	a	$\bar{c}_3$	–
s	s	s	a	a	a	c_4	–
s	a	s	s	a	a	c_5	–
a	s	s	s	s	s	–	l_0
s	s	s	a	s	s	–	$\bar{c}_1$
s	s	s	s	a	a	–	$\bar{c}_2$
s	a	s	a	s	s	–	$\bar{l}_1$
s	s	a	s	a	a	–	$\bar{l}_2$

The copy fidelity verification and qubit mapping strategy are described in Section 2.2.5.

2.2.3 Maximum Likelihood Estimation

Due to the presence of experimental noise in quantum measurements, appropriate error correction methods are required to ensure that the obtained results are consistent with physical conditions. To this end, we use the maximum likelihood estimation (MLE) method, which enables us to determine a state's parameters while maintaining its physicality. In QST, the MLE method is commonly used to solve similar problems [36]. However, optimization in the context of multicopy measurements is more complex than in standard quantum state tomography. In the classical QST approach, the likelihood function depends quadratically on the density matrix parameters, simplifying the optimization procedure. For multicopy observables, however, the likelihood function is a higher-order polynomial with respect to the elements of the single-copy density matrix, significantly complicating the numerical optimization process. The MLE approach we used for multiple copy measurements can be written in the following form [36]:

$$\rho_{\text{MLE}} = \text{argmax}_{\rho \geq 0} \sum_i n_i \log(p_i(\rho)) + \lambda \text{Tr}(\rho^2), \tag{2.19}$$

where n_i represents measurement counts, $p_i(\rho)$ are theoretical probabilities, and the second term enforces purity constraints. Optimization is performed while maintaining the following physical conditions:

$$\mathrm{Tr}(\rho) = 1 \quad \text{(normalization)}, \tag{2.20a}$$

$$\rho \geq 0 \quad \text{(positive semidefiniteness)}, \tag{2.20b}$$

$$\rho = \rho^\dagger \quad \text{(Hermitian)}. \tag{2.20c}$$

The computational efficiency of the procedure can be improved by taking advantage of the fact that the measured quantities are invariant under local unitary transformations. This property enables the reduction of the number of free parameters in the density matrix model by eliminating the degrees of freedom associated with phase factors. Consequently, the optimization problem can be formulated in the space of real matrices, significantly improving the convergence of the maximum likelihood estimation (MLE) algorithm.

Analysis of the experimental results confirms that the MLE method effectively reduces the impact of measurement noise, resulting in the reconstruction of states that satisfy physical requirements. This enables the practical implementation of multicopy measurements on modern quantum processors.

2.2.4 Artificial neural networks for quantum correlation measurement

Integrating artificial neural networks and SHAP analysis into our measurement protocol stems from specific requirements related to multicopy experiments. First, data from NISQ devices are characterized by a significant amount of measurement noise. Neural networks excel at recognizing patterns in error-prone data, which enables effective signal-to-noise separation. Our results show a substantial decrease in the influence of statistical and systematic errors compared to direct analytical calculations, where measurement uncertainties amplify through complex nonlinear expressions. Second, neural networks naturally approximate nonlinear relationships between singlet projection results and quantum correlation measures, eliminating the need for direct mathematical modeling and avoiding numerical instabilities inherent in polynomial root-finding algorithms. Third, implementing SHAP analysis enables the identification of measurement configurations that provide the most information and the optimization of the entire experimental protocol, reducing experimental overhead by 67% compared to quantum state tomography. Finally, the network architecture enables readaptation to changing hardware error characteristics through retraining without modifying fundamental theoretical assumptions. This adaptability to specific hardware characteristics makes the approach particularly suitable for diverse quantum computing platforms. Thanks to

neural networks, the method of measuring multiple copies, which is theoretically elegant but practically demanding, becomes an effective tool for quantifying quantum correlations [37].

The increased noise immunity of our approach based on artificial neural networks (ANNs), compared to direct analytical calculations, stems from several key factors. First, the neural network is trained on 5×10^5 randomly generated two-qubit states, learning robust mappings between projection results and quantum correlation measures. Training on a diverse set of states allows for better generalization under experimental noise conditions than directly evaluating nonlinear expressions, which can be numerically unstable when projection coefficients differ from ideal values. This extensive training set covers diverse noise conditions encountered in real quantum hardware, enabling the network to learn smooth mappings less sensitive to input perturbations than direct analytical methods.

SHAP analysis reveals that the network automatically identifies the five most informative measurements that have the greatest impact on predicting negativity and Bell nonlocality measures. Using only these five measurements instead of all twelve reduces total experimental effort and potentially increases noise resilience by reducing the number of required circuits. Unlike in analytical calculations, where measurement errors propagate directly through complex polynomial expressions, neural networks learn smooth nonlinear mappings that are less sensitive to small input value perturbations. A network architecture with ReLU activations and $L2$ regularization ($\lambda = 10^{-5}$) provides robustness to input noise through learned representations. Our experimental results demonstrate that this approach achieves higher accuracy under realistic NISQ device noise conditions.

SHAP analysis for optimal measurement selection

A key innovation of our approach is the systematic identification of optimal measurement configurations via SHAP (SHapley Additive exPlanations) analysis. This procedure has revealed that only five key projections are necessary for accurately quantifying quantum correlations, significantly reducing the required experimental effort. In our study, SHAP analysis quantifies the impact of each measurement configuration on the value of quantum correlations. The SHAP value of a feature (measurement configuration) j is calculated as [38]:

$$\phi_j = \sum_{S \subseteq F \setminus \{j\}} \frac{|S|!(|F| - |S| - 1)!}{|F|!} [f_S(\{j\}) - f_S(\emptyset)], \tag{2.21}$$

where ϕ_j represents the SHAP value of feature j, quantifying its contribution to the model's prediction. The symbol F denotes the set of all measurement configurations, while f_S is the model's prediction using the subset S of measurements, and $f_S(\emptyset)$ represents the prediction without configuration j.

Intuitively, the SHAP procedure determines the average marginal contribution of a given measurement configuration, considering all possible combinations of remaining configurations. Figure 2.7 presents SHAP values for negativity and Bell's nonlocality measure. It clearly indicates which measurements have the greatest impact on prediction accuracy. For negativity ($\mathcal{N}$), the SHAP analysis determined the following key measurements: $\{l_1, \bar{c}_3, \bar{l}_2, \bar{l}_1, c_3\}$, while for nonlocality ($\mathcal{B}$) we obtained: $\{l_1, c_2, c_1, c_3, c_5\}$.

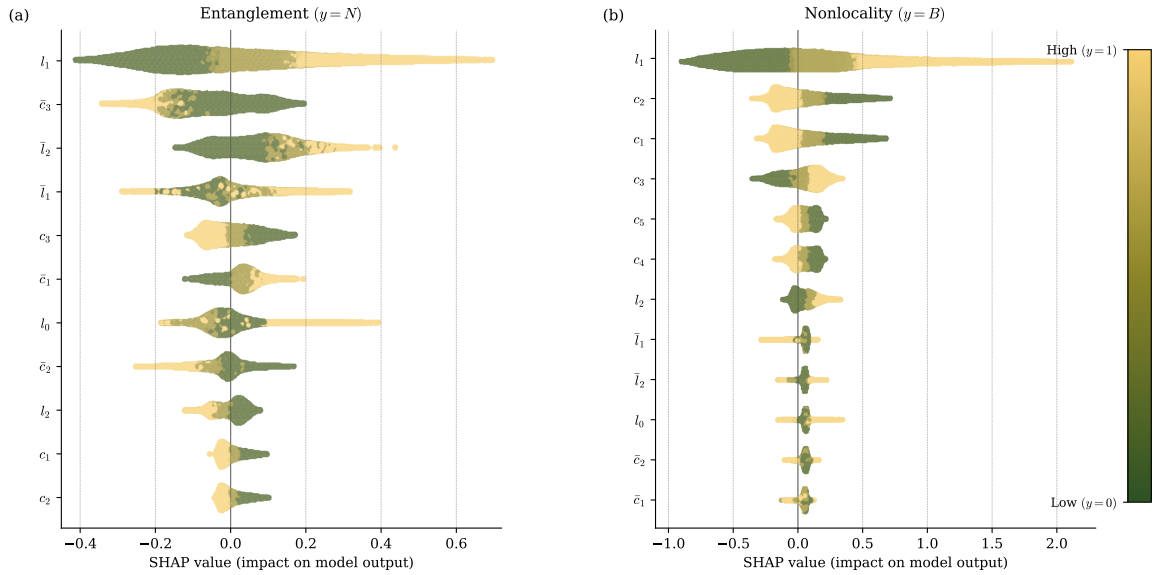


Fig. 2.7 SHAP value analysis for quantum correlation measurements. Results shown for (a) negativity $\mathcal{N}$ and (b) nonlocality measure $\mathcal{B}$, computed from 5×10^5 random input states. The impact strength indicates each projection's contribution to the final measurement outcome.

The different SHAP patterns for negativity and Bell nonlocality reflect these measures' different physical requirements: entanglement detection primarily relies on preserving local coherence (hence the dominance of l_1) combined with correlations between subsystems ($\bar{l}_2, \bar{l}_1$), while Bell nonlocality depends on the local statistics of measurements used to evaluate the CHSH inequality. This exposes chain projections (c_1, c_2, c_4, c_5) that capture the required measurement correlations.

SHAP analysis revealed that only five measurements are necessary to estimate both entanglement and Bell nonlocality. This represents a significant reduction in measurement requirements, as only 15 measurements are required for full QST and 12 for complete characterization of invariants, representing a 67% reduction in measurement requirements.

Neural Network architecture and training

Based on SHAP analysis, we designed a neural network architecture with the following characteristics: The input layer consists of five-dimensional vectors constructed from experimental measurement results for specific quantum states and projections. Each input node represents a specific configuration of data obtained from singlet projection measurements. The five hidden layers each contain nine neurons with ReLU activation functions. This architecture effectively detects nonlinear relationships in the data without the need for extensive hyperparameter optimization. Depending on the variant, the output layer generates negativity or a Bell nonlocality measure using a linear activation function to produce continuous predictions across the entire range of possible quantum correlations.

The following parameter space was explored in the hyperparameter optimization process:

(i) **Network Architecture**:

- Hidden layer sizes: $(9, 9, 9, 9, 9)$ for the negativity and Bell's nonlocality measure predictions;
- Number of layers: 5 fully connected layers;
- Activation functions: ReLU;

(ii) **Training Parameters**:

- Learning rate: $\alpha = 10^{-5}$;
- Optimization solver: adam;
- Maximum iterations: 2000;
- Batch size: determined internally by the solver;
- Warm start: `True`;

(iii) **Regularization strength**:

- $\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{pred}} + \lambda_1 \|\mathbf{w}\|_2^2 + \lambda_2 \|\nabla \mathcal{L}\|_2^2$;
- $\lambda_1 = 10^{-5}$ is the $L2$ regularization parameter;
- $\lambda_2 = 0$ (no additional gradient–based penalty or dropout layer);
- $\|\mathbf{w}\|_2^2$ is the squared $L2$ norm of the weight parameters.

The data was divided into training and test sets (75% and 25%, respectively). The quality of the model was evaluated using the built-in `MLPRegressor` scoring mechanism. A separate validation set was not created; instead, performance on the test set was monitored after the

model stabilized. Due to the large size of the dataset, which provided sufficient coverage of the measurement results space, no data augmentation techniques were employed. The model's fit was assessed using the coefficient of determination (R^2), and the training procedure used the default mean-squared error (MSE) objective function in `MLPRegressor` [39]. Specifically:

$$\mathcal{L} = \|y_{\text{pred}} - y_{\text{true}}\|^2 + \lambda_1 \|\mathbf{w}\|_2^2, \tag{2.22}$$

where $\lambda_1 = 10^{-5}$ provides $L2$ regularization to mitigate overfitting. An appropriate subset of features and labels for negativity and nonlocality measures was used to train each model independently.

The experimental results demonstrate that the neural network–based approach significantly improves the precision and stability of quantum correlation measurements in the presence of noise compared with direct calculations using analytical expressions. A neural network's ability to correct systematic errors in the measurement process is particularly valuable in the context of NISQ processors.

The superior performance stems from the network learning smooth, nonlinear mappings that are less sensitive to small input perturbations than direct analytical inversion. In the latter, measurement errors propagate through complex polynomial expressions. The ReLU architecture with $L2$ regularization further suppresses high-frequency noise through learned representations.

2.2.5 Experimental validation on NISQ devices

Werner and Horodecki states as testbeds

We evaluated our approach on two families of two-qubit mixed states. Werner and Horodecki states are useful for verifying the accuracy of measuring the entanglement of two qubits. This is because their entanglement and nonlocality properties are well-defined and change systematically with the mixing parameter p.

Werner states can be prepared from Bell states that have undergone a depolarization process:

$$\rho_W(p) = p|\Psi^-\rangle\langle\Psi^-| + (1-p)\frac{\mathbb{I}}{4}. \tag{2.23}$$

Horodecki states correspond to a Bell state that has undergone an amplitude-damping channel:

$$\rho_H(p) = p|\Psi^-\rangle\langle\Psi^-| + (1-p)|00\rangle\langle 00|. \tag{2.24}$$

Here $|\Psi^-\rangle$ is the singlet state and I the identity matrix.

Practical circuit implementation

We implemented our measurement protocol using quantum circuits and controlled interference operations on real quantum processors. On average, each singlet projection measurement requires 12 gates. To reliably prepare multiple copies of the states, we verified the similarity of the quantum state copies by estimating their mutual fidelity overlap:

$$F_{\text{copy}} = \text{Tr}(\rho_1 \rho_2) \geq 1 - \epsilon, \tag{2.25}$$

where ϵ is the accepted fidelity threshold. For the *ibm_hanoi* processor, we achieved $\epsilon \leq 0.02$, corresponding to an effective copy fidelity exceeding 98%. Given the heterogeneous error characteristics of different physical qubit configurations:

$$M_i = \{(q_1, q_2, q_3, q_4) | (q_j, q_k) \in E \text{ for required connections}\}, \tag{2.26}$$

where E corresponds to the set of available connections in the processor, we averaged measurement results across multiple qubit mappings with weights determined by their respective error characteristics:

$$w_i = \frac{\exp\left(-\sum_j \epsilon_j^{(i)}\right)}{\sum_k \exp\left(-\sum_j \epsilon_j^{(k)}\right)}, \tag{2.27}$$

where $\epsilon_j^{(i)}$ denotes individual error coefficients for mapping configuration i. The final measurement result is calculated as:

$$\langle O \rangle = \sum_i w_i \langle O \rangle_i. \tag{2.28}$$

To ensure reliable measurements in the presence of noise, we employed several error mitigation techniques. First, we performed zero-noise extrapolation, in which each quantum circuit was run at multiple controlled noise levels. Systematic noise scaling enabled us to extrapolate the results to the theoretical limit of zero noise, thereby reducing the impact of coherent errors.

The second technique focused on mitigating measurement errors. We characterized the measurement processes using calibration circuits, which enabled us to determine a confusion matrix for each qubit. Based on this matrix, we applied Bayesian correction, taking into account the known probabilities of misreading the base states.

For selected measurements, post-selection based on quantum state purity was additionally applied. We verified the purity of the prepared state copies by measuring the second moment

$\mu_2 = \mathrm{Tr}[(\rho^{T_A})^2]$. We rejected cases in which the measured purity deviated by more than $\epsilon = 0.02$ from the theoretical value. Although this approach reduced the measurement statistics, it significantly improved the quality of the retained data.

Combining these three error mitigation techniques enabled us to obtain reliable quantum correlation results, even in cases of high noise levels where traditional unmitigated methods would lead to significant systematic errors.

IBM Quantum hardware platform

We conducted our experiments using the IBM Quantum platform [40] and Qiskit [41], which provides access to superconducting quantum processors in the cloud. IBM Quantum devices are based on fixed-frequency transmon qubits [42]. These are superconducting circuits that operate at temperatures around millikelvins ($\sim$ 15 mK). Transmons are a variant of the Cooper-pair box qubit design that achieves reduced sensitivity to charge noise by operating in the high charge energy range. The qubits are coupled using fixed-frequency microwave resonators, which enable two-qubit gate operations through controlled interactions.

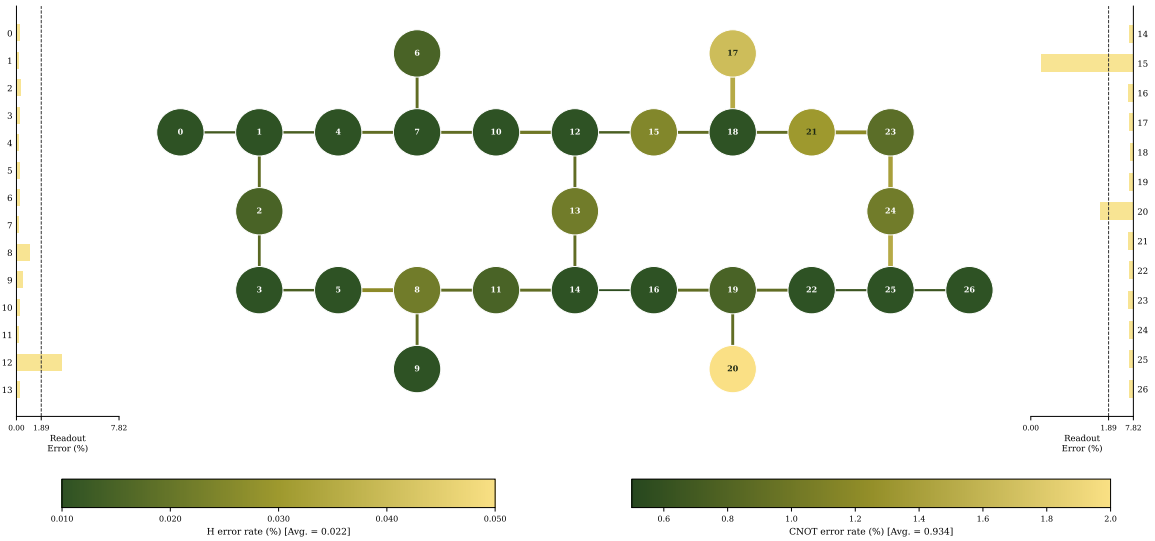


Fig. 2.8 **Error characterization of the *ibm_hanoi* processor.** Calibration data shown for the relevant qubit subset 1-5, 7, 10, 12-14, 16, 19 used in our experiments. Error rates and coherence times directly impact measurement fidelity and inform our qubit mapping optimization strategy.

We specifically used the *ibm_hanoi* processor for the results presented. It is a 27-qubit device with a heavy-hexagonal network topology that ensures connectivity with the nearest neighbor. During our experiments, the processor exhibited the following performance characteristics: quantum volume of 64, average CNOT error rate of 9.34×10^{-2}, average

single-qubit readout error of 1.89×10^{-2}, and T_2 dephasing times of approximately 100 μs. These specifications are representative of current NISQ-era devices, in which gate fidelity remains below error tolerance thresholds, and decoherence timescales limit circuit depth.

The distribution of errors in *ibm_hanoi*, with error rates varying by up to a factor of three across different qubits and two-qubit gates, required the use of various optimization strategies for qubit mapping, as described in the previous section. Figure 2.8 shows detailed calibration data for the subset of qubits (1-5, 7, 10, 12-14, 16, and 19) used in our measurements. These qubits were selected based on the connectivity patterns required for implementing singlet projections and their favorable error characteristics relative to the device average.

Experimental results

Figure 2.9 presents experimental results for negativity $\mathcal{N}$ and Bell nonlocality measure $\mathcal{B}$ as functions of the mixing parameter p, comparing three measurement approaches: panels (a) and (b) show quantum state tomography, panels (c) and (d) show multicopy estimation with all 12 projections, and Fig. 2.10 shows the ANN approach with only 5 optimized measurements.

Results are shown for both ideal simulations (● Werner, ● Horodecki on `qasm_simulator`) and hardware data ($+$ Werner, $\times$ Horodecki on *ibm_hanoi*). Standard deviations were estimated by simulating 10^5 experiments with noise models based on processor calibration data.

Shot noise analysis

Figure 2.11 presents the standard deviations of negativity $\mathcal{N}$ and nonlocality measure $\mathcal{B}$ as functions of shot count. The analysis compares QST (panels (a),(b)) with MCE (panels (c),(d)) for both Werner and Horodecki states.

As expected, statistical fluctuations decrease with increasing shot number. However, the rate of improvement depends critically on the measurement protocol. The MCE approach can potentially achieve similar or improved accuracy with fewer shots compared to QST, because it employs a smaller set of more informative measurements. This translates directly to reduced experimental time and computational resources.

For both state families the standard deviation scales as $\sigma \propto 1/\sqrt{N_{\text{shots}}}$, with MCE requiring fewer total shots than QST to achieve equivalent precision, and the ANN approach reducing shot requirements further through learned noise suppression.

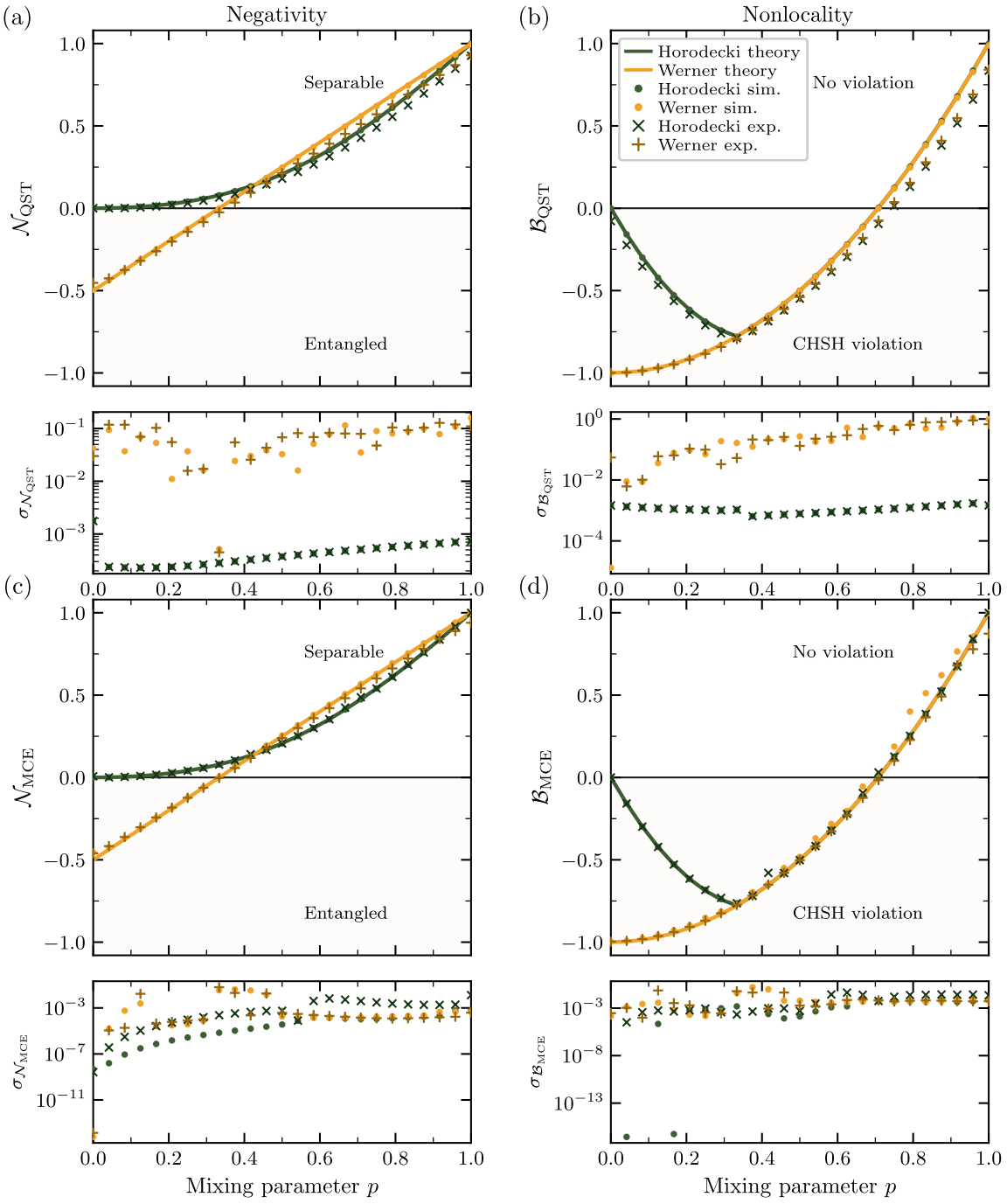


Fig. 2.9 **Experimental quantification of entanglement (left column) and nonlocality (right column) for Werner and Horodecki states.** Negativity $\mathcal{N}$ and Bell nonlocality measure $\mathcal{B}$ as functions of mixing parameter p. Solid curves: theoretical predictions for ideal states. Results obtained using (a),(b) quantum state tomography and (c),(d) multicopy estimation with all 12 projections. Symbols: ● Werner states, ● Horodecki states (ideal simulation on `qasm_simulator`); + Werner states, × Horodecki states (hardware data from *ibm_hanoi*). Standard deviations σ estimated by simulating 10^5 experiments with noise models from calibration data.

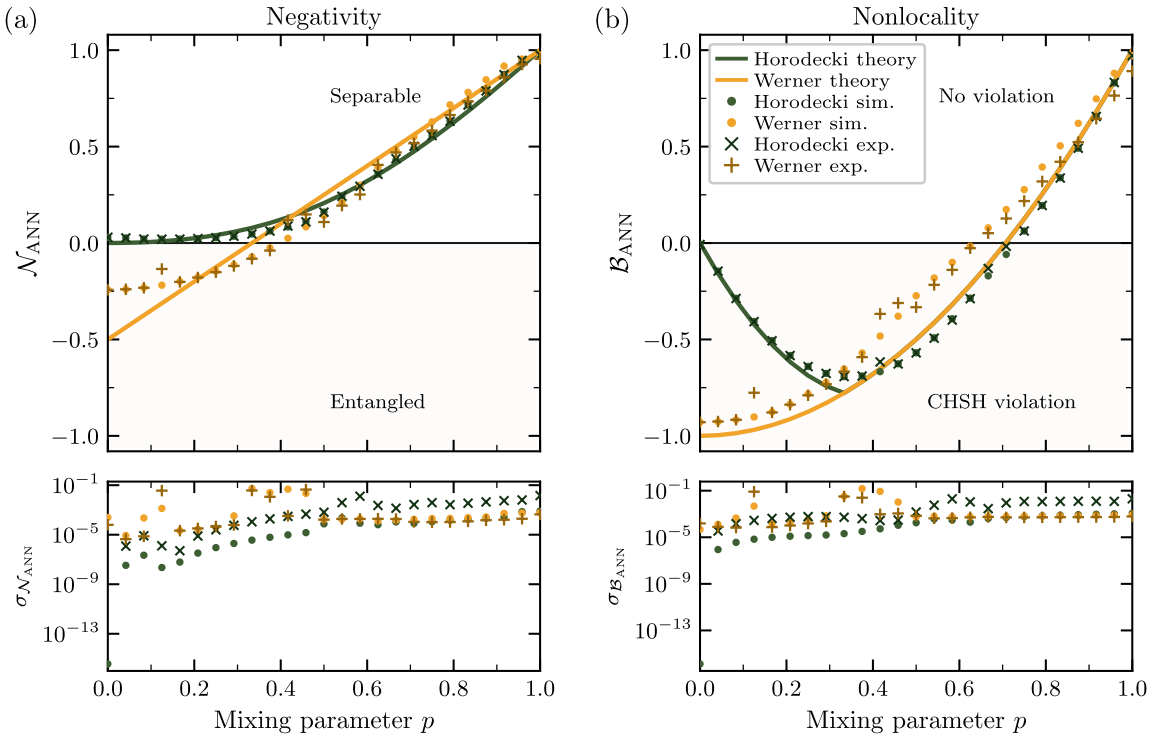


Fig. 2.10 Same as Fig. 2.9, but for the ANN trained on random input states using only the five most significant projections identified by SHAP analysis.

Comparison with Randomized Measurements

Recent research results have shown that random measurements can effectively estimate negativity using only single-copy operations [43, 44]. However, achieving an accuracy of ϵ in random measurement protocols requires $\mathcal{O}(1/\epsilon^2)$ measurements, which comes with a large upfront cost. Scaling the system size is beneficial, even with just a few measurement samples.

In comparison to existing methods [43, 44], our multi-copy neural network approach requires only five optimized measurement configurations and provides good precision with moderate sample sizes. While this approach necessitates the reliable preparation of multiple copies, it offers a potentially more scalable solution. According to our experimental results, our method outperforms random measurements in terms of negativity estimation accuracy when the total number of measurements is equivalent. This advantage is particularly noticeable in the presence of noise because neural network processing increases noise resilience.

Resource requirements

For a two-qubit system, QST requires 15 different measurement settings while MCE with ANN requires only 5 optimized measurements, which is a 67% reduction. Although the average

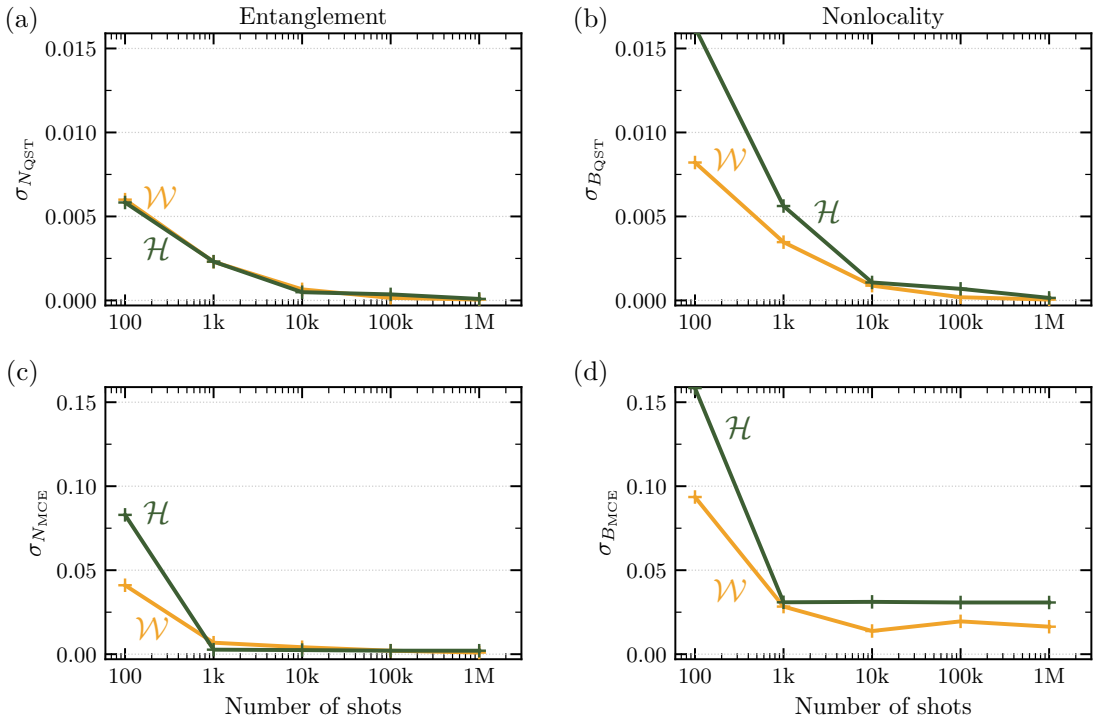


Fig. 2.11 **Measurement precision analysis.** Standard deviations of (a) and (c) negativity $\mathcal{N}$ and (b) and (d) nonlocality measure $\mathcal{B}$ as functions of shot count. The results are shown for the (a),(b) QST and (c) and (d) MCE methods, applied to Werner ($\mathcal{W}$) and Horodecki ($\mathcal{H}$) states.

circuit depth is 5 gates per QST measurement versus 12 for MCE, the smaller configuration count more than compensates: $15 \times 5 = 75$ total gates for QST against $5 \times 12 = 60$ for MCE, a 20% total reduction in gate overhead.

Moreover, in terms of noise resilience, our protocol maintains approximately twice the fidelity of standard QST at equivalent noise levels:

$$F_{\mathrm{QST}}(\epsilon) = F_0 e^{-2\alpha\gamma\epsilon} \quad \text{vs} \quad F_{\mathrm{ANN}}(\epsilon) = F_0 e^{-\alpha\gamma\epsilon}. \tag{2.29}$$

Here ϵ denotes noise strength, γ the characteristic decay rate of measurement fidelity with noise, and F_0 the zero-noise baseline; both α and γ are fitted from the shot-noise analysis shown in Fig. 2.11. The factor-of-two difference in the exponent reflects the reduced error-propagation sensitivity of the learned nonlinear mapping relative to direct polynomial inversion.

Noise resilience analysis

We model the effect of noise on negativity estimation as

$$N_{\text{measured}}(\epsilon) = N_{\text{true}} \cdot e^{-\beta\gamma\epsilon} + \delta(\epsilon), \tag{2.30}$$

where ϵ represents noise strength, γ is the decay coefficient, and $\delta(\epsilon)$ accounts for systematic bias. The MCE approach demonstrates reduced noise sensitivity relative to standard QST for three reinforcing reasons. First, selecting the most informative measurement configurations limits noise accumulation to the channels that carry the most signal. Second, operating in canonical form means only parameters relevant to nonlocal features are tracked, so calibration errors do not compound across redundant degrees of freedom. Third, the trained neural network recognises typical noise patterns and partially compensates for them without any explicit noise model.

Quantitatively, the expected estimation error scales as

$$E_{\text{QST}} \propto \sqrt{d^4}\,\epsilon \quad \text{vs} \quad E_{\text{MCE}} \propto (d^2-1)\,\epsilon, \tag{2.31}$$

where d is the system dimension and ϵ is the single-measurement error rate. The advantage of our approach therefore grows with system size, becoming particularly pronounced for multi-qubit systems where QST is prohibitively expensive.

Scalability analysis

For a general two-qudit system of dimension d, the density matrix contains $1 + 2(d^2-1) + (d^2-1)^2$ real parameters in the most general Fano expansion, but local unitary equivalence reduces the correlation tensor to d^2-1 independent elements, giving $1+3(d^2-1)$ parameters in canonical form. QST requires $\mathcal{O}(d^4)$ measurements for such systems, equivalent to $\mathcal{O}(4^n)$ for n qubits. MCE requires only $\mathcal{O}(d^2-1)$ configurations for full invariant characterisation, reducing to $\mathcal{O}(2^n)$. Both methods are exponential, but the exponent base is halved, which is decisive for near-term devices of moderate size. For multipartite systems, QST scales as $\mathcal{O}(d^{2k})$ for k qudits, while a hierarchical MCE extension treating pairwise correlations would scale as $\mathcal{O}(k^2(d^2-1)^2)$.

Potential extensions to more complex quantum states

While the experimental demonstration focuses on Werner and Horodecki states, the method extends naturally to a broader family of targets. For general two-qubit states the local unitary invariants measured form a complete basis for all nonlocal properties, so performance is not

limited to states with special symmetry. The same invariant structure extends directly to qudit systems: higher dimension increases the number of configurations, but the fundamental multicopy approach is unchanged. For multipartite systems, a hierarchical strategy detects pairwise entanglement first and then higher-order correlations, retaining the exponential advantage over QST. Finally, the neural network component can be retrained to distinguish between different noise characteristics, extending noise resilience beyond depolarizing noise and amplitude damping without modifying the measurement circuits.

2.2.6 Conclusions and outlook

The multicopy estimation approach combined with neural network processing achieves superior performance over standard quantum state tomography: measurement settings are reduced from 15 to 5 (a 67% reduction), and noise robustness is enhanced through maximum likelihood estimation and learned nonlinear mappings. The scaling advantage reduces measurement requirements from $\mathcal{O}(4^n)$ to $\mathcal{O}(2^n)$, providing decisive savings for near-term devices of moderate size. The experimental demonstration on IBM quantum hardware confirms that multicopy measurements can be reliably implemented in practice, making this a viable tool for entanglement quantification on NISQ processors.

Several directions remain open. Circuit constructions for systems larger than two qubits incur increased depth and error rates; more efficient decompositions and hardware-specific optimisations are needed before the approach scales beyond the two-qubit case demonstrated here. Preparing multiple identical copies is challenging at scale, and adaptive protocols that reduce the required number of copies are a natural next step. Quantifying other nonlinear state properties, such as entropy measures, coherence quantifiers, and multipartite correlations, requires only retraining of the neural network component and no modification to the underlying measurement circuits. This makes these properties straightforward targets for future work.

2.3 Algebraic classification of quantum states: Gröbner basis decision trees

The multicopy estimation method of the preceding section provides efficient access to partial-transpose moments μ_k. The next question is how to convert those moments into a negativity value with the minimum number of measured configurations. The answer depends on the eigenvalue structure of ρ^{T_A}: states with degenerate eigenvalue spectra satisfy polynomial constraints among their moments, and these constraints select simplified negativity formulas that bypass the full quartic eigenvalue problem. Gröbner basis theory

formalizes this observation into an algebraic decision tree that automatically identifies the correct computational strategy from partially observed moment data. This reduces the average measurement overhead to one to four configurations for two-qubit systems and one to five for qubit-qutrit systems, compared to 16 and 36 for full tomography, respectively.

Beyond computational efficiency, the algebraic structure has a deeper role: the same moment differences $C_k = \mu_k - \mathcal{I}_k$ that define the Gröbner conditions turn out to encode scalar spin chirality operators across state copies, as will be shown in Section 2.4. The current section establishes the algebraic scaffolding; the next provides its physical interpretation.

2.3.1 Three families of matrix invariants

The partial-transpose moments $\mu_k = \mathrm{Tr}[(\rho^{T_A})^k]$ can be expressed in terms of three complementary families of invariants, each accessible through controlled-SWAP circuits.

Standard invariants $\mathcal{I}_k = \mathrm{Tr}[\rho^k]$ characterise the eigenvalue spectrum of ρ itself: $\mathcal{I}_2$ is the purity, equal to 1 for pure states and $1/d$ for the maximally mixed state in dimension d.

Realignment invariants $R_k = \mathrm{Tr}[(R(\rho))^k]$ probe nonclassical correlations through the realignment map $R(\rho)_{ik,jl} = \rho_{ij,kl}$, which rearranges density matrix indices in a way that is algebraically independent of partial transposition.

Mixed invariants $M_k = \mathrm{Tr}[(\rho \cdot R(\rho))^k]$ encode the interplay between local and nonlocal structure through a product of ρ with its realignment. These higher-order quantities carry information not present in either $\mathcal{I}_k$ or R_k alone.

For two-qubit systems the partial-transpose moments can be related to these families through index-crossing combinatorics. A key identity, proved in Section 2.4, is that the second-order moments coincide exactly:

$$\mu_2 = \mathcal{I}_2, \tag{2.32}$$

so a single purity circuit measures both μ_2 and $\mathcal{I}_2$. At third and fourth order the chirality correction $C_k = \mu_k - \mathcal{I}_k$ is nonzero:

$$\mu_3 = \mathcal{I}_3 + C_3, \tag{2.33}$$

$$\mu_4 = \mathcal{I}_4 + C_4, \quad C_4 = 4R_4 + 2M_2. \tag{2.34}$$

Equation (2.34) is the key relation: M_2 appears only at fourth order, because crossing patterns that mix all four copies first become possible there. This is precisely the order at which the chirality-chirality correlator $C_4 = \mu_4 - \mathcal{I}_4$ acquires its nontrivial content (as developed in Section 2.4).

All three families are measurable via controlled-SWAP circuits on multiple state copies. The R_2 circuit requires two copies and a SWAP on subsystem A only, while the M_2 measurement requires four copies and a four-term combination involving SWAP and $Y \otimes Y$ operations. Each circuit reads out a single ancilla qubit, maintaining the uniform interface of the moment measurement protocol.

2.3.2 Gröbner basis conditions for eigenvalue degeneracy

For a two-qubit system, ρ^{T_A} has four eigenvalues $\{\lambda_i\}$ satisfying $\sum_i \lambda_i = 1$. The discriminant $\mathcal{D}(\mu_2, \mu_3, \mu_4)$ of the characteristic polynomial vanishes if and only if at least two eigenvalues coincide. For the full quartic this is a degree-6 polynomial in the moments, but specific degeneracy *types* are characterized by much simpler conditions, derived via resultant elimination from parametric eigenvalue equations.

For eigenvalues with a triple root of $(\alpha, \alpha, \alpha, \beta)$, which includes Bell states, Werner states, and the maximally mixed state, the degeneracy condition involves only μ_2 and μ_3:

$$\mathcal{G}_2 \;=\; 16\mu_2^3 - 39\mu_2^2 + 72\mu_2\mu_3 + 12\mu_2 - 48\mu_3^2 - 12\mu_3 - 1 \;=\; 0. \tag{2.35}$$

For eigenvalues with a two-pair degeneracy $(\alpha, \alpha, \beta, \beta)$, the condition is linear:

$$\mathcal{G}_1 \;=\; 6\mu_2 - 8\mu_3 - 1 \;=\; 0. \tag{2.36}$$

Since both conditions involve only μ_2 and μ_3, not μ_4, the degeneracy type can be determined from two measurement circuits before executing the third. The geometric interpretation is a stratification of the degeneracy locus:

$$V(\mathcal{D} = 0) \;=\; \underbrace{V(\mathcal{G}_2)}_{\text{triple+}} \;\cup\; \underbrace{\big(V(\mathcal{G}_1) \cap V(\mathcal{D})\big)}_{\text{two-pair}} \;\cup\; \{\text{simple pairs}\}. \tag{2.37}$$

Analogous conditions exist for qubit–qutrit systems (six eigenvalues). The two-triple degeneracy $(\alpha, \alpha, \alpha, \beta, \beta, \beta)$ satisfies:

$$\mathcal{G}_1^{(2\times3)} \;=\; 9\mu_2 - 18\mu_3 - 1 \;=\; 0, \tag{2.38}$$

and the quintuple degeneracy $(\alpha, \alpha, \alpha, \alpha, \alpha, \beta)$:

$$\mathcal{G}_2^{(2\times3)} \;=\; 96\mu_2^3 - 93\mu_2^2 + 180\mu_2\mu_3 + 18\mu_2 - 180\mu_3^2 - 20\mu_3 - 1 \;=\; 0. \tag{2.39}$$

Connection to moment invariants

The Gröbner conditions can be re-expressed in terms of the physical invariants $(\mathcal{I}_k, R_k)$ using $\mu_2 = \mathcal{I}_2$ and $\mu_3 = \mathcal{I}_3 + C_3$. For $\mathcal{G}_1$, substituting directly:

$$\mathcal{G}_1 = 6\mathcal{I}_2 - 8(\mathcal{I}_3 + C_3) - 1 = 0, \tag{2.40}$$

which shows that the two-pair degeneracy condition constrains a linear combination of purity and the chirality correction C_3. This representation makes the physical content clearer: $\mathcal{G}_1 = 0$ is satisfied when the purity-chirality balance takes a very specific value consistent with a two-pair eigenvalue pattern. For $\mathcal{G}_2$, the cubic dependence on $\mu_2 = \mathcal{I}_2$ means the triple-degeneracy surface is genuinely nonlinear in the invariant space.

2.3.3 Decision trees for adaptive negativity computation

The Gröbner conditions define a binary decision tree (Fig. 2.12) that classifies states into four computational branches, each admitting a progressively simplified negativity formula.

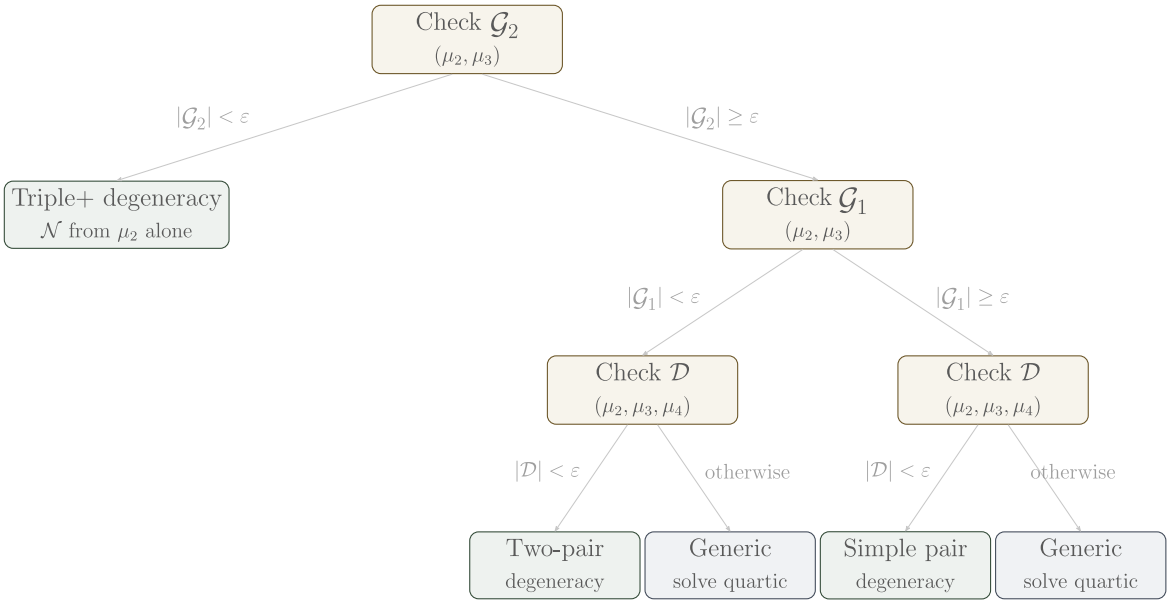


Fig. 2.12 **Decision tree for two-qubit negativity computation.** Gold nodes test algebraic conditions derived from Gröbner basis analysis; dark-green nodes return closed-form negativity formulas for degenerate cases; slate-blue nodes invoke the full quartic solve via Newton–Girard identities. The first two checks ($\mathcal{G}_2$ and $\mathcal{G}_1$) require only moments μ_2 and μ_3; μ_4 is measured only when the discriminant $\mathcal{D}$ must be evaluated.

The four branches and their negativity formulas are the following.

Triple-plus degeneracy ($\mathcal{G}_2 = 0$)

States satisfying $\mathcal{G}_2 = 0$– Bell states, Werner states, and the maximally mixed state– have at least three equal partial-transpose eigenvalues λ with $\mu = 1 - 3\lambda$ as the fourth. Negativity follows from μ_2 alone:

$$\mathcal{N}_{\text{triple}} = \max\left\{0,\ \frac{\sqrt{12\mu_2 - 3} - 1}{4}\right\}. \tag{2.41}$$

For Bell states ($\mu_2 = 1$) this gives $\mathcal{N} = 1/2$; for the maximally mixed state ($\mu_2 = 1/4$) it gives $\mathcal{N} = 0$. This branch requires a single circuit.

Two-pair degeneracy ($\mathcal{G}_1 = 0$, $\mathcal{G}_2 \neq 0$, $\mathcal{D} = 0$)

States with eigenvalues $(\lambda, \lambda, \mu, \mu)$ and $2\lambda + 2\mu = 1$ satisfy:

$$\mathcal{N}_{\text{two-pair}} = \max\left\{0,\ \frac{\sqrt{4\mu_2 - 1} - 1}{2}\right\}. \tag{2.42}$$

This requires two circuits (μ_2 and μ_3 to check $\mathcal{G}_1$, then $\mathcal{D}$ as a consistency check using only those two moments since $\mathcal{D}|_{G_1=0}$ reduces to a function of μ_2 only).

Simple pair degeneracy and generic case

States with exactly one repeated eigenvalue require a reduced cubic (three circuits: μ_2, μ_3, μ_4), while fully generic states require the full quartic solution via Newton–Girard identities (also three circuits). In practice these two branches merge computationally.

Efficiency summary

Monte Carlo sampling over 10^5 Haar-random two-qubit states confirms that 95.3% have four distinct eigenvalues ($\mathcal{D} \neq 0$) and 4.7% exhibit simple pair degeneracy; both require three measurements. Only specially structured states (Bell, Werner, maximally mixed) fall into the one- or two-circuit branches. The resulting average measurement count is 2.15 configurations, corresponding to a 119-fold reduction versus full tomography (256 configurations for two-qubit QST), or equivalently a $5.3\times$ reduction in measurement *settings* (since each configuration measures a global scalar from a single circuit rather than a local expectation value). For qubit–qutrit systems the decision tree spans 1–5 configurations with an average of 2.8, giving a 260-fold reduction versus full tomography (729 configurations) or $7.2\times$ in settings.

Table 2.2 Measurement efficiency of the Gröbner decision tree versus full tomography and classical shadows. "Settings" counts distinct circuit configurations; "configs" counts measurement outcomes needed for tomography.

System	Tomography	Decision tree	Gain (settings)	Gain (configs)
2×2	16 settings / 256 meas.	1–3 settings	$5.3\times$	$\sim 119\times$
2×3	36 settings / 729 meas.	1–5 settings	$7.2\times$	$\sim 260\times$

Noise robustness

Under measurement noise the exact conditions $G_i = 0$ become approximate tests $|G_i| < \epsilon$, where $\epsilon = 3\sigma_G$ and σ_G is the estimated standard deviation of the computed Gröbner polynomial propagated from moment uncertainties. At this tolerance the misclassification probability per decision node is $\mathrm{erfc}(3/\sqrt{2}) \approx 0.3\%$. Misclassification routes a state into the wrong branch, causing a loss of efficiency (more circuits executed) but *not* an error in the final negativity: the generic quartic solver is always correct, and every branch eventually terminates there as a fallback. The tree was tested on 500 randomly sampled two-qubit states under depolarizing noise: classification accuracy remains above 95% for noise strengths $p \lesssim 0.065$, which encompasses the typical operating range of current superconducting processors ($p \approx 0.04$–0.08) [6].

2.3.4 Physical content of the invariant decomposition: bridge to chirality

The three-family decomposition (2.33)–(2.34) has a physical consequence that motivates the next section. The fourth-order correction $C_4 = \mu_4 - \mathcal{I}_4 = 4R_4 + 2M_2$ mixes realignment and mixed invariants. Of these, M_2 carries the qualitatively new physics: it is the first invariant to involve the product $\rho \cdot R(\rho)$, capturing how the state correlates with its own realignment.

The measurement operator for M_2 decomposes, via the singlet projector algebra of four-copy space, into scalar spin chirality operators:

$$M_2 \;\propto\; \langle \chi_A \chi_B \rangle + \langle R_A R_B \rangle + (\text{trace terms}), \tag{2.43}$$

where $\chi_A = \mathbf{S}_{A_1} \cdot (\mathbf{S}_{A_2} \times \mathbf{S}_{A_3})$ is the scalar spin chirality of A-qubits from three state copies, and R_A, R_B are ring-exchange operators on four copies. The scalar spin chirality is the order parameter for chiral spin liquids and the topological Hall effect in condensed matter [45, 46]; here it appears among qubits from *different copies* of the same state.

This connection motivates the chirality witness

$$Q^{(\mathrm{irr})} = |\langle \chi_A \chi_B \rangle| + |\langle R_A R_B \rangle|, \tag{2.44}$$

which takes absolute values to ensure detection regardless of the sign of the correlation [23]. For all four Bell states $Q^{(\mathrm{irr})} \approx 0.234$, while for pure product states $Q^{(\mathrm{irr})} = 0$. The separable bound, derived symbolically for rank-dependent mixtures, is:

$$\begin{aligned}
&\beta_{\mathrm{sep}}^{(1)} = 0 \quad \text{(product states)},\\
&\beta_{\mathrm{sep}}^{(2)} = \tfrac{1}{256} \approx 0.0039 \quad \text{(rank-2 separable)},\\
&\beta_{\mathrm{sep}}^{(3)} = \tfrac{23}{432} \approx 0.053 \quad \text{(rank-3, tight bound)},\\
&\beta_{\mathrm{sep}}^{(4)} \approx 0.037 = \tfrac{1}{27} \quad \text{(rank-4, equals the } C_4 \text{ separable bound)}.
\end{aligned} \tag{2.45}$$

The rank-3 bound is the maximum across all separable states; it is achieved by equal-weight mixtures of three product states with mutually unbiased Bloch vectors on each subsystem. The non-monotonic hierarchy $\beta^{(2)} < \beta^{(4)} < \beta^{(3)}$ reflects a geometric constraint: rank-2 separable states cannot generate chirality (two Bloch vectors are coplanar), while rank-4 states are over-constrained by the full-rank condition. The detection criterion $Q^{(\mathrm{irr})} > 23/432 \approx 0.053$ therefore certifies entanglement with zero false positives for any separable state of any rank.

Table 2.3 Comparative detection thresholds for two-qubit entanglement criteria, expressed as functions of concurrence C for pure states and mixing parameter p for Werner states $\rho_W(p) = p|\Phi^+\rangle\langle\Phi^+| + (1-p)\mathbb{I}/4$.

Criterion	Pure states threshold	Werner threshold	Detection rate (rank-1)
$\|R\|_1 > 1$ (realignment)	$C > 0$ (all)	$p > 1/3$	100%
$Q^{(\mathrm{irr})} > 23/432$ (chirality)	$C > 0.48$	$p > 0.72$	52%
CHSH ($\mathcal{B}_{\max} > 2$)	$C > 1/\sqrt{2}$	$p > 1/\sqrt{2}$	100% (pure only)

The realignment trace norm $\|R\|_1 > 1$ achieves the strongest detection (all pure entangled states, 85–100% of mixed entangled states depending on rank), while $Q^{(\mathrm{irr})}$ provides physical interpretation at the cost of a higher threshold. Crucially, $Q^{(\mathrm{irr})}$ is the witness that makes the connection between the Gröbner algebraic classification and the condensed-matter physics of spin chirality explicit: the same mathematical object that classifies eigenvalue degeneracies also quantifies the handedness of multi-copy spin configurations.

Section 2.4 develops this connection in full, showing that the chirality-chirality correlator $C_k = 8\,\mathrm{Tr}[\Omega_A \Omega_B \rho^{\otimes k}]$ is not merely analogous to spin chirality but *identical* in algebraic structure, with the indices labelling state copies rather than lattice sites.

2.4 Multi-copy chirality reveals quantum entanglement

The characterization methods developed in the preceding sections of this chapter operate within the framework of single-copy measurements: each circuit acts on one copy of the quantum state, and the entanglement measures are inferred from the resulting probabilities. This section takes a different approach, asking what additional information becomes accessible when multiple simultaneous copies of the state are measured jointly. The answer turns out to be qualitatively richer than a simple increase in precision.

Two distinct forms of entanglement are completely hidden from any single-copy measurement on an *unknown* state. The first is the entanglement of bound entangled states, which satisfy the positive partial transpose (PPT) condition $\rho^{T_A} \geq 0$ by definition [47]. The Peres–Horodecki criterion [7, 8], which detects entanglement through a negative eigenvalue of ρ^{T_A}, is therefore structurally blind to this entire class, as is the full hierarchy of partial-transpose moment inequalities [48, 49] built on the same spectral basis. The second form is subtler: certain correlations in the joint statistics of multiple replicas of a state cannot be expressed as $\mathrm{Tr}[O\rho]$ for any single-copy observable O [31, 32]. For a fully reconstructed state these correlations can be computed classically, but without prior state tomography they are genuinely inaccessible to single-copy measurements.

Both forms become accessible through partial-transpose moments $\mu_k = \mathrm{Tr}[(\rho^{T_A})^k]$, measurable via controlled-SWAP circuits [50, 51] or randomized measurements [52, 53]. The scaling of required moments with system dimension follows from the number of partial-transpose eigenvalues: three moments $\{\mu_2, \mu_3, \mu_4\}$ for two-qubit systems, five for qubit–qutrit, eight for qutrit–qutrit. Critically, each circuit reads out a single ancilla qubit regardless of system dimension, providing a uniform measurement interface across all cases. Despite the practical utility of these moments in estimating negativity via Newton–Girard identities [54], the physical content of the difference $C_k = \mu_k - \mathcal{I}_k$ between partial-transpose and purity moments had remained an open question [54, 31, 32].

Section 2.4.2 shows that this difference decomposes exactly as a *chirality-chirality correlator*: the scalar spin chirality of multi-copy configurations on subsystem A, correlated with the same on subsystem B. This result simultaneously provides a geometric interpretation of the measurement, a bridge to condensed-matter physics through the same operator that governs chiral spin liquids [45] and the topological Hall effect [46], and a practical route to detecting bound entanglement through realignment matrix spectral features accessed by the same circuit infrastructure.

2.4.1 Physical picture: multi-copy chirality as a condensed-matter bridge

The scalar spin chirality $\chi_{ijk} = \mathbf{S}_i \cdot (\mathbf{S}_j \times \mathbf{S}_k)$ is a well-known quantity in condensed-matter physics. It measures whether three spin vectors span a three-dimensional volume or remain coplanar: geometrically, it equals the signed volume of the parallelepiped they define, or equivalently half the solid angle $\omega = 2|\chi|$ their unit vectors subtend on the Bloch sphere. It is the order parameter for chiral spin liquids (Kalmeyer–Laughlin states) [45], drives the topological Hall effect [46], and characterises topological phases with broken time-reversal symmetry. In all these condensed-matter applications, the indices i, j, k label physical spins at lattice sites.

The central result of this section is that the *same mathematical operator* appears in multi-copy entanglement measurements, but the indices no longer label lattice sites: they label qubits drawn from *different copies* of a bipartite quantum state. This reinterpretation changes the physical context entirely while preserving the algebraic structure. A product state generates multi-copy spin configurations that remain coplanar across copies ($\chi = 0$), because the A and B subsystems carry no correlated orientational information. An entangled state breaks this coplanarity: the A-copy spins and B-copy spins develop correlated three-dimensional volume, generating $\chi \neq 0$ jointly. The chirality correction $C_4 = \mu_4 - \mathcal{I}_4$ is precisely the expectation value of the product of chirality operators on both subsystems, measured jointly across four copies.

Figure 2.13 summarises the operational consequences of this picture: panel (a) maps the chirality landscape across 10^7 random states in 2×2 and 2×3 systems, projected onto the purity–negativity–$|C_4|$ space and bounded from below by the separability threshold $|C_4| = 1/27$, while panel (b) shows the resulting detection architecture in which the controlled-SWAP infrastructure simultaneously accesses the NPT region via partial-transpose moments and chirality, and the bound entangled region via the multichannel criterion built on the non-Hermiticity gap D_k. The parametric family $|\psi(\theta)\rangle = \cos(\theta/2)|00\rangle + \sin(\theta/2)|11\rangle$ provides a useful reference for these regimes: along it, $C_4(\theta) = \mu_4 - \mathcal{I}_4$ vanishes at the product state ($\theta = 0$) and reaches $C_4 = -3/4$ at the Bell state ($\theta = \pi/2$).

The connection to condensed matter is structural rather than superficial: both phenomena probe whether spin vectors span a three-dimensional volume or collapse to a plane. That scalar spin chirality is itself a witness of genuine tripartite entanglement measurable via Hadamard test circuits [55] deepens this bridge: the multi-copy protocol probes bipartite chirality correlations across state replicas using the same controlled-SWAP infrastructure that detects tripartite entanglement in lattice spin models.

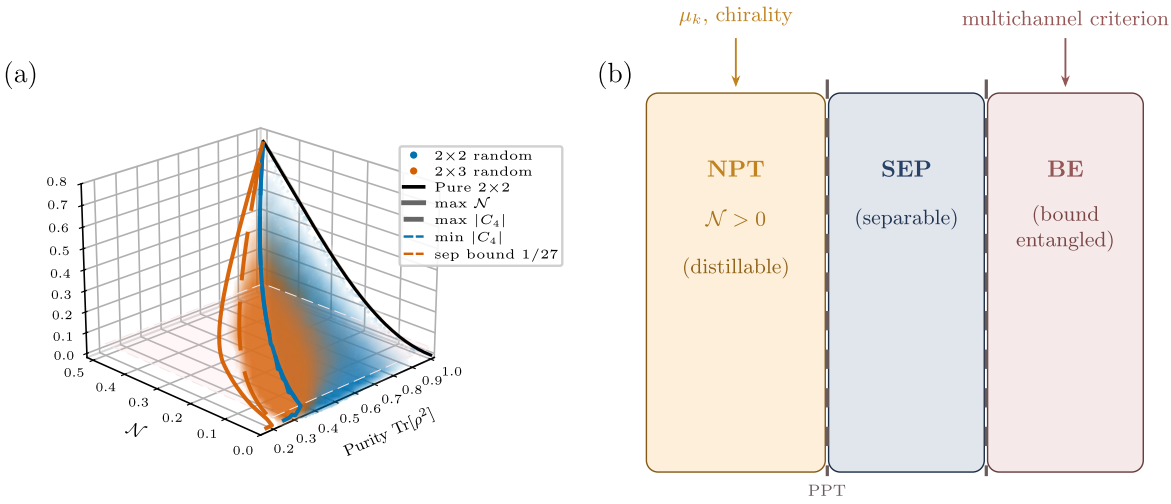


Fig. 2.13 **Chirality across quantum state copies reveals hidden entanglement.** (a), Chirality landscape: purity $\mathrm{Tr}[\rho^2]$ versus negativity $\mathcal{N}$ versus $|C_4|$ for 10^7 random states (2×2 blue, 2×3 orange). Black curve: pure 2×2 states. Dashed line: separable bound $|C_4| = 1/27$. (b), Detection architecture: controlled-SWAP circuits access negativity via partial-transpose moments μ_k (NPT region), the non-Hermiticity gap D_k (separable vs. bound entangled), and the chirality fingerprint C_k (bound entanglement).

The transition from flat to volumetric spin patterns is the geometric signature of entanglement exposed by multi-copy measurements, and it is completely invisible to single-copy measurements on unknown states. This is what makes the chirality correlator an intrinsically multi-copy quantity: for a known state it can be computed from the density matrix, but no single-copy measurement protocol can extract it from an unknown state without first performing complete state reconstruction. For mixed separable states, the chirality correction $C_k = \mu_k - \mathcal{I}_k$ can be nonzero, because μ_k and $\mathcal{I}_k$ are nonlinear in ρ and do not decompose linearly over convex mixtures; the tight bounds $|C_3| \leq 1/36$ and $|C_4| \leq 1/27$ established in Section 2.4.3 quantify the maximum contribution from separable states.

Multi-copy chirality is conceptually distinct from all standard forms of quantum correlation: Bell nonlocality, quantum steering, and entanglement itself are each detectable from a single copy of the state, while multi-copy chirality requires at least three copies and is genuinely absent from single-copy measurement statistics (Table 2.4). Standard correlations are properties of individual states, whereas multi-copy chirality is a property of the ensemble of copies—a collective quantum signature that only exists in joint measurement statistics.

Table 2.4 Comparison of nonclassical correlation types.

Correlation type	Copies required	Detection method
Bell nonlocality	1	CHSH inequality
Quantum steering	1	Steering inequalities
Entanglement	1	Witnesses, PPT criterion
Multi-copy chirality	3+	$C_4 = \mu_4 - \mathcal{I}_4 \neq 0$

2.4.2 Algebraic structure of the chirality decomposition

Permutation trace representation

Both μ_k and $\mathcal{I}_k = \mathrm{Tr}[\rho^k]$ admit a unified representation as permutation traces on the k-copy Hilbert space. With $\sigma = (12\cdots k)$ the cyclic permutation and subscripts denoting action on subsystem A or B:

$$\mu_k = \mathrm{Tr}\Big[(\sigma_A^{-1} \otimes \sigma_B)\, \rho^{\otimes k}\Big], \tag{2.46}$$

$$\mathcal{I}_k = \mathrm{Tr}\Big[(\sigma_A \otimes \sigma_B)\, \rho^{\otimes k}\Big]. \tag{2.47}$$

The difference $C_k \equiv \mu_k - \mathcal{I}_k$ is governed by the single operator $\Delta \equiv \sigma^{-1} - \sigma$:

$$C_k = \mathrm{Tr}\Big[(\Delta_A \otimes \sigma_B)\, \rho^{\otimes k}\Big]. \tag{2.48}$$

Because σ is unitary with $\sigma^\dagger = \sigma^{-1}$, the operator Δ satisfies $\Delta^\dagger = -\Delta$: it is anti-Hermitian. This anti-Hermiticity will be the key to rewriting C_k as a manifestly real chirality-chirality correlator.

For two-qubit systems the cyclic permutation decomposes via adjacent SWAP operators $S_{mn} = \frac{1}{2}(I + g_{mn})$ with $g_{mn} = \boldsymbol{\sigma}_m \cdot \boldsymbol{\sigma}_n = 2S_{mn} - I$. The fundamental identity connecting the SWAP algebra to chirality is that pairs sharing one index satisfy

$$[g_{ij},\ g_{jk}] = -16i\, \chi_{ijk}, \tag{2.49}$$

where $\chi_{ijk} = \mathbf{S}_i \cdot (\mathbf{S}_j \times \mathbf{S}_k) = \frac{1}{8}\sum_{abc} \varepsilon_{abc}\, \sigma_i^a \sigma_j^b \sigma_k^c$ is the scalar spin chirality operator. Pairs that do not share an index commute: $[g_{ij}, g_{mn}] = 0$ when $\{i,j\} \cap \{m,n\} = \emptyset$. Equivalently, in terms of singlet projectors $P_{ij}^- = \frac{1}{4} - \mathbf{S}_i \cdot \mathbf{S}_j$:

$$[P_{ij}^-,\ P_{jk}^-] = -i\, \chi_{ijk}, \tag{2.50}$$

so chirality arises whenever two controlled SWAPs share a copy index. An important symmetry property is that C_4 is invariant under local unitary (LU) transformations $\rho \to (U_A \otimes U_B)\rho(U_A \otimes U_B)^\dagger$: purity moments $\mathcal{I}_k$ are LU-invariant since the eigenvalues of ρ are unchanged, and μ_k are LU-invariant because partial transposition transforms as $(U_A^* \otimes U_B)\rho^{T_A}(U_A^T \otimes U_B^\dagger)$, which is unitarily equivalent to ρ^{T_A} and therefore has the same spectrum.

Explicit chirality decompositions at orders $k = 2, 3, 4$

The commutator identities (2.49)–(2.50) yield explicit chirality decompositions of Δ at each order.

Second order ($k = 2$). Since $\sigma = \sigma^{-1} = S_{12}$, we have $\Delta = 0$ and therefore:

$$\mu_2 = \mathcal{I}_2, \tag{2.51}$$

for all bipartite states regardless of dimension. This was already noted in Section 2.3 as a practical simplification (one circuit serves for both invariants); here its algebraic origin is explicit: $\mathrm{Tr}[(\rho^{T_A})^2] = \sum_{ijkl} \rho_{kj,il}\rho_{il,kj} = \sum_{ijkl} |\rho_{ij,kl}|^2 = \mathrm{Tr}[\rho^2]$. This dimension-independent identity means a single purity circuit measures both μ_2 and $\mathcal{I}_2$, reducing the total number of required configurations by one.

Third order ($k = 3$). With $\sigma = S_{12}S_{23}$ and $\sigma^{-1} = S_{23}S_{12}$, the difference is a commutator:

$$\Delta_3 = [S_{23},\ S_{12}] = \tfrac{1}{4}[g_{23},\ g_{12}] = 4i\,\chi_{123}, \tag{2.52}$$

giving the intermediate form:

$$C_3 = 4i\,\mathrm{Tr}\Big[\chi_A^{(123)}\,\sigma_B\,\rho^{\otimes 3}\Big]. \tag{2.53}$$

Fourth order ($k = 4$). With $\sigma = S_{12}S_{23}S_{34}$ and $\sigma^{-1} = S_{34}S_{23}S_{12}$, the expansion $8\Delta_4 = 8(\sigma^{-1} - \sigma)$ gives:

$$\begin{aligned} 8\Delta_4 &= (I{+}g_{34})(I{+}g_{23})(I{+}g_{12}) - (I{+}g_{12})(I{+}g_{23})(I{+}g_{34}) \\ &= [g_{23}, g_{12}] + [g_{34}, g_{23}] + g_{34}[g_{23}, g_{12}] + g_{12}[g_{34}, g_{23}] \\ &= 16i\Big\{(I{+}g_{34})\chi_{123} + (I{+}g_{12})\chi_{234}\Big\} \\ &= 32i\Big(S_{34}\,\chi_{123} + S_{12}\,\chi_{234}\Big), \end{aligned} \tag{2.54}$$

where the third line uses $[g_{12}, g_{34}] = 0$ (disjoint pairs) and $I + g_{mn} = 2S_{mn}$.

To simplify $\Omega = S_{34}\chi_{123} + S_{12}\chi_{234}$, one evaluates the products $g_{mn}\chi_{ijk}$ via the Pauli algebra identity $\sigma_i^a\sigma_i^d = \delta_{ad}I + i\sum_e \varepsilon_{ade}\sigma_i^e$ and the Levi-Civita contraction $\sum_d \varepsilon_{bcd}\varepsilon_{ade} = \delta_{ba}\delta_{ce} - \delta_{be}\delta_{ca}$. When index 3 is shared between g_{34} and χ_{123}, this yields:

$$g_{34}\,\chi_{123} = \chi_{124} + \tfrac{i}{8}(g_{14}g_{23} - g_{13}g_{24}), \tag{2.55}$$

$$g_{12}\,\chi_{234} = \chi_{134} + \tfrac{i}{8}(g_{13}g_{24} - g_{14}g_{23}). \tag{2.56}$$

The imaginary cross-terms cancel upon addition, and collecting all four contributions:

$$\begin{aligned}\Omega &= \tfrac{1}{2}(I + g_{34})\chi_{123} + \tfrac{1}{2}(I + g_{12})\chi_{234} \\ &= \tfrac{1}{2}(\chi_{123} + \chi_{234}) + \tfrac{1}{2}(\chi_{124} + \chi_{134}) \\ &= \tfrac{1}{2}\sum_{i<j<k}\chi_{ijk},\end{aligned} \tag{2.57}$$

where the sum runs over all $\binom{4}{3} = 4$ chirality triples on four copies. This shows that all four triples contribute with equal weight $1/2$, a consequence of the permutation symmetry of the moment relations. The fourth-order correction is therefore:

$$C_4 = 2i\,\mathrm{Tr}\Big[\Big(\sum_{i<j<k}\chi^A_{ijk}\Big)\,\sigma_B\,\rho^{\otimes 4}\Big]. \tag{2.58}$$

The chirality-chirality correlator form

The intermediate expressions (2.53) and (2.58) still involve the non-Hermitian cyclic permutation σ_B, making the reality of C_k non-obvious. Taking the complex conjugate and using $\Delta_A^\dagger = -\Delta_A$ and $\sigma_B^\dagger = \sigma_B^{-1}$ gives $C_k^* = \mathrm{Tr}[(-\Delta_A \otimes \sigma_B^{-1})\rho^{\otimes k}]$. Since C_k is real ($C_k = C_k^*$), averaging:

$$C_k = \tfrac{1}{2}(C_k + C_k^*) = \tfrac{1}{2}\,\mathrm{Tr}\Big[\Delta_A \otimes (\sigma_B - \sigma_B^{-1})\,\rho^{\otimes k}\Big] = -\tfrac{1}{2}\,\mathrm{Tr}\Big[(\Delta_A\Delta_B)\,\rho^{\otimes k}\Big]. \tag{2.59}$$

The operator $\Delta_A\Delta_B$ is Hermitian because Δ_A and Δ_B act on disjoint Hilbert spaces (so they commute) and $(\Delta_A\Delta_B)^\dagger = \Delta_B^\dagger\Delta_A^\dagger = (-\Delta_B)(-\Delta_A) = \Delta_A\Delta_B$.

Defining Ω_k as the Hermitian operator with $\Delta_k = 4i\,\Omega_k$ ($\Omega_3 = \chi_{123}$, $\Omega_4 = \frac{1}{2}\sum_{i<j<k}\chi_{ijk}$), substituting $\Delta_A\Delta_B = (4i\,\Omega_A)(4i\,\Omega_B) = -16\,\Omega_A\Omega_B$ into Eq. (2.59) yields the central identity:

$$C_k = 8\,\mathrm{Tr}\Big[\Omega_A\,\Omega_B\,\rho^{\otimes k}\Big]. \tag{2.60}$$

Written out explicitly:

$$C_3 = 8\,\mathrm{Tr}\Big[\chi_A\,\chi_B\;\rho^{\otimes 3}\Big], \tag{2.61}$$

$$C_4 = 8\,\mathrm{Tr}\Big[\Omega_A\,\Omega_B\;\rho^{\otimes 4}\Big], \tag{2.62}$$

where $\Omega = \frac{1}{2}\sum_{i<j<k}\chi_{ijk}$. The difference between partial-transpose and purity moments is thus a chirality-chirality correlator: the Hermitian chirality operator on subsystem A correlated with the corresponding operator on subsystem B, measured across k state copies. For product states both chirality factors vanish independently (coplanar spin configurations), giving $C_k = 0$. For pure states, entanglement breaks this coplanarity simultaneously on both subsystems, generating $C_k \neq 0$. For mixed separable states, C_k can be nonzero due to cross-terms in the multi-copy expansion, but is bounded by $|C_3| \leq 1/36$ and $|C_4| \leq 1/27$ (Section 2.4.3).

This identity answers the open question of what multi-copy correlations distinguish partial transposition from purity [54, 31, 32]: they are chirality-chirality correlators, connecting multi-copy entanglement detection to the physics of topological spin systems [45, 46].

2.4.3 Geometric interpretation and chirality-negativity relation

Correlation tensor and signed volume

For a general two-qubit state in the Fano decomposition (introduced in Section 1.3.2 using the notation $\hat{\beta}_{ij}$; here written with the convention $T_{ij} \equiv \beta_{ij}$ to match the literature on chirality):

$$\rho = \tfrac{1}{4}\Big(I\otimes I + \mathbf{r}\cdot\boldsymbol{\sigma}\otimes I + I\otimes\mathbf{s}\cdot\boldsymbol{\sigma} + \sum_{ij} T_{ij}\,\sigma_i\otimes\sigma_j\Big), \tag{2.63}$$

with local Bloch vectors $\mathbf{r}$, $\mathbf{s}$ and correlation tensor $T_{ij} = \mathrm{Tr}[\rho(\sigma_i\otimes\sigma_j)]$, the chirality correction carries a precise geometric interpretation. The three columns $\mathbf{T}_x$, $\mathbf{T}_y$, $\mathbf{T}_z$ of T encode spin–spin correlations. For product states these columns are coplanar ($\det T = 0$); for entangled states they span three-dimensional space ($|\det T| > 0$).

Proposition 1 (C_4 and $\det(T)$ for states with vanishing Bloch vectors)**.** *For any two-qubit state with* $\mathbf{r} = \mathbf{s} = \mathbf{0}$*, which includes all Bell-diagonal states and their local unitary rotations:*

$$C_4 = \tfrac{3}{4}\det(T). \tag{2.64}$$

Sketch Any state with $\mathbf{r} = \mathbf{s} = \mathbf{0}$ has $\rho = \frac{1}{4}(I + \sum_{ij} T_{ij}\sigma_i\otimes\sigma_j)$. The singular value decomposition $T = U\,\mathrm{diag}(t_1,t_2,t_3)\,V^{\mathsf{T}}$ with $U, V \in \mathrm{SO}(3)$ can be implemented by local

unitaries $U_A, U_B \in \mathrm{SU}(2)$ via the standard $\mathrm{SU}(2) \to \mathrm{SO}(3)$ homomorphism, transforming ρ to a Bell-diagonal state $\rho' = \frac{1}{4}(I + \sum_i t_i\, \sigma_i \otimes \sigma_i)$. Since C_4 is LU-invariant and $\det(T) = t_1 t_2 t_3$ is preserved by $\mathrm{SO}(3)$ rotations on both sides, it suffices to verify the formula for Bell-diagonal states, where direct eigenvalue calculation confirms $C_4 = \mu_4 - \mathcal{I}_4 = \frac{3}{4} t_1 t_2 t_3$. Numerical verification across 10^3 random Bell-diagonal states confirms agreement to machine precision.

Chirality thus measures the signed volume of the correlation parallelepiped, with the coefficient $3/4$ fixed by the Bell state value ($\det T = -1$, $C_4 = -3/4$). The sign of $\det(T)$ distinguishes left-handed from right-handed correlations. The CHSH criterion [23] depends on the two largest eigenvalues of $T^{\mathsf{T}} T$, while chirality depends on all three through $\det(T)$, making them complementary probes of the full correlation tensor structure. For Bell states $|\det(T)| = 1$ (maximum volume).

For states with nonzero Bloch vectors, the exact relation (2.64) receives corrections expressible in terms of the fundamental two-qubit LU invariants

$$\{|\mathbf{r}|^2, |\mathbf{s}|^2, \mathrm{Tr}[T^{\mathsf{T}} T], \det(T), \mathbf{r}^{\mathsf{T}} T^{\mathsf{T}} T \mathbf{r}, \mathbf{s}^{\mathsf{T}} T T^{\mathsf{T}} \mathbf{s}, \mathbf{r}^{\mathsf{T}} T \mathbf{s}\}. \tag{2.65}$$

The correction $C_4 - \frac{3}{4}\det(T)$ vanishes for $\mathbf{r} = \mathbf{s} = \mathbf{0}$ and scales as $O(|\mathbf{r}|^2 + |\mathbf{s}|^2)$ for small Bloch vectors, but involves a nonlinear combination of invariants without a simple closed form. For the experimentally relevant Bell-product mixtures $\rho(p) = (1-p)|\Phi^+\rangle\langle\Phi^+| + p|00\rangle\langle 00|$ (where $\mathbf{r} = \mathbf{s} = (0,0,p)^T$), numerical evaluation gives corrections of order $|C_4 - \frac{3}{4}\det(T)| \leq 0.016$ across the full range $p \in [0,1]$.

Chirality–negativity relation for pure states

Negativity and chirality probe entanglement through complementary mechanisms. Negativity detects inseparability through the spectrum of ρ^{T_A}; chirality detects it through collective multi-copy interference patterns invisible to any single-copy observable. Both vanish for product states and reach extremal values at Bell states ($\mathcal{N} = 1/2$, $C_4 = -3/4$). For the parametric family $|\psi(\theta)\rangle = \cos(\theta/2)|00\rangle + \sin(\theta/2)|11\rangle$, both are monotone functions of the entanglement angle θ:

$$\mathcal{N}(\theta) = \frac{\sin\theta}{2}, \qquad C_4(\theta) = -\sin^2\theta \left(1 - \frac{\sin^2\theta}{4}\right). \tag{2.66}$$

For any pure two-qubit state, a direct computation using the partial-transpose eigenvalues $\cos^2(\theta/2)$, $\sin^2(\theta/2)$, $\pm\frac{1}{2}\sin\theta$ and $\mathcal{N} = \frac{1}{2}\sin\theta$ establishes the closed-form relation:

$$C_4 = -4\mathcal{N}^2(1 - \mathcal{N}^2). \tag{2.67}$$

This can be verified for the Bell state: $C_4 = -4 \cdot \frac{1}{4}(1 - \frac{1}{4}) = -\frac{3}{4}$. The relationship reveals the geometric content of the negativity–chirality connection: $|C_4|$ is a quartic polynomial in $\mathcal{N}$ that vanishes at $\mathcal{N} = 0$ (product states) and, were it physical, at $\mathcal{N} = 1$. Its maximum over all states is at $\mathcal{N} = 1/\sqrt{2}$, giving $|C_4| = 1$. For physical two-qubit states where $\mathcal{N} \leq 1/2$, the function increases monotonically from 0 at the product state to $3/4$ at the Bell state.

The relation (2.67) can be inverted to recover negativity directly from a chirality measurement, without going through the Newton–Girard reconstruction:

$$\mathcal{N} = \sqrt{\frac{1 - \sqrt{1 + C_4}}{2}}. \tag{2.68}$$

The third-order correction admits an even simpler pure-state form, following from $\mu_3 = \frac{1}{4}(1 + 3\cos^2\theta)$ and $\mathcal{I}_3 = 1$:

$$C_3 = -3\mathcal{N}^2. \tag{2.69}$$

Since for pure two-qubit states the negativity equals half the concurrence [56], $\mathcal{N} = \mathcal{C}/2$, the chirality correction also relates to the concurrence:

$$C_4 = -\mathcal{C}^2\left(1 - \frac{\mathcal{C}^2}{4}\right). \tag{2.70}$$

For pure two-qubit states, $C_4 \neq 0$ if and only if the state is entangled, and the same completeness holds for pure qubit–qutrit systems. The forward direction follows from (2.67): $C_4 < 0$ whenever $\mathcal{N} > 0$. The reverse direction holds because for product states the partial transpose factors as $\rho^{T_A} = \rho_A^T \otimes \rho_B$, transposition preserves eigenvalues, and therefore $\mu_k = \mathcal{I}_k$, giving $C_4 = 0$.

Theorem 1 (Pure-state completeness)**.** *For pure* 2×2 *and* 2×3 *states:*

$$C_4 \neq 0 \quad \Longleftrightarrow \quad |\psi\rangle \textit{ is entangled.} \tag{2.71}$$

In higher dimensions ($d_A d_B > 6$), bound entangled states exist that are entangled but have $\rho^{T_A} \geq 0$, so $\mathcal{N} = 0$ and C_4 reflects only the nonlinearity of the moment functions rather than entanglement. Detecting bound entanglement requires the realignment spectral features of Section 2.4.5. Table 2.5 summarizes the scope of C_4 as an entanglement indicator.

Separable bounds and mixed-state limitations

An important limitation applies to mixed entangled states: $|C_4|$ can fall below the separable bound $1/27$, so chirality is a faithful entanglement witness only for pure states. For Werner

Table 2.5 Scope of C_4 as entanglement indicator.

System	Pure states: $C_4 \neq 0 \Leftrightarrow$ entangled	Mixed states: $\mathcal{N} > 0 \Leftrightarrow$ entangled
2×2	✓	✓(Peres–Horodecki)
2×3	✓	✓(Peres–Horodecki)
3×3+	✓	$\times$ (bound entanglement)

states $\rho_W(p) = p|\Psi^-\rangle\langle\Psi^-| + (1-p)\mathbb{I}/4$, the local Bloch vectors vanish and $\det(T) = -p^3$, giving $C_4(p) = -\frac{3}{4}p^3$. This can also be obtained directly from the partial-transpose eigenvalues $(1+p)/4$ (triple) and $(1-3p)/4$ (singlet):

$$\mu_4(p) = \frac{3(1+p)^4 + (1-3p)^4}{256}, \qquad \mathcal{I}_4(p) = \frac{(1+3p)^4 + 3(1-p)^4}{256}, \tag{2.72}$$

confirming $C_4(p) = -3p^3/4$. The magnitude $|C_4|$ exceeds $1/27$ only for $p \gtrsim 0.37$, while entanglement onset is at $p > 1/3$. In this gap chirality alone cannot certify entanglement; the Werner state at the separability boundary has $|C_4(1/3)| = 1/36 < 1/27$, well within the separable bound. The Werner states do not achieve the extremal separable bound: the extrema $C_4 = \pm 1/27$ are achieved by the rank-3 MUB mixtures, not by Werner states. For mixed states in general, once all moments $\{\mu_2, \mu_3, \mu_4\}$ are measured, C_4 is fully determined and negativity from the Newton–Girard reconstruction provides the definitive test. The value of chirality for mixed states lies in its geometric interpretation as a multi-copy handedness correlator, not as a stand-alone detection criterion.

For any two-qubit separable state the chirality correction satisfies $|C_3| \leq 1/36$ and $|C_4| \leq 1/27$. Both bounds are tight, achieved by rank-3 equal mixtures of three mutually unbiased basis (MUB) product states. To see why: product states have $C_k = 0$ because their partial transpose factors as $\rho^{T_A} = \rho_A^T \otimes \rho_B$ and transposition preserves eigenvalues. Rank-2 separable states also satisfy $C_k = 0$, since mixtures of two product states preserve this factorization structure. For rank-3 states the cross-terms in the multi-copy expansion acquire a nonzero chirality contribution; the MUB phase structure maximizes the constructive interference, yielding:

$$\rho_\pm = \tfrac{1}{3}\Big(|z_+, z_\pm\rangle\langle z_+, z_\pm| + |x_+, x_\pm\rangle\langle x_+, x_\pm| + |y_+, y_\pm\rangle\langle y_+, y_\pm|\Big), \tag{2.73}$$

with $C_4(\rho_+) = +1/27$ and $C_4(\rho_-) = -1/27$. Entangled pure states have $|C_4| \in (0, 3/4]$, while all separable states satisfy $|C_4| \leq 1/27 \approx 0.037$, a factor of 20 smaller than the Bell state value (Table 2.6).

Table 2.6 Chirality correction ranges for two-qubit state classes.

State class	C_3 range	C_4 range	Example at extremum
Product	0	0	$\lvert 00\rangle$
Rank-2 separable	0	0	$\frac{1}{2}(\lvert 00\rangle\langle 00\rvert + \lvert 11\rangle\langle 11\rvert)$
General separable	$[-\frac{1}{36},\ \frac{1}{36}]$	$[-\frac{1}{27},\ \frac{1}{27}]$	MUB mixtures $\rho_{\pm}$
Entangled pure	$(-\frac{3}{4},\ 0)$	$(-\frac{3}{4},\ 0)$	Bell state: $C_4 = -3/4$

2.4.4 Negativity reconstruction and measurement circuits

Newton–Girard identities and Ferrari reconstruction

Once moments $\{\mu_2, \mu_3, \mu_4\}$ are measured, negativity is recovered from the characteristic polynomial of ρ^{T_A} via the Newton–Girard identities. The elementary symmetric polynomials $e_1, \ldots, e_4$ of the eigenvalues are:

$$\begin{aligned} e_1 &= 1, \\ e_2 &= \tfrac{1}{2}(1 - \mu_2), \\ e_3 &= \tfrac{1}{6}(1 - 3\mu_2 + 2\mu_3), \\ e_4 &= \tfrac{1}{24}(1 - 6\mu_2 + 8\mu_3 + 3\mu_2^2 - 6\mu_4). \end{aligned} \tag{2.74}$$

Eigenvalues follow via Ferrari's method and negativity is $\mathcal{N} = \sum_{\lambda_i<0} |\lambda_i|$. Since $\mu_2 = \mathcal{I}_2$, the coefficient e_2 depends only on purity, confirming that the second-order circuit already exhausts all information available at that order.

The reconstruction is exact generically but requires care when eigenvalues are degenerate. The degeneracy conditions are characterized by the discriminants

$$\mathcal{G}_1 = 2\mu_3 - 3\mu_2 + 1 = 0, \tag{2.75}$$

$$\mathcal{G}_2 = \mu_4 - \tfrac{1}{3}(3\mu_2^2 - 2\mu_3 - 1) = 0. \tag{2.76}$$

Although notationally similar to the Gröbner polynomials $\mathcal{G}_1$, $\mathcal{G}_2$ of Section 2.3.2, these are distinct objects with different functional forms; the calligraphic $\mathcal{G}$ is used throughout this subsection to avoid confusion. Monte Carlo sampling over 10^5 Haar-random two-qubit states confirms that 95.3% require the full three configurations and 4.7% exhibit simple pair degeneracy; both cases are covered by three measurements, and the efficiency gain over full tomography is $16/3 \approx 5.3\times$ for 2×2 systems and $36/5 = 7.2\times$ for 2×3 systems.

For 2×3 systems, five moments $\{\mu_2, \ldots, \mu_6\}$ are needed in general; the analogous Gröbner degeneracy conditions $\mathcal{G}_1^{(2\times3)}$ and $\mathcal{G}_2^{(2\times3)}$ are given in Section 2.3.2 (Eqs. (2.38)–

(2.39)). Under depolarizing noise of strength η, Monte Carlo simulation gives $\mathrm{RMSE}_{2\times2} = (0.245 \pm 0.004)\eta$ and $\mathrm{RMSE}_{2\times3} = (0.219 \pm 0.002)\eta$. Depolarizing noise consistently underestimates entanglement, since $\mathcal{N}(\rho_{\mathrm{noisy}}) \approx (1-\eta)\mathcal{N}(\rho)$, which ensures the protocol never falsely reports entanglement in separable states within this noise model.

Controlled-SWAP circuits for moment measurement

All moments μ_k are measured through the Hadamard test: k copies of ρ with a controlled permutation conditioned on a single ancilla qubit, yielding $\mu_k = 2p_0 - 1$ from the ancilla measurement outcome. The circuits for μ_2, μ_3, μ_4 are shown in Figs. 2.14–2.16.

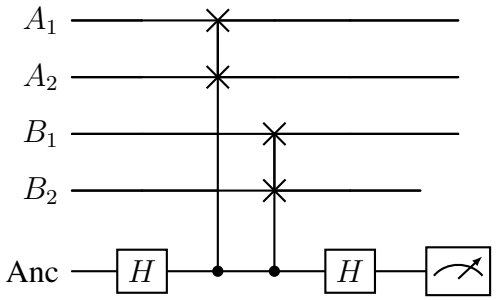


Fig. 2.14 **Circuit $\mathcal{C}_{\mu_2}$ for the second partial transpose moment.** Outcome: $\mu_2 = 2p_0 - 1$. Since $\mu_2 = \mathcal{I}_2$, this single circuit measures both the second partial-transpose moment and the purity.

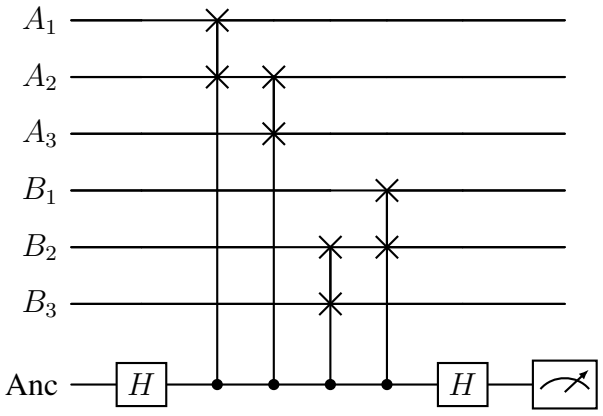


Fig. 2.15 **Circuit $\mathcal{C}_{\mu_3}$ for the third partial transpose moment.** The ancilla controls the A-chain (anticyclic, rows 1–3) and the B-chain (cyclic, rows 4–6), realizing the cycle–anticycle decomposition of $\mu_3 = \mathrm{Tr}[(\sigma_A^{-1} \otimes \sigma_B)\rho^{\otimes3}]$.

Table 2.7 summarizes circuit resources.

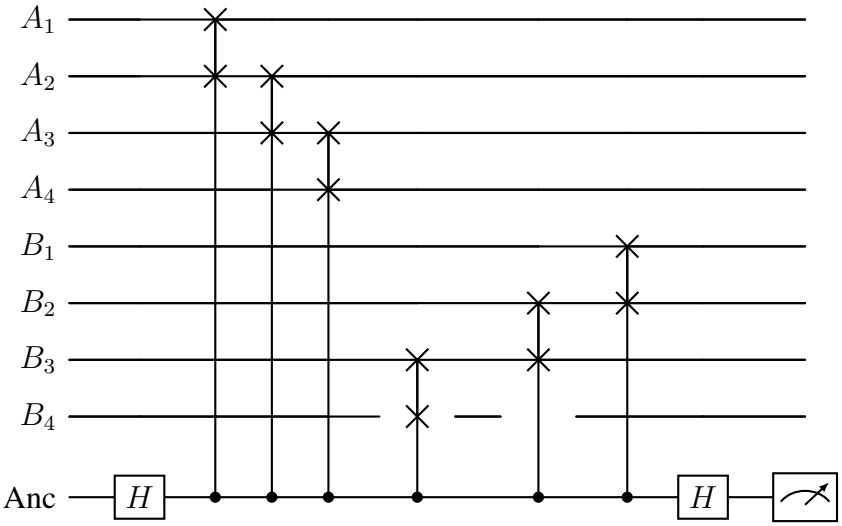


Fig. 2.16 **Circuit $\mathcal{C}_{\mu_4}$ for the fourth partial transpose moment.** The A-subsystem (rows 1–4) implements the forward cycle $(A_1A_2A_3A_4)$; the B-subsystem (rows 5–8) implements the reverse cycle $(B_4B_3B_2B_1)$, realizing $\mu_4 = \mathrm{Tr}[(\sigma_A^{-1} \otimes \sigma_B)\rho^{\otimes 4}]$.

Table 2.7 Quantum circuit resources (one ancilla included in qubit count).

Invariant	Copies	Qubits (2×2)	Qubits (2×3)	cSWAPs
$\mathcal{I}_2 = \mu_2$	2	5	7	2
μ_3	3	7	10	4
μ_4	4	9	13	6
Σ_2, G_2	2	5	9	2
Σ_4, G_4	4	9	17	4–6

The 2×3 qubit counts assume each qutrit encoded in two physical qubits with the computational subspace $\{|00\rangle, |01\rangle, |10\rangle\} \cong \{|0\rangle, |1\rangle, |2\rangle\}$, excluding $|11\rangle$. Each circuit uses a single ancilla qubit regardless of system dimension, which is the key practical advantage over singlet-projection-based schemes that have no direct counterpart for qutrits.

Realignment circuits and the universal identity $S_2 = \mathcal{I}_2 = \mu_2$

The realignment singular-value moments Σ_k and eigenvalue moments G_k are measured through two distinct circuit structures. For $\Sigma_2 = \mathrm{Tr}[R^\dagger R]$, a standard two-copy SWAP test on the full system suffices, because the Frobenius norm is preserved under reshuffling: $\Sigma_2 = \|\rho\|_F^2 = \mathrm{Tr}[\rho^2] = \mathcal{I}_2$. Since $\mu_2 = \mathcal{I}_2$ holds for any bipartite state (Section 2.3), all three

second-order invariants coincide:

$$S_2 = \Sigma_2 = \mathcal{I}_2 = \mu_2, \tag{2.77}$$

and are therefore obtained from a single circuit at no additional measurement cost.

For $G_2 = \mathrm{Re}\,\mathrm{Tr}[R(\rho)^2]$, the permutation structure differs: only A-indices are exchanged between copies while B remains fixed (cyclic permutation Π_A only): For Σ_4 at fourth

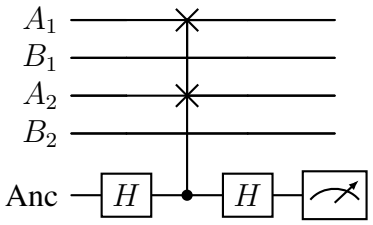


Fig. 2.17 **Circuit for $G_2 = \mathrm{Re}\,\mathbf{Tr}[R(\rho)^2]$.** The ancilla-controlled SWAP acts on subsystem A only; B qubits are unchanged. Outcome: $G_2 = 2p_0 - 1$. Compare with $\mathcal{C}_{\mu_2}$ (Fig. 2.14) where the ancilla controls SWAPs on *both* subsystems.

order, the permutation pairs copies symmetrically: $\Pi_A^{(4)} = S_{12}^{(A)} S_{34}^{(A)}$ (pairwise swaps on A) and $\Pi_B^{(4)} = S_{14}^{(B)} S_{23}^{(B)}$ (cross swaps on B), whereas for μ_4 the structure is cycle-anticycle. This difference in permutation pattern is the operational signature that Σ_k and μ_k probe complementary structures of the density matrix.

Extension to qutrit systems: cyclic asymmetry

The chirality framework of Section 2.4 uses the Pauli algebra of SU(2) and the commutator identity $[g_{ij}, g_{jk}] = -16i\,\chi_{ijk}$ as its algebraic engine. For qutrit subsystems the Pauli matrices are replaced by the Gell-Mann generators $\{\lambda^a\}_{a=1}^8$ of SU(3), and the SWAP operator decomposes as:

$$S_{ij}^{(\mathrm{qutrit})} = \tfrac{1}{3}\left(I_9 + \tfrac{1}{2}\sum_{a=1}^{8} \lambda_i^a \otimes \lambda_j^a\right), \tag{2.78}$$

with factor $1/3$ instead of $1/2$ for qubits, reflecting the three-dimensional local space. The commutator of overlapping generator products yields a *cyclic asymmetry operator* $\mathcal{X}_{ijk}$, built from the SU(3) structure constants f^{abc}, that plays the role of χ_{ijk} in the qutrit chirality decomposition; it measures cyclic asymmetry in SU(3) space rather than geometric handedness in SU(2).

The key qualitative consequence is that the chirality-chirality correlator prefactor is substantially smaller for qutrits than for qubits. This reflects a weaker SU(3) cyclic asymmetry signal, which is the same mechanism that makes SU(3) chiral condensates harder to probe

than their SU(2) counterparts. The $C_4 \neq 0$ entanglement witness nonetheless remains valid for pure qutrit–qutrit states with zero false positives, by the same algebraic argument as Theorem 1.

Circuit resources scale as $6k + 1$ qubits for k copies of a 3×3 state under the two-qubit-per-qutrit encoding, making C_3 (19 qubits) and C_4 (25 qubits) accessible on current large processors. For the experimentally validated 2×3 case, the overhead is more modest (Table 2.7), with hardware demonstration on *ibm_torino* confirming $7.2\times$ efficiency over full tomography.

2.4.5 Bound entanglement detection via realignment spectral features

Realignment map, non-Hermiticity gap, and spin-vector interpretation

When both negativity and chirality vanish, bound entanglement can still be present in systems of dimension $d_A d_B \geq 9$. Detecting it requires moving outside the partial-transpose framework entirely. The realignment map (defined in Section 1.3.2 of Chapter 1) [12, 13]

$$R(\rho)_{(i,k),(j,l)} = \rho_{(i,j),(k,l)} \tag{2.79}$$

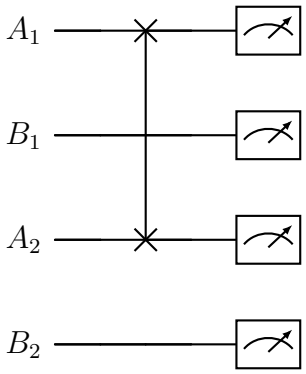
reshuffles density matrix elements in a way that is mathematically independent of partial transposition: the singular values and eigenvalues of $R(\rho)$ carry information about the state's entanglement structure that the PPT criterion misses by construction. Three measurable quantities are defined from the spectrum of $R(\rho)$ (note: G_k here denotes realignment eigenvalue moments and is distinct from the Gröbner degeneracy polynomials G_1, G_2 of Section 2.3.2):

$$\Sigma_k = \sum_i \sigma_i^k, \qquad G_k = \mathrm{Re}\,\mathrm{Tr}[R^k], \qquad D_k = \Sigma_k - G_k. \tag{2.80}$$

The non-Hermiticity gap D_k has a direct spin-physics interpretation. Expanding the state in the Gell-Mann basis, $R(\rho)$ is equivalent (up to normalisation) to the A–B spin correlation tensor $T_{\alpha\beta} = \langle \Lambda_\alpha^A \otimes \Lambda_\beta^B \rangle$. Its antisymmetric part is proportional to the cross-product $\langle \mathbf{J}_A \times \mathbf{J}_B \rangle$, and D_k measures this antisymmetric component. The PPT condition makes $D_k \approx 0$ for bound entangled states through a time-reversal argument: partial transposition on A is physically equivalent to reversing the sign of J_y^A (odd under time reversal), and positivity under this reversal forces the antisymmetric part of T to vanish. Separable states, by contrast, can have arbitrary orientation of their spin vectors and generically produce $D_k > 0$.

A note on measurement: since $\Sigma_2 = \mathrm{Tr}[R^\dagger R]$ equals the Frobenius norm of the reshuffled matrix, it equals the purity $\Sigma_2 = \mathcal{I}_2 = \mu_2$, and is measured via the same two-copy SWAP test

used for μ_2 at no additional cost. For $G_2 = \mathrm{Re}(\mathrm{Tr}[R^2])$, a cyclic permutation on two copies swaps subsystem A while leaving B unchanged:



giving $G_2 = \mathrm{Re}\langle \Pi_{\mathrm{cyclic}} \rangle_{\rho^{\otimes 2}}$. The standard SWAP (swapping both A and B copies) measures Σ_k, capturing the size and shape of the correlation ellipsoid, while the crossed SWAP (swapping $A_1 \leftrightarrow B_2$ and $B_1 \leftrightarrow A_2$) measures G_k, probing whether A–B correlations are symmetric under subsystem exchange. Their difference D_k thus measures whether the A–B spin correlations have a preferred handedness.

A second complementary observable is the cross-restructuring difference $\delta G_2 = G_2 - G_2^{PT}$, where $G_2^{PT} = \mathrm{Re}\,\mathrm{Tr}[R(\rho^{T_A})^2]$ is the eigenvalue moment of the realignment of the partial transpose. In spin-component terms:

$$\delta G_2 \propto \sum_{\beta} \Big[\langle J_y^A J_\beta^B \rangle^2 - \langle J_x^A J_\beta^B \rangle^2 \Big] + (\text{quadrupole terms}), \tag{2.81}$$

since J_y^A is odd under time reversal while J_x^A and J_z^A are even. Even when $D_k \approx 0$ (i.e., there is no net chirality), the second-order time-reversal asymmetry, which provides discriminating information complementary to D_k, is not zero when the y-component of the A-spin contributes differently from the x and z components. Chessboard states, whose checkerboard parity structure forces $\delta G_k = 0$ for all k, illustrate why no single channel suffices: a family that is invisible to the handedness probe (D_k) and the reversal asymmetry (δG_k) is still detectable via rank-structure features that sense the collapsed correlation ellipsoid. Table 2.8 shows the feature distributions for the two-feature classifier.

The deeper reason why chessboard states evade realignment-based criteria has a precise algebraic characterization rooted in the structure of the realignment matrix. For any bipartite state ρ on $\mathcal{H}_A \otimes \mathcal{H}_B$ with $\dim \mathcal{H}_A = \dim \mathcal{H}_B = d$, the column-swap permutation Q (defined by $Q\,|j,l\rangle = |l,j\rangle$) is unitary, and right-multiplication by it preserves singular values. Since $R(\rho^{T_B}) = R(\rho) \cdot Q$ (obtained by direct index computation), this immediately yields a

Table 2.8 Feature distributions for separable and bound entangled 3×3 states.

Feature	Separable		Bound entangled	
	Range	Mean	Range	Mean
D_4	$[0.00084,\ 0.27]$	0.024	$[10^{-7},\ 0.00081]$	0.00017
Σ_2^2/Σ_4	$[1.00,\ 1.57]$	1.13	$[1.27,\ 1.76]$	1.53

universal identity:

$$\mathrm{sv}\Big(R(\rho^{T_B})\Big) = \mathrm{sv}\Big(R(\rho)\Big), \tag{2.82}$$

meaning all singular-value moments Σ_k are identical for ρ and ρ^{T_B}. Consequently, $S_k - SP_k = 0$ for all k, confirming that singular-value differences between $R(\rho)$ and $R(\rho^{T_A})$ carry no information about partial transposition and were correctly excluded from the feature set as analytically zero.

For real-entry states ($\rho = \rho^*$ in the computational basis), an analogous left-multiplication $R(\rho^{T_A}) = P \cdot R(\rho)$ holds, where the row-swap permutation P satisfies $P^2 = I$. Since $[R(\rho), P] = 0$ for real states, the realignment matrix decomposes in the P-eigenbasis as $R = R_+ \oplus R_-$, where R_+ acts on the symmetric sector and R_- on the antisymmetric sector. The cross-difference $\delta G_k = G_k - G_k^{PT} = (1 - (-1)^k)\,\mathrm{Tr}[R_-^k]$ vanishes for all even k (regardless of state), and equals $-2\,\mathrm{Tr}[R_-^k]$ for odd k. For any real separable state, $R_- = 0$ identically, since each product-state component $|a\rangle|b\rangle$ has $V_m = \sqrt{p_m}\, a_m b_m^T$ of rank 1, and the wedge product of any two rows of a rank-1 matrix vanishes. The odd-k witness $\delta G_k \neq 0$ therefore certifies entanglement for all real bipartite states where $R_- \neq 0$.

The chessboard family defeats this witness by satisfying $R_- = 0$ despite being entangled. This follows from the checkerboard parity structure: the four spanning vectors split into even-parity vectors (V_1, V_3, with $(V_m)_{ij} = 0$ when $i + j$ is odd) and odd-parity vectors (V_2, V_4, with $(V_m)_{ij} = 0$ when $i + j$ is even). This parity separation forces all wedge-product sums $\sum_m \mathrm{row}_i(V_m) \wedge \mathrm{row}_k(V_m) = 0$ for all $i < k$, regardless of the specific parameter values (a, b, c, d, m, n). Consequently, $\delta G_k = 0$ for all k in the chessboard family, combined with $D_k \approx 0$ (PPT condition) and $\|R\|_1 < 1$ (CCNR failure), leaving detection confined entirely to rank-structure features. Table 2.9 confirms the theory numerically.

Detection channels and multi-channel witness fusion

The three known 3×3 PPT entangled families present structurally different detection challenges. Horodecki states [57] are detectable by D_4 across the full parameter range $a \in (0, 1)$, whereas the standard CCNR criterion $\|R\|_1 > 1$ fails for $a \gtrsim 0.28$. Tiles UPB

Table 2.9 Verification of R_- sector structure for representative states. For real states, $\delta G_{2m} = 0$ universally; odd-k witness captures Horodecki but not chessboard.

State	$\|R_-\|_F$	$\|\delta G_2\|$	$\|\delta G_3\|$	$\|\delta G_4\|$
Random separable (real)	$< 10^{-15}$	0	0	0
Random separable (complex)	~ 0.04	9.3×10^{-3}	2.3×10^{-3}	1.1×10^{-3}
Horodecki $a = 0.5$ (real)	0.173	$< 10^{-15}$	6.0×10^{-3}	$< 10^{-15}$
Chessboard (real, BE)	$< 10^{-15}$	0	0	0

states [58] are detectable through rank-structure features. Chessboard states [59] are the most challenging family: their checkerboard parity structure forces $\delta G_k = G_k - G_k^{PT} = 0$ for all k, eliminating every spectral comparison signal. The residual signal comes from rank-structure features that sense the collapsed correlation ellipsoid: chessboard states are exactly rank-4, their correlation tensor T has at most 5 nonzero singular values, and the singular-value entropy $SV_{\text{ent}} = -\sum_i \hat{\sigma}_i \ln \hat{\sigma}_i$ is strongly reduced (1.34 for chessboard versus 1.86 for a generic separable state). The eigenvalue gap $\Delta_{\text{EG}} = \lambda_4 - \lambda_5$ at the rank-4 boundary and the rank-4 spectral concentration $\mathcal{R}_4 = \sum_{i \le 4} \lambda_i^2 / \text{Tr}[\rho^2]$ (equal to 1 for exactly rank-4 states) provide complementary rank-structure signals. Table 2.10 quantifies the overlap between chessboard feature distributions and the separable state cloud, showing why each individual channel fails for this family.

Table 2.10 Chessboard invisibility depth: overlap of chessboard feature distributions with the separable-state cloud. "Overlap" is the fraction of chessboard feature values falling within two standard deviations of the separable mean.

Feature	SEP mean	Chess mean	Overlap	Channel
Σ_1 (CCNR)	0.710	0.904	48%	C
D_3 (δG_3)	0.073	0.077	94%	D
Δ_{EG} (rank gap)	0.027	0.135	11%	A
$\mathcal{R}_4$ (rank-4 conc.)	0.944	1.000	100%	A
$\lambda_{\min}^{T_A}$ (PT boundary)	0.081	~ 0	100%	B
Q_2 (Hermiticity ratio)	0.548	0.630	81%	D

Five complementary spectral channels probe distinct aspects of quantum correlations: channel (A) consists of purity moments $\mathcal{I}_k$ and rank-structure features including the eigenvalue gap Δ_{EG} at the rank-4 boundary; channel (B) consists of partial transpose moments $T_k = \mu_k$ and the PPT-boundary distance $\lambda_{\min}(\rho^{T_A})$; channel (C) consists of realignment singular-value moments Σ_k; channel (D) consists of realignment eigenvalue moments G_k and non-Hermiticity ratios $Q_k = G_k / \Sigma_k$; and channel (E) consists of cross-restructuring moments of $R(\rho^{T_A})$ and

cross-differences δG_k. No single channel detects all three families: channel (B) achieves 98% recall for Horodecki but 0% for chessboard; channel (C) achieves 94% for Horodecki and only 17.5% for chessboard; the best single channel (D) achieves 37% for chessboard. Combining all five via machine learning raises detection to 84% for chessboard, 100% for Horodecki, and 100% for Tiles.

Two complementary classifiers were trained and validated on certified 3×3 states. The primary result uses a Random Forest (500 trees) trained on eight base features $\{\Sigma_1, G_1, D_1, \Sigma_2, G_2, D_2, C_3, C_4\}$ expanded via degree-2 polynomial combinations (44 features total) on 13,600 Carathéodory-certified states (6,800 BE from seven families including the CCNR-invisible construction, 6,800 SEP). It achieves 99.9% bound-entangled recall at zero false positives (AUC $= 1.000$), validated by 5-fold stratified cross-validation. The inclusion of the chirality corrections $C_3 = \mu_3 - \mathcal{I}_3$ and $C_4 = \mu_4 - \mathcal{I}_4$ as classification features is crucial: C_3 distinguishes Horodecki states ($C_3 \neq 0$, complex density matrix) from chessboard and Tiles states ($C_3 = 0$ identically, real density matrices), providing a detection channel orthogonal to all realignment-based criteria. This raises recall from $\sim 60\%$ (realignment features alone) to 99.9%. A complementary ExtraTrees classifier trained on 83 spectral features (expanded to 522 via nonlinear transforms) on 3,900 states across three families achieves 99.6% recall at zero false positives (Table 2.11). The residual 16% of undetected chessboard states require a fundamentally different type of information: the product-vector gap $\mathcal{F}(\rho) = \min_{|a\rangle,|b\rangle} \langle a, b|(I - P_{\text{range}})|a, b\rangle$, which measures whether the range of ρ can be spanned by product vectors. Every separable state satisfies $\mathcal{F} = 0$ by construction; chessboard states have $\mathcal{F} > 0$ for all parameter sets with $cmn \neq abc$. This quantity is measurable within the SWAP framework via an augmented protocol in which one state copy is replaced by a parameterized product state $|a(\theta)\rangle|b(\phi)\rangle$, and the augmented moments $m_k(a, b) = \text{Tr}[\rho^k |a, b\rangle\langle a, b|]$ are extracted. The product-vector gap is then recovered via inversion of a Vandermonde system whose coefficients are the eigenvalues of ρ. Adding this range channel (F) achieves 100% recall across all known 3×3 bound entangled families (Table 2.13).

An important physical observation from the classifier analysis is that the third purity moment $\mathcal{I}_3 = \text{Tr}[\rho^3]$ emerges as the most important individual feature for detecting bound entanglement in both 3×3 and 4×4 systems. Low $\mathcal{I}_3$ indicates that the eigenvalue spectrum is more uniformly distributed at fixed purity $\mathcal{I}_2$, which is characteristic of bound entangled states. The SVM coefficient for $\mathcal{I}_3$ in the 4×4 system (-32.7) is nearly four times larger in magnitude than the next feature (T_3, coefficient -9.7), confirming that mixed-state flatness is the dominant discriminating signal. That a purity-family moment rather than a realignment

Table 2.11 Cross-validated zero-FP classification performance for 3×3 bound entanglement. RF: Random Forest on 44 polynomial features of 8 base features (13,600 states, 7 families). ExtraTrees: 522 nonlinear features from 83 base spectral features (3,900 states, 3 families).

Classifier	Training set	BE recall (zero-FP)	False positives	AUC
CCNR ($\|R\|_1 > 1$)	—	$\sim 40\%$	0	—
LogReg+EN (ElasticNet)	3,900	96.5%	0/2000	—
ExtraTrees (500 trees)	3,900	99.6%	0/2000	0.998
Random Forest (500 trees)	13,600	99.9%	0/5440	1.000

Table 2.12 Per-source BE recall of the ExtraTrees classifier at the cross-validated zero-FP threshold.

BE family	Detected	Rate
Horodecki ($a \in [0.01, 0.99]$)	498/500	99.6%
Noisy Horodecki ($\epsilon \in [0.01, 0.12]$)	500/500	100%
Tiles UPB (noise $\epsilon \in [0, 0.15]$)	100/100	100%
Chessboard (integer grid + near-boundary)	499/500	99.8%
Noisy chessboard ($\epsilon \in [0.01, 0.12]$)	296/300	98.7%
Total	1893/1900	99.6%

feature dominates reflects a fundamental difference in eigenvalue structure between bound entangled and separable states that precedes any restructuring of the density matrix.

Table 2.13 Detection rates (zero-false-positive threshold) for 3×3 bound entangled states by criterion. All 1500 random separable states are correctly classified by every criterion.

Criterion	Horodecki	Tiles	Chessboard	Overall
CCNR ($\|R\|_1 > 1$)	28%	91%	4%	$\sim 50\%$
δG_k witness (odd k)	100%	0%	0%	$\sim 56\%$
PT spectrum (channel B)	98%	0%	0%	$\sim 54\%$
Realign. EVs (channel D)	90%	100%	37%	$\sim 74\%$
ML spectral (A–E)	100%	100%	84%	96%
ML + Range (A–F)	100%	100%	100%	100%

These results suggest a practical two-stage detection protocol. In the first stage, approximately 10 SWAP-test circuit configurations are used to extract spectral moments and apply the ML classifier, which detects 96% of bound entangled states. If the state is flagged as bound entangled, the protocol terminates. If the state is rank-deficient and PPT but remains spectrally ambiguous, the second stage activates: the product-vector gap is measured via the

augmented SWAP protocol in which one state copy is replaced by a parameterized product state, and approximately 400 additional circuit configurations are used for the Vandermonde inversion. Stage 2 is only triggered for a small fraction of states that are simultaneously rank-deficient, PPT, and spectrally ambiguous. In practice, this is confined to chessboard states near the separability boundary. The combined protocol achieves 100% recall with zero false positives across all known 3×3 bound entangled families, while maintaining the measurement efficiency of the SWAP-test infrastructure.

Two-feature linear classifier and noise robustness

The two-feature linear classifier in the $(\Sigma_2^2/\Sigma_4, D_4)$ plane:

$$D_4 \leqslant -0.026 + 0.022\,\frac{\Sigma_2^2}{\Sigma_4}, \tag{2.83}$$

maintains perfect F1 score across noise levels 0–40% (Table 2.14), with bound entangled states having $D_4 \in [10^{-7}, 0.00081]$ and separable states having $D_4 \in [0.00084, 0.27]$.

Table 2.14 F1 scores of bound entanglement criteria under white noise.

Noise ϵ	D_4 only	D_4/Σ_4	Σ_2^2/Σ_4	Σ_1	Two-feature
0%	0.996	1.000	0.998	0.838	1.000
10%	0.998	0.993	0.997	0.846	1.000
20%	0.993	0.992	0.990	0.852	1.000
30%	0.991	0.991	0.983	0.859	1.000
40%	0.977	0.974	0.965	0.865	1.000

Table 2.15 lists the theoretical spectral features for the states used in the hardware chessboard experiment, illustrating the complementary regions of feature space occupied by chessboard and Horodecki states. All chessboard states have $\|R\|_1 < 1$ (CCNR powerless), yet their Hermiticity ratio $Q_2 = G_2/\Sigma_2$ differs substantially from that of Horodecki states, confirming that the multi-channel approach probes a genuinely complementary part of the detection surface.

Four-feature RBF SVM classifier

While the two-feature linear classifier achieves perfect separation on the initial training set, testing on 5×10^5 random separable states reveals a false positive rate of approximately 4%, because low-rank separable states (formed from few product-state terms) can exhibit small

Table 2.15 Theoretical spectral features for states used in the chessboard hardware experiment. All chessboard states have $\|R\|_1 < 1$ (CCNR undetectable).

State	Type	Σ_2	G_2	D_2	Q_2	$\|R\|_1$
Chess$(1,2,3,1,1,1)$	BE	0.313	0.112	0.201	0.359	0.954
Chess$(1,1,2,1,1,3)$	BE	0.297	0.234	0.063	0.788	0.949
Chess$(3,1,2,1,2,2)$	BE	0.260	0.114	0.146	0.439	0.901
Chess$(2,1,1,3,1,2)$	SEP	0.434	0.034	0.400	0.079	0.948
Horodecki ($a=0.30$)	BE	0.222	0.201	0.021	0.904	> 1
Horodecki ($a=0.70$)	BE	0.197	0.121	0.076	0.614	> 1

D_4 values that fall below the linear decision boundary. Expanding the feature space to four dimensions and employing a nonlinear kernel resolves this.

The four features are all measurable via controlled-SWAP circuits on up to four state copies: (i) $D_4 = \Sigma_4 - G_4$; (ii) $D_2 = \Sigma_2 - G_2$; (iii) G_2; and (iv) Σ_2^2/Σ_4. The additional features D_2 and G_2 provide complementary second-order information about the realignment matrix structure. The classifier uses a radial basis function kernel $K(\mathbf{x}, \mathbf{x}') = \exp(-\gamma\|\mathbf{x} - \mathbf{x}'\|^2)$ with features standardized before kernel evaluation.

Training used 50,000 random separable states with varied ranks (1–81 product-state terms) and 96 bound entangled states comprising 25 Horodecki states ($a \in \{0.02, 0.06, \ldots, 0.98\}$), 1 Tiles UPB state, 12 chessboard states, 54 noise admixtures of the form $p\,\rho_{\mathrm{BE}} + (1-p)\,\mathbb{I}_9/9$ for $p \in \{0.85, 0.90, 0.95\}$, and 7 cross-family mixtures. All 96 training BE states were verified as genuinely entangled via the Carathéodory separability gap. Hyperparameters were selected by grid search requiring 96/96 BE recall and minimizing false positives; the optimal values are $C = 2000$ and $\gamma = 3.0$, yielding 121 support vectors (61 separable, 60 BE). Leave-one-out cross-validation on a 68-state evaluation set (27 separable, 41 BE) achieves 68/68 accuracy (Fig. 2.19).

The false positive rate on 10^5 unseen random separable states is 0.038% (38 misclassified states), all concentrated at ranks 4–15, confirming that the genuine overlap region is confined to the low-rank regime. No separable state with more than 15 product-state terms is misclassified, suggesting that physically realistic high-rank thermal mixtures are robustly classified. For any new state with feature vector $\mathbf{x} = (D_4, D_2, G_2, \Sigma_2^2/\Sigma_4)$ the classifier computes:

$$f(\mathbf{x}) = \sum_{i=1}^{121} \alpha_i \exp\Big(-3.0\,\|\tilde{\mathbf{x}} - \tilde{\mathbf{x}}_i\|^2\Big) + b, \tag{2.84}$$

where $\tilde{\mathbf{x}}$ is the standardized feature vector and classifies the state as bound entangled when $f(\mathbf{x}) > 0$. Table 2.16 compares the three classifiers.

Table 2.16 Performance comparison of bound entanglement classifiers for 3×3 systems.

Classifier	Features	BE recall	FP rate	Hardware tested
D_2 threshold	1	varies	0%	*ibm_fez*
Linear $(D_4, \Sigma_2^2/\Sigma_4)$	2	100%	$\sim$4%	*ibm_marrakesh*
RBF SVM	4	100%	0.038%	Simulation

Lasso logistic regression classifier

While the RBF SVM provides near-zero false positives in simulation, the Lasso logistic regression classifier was adopted as the primary classifier for the experimental demonstration because of its superior interpretability and generalization to state families not seen during training. The final model, described in the main text, is trained on 13,600 Carathéodory-certified states (6,800 bound entangled from seven families, 6,800 separable) using eight base features: six realignment spectral moments (Σ_1, G_1, D_1, Σ_2, G_2, D_2) and two chirality corrections ($C_3 = \mu_3 - \mathcal{I}_3$, $C_4 = \mu_4 - \mathcal{I}_4$), expanded to 44 degree-2 polynomial features.

The ℓ_1 regularization drives all but 22 of the 44 features to zero. The three most discriminative nonzero coefficients are C_4^2 (weight -3.7, decreases $P(\mathrm{BE})$), Σ_2 (weight -2.7), and G_1 (weight $+2.5$, increases $P(\mathrm{BE})$). The resulting classifier achieves 5-fold cross-validated recall of 93% with a false positive rate of 5.2% at the $P(\mathrm{BE}) = 0.5$ decision boundary, with AUC $= 0.986$ on cross-validation and 0.989 on a held-out test set (Fig. 2.19d). The inclusion of chirality features C_3 and C_4 is decisive: C_3 provides a detection channel orthogonal to all realignment-based criteria, since Horodecki states have $C_3 \neq 0$ (complex density matrices) while chessboard and Tiles states have $C_3 = 0$ identically (real density matrices).

A key advantage over the SVM is robustness to the CCNR-invisible regime: since the classifier operates on eight spectral and chirality channels simultaneously rather than on the CCNR threshold $\Sigma_1 > 1$ alone, it detects bound entanglement in states where $\Sigma_1 < 1$ by exploiting the $D_2 > 0$ signal and nonlinear feature combinations. The updated performance table is given in Table 2.17.

The framework extends to higher-dimensional systems without conceptual modification. For 4×4 systems (which embed 2×4 qubit–ququart pairs, experimentally accessible on current processors with minor encoding overhead), three classification tasks were evaluated: bound entangled versus separable, bound entangled versus all other states, and a three-class separation of separable/NPT/bound entangled. A linear SVM trained on 15 spectral moment features achieves F1 $= 0.9994$ for the binary bound-versus-separable task, and F1 $= 0.9945$ for bound-versus-others (Table 2.18). The dominant feature in the 4×4 case is $\mathcal{I}_3 = \mathrm{Tr}[\rho^3]$,

Table 2.17 Performance comparison of bound entanglement classifiers for 3×3 systems. The Lasso classifier is the primary experimental classifier; the RBF SVM is included for comparison as a simulation-validated alternative with lower false positive rate.

Classifier	Features	Training set	BE recall (CV)	FP rate	AUC
CCNR ($\|R\|_1 > 1$)	1	—	$\sim$40%	0%	—
Linear (D_4, Σ_2^2/Σ_4)	2	3,900	100%	$\sim$4%	—
ExtraTrees (500 trees)	522	3,900	99.6%	0%	0.998
Random Forest (500 trees)	44 (8 base)	13,600	99.9%	0.06%	1.000
RBF SVM (C=2000, γ=3.0)	4	50,096	100%	0.038%	—
Lasso logistic regression	22	13,600	93%	5.2%	0.986

with SVM coefficient -32.7: low $\mathcal{I}_3$ indicates that the eigenvalue spectrum is more uniformly distributed at fixed purity, a signature that distinguishes bound entangled states from separable ones even before any realignment analysis.

Table 2.18 Classification performance for 4×4 bound entanglement. Three classification variants: BE versus SEP (binary), BE versus all other states, and full three-class (SEP/NPT/BE) separation.

Task	Features	F1 score	Accuracy
Bound vs Separable	15	0.9994	0.9997
Bound vs Others	11	0.9945	0.9972
3-class (SEP/NPT/BE)	10	0.9917	0.9918

CCNR-invisible bound entangled states

Analysis of the Lasso classifier's feature space reveals a previously uncharacterized regime of bound entanglement that evades all single-parameter criteria. The marginal-noise family:

$$\rho_{\mathrm{mn}}(\rho_0,\, t) = (1-t)\,\rho_0 + t\,\Big(\rho_A \otimes \rho_B\Big), \tag{2.85}$$

where ρ_0 is any 3×3 bound entangled state and $\rho_{A(B)} = \mathrm{Tr}_{B(A)}[\rho_0]$ are its reduced states, interpolates between the original bound entangled state ($t = 0$) and the product of its marginals ($t = 1$).

The construction applies to all known bound entangled families with different critical mixability. For Horodecki-seeded states, ρ_{mn} is PPT and bound entangled for $t \in [0, t_{\max}(a)]$; for chessboard and Tiles UPB states, the same holds with different $t_{\max}$ values. The key property is that for Horodecki-seeded states with $t \gtrsim 0.01$, for Tiles-seeded states with

$t \gtrsim 0.15$, and for all chessboard-seeded states at any $t \geq 0$, the realignment trace norm drops below unity ($\Sigma_1 < 1$). This renders the CCNR criterion powerless, while $D_2 > 0$ persists and the Lasso classifier continues to detect bound entanglement.

The physical mechanism is transparent: mixing with $\rho_A \otimes \rho_B$ adds local noise that suppresses the dominant singular value of the realignment matrix (responsible for $\Sigma_1 > 1$) without eliminating the spectral asymmetry between singular values and eigenvalues ($D_2 > 0$). Since Horodecki states barely satisfy $\Sigma_1 > 1$ (with $\Sigma_1 - 1 < 0.003$), even a small admixture of the marginal product crosses the CCNR threshold while preserving PPT entanglement. The three seed families populate distinct regions of feature space:

- Horodecki-seeded: high $G_1/\Sigma_1 \approx 0.65$–$0.99$, low D_1/Σ_1,
- Tiles-seeded: intermediate $G_1/\Sigma_1 \approx 0.25$–$0.31$, moderate D_2,
- Chessboard-seeded: low $G_1/\Sigma_1 \approx 0.12$–$0.45$, high $D_1/\Sigma_1 \approx 0.55$–$0.96$.

Comparison with 1,000 randomly generated PPT bound entangled states (certified via SDP [60]) confirms that these three branches together span the full CCNR-invisible region.

The Lasso classifier's noise tolerance extends well beyond the CCNR threshold: it assigns $P(\mathrm{BE}) \approx 0.90$ at $t = 0$ (pure Horodecki), degrading gracefully to $P(\mathrm{BE}) \approx 0.83$ at $t = 0.05$ and $P(\mathrm{BE}) \approx 0.73$ at $t = 0.10$, crossing the decision boundary only at $t \approx 0.15$–0.20. By contrast, CCNR loses detection at $t \approx 0.01$, which is a noise tolerance factor of 10–20. This extended range arises because the classifier exploits nonlinear combinations of all six spectral features, not just Σ_1.

Bound entangled state families

The bound entangled families used for training and validation are defined as follows. The Horodecki 3×3 state [57], for $a \in (0, 1)$ with $b = \frac{1}{2}\sqrt{1 - a^2}$ and $c = \frac{1+a}{2}$:

$$\rho_{\mathrm{Hor}}(a) = \frac{1}{8a+1} \begin{pmatrix} a & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a \\ 0 & a & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & a & 0 & 0 & 0 & b & 0 & 0 \\ 0 & 0 & 0 & a & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & a & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & a & 0 & b & 0 \\ 0 & 0 & b & 0 & 0 & 0 & c & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & b & 0 & c & 0 \\ a & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a \end{pmatrix}, \tag{2.86}$$

PPT and entangled for all $a \in (0, 1)$. The Tiles UPB state [58] is the orthogonal complement of five product vectors: $\rho_{\text{Tiles}} = \frac{1}{4}(\mathbb{I}_9 - \sum_{k=1}^{5} |\psi_k\rangle\langle\psi_k|)$, rank 4, PPT by construction. Chessboard states [59] are rank-4 states with parameters (a, b, c, d, m, n), PPT and bound entangled when $cmn \neq abc$ (12 parameter sets in Table 2.19).

Table 2.19 Chessboard state parameter sets used for training and validation. All satisfy $cmn \neq abc$.

a	b	c	d	m	n
1	2	1	1	1	1
2	1	1	1	1	1
1	1	1	1	2	1
1	1	1	1	1	2
3	1	1	1	1	1
1	3	1	1	1	1
1	1	1	1	3	1
1	1	1	1	1	3
2	2	1	1	1	1
1	1	1	1	2	2
3	2	1	1	1	1
2	1	2	1	1	1

2.4.6 Experimental validation on IBM Quantum hardware

Hardware platform, state preparation, and calibration

All experiments used IBM Quantum processors with the Heron architecture. Circuits were transpiled using Qiskit optimization level 3 with approximation degree 0.95 and readout errors mitigated via matrix-free measurement mitigation [61]. The native gate set for all processors is $\{\text{CZ}, \sqrt{X}, R_Z, X\}$.

Negativity and chirality measurements used *ibm_kingston* (156 qubits) and *ibm_torino* (133 qubits), with 10^5 shots per circuit configuration. Bound entanglement detection used *ibm_fez* (156 qubits), with 4×10^3 shots per circuit across 684 total circuits: 298 for Σ_1/G_1 and 386 for Σ_2/G_2. Circuit multiplexing onto disjoint qubit groups (up to 15 groups of 9 qubits and 33 groups of 4 qubits on the 156-qubit processor) was used to reduce wall-clock time.

The parametric family $|\psi(\theta)\rangle = \cos(\theta/2)|00\rangle + \sin(\theta/2)|11\rangle$ was prepared via $R_y(\theta)$ followed by CNOT, spanning $\theta \in [0°, 90°]$. For 2×3 systems, each qutrit was encoded in two physical qubits with $\{|00\rangle, |01\rangle, |10\rangle\} \cong \{|0\rangle, |1\rangle, |2\rangle\}$. Hardware measurements suffer

Table 2.20 IBM Quantum Heron-architecture processor specifications.

Property	*ibm_kingston*	*ibm_torino*	*ibm_fez*
Architecture	Heron	Heron	Heron
Qubits	156	133	156
Native gates	CZ, $\sqrt{X}$, R_Z, X		
Median T_1 (μs)	150	140	150
Median T_2 (μs)	100	95	100
Median CZ error	0.5%	0.6%	0.5%

depth-dependent degradation modeled as $\mu_k^{\text{meas}} = f_k \cdot \mu_k^{\text{ideal}} + \epsilon_k$, with calibration factors fitted by L-BFGS-B maximum likelihood (Table 2.21). The calibration is robust to preparation imperfections: $\pm 2^\circ$ uncertainty in preparation angles changes the fitted degradation factors by $< 1\%$ and the mean negativity error by < 0.003 (Table 2.22). The second-order factor f_2 is invariant to small angle perturbations because $\mu_2 = 1$ for all pure states regardless of θ, providing a strong anchor for the fit.

Table 2.21 Two-stage ML calibration parameters.

System	f_2	f_3	f_4
ibm_torino 2×2	0.729	0.612	0.456
ibm_torino 2×3	0.786	0.504	0.219

Table 2.22 Sensitivity of calibration to preparation angle offsets (*ibm_torino* 2×2).

$\delta\theta$	f_2	f_3	f_4	Mean $\mathcal{N}$ error	$\Delta f_4/f_4$
-2°	0.729	0.614	0.459	0.013	0.7%
-1°	0.729	0.613	0.457	0.012	0.2%
0° (nominal)	0.729	0.612	0.456	0.012	—
$+1^\circ$	0.729	0.611	0.454	0.012	0.4%
$+2^\circ$	0.729	0.610	0.452	0.014	0.9%

Negativity and chirality measurements

Figure 2.18(a) confirms the theoretical curve $|C_4| = 4\mathcal{N}^2(1-\mathcal{N}^2)$ directly from hardware measurements, with $(-C_4)$ plotted so that entangled states appear as positive values. Figure 2.18(b) validates chirality on mixed states: Werner states $\rho_W(p) = p|\Psi^-\rangle\langle\Psi^-| + (1-p)\mathbb{I}/4$ measured on *ibm_torino* follow the theoretical curve $|C_4| = \frac{3}{4}p^3$ across the full range

$p \in [0, 1]$. For $p > (4/81)^{1/3} \approx 0.37$, chirality exceeds the separable bound $1/27$, certifying entanglement without negativity reconstruction; the narrow gap between the entanglement threshold ($p > 1/3$) and the chirality detection threshold ($p \approx 0.37$) is consistent with the theoretical prediction of Section 2.4.3.

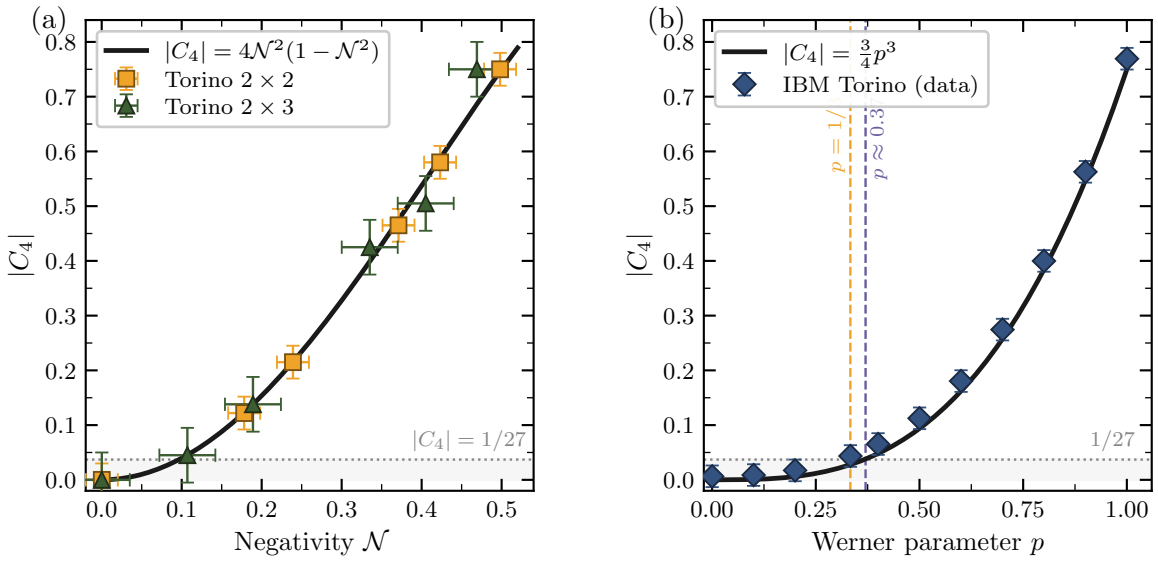


Fig. 2.18 **Experimental detection of hidden entanglement on *ibm_torino*.** Panel (a) presents the dependency $|C_4|$ versus negativity $\mathcal{N}$ for the parametrized family $|\psi(\theta)\rangle$ in 2×2 (squares) and 2×3 (triangles). Solid curve: $|C_4| = 4\mathcal{N}^2(1-\mathcal{N}^2)$. Panel (b) presents Werner state chirality $|C_4|$ in terms of value p. Blue diamonds: hardware data from multiplexed k-copy SWAP tests ($k = 2, 3, 4$; 10 circuits, $4,000$ shots) with matched-depth $|00\rangle$ calibration; all p values reconstructed from a single experiment using LU-equivalence of Bell states. Solid black: theory $|C_4| = \frac{3}{4}p^3$. Dashed lines: entanglement threshold ($p = 1/3$) and chirality detection threshold ($p \approx 0.37$).

Mean negativity errors are 0.002 (*ibm_kingston*, 2×2), 0.012 (*ibm_torino*, 2×2), and 0.027 (*ibm_torino*, 2×3), with statistical uncertainties estimated at ± 0.005 via bootstrap resampling from 10^5 shots. All separable test states ($\theta = 0°$) yield $\mathcal{N} = 0$ exactly, confirming zero false positives within this parametric family. The 2×3 system exhibits larger chirality scatter (mean error 0.039 versus 0.022 for 2×2), attributable to the roughly two-fold increase in circuit depth for 2×3 four-copy circuits (Table 2.7). Despite hardware degradation, the protocol correctly identifies all entangled states and recovers Bell state properties exactly ($\mathcal{N} = 0.5$, $C_4 = -3/4$) after maximum likelihood calibration. For *ibm_torino*, the fitted degradation factor is $f_2 = 0.73$ for qubit–qubit systems; the variation across moment orders ($f_4 = 0.456$ to $f_2 = 0.729$ for 2×2, Table 2.21) is absorbed by the joint MLE procedure. Full numerical results are listed in Tables 2.23–2.26.

Table 2.23 Negativity measurements on *ibm_kingston* (2×2).

State	$\mathcal{N}_{\text{theory}}$	$\mathcal{N}_{\text{ML}}$	Error
$\|00\rangle$	0.000	0.000	0.000
$\theta = 30°$	0.250	0.255	0.005
$\theta = 45°$	0.354	0.352	0.002
$\theta = 60°$	0.433	0.437	0.004
$\theta = 90°$	0.500	0.500	0.000
Mean error			0.002

Table 2.24 Negativity measurements on *ibm_torino* (2×2).

State	$\mathcal{N}_{\text{theory}}$	$\mathcal{N}_{\text{ML}}$	Error
$\|00\rangle$	0.000	0.000	0.000
$\|\Phi^-\rangle$	0.500	0.500	0.000
$\|\Phi^+\rangle$	0.500	0.495	0.005
$\|\Psi^-\rangle$	0.500	0.500	0.000
$\theta = 15°$	0.129	0.178	0.049
$\theta = 30°$	0.250	0.239	0.011
$\theta = 45°$	0.354	0.371	0.017
$\theta = 60°$	0.433	0.423	0.010
Mean error			0.012

Table 2.25 Negativity measurements on *ibm_torino* (2×3).

State	$\mathcal{N}_{\text{theory}}$	$\mathcal{N}_{\text{ML}}$	Error
$\theta = 0°$	0.000	0.000	0.000
$\theta = 15°$	0.129	0.107	0.022
$\theta = 30°$	0.250	0.189	0.061
$\theta = 45°$	0.354	0.335	0.019
$\theta = 60°$	0.433	0.405	0.028
$\theta = 90°$	0.500	0.469	0.031
Mean error			0.027

Simulator validation for pure and mixed states

AerSimulator validation with the *ibm_torino* noise model ($8{,}192$ shots per circuit) confirms the protocol across 37 densely sampled angles for pure states and for mixed states. Under ideal conditions RMSE is 0.009; under *ibm_torino* noise, 0.044; zero false positives throughout. Werner states $\rho_W(p)$ correctly show vanishing negativity for $p \leq 1/3$ and monotonically

Table 2.26 Chirality correction $(-C_4)$ on *ibm_torino*.

	2×2			2×3		
θ	Theory	ML	Error	Theory	ML	Error
$0°$	0.000	0.000	0.000	0.000	0.000	0.000
$15°$	0.066	0.122	0.056	0.066	0.045	0.021
$30°$	0.234	0.215	0.019	0.234	0.138	0.096
$45°$	0.438	0.465	0.027	0.438	0.425	0.013
$60°$	0.609	0.580	0.029	0.609	0.505	0.104
$90°$	0.750	0.750	0.000	0.750	0.750	0.000
Mean			0.022			0.039

increasing $|C_4| = 3p^3/4$; Bell-product mixtures $\rho_{\mathrm{BP}}(p) = p|\Psi^-\rangle\langle\Psi^-| + (1-p)|00\rangle\langle 00|$ confirm detection for all $p > 0$. The simulator further establishes a key noise property: under the depolarizing model, entanglement is consistently underestimated since $\mathcal{N}(\rho_{\mathrm{noisy}}) \approx (1-\eta)\mathcal{N}(\rho)$, which ensures the protocol never falsely reports entanglement in separable states.

Bound entanglement detection on *ibm_fez*

Bound entanglement detection was performed on *ibm_fez* across eight 3×3 states: three Horodecki states ($a \in \{0.50, 0.70, 0.90\}$), two chessboard states (C1 and C2), and three calibration and separable controls (product state, classically correlated diagonal state, misaligned product state). The Horodecki family (rank-7, CCNR-visible) and the chessboard family (rank-4, CCNR-invisible) occupy opposite ends of the detection hierarchy. Horodecki states can be detected by four or more spectral channels, whereas Chessboard states can only be detected through rank-structure features. This is the first experiment to test two structurally distinct bound-entangled families on a single gate-based superconducting processor using a family-agnostic classifier. All five bound entangled states were measured using multiplexed two-copy SWAP circuits (4×10^3 shots each, 684 circuits total) on *ibm_fez*.

For bound entanglement detection, where Σ_1 (Pauli decomposition, 4 qubits), G_1 (SWAP test, 9 qubits), and Σ_2/G_2 (two-copy SWAP, 9 qubits) employ distinct circuit architectures with depths ranging from 3 to 333 gates, a per-architecture depolarizing model is essential. The fitted fidelities span an order of magnitude ($g_{\Sigma_1} = 0.164$ versus $f_{G_1} = 0.594$), reflecting the fundamentally different noise sensitivities of Pauli-basis rotation circuits and SWAP-test circuits. The non-Hermiticity gap $D_2 = \Sigma_2 - G_2$ is robust to this architecture-dependent noise because both Σ_2 and G_2 share the same SWAP-test circuit structure, so depolarization attenuates both equally and their difference preserves the physical signal.

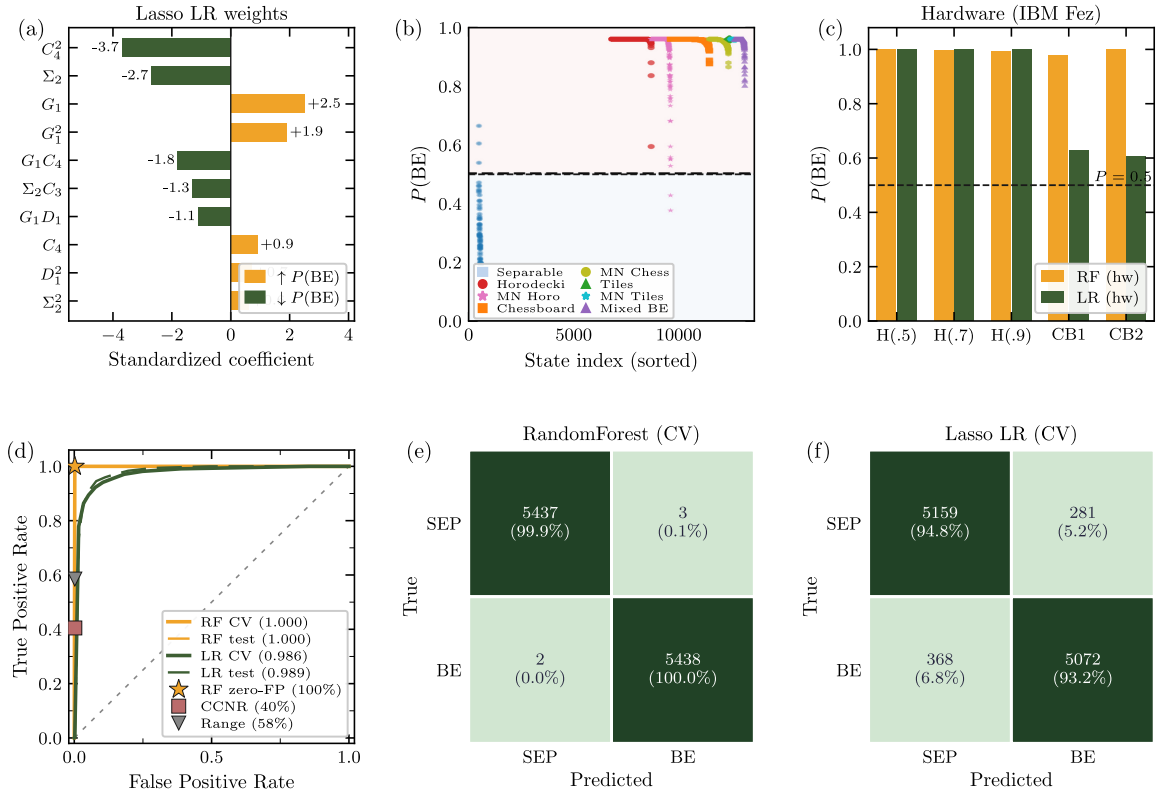


Fig. 2.19 **Bound entanglement detection via machine learning on *ibm_fez*.** (a), Standardized Lasso coefficients (22 nonzero out of 44 degree-2 polynomial features) trained on 13,600 certified states. Yellow bars: increase $P(\mathrm{BE})$; green bars: decrease it. (b), Random Forest score $P(\mathrm{BE})$ sorted by state index: separable (blue), Horodecki (red), marginal-noise Horodecki (pink), chessboard (orange), marginal-noise chessboard (yellow), Tiles (green), marginal-noise Tiles (light green), mixed BE (purple). (c), Hardware scores on *ibm_fez*: MLE-calibrated spectral features ($g_{\Sigma_1} = 0.164$, $f_{G_1} = 0.594$, $f_{\Sigma_2,G_2} = 0.306$; 684 circuits, 4×10^3 shots each). All five bound entangled states score above the decision boundary at $P = 0.5$. (d), ROC curves: RF CV (AUC $= 1.000$), RF test (1.000), Lasso CV (0.986), Lasso test (0.989). Analytical criteria at zero FP: CCNR 40%, range 58%. (e), Random Forest confusion matrix (CV): 99.96% recall, 0.06% FP rate. (f), Lasso LR confusion matrix (CV): 93.2% recall, 5.2% FP rate.

The per-circuit maximum likelihood estimation jointly fits three fidelity parameters and 20 physical features ($\Sigma_1, G_1, \Sigma_2, G_2$ for each of 5 test states) subject to spectral constraints ($\Sigma_k \geq G_k \geq 0$), with calibration states anchoring the fidelity factors.

Figure 2.19 shows the hardware results. All five tested bound entangled states are detected with both $D_1 > 0$ and $D_2 > 0$ (Table 2.27). The D_2 channel, which relies on two circuits with the same architecture, is free of cross-circuit calibration artifacts and provides a more reliable detection signal. Four of the five states are detected at a significance level of at $\geq 4.1\sigma$ via

D_2, with the weakest signal at 0.9σ for Horodecki $a = 0.90$. This is consistent with its nearly vanishing theoretical value, $D_2^{\text{thy}} = 0.001$. The D_1 channel shows positive values at higher significance (4.4–22.7σ), but the reconstructed Σ_1 values are systematically overestimated because the Pauli circuit fidelity $g_{\Sigma_1} = 0.164$ provides almost no dynamic range, making the apparent D_1 significance unreliable as a standalone criterion. Both chessboard states ($\Sigma_1 < 1$, undetectable by CCNR) are correctly identified through $D_2 > 0$, demonstrating detection of bound entanglement beyond the reach of the CCNR criterion on a gate-based superconducting processor. Both separable control states yield D_2 consistent with theory, confirming zero false positives.

Table 2.27 Bound entanglement detection on *ibm_fez* via per-circuit maximum likelihood estimation. Reconstructed spectral features with Cramér–Rao error bars from 4×10^3 shots per circuit (684 circuits total). All five BE states yield $D_2 > 0$; significance in units of σ. The D_2 channel is the primary detection signal because Σ_2 and G_2 share the same SWAP-test architecture, eliminating cross-architecture calibration artefacts. Theoretical values D_2^{thy} are given for comparison.

State	Type	D_1^{ML}	sig.	D_2^{ML}	sig.	D_2^{thy}
Horo ($a = 0.50$)	BE	$+1.60$	19.7σ	$+0.064 \pm 0.014$	4.6σ	0.018
Horo ($a = 0.70$)	BE	$+1.80$	22.7σ	$+0.095 \pm 0.014$	6.8σ	0.006
Horo ($a = 0.90$)	BE	$+1.41$	18.3σ	$+0.013 \pm 0.014$	0.9σ	0.001
Chess C1	BE	$+0.79$	6.9σ	$+0.093 \pm 0.023$	4.1σ	0.201
Chess C2	BE	$+0.52$	4.4σ	$+0.195 \pm 0.021$	9.2σ	0.063

A dense sweep of 19 Horodecki states ($a \in [0.05, 0.95]$) confirms that D_2 decreases monotonically with a and remains detectable for $a \lesssim 0.85$, establishing the practical detection boundary for this family. The chirality bias introduced by differential degradation of μ_4 and $\mathcal{I}_4$ remains bounded: for the *ibm_fez* data, $|f_{\mu_4} - f_{I_4}| \lesssim 0.05$, yielding a worst-case chirality bias of approximately 0.01. This is small compared to the entangled-state signal ($|C_4| \geq 0.066$ for $\theta \geq 15°$) but comparable to the separable bound $1/27 \approx 0.037$. Therefore, uncalibrated moment differences near the separability threshold require caution.

The classification accuracy scales with the feature budget. The ExtraTrees classifier, when trained on the full spectral feature set, achieves 99.6% cross-validated recall at zero false positives (see Section 2.4.5). Adding the range criterion increases this to 100% (see Fig. 2.19(e)). The pipeline for the primary Lasso classifier demonstrated on hardware is family-agnostic and requires no prior knowledge of the state structure. This is the first such demonstration on a gate-based superconducting processor.

2.4.7 Discussion and outlook

Three conceptually distinct layers of quantum correlations are probed by the methods developed in this section, each requiring progressively more measurement copies. The first layer is conventional entanglement, accessible through the Peres–Horodecki criterion and single-copy witnesses. The second layer is the structure revealed by multi-copy chirality: joint measurements on replicas of the state expose handedness patterns that vanish for product states and are genuinely invisible to any single-copy observable on an unknown state; pure-state chirality and negativity are related by the closed form $C_4 = -4\mathcal{N}^2(1-\mathcal{N}^2)$, and mixed-state negativity is reconstructible from three moment configurations via Newton–Girard identities. The third layer is bound entanglement behind the PPT barrier: its detection requires criteria algebraically independent of the partial-transpose spectrum, which the realignment non-Hermiticity gap provides through the same controlled-SWAP circuits.

What makes this progression structurally coherent is the identity $C_k = 8\,\mathrm{Tr}[\Omega_A\Omega_B\rho^{\otimes k}]$, which answers a previously open question about the physical content of partial-transpose moments: the difference $\mu_k - \mathcal{I}_k$ between partial-transpose and purity moments encodes chirality-chirality correlations across state copies. This is not merely a computational rewriting. Scalar spin chirality is the order parameter for chiral spin liquids in condensed matter physics. The fact that partial transposition, a quantum information operation defined on density matrices, probes the same geometric quantity reveals a structural connection between the two fields that goes beyond analogy. In condensed matter the chirality is measured among physical spins on a lattice; here it is measured among qubits from different copies of a bipartite state, with the same algebraic role and opposite physical context.

The bound entanglement results address a qualitatively harder problem than NPT detection. Previous hardware demonstrations relied on witnesses tailored to specific families [62–64], and all partial-transpose-based criteria are structurally blind to PPT states. Even the realignment criterion $\|R\|_1 > 1$ [12, 13] barely detects Horodecki states ($\|R\|_1 - 1 < 0.004$). The multi-channel approach performs witness fusion: combining complementary spectral channels with chirality features C_3 and C_4 into a composite nonlinear classifier that achieves 99.9% recall at zero false positives (Random Forest, simulation-validated on 13,600 certified states) and 93% recall with AUC 0.986 (Lasso logistic regression, hardware-demonstrated), compared with $\sim$40% for CCNR alone. Adding the range criterion reaches 100% across all known families.

The CCNR-invisible marginal-noise construction of Section 2.4.5 makes the limitation of single-criterion approaches concrete: states that satisfy $\Sigma_1 < 1$ and are therefore undetectable by CCNR remain detectable through $D_2 > 0$, with the Lasso classifier's noise tolerance extending 10–20$\times$ beyond the CCNR threshold. Recent complementary work on realignment

moment witnesses [65] and generalised realignment criteria [66] approaches this regime from a different angle; the multi-channel strategy demonstrated here shows that combining individually weak criteria achieves detection rates unattainable by any single criterion, including the new analytical ones.

Three limitations remain. The maximum likelihood calibration relies on states with known parameters; perturbation analysis shows robustness to $\pm 2^\circ$ preparation errors, but extension to fully blind calibration is an open problem. The hardware demonstration covers two of the three known bound entangled families (Horodecki and chessboard) plus the CCNR-invisible construction, with each hardware-tested family represented by multiple parameter values; extension to the Tiles family, denser parameter sweeps, and higher-dimensional systems (2×4, 4×4) are natural next steps. The efficiency gain ($5.3\times$ for 2×2, $7.2\times$ for 2×3) counts measurement configurations; compared with classical shadow tomography [53], the controlled-SWAP approach uses $O(1)$ configurations per moment at the cost of k coherent copies per circuit.

The scalability of the framework is favorable: a $d_A \times d_B$ system requires $d_A d_B - 1$ independent moments, so circuit depth and configuration count both grow linearly in $d_A d_B$, compared to the $O(d^2)$ scaling of full state tomography. The qubit overhead of multi-copy circuits is a genuine cost, though the circuit multiplexing strategy (packing multiple SWAP tests onto disjoint qubit regions of large processors) substantially mitigates it on current hardware. A basis-independent extension is possible: $D_k^{\min} = \Sigma_k - \max_{U_A,U_B} G_k[(U_A \otimes U_B)\rho(U_A \otimes U_B)^\dagger]$ eliminates the $\rho = \rho^*$ restriction on G_k measurements via variational optimisation of local unitaries. The framework extends naturally to multipartite systems, higher-order moments, and integration with randomized measurement and classical shadow techniques [53, 65].

The measurement infrastructure developed here— controlled-SWAP circuits, multi-copy protocols, and the chirality-chirality interpretation of partial-transpose moments— raises an immediate question of state generation. Given a target entangled state, can it be efficiently produced on hardware within the resource constraints that make the characterization protocols above tractable? This question is addressed in Chapter 3.

Chapter 3

Generating nonlinear systems: synergic quantum generative machine learning

A natural continuation of the characterization of nonlinear quantum systems discussed in the previous chapter is to consider the problem of generating quantum states with specific properties. Although collective methods allow us to effectively measure quantum correlations, a fundamental question remains: Can we not only measure and describe such systems, but also generate them? This issue is particularly important in the ongoing NISQ era [6], when computational resources are limited.

This chapter introduces a new approach to quantum generative learning: the Synergic Quantum Generative Network (SQGEN). Unlike the adversarial training paradigm, SQGEN replaces competition between the generator and discriminator with cooperation, which is optimized by minimizing a single cost function. This fundamental change significantly reduces hardware requirements (from $3n + 2$ to $n + 1$ qubits) and eliminates problematic hyperparameters. At the same time, SQGEN maintains comparable or even better performance than standard quantum adversarial networks (QGANs).

3.1 Quantum machine learning context

Quantum machine learning (QML) is an interdisciplinary field that combines machine learning algorithms with quantum computing [67]. This field considers, among other things, the potential advantages of using quantum computers for machine learning tasks over classical algorithms. In what specific applications and under what conditions might these advantages be realized? This question is particularly relevant in the context of NISQ devices, which have

limited computing power but offer unique opportunities due to the quantum nature of their calculations.

3.1.1 Classical vs quantum machine learning

Machine learning can be classified according to the nature of the input data and the methods used to process it. This classification yields four fundamental areas, as shown in Fig. 3.1. Classic machine learning (CC) involves processing classic data, which is represented as vectors in Euclidean spaces, using algorithms that run on classic computers. Neural networks, support vector machines, and random forests operate on this data through linear and nonlinear processes to learn patterns by optimizing cost functions. This approach's main limitations are the complexity of processing high-dimensional data, the scaling of computational costs with increasing data size, and the difficulty of representing certain types of correlations.

The second area, QC (quantum data, classical processing), covers situations in which data are in quantum states but are analyzed using classical algorithms. Examples include quantum state tomography, in which quantum measurements provide classical data that is then processed by reconstruction algorithms, and the methods of characterizing quantum correlations described in Chapter 2. In these cases, the quantum computer serves as a source of data, while the actual learning and inference are performed classically.

The third area, CQ (classical data, quantum processing), represents applications in which classical data are encoded into quantum states and processed by quantum circuits. Quantum kernel methods map classical data points to Hilbert space via quantum feature maps [68]. The Variational Quantum Eigensolver (VQE), which is used to solve classical optimization problems, also falls into this category [69]. The quantum advantage here comes from the ability to compute kernel functions in exponentially large feature spaces efficiently.

The fourth area, QQ (quantum computing and processing), is the most ambitious and potentially revolutionary. SQGEN operates in this area. Both the input data and the processing methods are fully quantum; there is no classical encoding or decoding stage. A quantum computer takes quantum states as input and produces quantum states as output. The entire learning process takes place in Hilbert space. This enables the direct manipulation of quantum correlations, entanglement, and other properties that lack classical analogues.

In the QQ domain, quantum machine learning uses superposition, entanglement, and interference to process quantum information. The quantum state of n qubits operates in a 2^n-dimensional Hilbert space, enabling the natural representation of exponentially large state spaces with a linear number of qubits. Unitary operations acting on these states implement transformations that preserve the quantum structure, including nonlinear correlations manifested by entanglement.

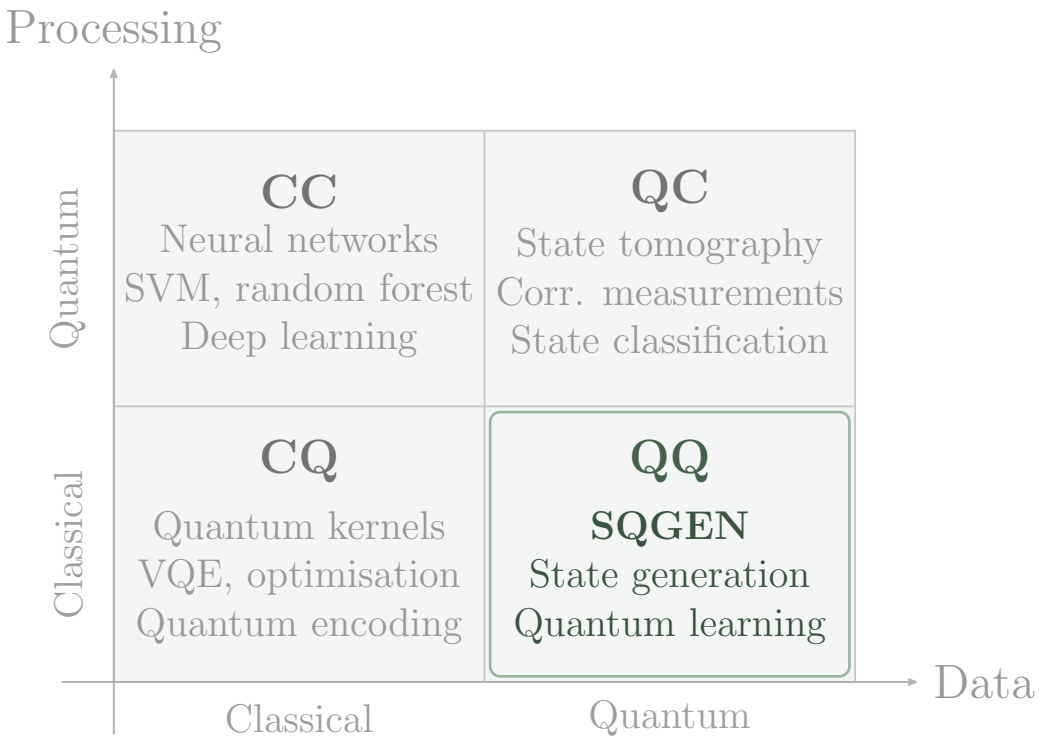


Fig. 3.1 **Taxonomy of machine learning paradigms according to the nature of data and processing.** The CC quadrant represents classical machine learning. The QC quadrant covers quantum data processed classically (e.g., state tomography). The CQ quadrant addresses classical data with quantum processing (e.g., quantum kernel methods). The highlighted QQ quadrant refers to fully quantum data and processing, the area in which SQGEN operates.

The differences between CC and QQ domains involve fundamental aspects of information processing. In classical machine learning, data can be copied; we can duplicate the training set and reuse the same samples multiple times in different parts of the algorithm. However, in the QQ domain, the no-cloning theorem prohibits copying unknown quantum states. This significantly alters the available learning strategies and is the primary reason why it is challenging to directly adapt classical algorithms, such as GANs, to the quantum domain. In quantum mechanics, measurement is probabilistic and irreversible— each measurement destroys the quantum state and provides only partial information about it. This contrasts with classical machine learning, where we can observe and analyze data without loss.

These fundamental differences lead to challenges and opportunities. The impossibility of copying quantum states complicates standard learning approaches, which assume access to the same data multiple times. This is the main challenge of QGAN, as discussed in

Section 3.2. Quantum superposition and entanglement enable new information processing strategies unavailable in classical systems. Quantum circuits implement kernel functions in exponentially large feature spaces using a linear number of gates, which is a potential source of computational advantage for certain classes of problems.

Classical generative algorithms (CC domain), such as GANs, VAE (Variational Autoencoders), and diffusion models, have achieved spectacular success in generating realistic images, text, and sound. However, these methods operate in a fundamentally classical space; they generate classical representations of data. In contrast, quantum generative learning in the QQ domain has a more ambitious goal: generating quantum states with desired properties. This has direct applications in preparing resource states for quantum communication, optimizing quantum protocols, and simulating quantum systems. As a representative of the QQ domain, SQGEN takes quantum states from a source and learns to generate similar states without the mediation of classical representations.

3.1.2 Generative adversarial networks (GANs) overview

Generative adversarial networks (GANs) are one of the most influential achievements in modern machine learning. Introduced by Goodfellow and colleagues in 2014 [70], GANs offer an elegant solution to the problem of learning probability distributions by framing it as a zero-sum two-player game between two neural networks.

The GAN architecture consists of two components. The generator G takes random noise z from a simple distribution (typically Gaussian or uniform) and transforms it into a sample $G(z)$ in the data space. Its goal is to learn a transformation such that the distribution $G(z)$ is indistinguishable from the actual data distribution p_{data}. The discriminator D takes a sample (real or generated) and evaluates the probability that it comes from the actual data distribution. Its goal is to maximize the ability to distinguish between real and generated samples.

GAN training can be formulated as solving a min-max problem:

$$\min_G \max_D V(D,G) = \mathbb{E}_{x\sim p_{data}}[\log D(x)] + \mathbb{E}_{z\sim p_z}[\log(1 - D(G(z)))], \tag{3.1}$$

where $V(D,G)$ is the value function. In this formulation, the discriminator tries to maximize V by correctly classifying real data as true ($D(x) \approx 1$) and generated data as false ($D(G(z)) \approx 0$). The generator tries to minimize V by producing data that the discriminator classifies as real.

In the theoretical Nash equilibrium of this process, the generator will learn to accurately reproduce the real data distribution ($p_G = p_{data}$), and the discriminator will be unable to distinguish between them ($D(x) = 1/2$ for all x). This elegant theory leads to a powerful

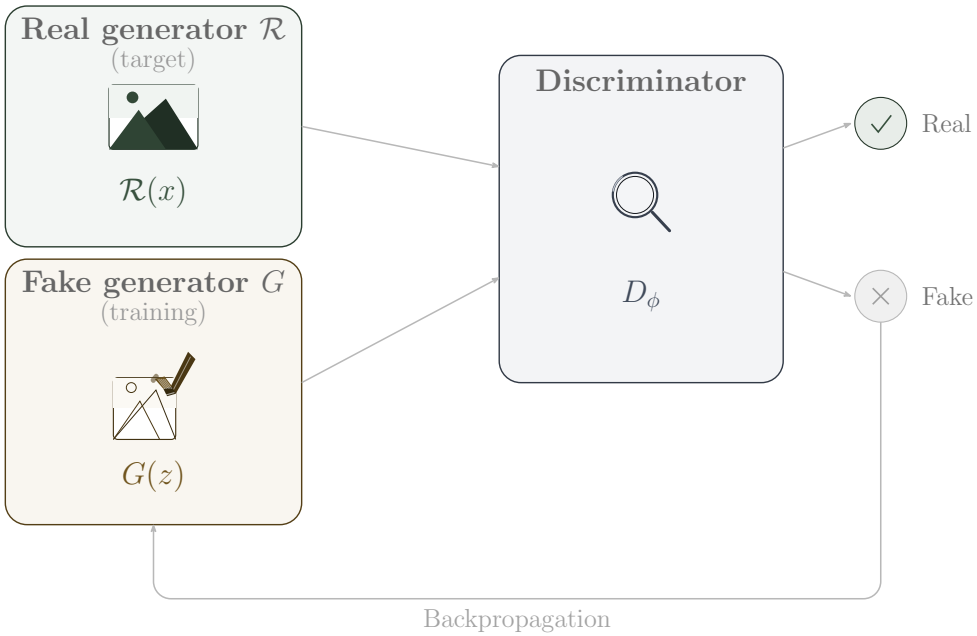


Fig. 3.2 **Schematic illustration of the Generative Adversarial Network (GAN) framework.** The real data generator $\mathcal{R}$ (target) produces authentic samples $\mathcal{R}(x)$, while the fake data generator G (training) attempts to produce indistinguishable counterfeits $G(z)$. The discriminator D_ϕ evaluates both streams and classifies each sample as real or fake. The generator is updated via backpropagation of the discrimination error, driving G to progressively close the gap between p_G and p_{data}.

learning algorithm rather than directly modeling the complex distribution p_{data}, GANs learn it indirectly through the dynamics of adversarial training.

3.1.3 Theoretical foundations of quantum generative learning

The key difference between SQGEN and QGAN is their respective strategies for discriminating quantum states. The traditional approach is based on von Neumann measurements in a suitably selected basis. In contrast, SQGEN uses a unitary transformation with an additional decision qubit. This difference has fundamental consequences for the efficiency of the learning process.

Consider the problem of distinguishing between two states: $|g\rangle$ representing the generator's output and $|r\rangle$ corresponding to the actual data. Regardless of the dimension of the Hilbert space, both states can be represented as unit vectors on a plane that define the angle β between them.

The standard approach to discrimination is to find a basis $|a\rangle$, $|b\rangle$ in which the states take the form:

$$|r\rangle = \cos(\pi/4 - \beta/2)\,|a\rangle + \cos(\pi/4 + \beta/2)\,|b\rangle\,, \tag{3.2}$$

$$|g\rangle = \cos(\pi/4 + \beta/2)\,|a\rangle + \cos(\pi/4 - \beta/2)\,|b\rangle\,. \tag{3.3}$$

The probability of correct discrimination through measurement in this basis is:

$$p_{a,b} = |\,\langle g|a\rangle\,|^2 |\,\langle r|b\rangle\,|^2 + |\,\langle g|b\rangle\,|^2 |\,\langle r|a\rangle\,|^2 = \frac{1+\sin^2\beta}{2}. \tag{3.4}$$

This strategy underlies QGAN training in which the optimization of the discriminator involves maximizing the angle β between discriminated states and finding the optimal measurement basis.

SQGEN implements discrimination differently. First, we introduce an additional decision qubit in the state $|0\rangle$, and then perform a controlled rotation $R_y(\theta)$ conditioned on the input state. The discriminator operator takes the form:

$$D = R_y(\theta) \otimes |\psi\rangle\,\langle\psi| + \mathbb{I} \otimes (\mathbb{I} - |\psi\rangle\,\langle\psi|), \tag{3.5}$$

where $|\psi\rangle$ is the internal reference state of the discriminator. Measuring the decision qubit in the Z basis yields a result of -1 with probability $p_r^{(-)} = \sin^2\theta |\,\langle r|\psi\rangle\,|^2$ for the actual state and $p_g^{(-)} = \sin^2\theta |\,\langle g|\psi\rangle\,|^2$ for the generated state.

The probability of optimal discrimination reaches the value $(1+\sin^2\beta)/2$ when $|\psi\rangle = |r\rangle$ or $|\psi\rangle = |g\rangle$ and $\sin\theta = 1$. The key difference is that SQGEN automatically sets the reference state $|\psi\rangle$ to $|r\rangle$ through the construction of a synergic circuit. We count only those events in which the state $|r\rangle$ collapses to $|\psi\rangle$ and the state $|g\rangle$ collapses to the orthogonal subspace $\mathbb{I} - |\psi\rangle\,\langle\psi|$.

Figure 3.3 illustrates the geometric differences between the two strategies. In the QGAN approach (panel a), the discriminator must learn to distinguish between states $|g\rangle$ and $|r\rangle$ with access to only one of them at a time. This requires additional computations to establish a reference frame for the discrimination process. The optimization process consists of finding a basis $|a\rangle$, $|b\rangle$ in which the projections of both states are maximally distinguishable.

Figure 3.3(b) shows that SQGEN performs discrimination by projecting onto the internal reference state $|\psi\rangle$. The discriminator optimizes the superposition $|\,\langle r|\psi\rangle\,|^2$, which automatically sets the state to $|\psi\rangle \to |r\rangle$. The synergic circuit design ensures that only events that satisfy the convergence condition of the actual and generated states are counted as successes.

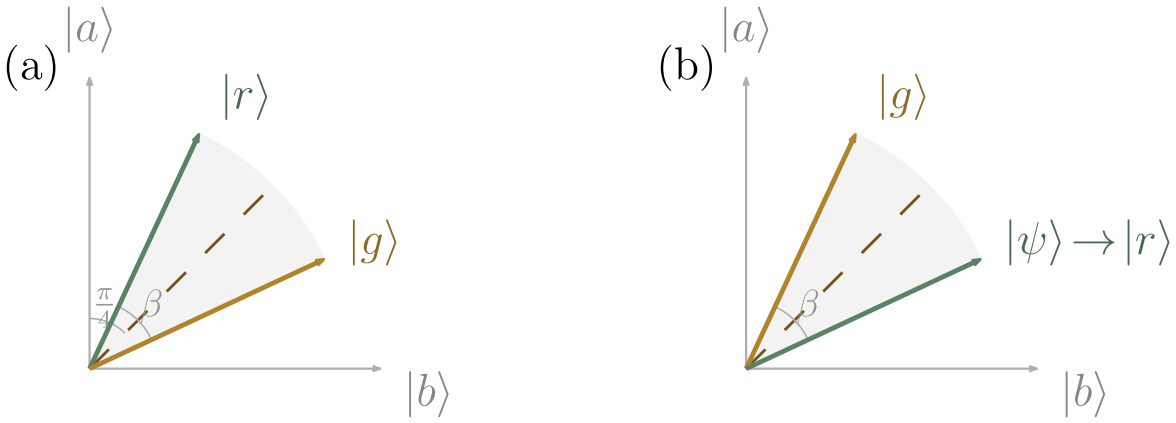


Fig. 3.3 **Geometric interpretation of state discrimination strategies.** (a) Standard measurement-based approach in basis $|a\rangle$, $|b\rangle$ used in QGAN. (b) Discriminator-based approach employed in SQGEN. The states to be discriminated are $|g\rangle$ and $|r\rangle$. The internal state of the discriminator associated with optimal discrimination is denoted as $|\psi\rangle$. The discriminator learns to find an optimal section of the Hilbert space supporting $|g\rangle$ and $|r\rangle$ where the overlap $|\langle r|\psi\rangle|^2$ is maximal.

This eliminates the alternating training of the generator and discriminator characteristic of QGAN.

The discrimination probability is maximized when both components of the network are operating optimally. The generator minimizes the difference between the states, and the discriminator maximizes its ability to distinguish between them. In SQGEN, these two processes occur simultaneously, expressing the synergic nature of the method.

The discriminator works optimally when the discrimination probability is maximized. Instead of maximizing the difference in the frequency of assigning real/fake labels to a sample provided by $\mathcal{R}$ or $\mathcal{G}$, as is done in standard GAN, we can maximize this probability. However, this probability decreases as the similarity between samples generated by $\mathcal{R}$ and $\mathcal{G}$ increases, similarly to the aforementioned frequency difference.

3.2 Limitations of QGANs

Quantum generative adversarial networks [10, 9] attempt to transfer the adversarial learning paradigm to the quantum domain. In a standard QGAN implementation, the generator G and discriminator D are parameterized quantum circuits, and their parameters are optimized in successive rounds. First, the discriminator is trained with fixed generator parameters, then the generator is trained with fixed discriminator parameters. While the process is conceptually elegant, it encounters fundamental limitations resulting from both the quantum nature of

information and the practical limitations of NISQ devices. Three main categories of these limitations are analyzed systematically: scalability issues, training instabilities, and hardware requirements.

3.2.1 QGAN architecture

A quantum generative adversarial network (QGAN) directly transfers the classical GAN idea to the quantum domain. This method determines the Nash equilibrium in a two-player game. In this game, the first player (generator $\mathcal{G}$) generates quantum states. The second player (discriminator $\mathcal{D}$) tries to distinguish whether a given state comes from the generator or from an external source $\mathcal{R}$.

The generator $\mathcal{G}$ takes as an input the standard state $|0\rangle^{\otimes n}$ and a random variable z_G representing the internal state of randomness. The output is the state $\sigma_{\lambda,z_G} = |\psi_{\lambda,z_G}\rangle \langle\psi_{\lambda,z_G}|$, where the parameter λ denotes the class label. We train the generator by minimizing the cross-entropy between its output and the actual states, while maximizing the probability that the discriminator will label the generated state as real.

The discriminator, denoted by $\mathcal{D}$ receives the quantum state (real ρ_{λ,z_R} or generated σ_{λ,z_G}), label λ, random variable z_D, and additional auxiliary space for complex quantum computations. The discriminator's task is to assign a value of 0 to states from the source $\mathcal{R}$ and a value of 1 to states from the generator $\mathcal{G}$. We train the discriminator by maximizing the difference between the probabilities p and q, where p is the frequency with which real states are correctly classified and q is the frequency with which generated states are incorrectly classified as real.

The cost functions for both components take the following form:

$$J_G = 1 - \sum_{z_G,z_R} p_\theta(z_G) p_R(z_R) \mathrm{Tr}(\sigma_{\lambda,z_G} \rho_{\lambda,z_R}), \tag{3.6}$$

$$J_D = 1 - \sum_{z_G,z_R,z_D} p_\theta(z_G) p_R(z_R) g(z_D) \cos^2(\theta_{z_D}(\sigma_{\lambda,z_G}) - \theta_{z_D}(\rho_{\lambda,z_R}) + \pi/2), \tag{3.7}$$

where p_θ and p_R are the probability distributions of the generator and source random variables, respectively; $g(z_D)$ is the distribution of the discriminator's internal state, and θ_{z_D} denotes the angle associated with discrimination in the z_D configuration.

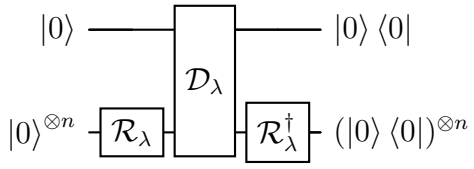
We formulate the QGAN problem as a min-max task:

$$\min_{\mathcal{G}} \max_{\mathcal{D}} \left[J_D(\mathcal{D},\mathcal{R}) - J_G(\mathcal{G},\mathcal{R}) \right], \tag{3.8}$$

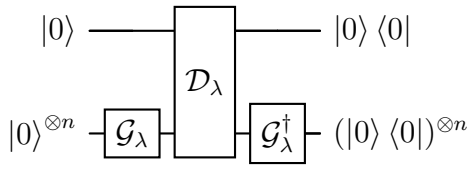
where the statistical distance between the outputs $\mathcal{G}$ and $\mathcal{R}$ is minimized with respect to the generator's strategy, while the distance between outputs $\mathcal{D}$ for states from $\mathcal{G}$ and $\mathcal{R}$ is maximized with respect to the possible strategies of the discriminator.

QGAN training requires three separate quantum circuits, as shown in Figure 3.4. Each circuit is used to calculate a different quantity needed for optimization.

(a) Evaluation of the discriminator on real data (calculation of p)



(b) Discriminator evaluation on generated data (calculation of q)



(c) Comparison of the source with the generator (calculation of F)

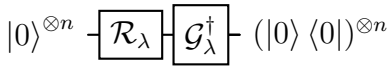


Fig. 3.4 **Circuits used for QGAN training.** (a) Circuit evaluating discriminator performance on real data generated by source $\mathcal{R}$, computing probability p of labeling data as real. (b) Circuit evaluating discriminator performance on generated data from $\mathcal{G}$, computing probability q of labeling data as real. (c) Circuit comparing distributions of real and generated data, computing fidelity F. These circuits are used in alternating rounds where either discriminator or generator parameters are optimized while keeping the other fixed.

The circuit in Fig. 3.4(a) evaluates the performance of the discriminator on real data. First, the source prepares the state. Then, the discriminator processes it by adding a decision qubit. Finally, we apply the inverse operation of the source. Measuring the decision qubit in the computational basis yields the probability that the actual state will be correctly recognized.

Fig. 3.4(b) shows an analogous circuit for generated states. The generator $\mathcal{G}_\lambda$ replaces the source $\mathcal{R}_\lambda$. The measurement gives the probability q that the discriminator will misclassify the generated state as real. During discriminator training, we maximize $|p - q|$, which forces the discriminator to distinguish between the two types of states as best as possible.

Fig. 3.4(c) shows a circuit that measures the fidelity, $F = \text{Tr}(\rho_\lambda \sigma_\lambda)$, between the average states produced by the source and the generator. After the source operation, the generator is used in the reverse direction, i.e., the generator is used as the inverse of the source operation. The frequency with which the state $|0\rangle^{\otimes n}$ is obtained on all generator qubits corresponds to the fidelity. During generator training, we maximize F while simultaneously maximizing q.

3.2.2 Scalability issues

A critical limitation of QGANs is that their qubit requirements dramatically increase with problem size. Training a generator for an n-qubit state, a standard QGAN requires a minimum of $3n + 2$ qubits. This number consists of: n qubits for the generator's output state, another n qubits for the state from the real data source R, another n qubits for a second copy of the source, which is needed for separate evaluations of the discriminator on real and generated states; and two auxiliary qubits for the discrimination mechanism.

This required duplication of the source R follows directly from the no-cloning theorem. In a classical GAN, we can use the same data sample multiple times: once to train the discriminator and again to compare it with the generator. In the quantum domain, a state cannot be copied, so we need independent calls to the source R for each of these purposes. For the relatively simple task of generating a six-qubit GHZ state, QGAN requires 20 qubits.

This scaling has dramatic consequences for NISQ devices [71]. Modern quantum processors typically offer 27–127 qubits, but not all of them are fully connected, and the quality of the qubits and gates varies significantly. The $3n + 2$ pattern means that, even for relatively small states of eight to ten qubits, the full implementation of QGAN begins to encounter the limitations of the available hardware. Each additional qubit introduces additional sources of decoherence and gate errors, degrading the overall quality of the computation.

In addition to the number of qubits, the scaling problem concerns the number of parameters to be optimized. Commonly used universal circuit ansatze, such as Möttönen's decomposition [72], require 4^n parameters to represent any unitary operation on n qubits. For $n = 6$, this yields 4096 parameters for the generator alone, plus thousands more for the discriminator. Optimization in such a high-dimensional parameter space can get stuck in local minima and requires a large number of cost function evaluations. Without advanced optimization techniques, it becomes practically infeasible.

The depth of the quantum circuit also scales unfavorably. For a standard QGAN, evaluating the cost function requires three circuits: one to compare R and G, one to evaluate D on real states, and one to evaluate D on generated states. The depth of each of these circuits increases with n, and the total execution time and accumulation of decoherence errors depend on their

combined depth. For n = 6, the depth of these circuits is typically in the hundreds or thousands of gate layers, which exceeds the practical coherence limits for modern superconducting qubits.

3.2.3 Training instabilities

The second fundamental problem with QGANs is the instability of the training process. This problem is inherited from classical GANs, but it is exacerbated by the quantum nature of the computations. In classical GANs, balancing the learning rates of the generator and discriminator requires careful tuning of numerous hyperparameters. In the quantum context, this problem becomes even more pronounced for several reasons.

The key hyperparameter in QGAN is the number of optimization iterations for the discriminator k_D relative to the number of iterations for the generator k_G in each training epoch [73, 74]. If k_D is too large, the discriminator becomes too good too quickly, begins to perfectly distinguish real states from generated ones, and provides the generator with a gradient signal close to zero. This stops the generator from learning, which is known as the vanishing gradients problem. Conversely, if k_D is too small, the discriminator does not provide a strong enough learning signal, allowing the generator to cheat the discriminator without learning the data distribution.

The numerical experiments presented below demonstrate that finding the optimal ratio, k_D/k_G, required considerable trial and error. Even worse, this optimal ratio depends on the size of the problem (n), the structure of the state to be learned, the circuit ansatz chosen, the optimization method (gradient descent vs. gradient-free), and the initial configuration of the parameters. Thus, for each new problem, the hyperparameters must be retuned, dramatically increasing the total computational cost.

Additional hyperparameters that require tuning include learning rates (α_D for the discriminator and α_G for the generator), which must be balanced with the number of iterations, as well as the weights of the various components of the cost function. For the discriminator, a typical optimization involves a linear combination of terms proportional to $|p - q|$, $p + q$, and sometimes additional regularization terms. Other hyperparameters include the learning rate scheduling strategy during training and the choice of gradient calculation method (parameter shift rule versus finite differences).

The mode collapse problem also occurs in QGAN. In some cases, when $n = 6$, the generator learned to produce only a subset of the possible states instead of the entire set. For GHZ states, in which the source produces a single state deterministically, the generator sometimes settles for an approximate solution instead of accurately reproducing the target

state. To circumvent this problem, the discriminator cost function was modified by adding a term proportional to $1 - p$, which introduces an additional hyperparameter to tune.

Quantum noise further destabilizes training. The probabilistic nature of quantum measurement introduces Poisson noise into the evaluation of the cost function. This means that, even with the same parameter configuration, successive evaluations yield slightly different results. The gradient calculated using such noisy measurements is also noisy, which can lead to erroneous optimization steps. In practice, a significant number of measurements (shots), typically 10^4 or more, are required for each evaluation to reduce the noise below an acceptable level, dramatically increasing computation time.

3.2.4 Hardware requirements

The third critical limitation is the hardware requirements necessary to implement QGAN on NISQ devices. In addition to the aforementioned $3n + 2$ qubits, there are several other technical requirements that limit QGAN's practical applicability.

In real quantum processors, qubit connectivity is limited; qubits can only perform two-qubit operations with specific neighbors according to the hardware topology. IBM Quantum processors typically have a heavy hexagonal topology, where each qubit is connected to two to three neighbors [40]. However, QGAN algorithms often require operations between arbitrary pairs of qubits. This necessitates decomposition into sequences of SWAP gates to "move" states between distant qubits. Each SWAP gate adds three CNOT gates to the circuit, which dramatically increases depth and error accumulation.

The quality of quantum gates in NISQ devices varies. The typical fidelity of single-qubit gates is 99.9%, whereas the typical fidelity of two-qubit gates is 98%-99%. While this may sound impressive, the accumulation of errors in a circuit consisting of hundreds of gates leads to dramatic degradation. For example, in a 300-gate circuit where half of the gates are two-qubit, the final state fidelity can drop below 50%, rendering the results virtually useless.

The coherence times, T_1 (lifetime) and T_2 (dephasing time), of superconducting qubits are usually $100\breve{}200\mu$s. A single-qubit gate takes about $20\breve{}50n$s, and a two-qubit gate takes $200\breve{}500n$s. For deep QGAN circuits, the execution time can approach the coherence time, which introduces significant decoherence errors. The longer the circuit, the more quantum information is lost through interactions with the environment.

Although measurement error mitigation is necessary, it is only partially effective. IBM processors typically have a measurement fidelity of 97-99% for single qubits, but errors accumulate for multi-qubit circuits. The standard procedure for mitigation involves calibrating the confusion matrix and reversing its effect [61, 75]. However, this method only works well

for relatively small circuits and low noise levels. For deep QGAN circuits, where noise is already significant, this procedure may be insufficient.

The bandwidth and availability of quantum computers also pose practical limitations. Computations requiring millions of measurements per training epoch, repeated for dozens of epochs, can consume hours or days of processor time. In cloud systems such as IBM Quantum, access to processors is limited, and the waiting queue can be long. These factors make experimental validation and iterative refinement of QGAN algorithms time-consuming and costly.

The three categories of limitations: scalability issues, training instability, and hardware requirements converge, making standard QGANs difficult to implement on modern NISQ devices. Each limitation is significant on its own, and together, they create a strong motivation to seek alternative approaches to quantum generative learning. The synergic quantum generative network, which will be presented in the following sections, addresses these limitations with a fundamentally different approach.

3.3 Synergic approach concept and theory

3.3.1 Collaboration vs competition paradigm

The synergic approach, introduced in [22], fundamentally changes this dynamic. Rather than competing, the generator and discriminator work together to achieve a shared success condition. The discriminator is optimal when it correctly identifies real states, and the generator is optimal when it produces states that are identical to real ones. The key observation is that these two conditions are complementary, not contradictory. When the generator produces perfect copies of real states, the discriminator can perform optimally in identifying real states because they are indistinguishable from the generated ones. When the discriminator operates optimally, it provides the generator with precise information about the characteristics of real states.

This synergy can be achieved through a well-designed quantum circuit. The reversibility of unitary transformations is exploited here: the discriminator D is a unitary operation that can be coupled with $D^\dagger$. The key innovation is placing a conditional negation gate (CNOT) between these operations. This creates a structure where the probability of success (specific state detection) is the product of generation and discrimination quality. Mathematically, the cost function takes the form $J \propto \cos^2\theta \cdot \mathrm{Tr}(\sigma\rho)$, where the first factor describes the discriminator and the second factor describes the generator.

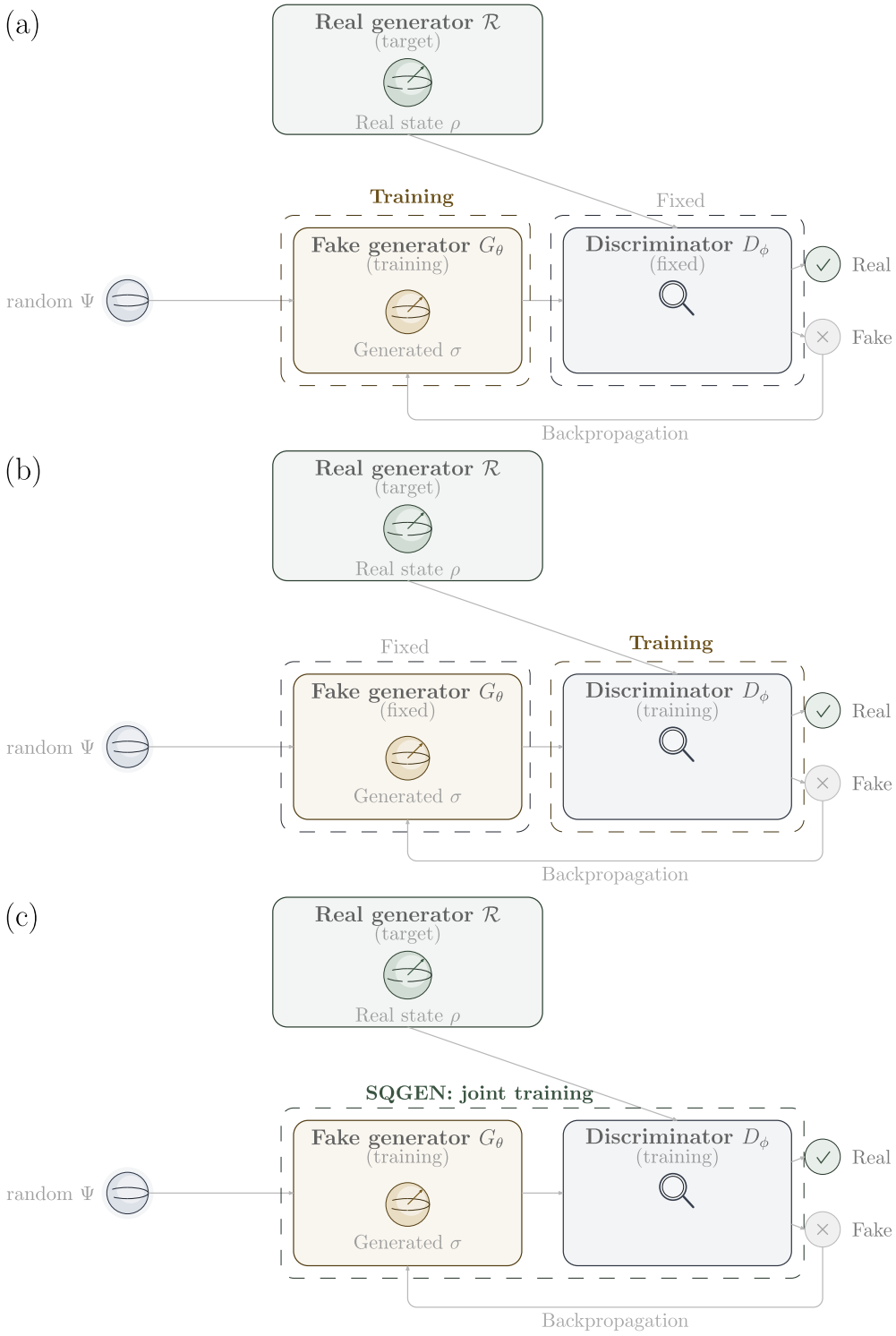


Fig. 3.5 Panels (a) and (b) depict the training processes of the generator and discriminator, respectively, in the QGAN method. This approach involves alternating iterative training between panels (a) and (b) while maintaining fixed generator or discriminator parameters. In contrast, the SQGEN method, as demonstrated in panel (c), involves cooperative interaction between the generator and discriminator during learning. Thus, the training process in SQGEN differs from that in QGAN.

This design naturally synchronizes learning. If the generator produces weak states ($\mathrm{Tr}(\sigma\rho)$ low), even an optimal discriminator will give a high value of J. Conversely, if the discriminator performs poorly ($\cos^2\theta$ is low), even a perfect generator will not minimize J. The J function is minimized only when both components perform well, hence the name "synergic" learning. Improving J by optimizing only one component while fixing the other is not possible; simultaneous optimization of both is necessary.

The practical consequences of this paradigm shift are profound. It eliminates the need to decide how many optimization iterations to devote to the discriminator versus the generator because both are optimized at every step. There is no need to balance learning rates for two competing goals because there is only one goal. The problem of vanishing gradients due to an overly good discriminator is also eliminated because a better discriminator improves the learning signal for the generator instead of eliminating it. Additionally, the mode collapse problem is mitigated. If the generator starts producing only a limited subset of states, the discriminator will automatically detect this because other states will be misclassified, providing a strong correction signal.

3.3.2 Cost function formulation

Formulating the cost function for SQGEN requires a careful analysis of quantum state discrimination theory and relative entropy. First, we must ask a fundamental question: How do we measure the similarity between quantum state assemblages, or sets of states and their probabilities?

For two quantum states $\sigma_{\lambda,z_G} = |\psi_{\lambda,z_G}\rangle\langle\psi_{\lambda,z_G}|$ (generated) and ρ_{λ,z_R} (real), a natural measure of their similarity is relative entropy (also known as Kullback-Leibler divergence) [76]:

$$S(\sigma_{\lambda,z_G}||\rho_{\lambda,z_R}) = \mathrm{Tr}(\sigma^2_{\lambda,z_G}) - \mathrm{Tr}(\sigma_{\lambda,z_G}\log\rho_{\lambda,z_R}). \tag{3.9}$$

This quantity is always nonnegative and only becomes zero when $\sigma = \rho$, making it an ideal candidate for the generator's cost function.

Direct calculation of relative entropy is difficult for practical calculations on a quantum computer due to the logarithm of the operator. However, it can be expanded into a Newton-Mercator series:

$$S(\sigma||\rho) = \langle 1-\rho\rangle + \frac{1}{2}\langle(1-\rho)^2\rangle + \frac{1}{3}\langle(1-\rho)^3\rangle + \ldots, \tag{3.10}$$

where $\langle A\rangle = \mathrm{Tr}(\sigma A)$ denotes the expected value of the operator A in the state σ.

For pure quantum states, the first nonlinear term of this expansion yields the linear entropy:

$$S_L(\sigma||\rho) = 1 - \text{Tr}(\sigma\rho) = 1 - |\langle\psi_G|\psi_R\rangle|^2, \tag{3.11}$$

which is directly measurable using the SWAP test or similar interferometric techniques [77]. The value $\text{Tr}(\sigma\rho)$ is the square of the overlap between states and is naturally interpreted as the probability of projecting one state onto another.

However, the use of linear entropy alone is insufficient for proper generator training. For example, a source that provides states $|0\rangle$ and $|1\rangle$ with equal probability has an average state of $\rho = (|0\rangle\langle 0| + |1\rangle\langle 1|)/2 = \mathbb{I}/2$. A generator that produces $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ and $|-\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$ with equal probabilities also has an average state $\sigma = \mathbb{I}/2$. Both assemblages give $\text{Tr}(\sigma\rho) = 1/2$, even though they are fundamentally different: the first operates in the basis $\{|0\rangle, |1\rangle\}$ while the second operates in $\{|+\rangle, |-\rangle\}$.

This example shows that linear entropy depends only on the average state without taking the assemblage's structure into account. A discrimination mechanism is needed to distinguish between these cases. Therefore, a discriminator is introduced that operates at the level of individual states, not just their averages.

In SQGEN, the discriminator D is a unitary operation that acts on a space consisting of a decision qubit initialized in the state $|0\rangle$ and working qubits that contain the processed state. Given an input state, $|\psi\rangle$, the discriminator performs a controlled rotation on the decision qubit:

$$D = \mathbb{I} \otimes U_{z_D}|0\rangle\langle 0|U_{z_D}^\dagger + R_y(\theta) \otimes (\mathbb{I} - U_{z_D}|0\rangle\langle 0|U_{z_D}^\dagger), \tag{3.12}$$

where U_{z_D} is a parameterized unitary transformation that represents the internal state of the discriminator, and $R_y(\theta) = \cos\theta\mathbb{I} + i\sin\theta Y$ is a rotation around the y axis.

A key property of this discriminator is that the classification probability depends on the overlap between the input state and the discriminator's reference state, which is given by $|\phi\rangle = U_{z_D}|0\rangle$. For the state $|\psi\rangle = \alpha|\phi\rangle + \sqrt{1-\alpha^2}|\phi^\perp\rangle$, the probability of recognition as "real" is:

$$p(\alpha) = |\langle 0|\langle\psi|D|0\rangle|\psi\rangle|^2 = |(1-\alpha^2)\cos\theta + \alpha^2|^2. \tag{3.13}$$

For $\theta = \pi/2$ (the discriminator regime), we have $p(1) = 1$ for a perfect match and $p(0) = 0$ for an orthogonal state. This automatically sets the discriminator reference state $|\phi\rangle$ to the actual state from the assemblage—the discriminator "learns" to recognize the characteristics of actual states.

The synergic cost function combines the generator's linear entropy and the discriminator's discrimination quality into a single expression:

$$J = 1 - \sum_{z_G, z_R, z_D} g(z_D) p(z_G) q(z_R) \cos^2(\theta_{z_D}) \mathrm{Tr}(\sigma_{\lambda, z_G} \rho_{\lambda, z_R}). \tag{3.14}$$

In order to improve the readability of the graphs and ensure normalization $J \in [0, 1]$, we use an equivalent form of the cost function with a factor of 2:

$$J = 1 - 2 \sum_{z_G, z_R, z_D} g(z_D) p(z_G) q(z_R) \cos^2(\theta_{z_D}) \mathrm{Tr}(\sigma_{\lambda, z_G} \rho_{\lambda, z_R}). \tag{3.15}$$

The cost function can be interpreted as the probability that the assemblages $\{\sigma_{\lambda,z_G}, p(z_G)\}$ and $\{\rho_{\lambda,z_R}, q(z_R)\}$ are distinguishable for a given discriminator configuration. This quantity is minimized when both the generator and the discriminator are optimized simultaneously. However, if we optimize only the generator or only the discriminator, there is always a possibility to improve J by optimizing the other component.

It is worth noting that instead of minimizing J, we can equivalently maximize:

$$1 - J = 2 \sum_{z_G, z_R, z_D} g(z_D) p(z_G) q(z_R) \cos^2(\theta_{z_D}) \mathrm{Tr}(\sigma_{\lambda, z_G} \rho_{\lambda, z_R}). \tag{3.16}$$

This function can be measured directly in a single quantum circuit by connecting the discriminator $\mathcal{D}$ to its inverse $\mathcal{D}^\dagger$ with a conditional negation gate (CNOT) between them, as shown in Fig. 3.7.

This formulation has a key synergy property. The following properties refer to the original cost function (3.14); the factor-of-2 normalization (3.15) rescales all quantities accordingly.

- When the discriminator is inactive ($\theta_{z_D} = 0$), J reduces to the linear entropy $J_G = 1 - \mathrm{Tr}(\sigma_\lambda \rho_\lambda)$ between the mean states;
- When the generator is perfect ($\mathrm{Tr}(\sigma_{\lambda, z_G} \rho_{\lambda, z_R}) = 1$ for all z_G, z_R), J reduces to the optimal discriminator cost function $J_D^* = 1 - \sum_{z_R, z_D} q(z_R) g(z_D) \cos^2(\theta_{z_D})$;
- In the general case, J is minimized only when both components operate optimally.

To enable visualization and comparison with QGAN, the following auxiliary quantities are defined: the probability p of correctly classifying the actual state, the probability q of misclassifying the generated state as actual, and the fidelity $F = \mathrm{Tr}(\sigma_\lambda \rho_\lambda)$ between the mean states. Here, g, p, and q represent the probability distributions of the discriminator, generator, and source, respectively. The probability p is related to the purity of the unknown set of states

$\{\rho_{\lambda,z_R}, q(z_R)\}$- if for some z_R we have $p(z_G) = 1$ and ρ_{λ,z_R} is a pure state, then the entire set is pure.

3.3.3 Quantum circuit and kernel interpretation

To develop a deeper understanding of SQGEN, it is necessary to analyze its mathematical foundations and its relationships with other machine learning methods. Two analytical perspectives are presented: the first interprets the construction through the lens of unitary reversibility and post-selection, while the second draws an analogy to kernel methods in machine learning.

Both the generator $\mathcal{G}$ and the unitary part of the discriminator U_{z_D} from equation (3.12) are implemented using Möttönen's universal decomposition [72] shown in Figure 3.6. This approach allows for the implementation of any n-qubit unitary transformation using controlled rotations around the Y and Z axes.

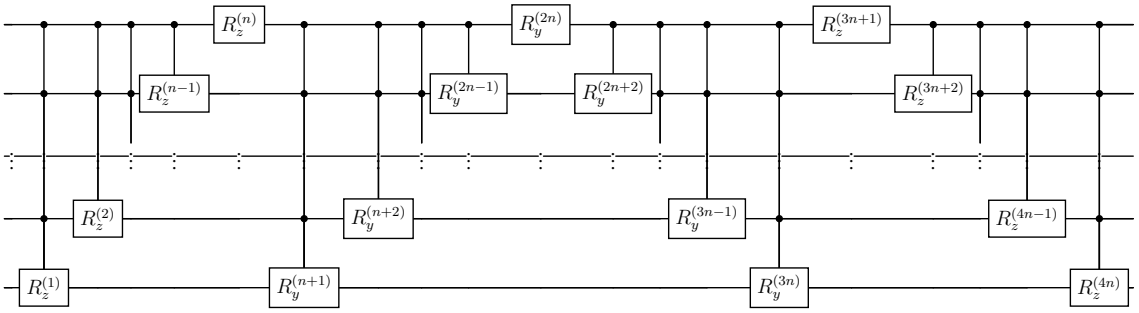
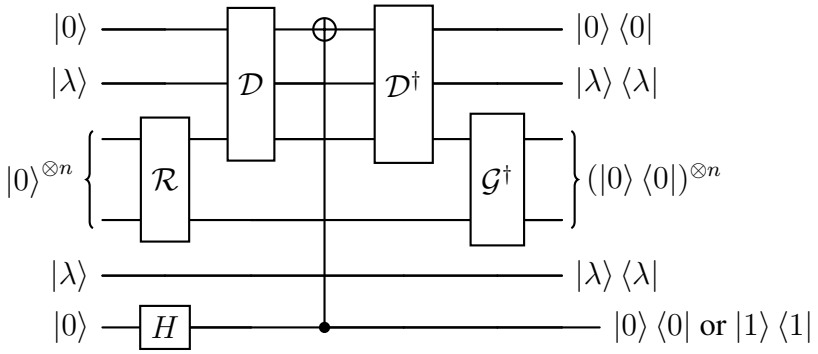


Fig. 3.6 **Möttönen's universal decomposition circuit for an n-qubit unitary.** The circuit uses 4^n parameters (rotation angles) arranged in alternating blocks of controlled R_z and R_y rotations. For each qubit layer, the controlled rotations may depend on the state of multiple control qubits, realizing any n-qubit unitary transformation exactly.

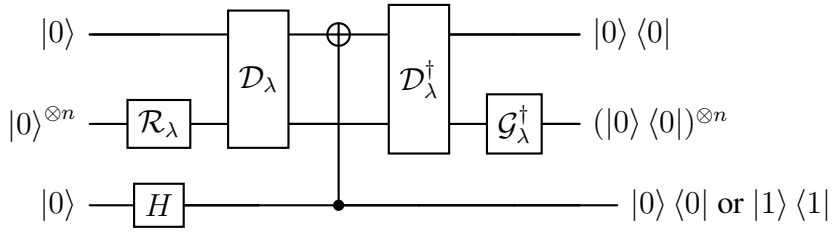
The decomposition uses a hierarchical structure of controlled rotations. For n qubits, 4^n parameters describing the rotation angles are required. Initially, we apply rotations R_z to all qubits, then a block of controlled rotations R_y, another block of rotations R_z, again R_y, and finally rotations R_z. Each controlled rotation may depend on the state of multiple control qubits. In practice, this requires decomposition into CNOT two-qubit gates and single-qubit rotations on a real quantum processor.

Figure 3.7 shows the complete circuit for the synergic generative learning protocol. In the simplest version (panel a), the parameter λ denoting the generator output class controls the choice of operations for both the real data source $\mathcal{R}$ and the generator $\mathcal{G}$. The state $|\lambda\rangle$ does not change during these operations.

(a) Full circuit with explicit class label



(b) Simplified circuit for classical label



(c) SWAP test variant

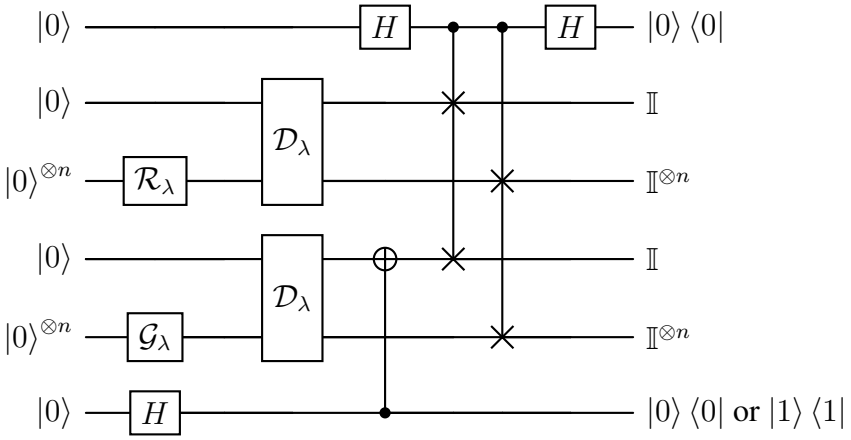


Fig. 3.7 **Synergic quantum generative learning protocol circuits.** (a) Full circuit with explicit class label $|\lambda\rangle$ controlling both real data source $\mathcal{R}$ and generator $\mathcal{G}$. (b) Simplified circuit where label is treated as classical control parameter. The conditional X gate between discriminator and its inverse ensures synergic training. (c) Alternative implementation inspired by SWAP test, where measured quantity depends on joint projection rates of first and last qubits. Depending on measurement outcome of bottom qubit, we either include or exclude discriminator operation, allowing measurement of either full cost function J or source-generator fidelity F.

For the classical label λ, the circuit is simplified to the form shown in Fig. 3.7(b). The source $\mathcal{R}_\lambda$ prepares the actual state, then the discriminator D_λ performs a controlled rotation on the decision qubit, the conditional gate X flips the state of this qubit, and then we apply the inverse discriminator $D_\lambda^\dagger$ and the inverse generator $\mathcal{G}_\lambda^\dagger$. The probability of simultaneous detection of the state $|0\rangle$ in the decision qubit and the states $|0\rangle$ in all generator qubits is proportional to the value of the cost function (3.15).

An alternative implementation (Fig. 3.7(c)) uses a construction inspired by the SWAP test. A multi-qubit controlled SWAP gate decomposes into n standard Fredkin gates, increasing the circuit's depth but potentially reducing execution time on a quantum processor through parallel operations. Measuring the lower qubit enables switching between measuring the full cost function J (when the qubit is in state $|1\rangle$) and measuring only the fidelity F between the source and generator (when the qubit is in state $|0\rangle$).

Due to Möttönen's universal parameterization, the circuit depth scales as $O(4^n)$. For $n = 6$, the depth reaches 5385 layers, exceeding the capabilities of modern NISQ devices. Practical implementations for larger systems require transitioning to more efficient ansatze, such as hardware-efficient ansatze [78]. However, this comes at the cost of losing the guarantee of universality.

This construction has deep implications. First, the entire learning process reduces to maximizing quantum probability, which is a natural quantity that can be accessed directly by measurement. Second, post-selection, or conditioning on the detection of the decision qubit in the state $|0\rangle$, automatically eliminates cases in which the discriminator does not function correctly. This focuses learning on relevant configurations. Third, the reversibility of the design, or the ability to "undo" the discriminator using $D^\dagger$, is crucial for synergy because it allows for the simultaneous evaluation of the quality of generation and discrimination in a single experiment.

Assessing the statistical resources required by SQGEN, one can show that estimating the cost function J to within accuracy ε with success probability at least $1 - \delta$ requires at most:

$$N = \mathcal{O}\left(\frac{\log(1/\delta)}{\varepsilon^2}\right) \tag{3.17}$$

measurements. This bound follows directly from the Hoeffding inequality applied to the Bernoulli random variable describing the outcome of a single circuit measurement. Because SQGEN requires only one measurement circuit per sample [22], the bound (3.17) applies to the full learning algorithm without any additional multiplicative constants arising from the number of separately evaluated circuits. This contrasts with standard QGAN training, where

each epoch requires independent evaluations of the discriminator and generator circuits, introducing an additive overhead in the total measurement count.

3.4 Hardware requirements and hyperparameter reduction

One of the most significant advantages of SQGEN over standard QGANs is its ability to reduce hardware requirements and the number of hyperparameters that need to be tuned. In this section, we conduct a systematic analysis of qubit efficiency and circuit depth optimization. We also provide a direct comparison of computational resources between SQGEN and classical QGANs.

3.4.1 Qubit efficiency analysis

The fundamental measure of quantum efficiency is the number of qubits needed to solve a given problem. The SQGEN circuit uses one decision qubit for the discrimination mechanism, n working qubits for the generated and actual states, and optionally one additional qubit for direct monitoring of fidelity F. This gives a total of $n + 1$ qubits for basic functionality (computing J), or $n + 2$ qubits if we want to track fidelity as well. In contrast, a standard QGAN requires $3n + 2$ qubits: n qubits for the generator, n qubits for the first copy of the source (for comparison with the generator), n qubits for the second copy of the source (for training the discriminator on actual states), and two auxiliary qubits.

Reducing from $3n + 2$ to $n + 1$ qubits results in a savings of approximately 67% for large n. This reduction has key practical implications for NISQ devices. First, the problem is within the scope of available resources. Modern quantum processors offer 27–156 qubits. For SQGEN, this enables the generation of states with 20+ qubits. However, QGAN is limited to eight qubits on a 27-qubit processor. Second, qubit quality in NISQ devices is uneven; the fewer qubits used, the greater the chance of finding a high-quality subset.

Third, each additional qubit introduces additional sources of error. In a typical IBM processor, the fidelity of single-qubit gates is 99.9%, while the fidelity of two-qubit gates is between 98% and 99%. For an n-qubit circuit, the total error scales approximately as $(1 - f)n$, where f is the average gate fidelity. Therefore, reducing from $3n$ to n qubits leads to a threefold reduction in error accumulation, which can mean the difference between a working and nonworking algorithm in NISQ conditions.

Fourth, qubit connectivity also favors SQGEN. In IBM processors' hexagonal topology, each qubit is directly connected to only two or three neighbors. Operations between distant qubits require chains of SWAP gates, with each SWAP consisting of three CNOT gates. For a

QGAN with 3n qubits, there is a much higher probability that the required qubit pairs will not be directly connected than for an SQGEN with n qubits. This leads to deeper circuits and more errors.

3.4.2 Circuit depth optimization

Circuit depth, or the number of layers of gates, is a critical parameter that determines the feasibility of a quantum algorithm on NISQ devices. The deeper the circuit, the longer the execution time and the greater the accumulation of decoherence errors.

For the universal Möttönen ansatz used in our implementations, a single unitary operation on n qubits requires 4^n parameters and can be decomposed into a sequence of single-qubit rotations and multi-qubit controlled gates. After decomposing the multi-qubit controlled gates into two-qubit gates (typically CNOT gates) and single-qubit rotations, the circuit depth scales as approximately $O(4^n)$ for the generator G and the discriminator D.

However, the key difference between SQGEN and QGAN lies not in the depth of a single circuit but in the number of circuits required to evaluate the cost function in a single iteration. SQGEN requires executing a single circuit with the structure $D_\lambda - X - G_\lambda^\dagger - D_\lambda^\dagger$, where the depth is approximately $2 \cdot \text{depth}(D) + \text{depth}(G)$. QGAN requires the execution of three separate circuits:

- Comparison of R with G for calculating fidelity F: $\text{depth}(R) + \text{depth}(G)$;
- Evaluation of D on real states: $\text{depth}(R) + \text{depth}(D)$;
- Evaluation of D on generated states: $\text{depth}(G) + \text{depth}(D)$.

Therefore, the total depth for QGAN is $2 \cdot \text{depth}(R) + 2 \cdot \text{depth}(G) + 2 \cdot \text{depth}(D)$, which is greater than a single SQGEN circuit by a factor of approximately $2 - 3$, depending on the relative depths of R, G, and D.

Furthermore, these three QGAN circuits must be executed sequentially, whereas the SQGEN circuit is executed once. Additionally, each call to a separate circuit requires qubit initialization, calibration, and other overhead operations. In practice, this increases the total execution time well above the sum of the circuit depths alone.

These depths are significant for larger n, but they represent single evaluations of the cost function. For QGAN, the total cost of the three separate circuits was comparable or greater, even though the individual component circuits are shallower (e.g., the generator for $n = 2$ had depth $= 18$).

Though not used in the experiments presented here, an important method for optimizing the SQGEN algorithm is to employ a circuit based on the SWAP algorithm. The multi-qubit

controlled-SWAP gate can be decomposed into n standard Fredkin gates, which could reduce the depth but increase the Hilbert space. However, a cost-benefit analysis of this optimization is beyond the scope of this work.

3.4.3 Comparison of SQGEN and QGAN performance

We evaluated the effectiveness of two generative quantum learning approaches based on three criteria: the number of experiments required per training epoch, the depth of the implemented quantum circuits, and execution time. For GHZ states with an increasing number of qubits (n = 1, ..., 6), we performed the comparison using the BFGS optimization algorithm for 20 training epochs.

Table 3.1 summarizes the performance parameters of both methods. For QGAN, we distinguish between the generator and discriminator components, which we train alternately. In SQGEN, however, we optimize both elements simultaneously within a single circuit.

The number of experiments per training epoch shows a clear advantage for SQGEN. For $n = 1$, the SQGEN method requires 32.68 experiments, while QGAN needs a total of 150.03 (the sum of the discriminator and generator components). This disparity decreases as n increases—for $n = 6$, SQGEN performs 176.79 experiments compared to 295.94 for QGAN. This reduction is due to the synergic nature of SQGEN, which optimizes both components simultaneously without the need for alternating training.

The depth of the quantum circuit scales exponentially with the number of qubits in both methods. Möttönen's universal parameterization requires 4^n parameters for an n-qubit unitary circuit. For $n = 6$, the depth of the SQGEN circuit reaches 5385 layers, which in practice exceeds the capabilities of modern NISQ devices due to the accumulation of decoherence errors. QGAN divides this complexity between the discriminator (4159 layers) and the generator (979 layers), but the total depth remains comparable. For practical applications, it is necessary to use ansatze with more favorable scaling, such as hardware-efficient ansatze.

The execution times of the simulations demonstrate the competitiveness of both approaches for small systems ($n \leq 3$). For $n = 3$, SQGEN takes 52.87s per epoch, while QGAN takes 39.80s. As the size of the system increases, so does the difference in execution time. For example, SQGEN requires an average of 1558.58 seconds for $n = 6$, while QGAN requires 999.93s. The significant variance in execution times results from the sensitivity of the optimization algorithms to the choice of starting parameters. In some configurations, SQGEN achieved minimum execution times comparable to those of QGAN; in others, however, the convergence process was significantly prolonged.

The key difference between the two methods is the need to select hyperparameters. QGAN requires setting the number of training iterations for the discriminator and generator

Table 3.1 Comparison of QGAN and SQGEN performance for 20 training epochs of GHZ states of varying sizes. Simulations were performed using the BFGS algorithm. Execution times are averaged over 5 runs on a workstation equipped with Intel Xeon X5690 @ 3.47 GHz processor. The tabulated data corresponds to Fig. 3.8.

n	Feature	SQGEN	QGAN	
			Discriminator	Generator
1	Experiments per epoch	32.68	19.3	130.73
	Circuit depth	27	19	5
	Average time per epoch [s]	0.93	1.52	
	Range of time per epoch [s]	$0.88 - 1.06$	$1.33 - 1.77$	
2	Experiments per epoch	61.01	46.8	124.86
	Circuit depth	92	71	18
	Average time per epoch [s]	10.87	9.97	
	Range of time per epoch [s]	$10.08 - 11.87$	$8.65 - 11.92$	
3	Experiments per epoch	92	70.72	147.25
	Circuit depth	317	239	55
	Average time per epoch [s]	52.87	39.80	
	Range of time per epoch [s]	$51.14 - 55.34$	$34.73 - 44.90$	
4	Experiments per epoch	136.42	102	229.19
	Circuit depth	976	747	174
	Average time per epoch [s]	229.44	174.34	
	Range of time per epoch [s]	$62.03 - 470.08$	$135.01 - 288.03$	
5	Experiments per epoch	127.69	156.66	172.83
	Circuit depth	2404	1851	434
	Average time per epoch [s]	516.45	515.85	
	Range of time per epoch [s]	$172.25 - 628.61$	$300.55 - 925.95$	
6	Experiments per epoch	176.79	137.5	158.44
	Circuit depth	5385	4159	979
	Average time per epoch [s]	1558.58	999.93	
	Range of time per epoch [s]	$443.82 - 2621.82$	$718.19 - 1148.19$	

in each epoch, necessitating experimental optimization through trial and error. SQGEN optimizes both components simultaneously, eliminating this problem. For GHZ states, we identified optimal QGAN hyperparameter configurations, but this process required additional simulations not included in Table 3.1.

In some cases, both algorithms converged to suboptimal solutions. For $n > 1$, we observed situations in which neither SQGEN nor QGAN achieved perfect fidelity ($F = 1$) after 20 epochs. Sensitivity to initial conditions is a common challenge for both approaches and requires a random optimization restart strategy.

For learning small GHZ states ($n \leq 3$), QGAN exhibits slightly faster convergence times when hyperparameters are chosen appropriately. SQGEN is simpler to implement and requires fewer experiments, but it has deeper circuits and longer simulation times for larger systems. In the context of implementation on real quantum devices, the exponential increase in circuit depth is a limitation for both methods and requires a transition to a more efficient variational ansatz.

3.5 Performance analysis and benchmarking

We systematically evaluated the performance of SQGEN and compared it with that of classical QGANs. We performed a series of numerical simulations to generate GHZ multi-qubit states. GHZ states, such as $|\Psi_{GHZ}\rangle = (|0\rangle^{\otimes n} + |1\rangle^{\otimes n})/\sqrt{2}$, are a canonical example of maximum multiparty entanglement. They are also widely used to test quantum algorithms. In this section, we analyze the convergence of the learning process, the fidelity metrics, and the numerical simulation results for $n = 1, 2, 3, 4, 5, 6$ qubits.

3.5.1 Convergence analysis

The convergence dynamics of the learning process are a key indicator of the algorithm's practical usefulness. For both SQGEN and QGAN, the evolution of three quantities was tracked as a function of the training epoch: the probability p of the discriminator correctly classifying the real state, the probability q of the discriminator incorrectly classifying the generated state as real, and the fidelity $F = \text{Tr}(\sigma_\lambda \rho_\lambda)$ between the mean states of the generator and the source.

For small problem sizes ($n = 2$ or 3), both approaches converge at a similar rate. Typically, after 10–15 epochs, fidelity (F) reaches values above 0.95, and the probabilities (p and q) stabilize at levels indicating a well-functioning discriminator. The learning process is relatively smooth, without large oscillations or sudden jumps in cost function values.

The key difference lies in how p and q evolve. For QGAN, these two quantities are optimized in separate phases: first, the discriminator is trained to maximize $|p - q|$; then, the generator is trained to increase q and improve F. This leads to characteristic oscillations where p increases (the discriminator improves), then q increases (the generator catches up), then p increases again, and so on. The amplitude of these oscillations critically depends on the ratio k_D/k_G. An inappropriate choice of this ratio can lead to large oscillations or even divergence.

In SQGEN, both p and q are optimized simultaneously via a single cost function, J. This results in a smoother evolution, with both quantities increasing monotonically (with small fluctuations) and without large oscillations. Although SQGEN generally requires fewer epochs to stabilize the values of p and q, QGAN is slightly faster in achieving a high fidelity F.

For larger sizes ($n = 4, 5, 6$), the picture becomes more complicated. The parameter space has a dimension of 4^n (for Möttönen's ansatz), which, for $n = 6$, results in 4096 parameters for the generator alone. In such a high-dimensional space, it is easy to get stuck in local minima. There have been cases where:

- The QGAN for $n = 6$ converged to a suboptimal solution with $F \approx 0.8$ and did not improve with additional epochs. This problem was solved by modifying the discriminator cost function by adding a component proportional to $1 - p$, which introduces an additional hyperparameter.
- SQGEN also sometimes settled in local minima, but less frequently than QGAN. For $n = 6$, it converged to solutions with $F > 0.95$ in most trials without modifying the cost function.
- For both methods, proper parameter initialization is critical. Random initialization can lead to suboptimal solutions, whereas initialization close to the identity matrix (with parameters close to zero) yields better results.

A key advantage of SQGEN is that it converges without the need to tune hyperparameters that control the balance between components. For QGAN, a significant number of trials with different values of k_D/k_G and weights in the discriminator cost function were necessary to find a configuration that yielded stable convergence. The only hyperparameter for SQGEN was the choice of optimization method (BFGS) and its standard parameters, which worked consistently for all n.

3.5.2 Fidelity metrics

The fidelity, F, between the generated and actual states is the primary metric for evaluating the success of generative learning. For pure states, fidelity is defined as:

$$F = |\langle\psi_G|\psi_R\rangle|^2 = \text{Tr}(\sigma\rho), \tag{3.18}$$

where σ and ρ are the density matrices of the generated and actual states, respectively. A value of $F = 1$ indicates perfect agreement, and a value of $F = 0$ indicates that the states are orthogonal.

For GHZ states, which are maximally entangled, fidelity is an especially demanding metric. Small errors in state preparation can lead to a significant drop in fidelity due to entanglement's nonlinear nature.

3.5.3 Numerical simulations

We analyze the dynamics of the learning process for maximally entangled GHZ states of the form $|\Psi\rangle = (|0\rangle^{\otimes n} + |1\rangle^{\otimes n})/\sqrt{2}$, where $n = 1, 2, 3, 4, 5, 6$ determines the number of qubits. Due to the nonlocal nature of entanglement, these states naturally present a challenge to generative learning: the generator must learn to correlate all qubits simultaneously instead of preparing each one independently.

Figure 3.8 shows the evolution of three key quantities during 20 training epochs: the probability p of the discriminator recognizing the actual state, the probability q of misclassifying the generated state as actual, and the fidelity $F = \mathrm{Tr}(\rho\sigma)$ between the source state and the generated state. We performed the simulations without noise using the BFGS optimization algorithm.

For a two-qubit system ($n = 2$, panel a), both algorithms achieve a fidelity greater than 0.99 within the first five epochs. QGAN stabilizes the probabilities p and q slightly faster; the difference between $1 - p$ and $1 - q$ falls below 0.05 after only three epochs, whereas SQGEN requires about six epochs. The final fidelity is comparable for both methods. The values of $1-p$ and $1-q$ converge to zero, which means that the discriminator has learned to distinguish perfectly between real and generated states, while the generator produces indistinguishable states.

Increasing the system dimension to $n = 3$ (panel b) reveals the first differences, such as convergence. QGAN achieves a fidelity of $F \approx 0.98$ after about three epochs and maintains it stably for the remaining iterations. SQGEN shows a slower increase in fidelity - it reaches $F \approx 0.75$ after ten epochs and slowly improves to $F \approx 0.95$ in epoch 20. The discrimination probabilities behave similarly to those for $n = 2$, except that in SQGEN, the value of $1 - p$ stabilizes at around 0.25 instead of falling to zero.

For $n = 4$ (panel c), the differences become more pronounced. QGAN converges to a high fidelity of approximately 0.99 in about five epochs. The discriminator quickly learns to distinguish between states. Initially, the difference between $1 - p$ and $1 - q$ is large for the first few epochs. Then, as the generator improves its performance, both values decrease. SQGEN shows much slower convergence in this configuration. Fidelity increases gradually from $F \approx 0.5$ to $F \approx 0.85$ over 20 epochs. The value of $1 - p$ remains at around 0.25, suggesting that the discriminator does not achieve optimal performance with this number of qubits in a given number of epochs.

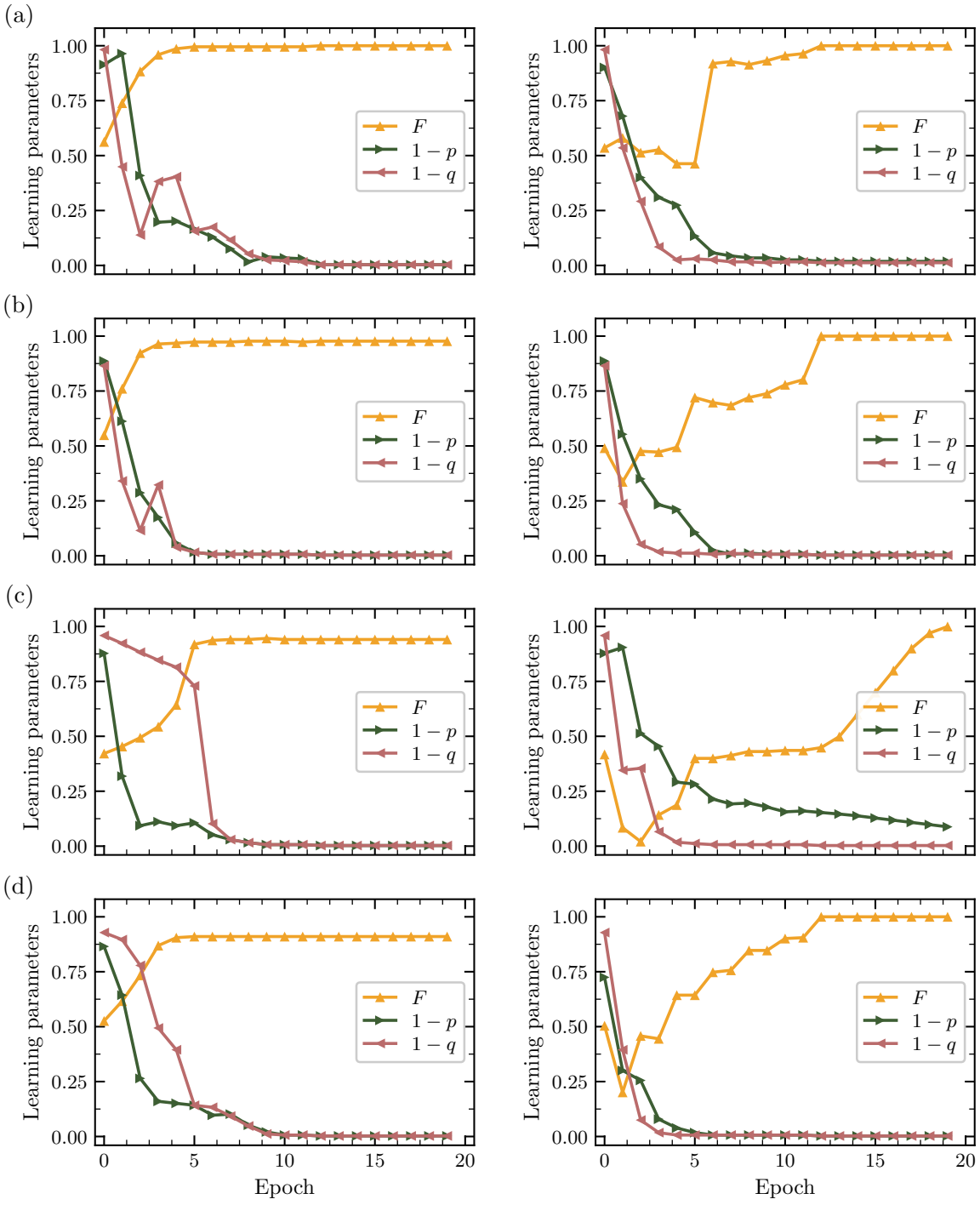


Fig. 3.8 **Comparison of QGAN (left column) and SQGEN (right column) training dynamics for generating n-qubit GHZ states.** Rows correspond to (a) $n = 2$, (b) $n = 3$, (c) $n = 4$, (d) $n = 5$. Yellow triangles show fidelity F between generated and real states, green triangles show $1-p$, where p is the probability of a source state being recognized as real by the discriminator. Red triangles show $1 - q$ which correspond to the probability q of a generated state being recognized. Noiseless simulations performed using *statevector_simulator* with BFGS optimizer for 20 training epochs.

The most complex case, $n = 5$ (panel d), shows that both algorithms achieve high final fidelity $F \approx 0.99$, but with different dynamics. QGAN converges rapidly in the first 4 epochs, while SQGEN shows a more gradual increase throughout the training period. Notably, the differences in discrimination probabilities are smaller in this case than in the $n = 3$ and $n = 4$ cases, potentially due to the specific configuration of the initial parameters used in the simulation.

3.6 Implementation on real quantum processors

Implementing SQGEN on real quantum processors enables us to evaluate the method's feasibility in the NISQ era and pinpoint the primary sources of errors that hinder learning performance. We conducted a proof-of-principle experiment with the simplest non-trivial example: a single-qubit generator of Pauli matrix basis states.

3.6.1 Experimental setup on *ibmq_manila*

We conducted our experiments on the *ibmq_manila* processor, which was one of the five-qubit IBMQ processors based on the Falcon r5.11L architecture with a quantum volume of 32. The processor is based on superconducting transmon qubits. Table 3.2 contains the qubit parameters measured during calibration at the time of the experiment.

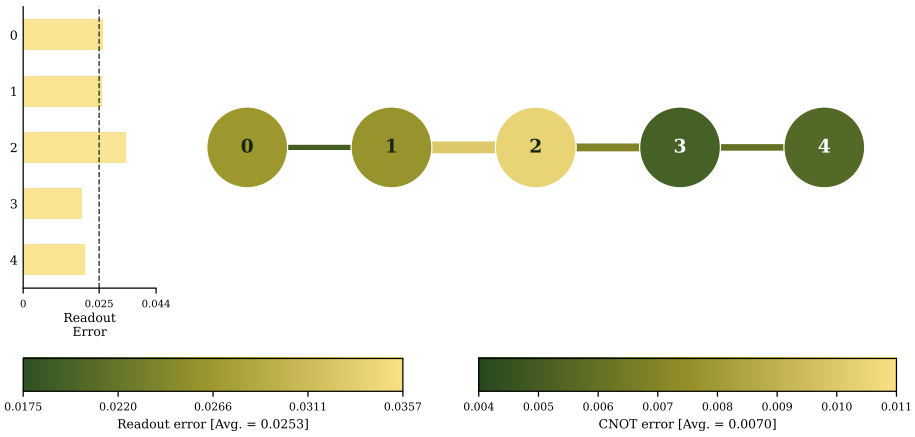


Fig. 3.9 **Calibration data of the *ibmq_manila* five-qubit processor.** The Hadamard gate H appears explicitly in the initialization circuits. Full calibration parameters, including relaxation times T_1, decoherence times T_2, gate error rates, and qubit frequencies, are summarized in Table 3.2.

The relaxation times T_1 range from 97.8 to 231.9 μs, while the phase decoherence times T_2 range from 24.1 to 108.8 μs. Qubit 2 is characterized by an especially short time $T_2 = 24.069$

Table 3.2 The parameters of the *ibmq_manila* processor's qubits, calibrated at the time of the experiment.

Qubit	T_1 [μs]	T_2 [μs]	Frequency [GHz]	Anharmonicity [GHz]	Readout error	CNOT error
0	97.769	108.828	4.9623	-0.345	0.0265	0_1: 0.005
1	178.823	76.082	4.838	-0.345	0.026	1_2: 0.010; 1_0: 0.005
2	130.887	24.069	5.037	-0.343	0.034	2_3: 0.007; 2_1: 0.010
3	231.879	73.581	4.951	-0.344	0.0194	3_4: 0.006; 3_2: 0.007
4	120.301	42.614	5.065	-0.342	0.0204	4_3: 0.006

μs, making it most susceptible to phase errors. Readout errors range from 1.94 to 3.4%, and CNOT two-qubit gate errors range from 0.5% to 1.0%. These parameters determine the maximum circuit depth that can be achieved while maintaining useful quantum coherence.

In the experiment, the source $\mathcal{R}$ randomly prepares one of the six basis states of the Pauli matrix: $|0\rangle$, $|1\rangle$, $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$, $|-\rangle = (|0\rangle - |1\rangle)/\sqrt{2}$, $|+i\rangle = (|0\rangle + i\,|1\rangle)/\sqrt{2}$, $|-i\rangle = (|0\rangle - i\,|1\rangle)/\sqrt{2}$. The parameter λ determines the basis ($\lambda \in \{x, y, z\}$), and the random variable $z_R \in \{0, 1\}$ selects one of the two states in this basis. The task of SQGEN is to learn how to prepare and recognize all six states using a discriminator.

To simplify the analysis, we focus on the special case $\lambda = x$ and $z_G = 0$, where the generator attempts to learn only the state $|+\rangle$, while the source $\mathcal{R}$ randomly produces $|+\rangle$ or $|-\rangle$. This enables us to isolate the fundamental learning process, eliminating the additional complexity of multi-class classification.

The physical implementation of the circuit in Figure 3.7(b) on the *ibmq_manila* processor requires transpilation to the processor's native gate set. The Falcon architecture physically implements controlled-phase (CZ) gates, single-qubit rotations, and measurements. The Qiskit compiler [41] automatically decomposes CNOT gates and multi-qubit controlled gates from Möttönen's ansatz into sequences of CZ gates and single-qubit rotations. For the three-qubit SQGEN circuit (decision qubit, generator qubit, auxiliary qubit for the discriminator), the final depth is 27 layers.

We used the Nelder-Mead algorithm instead of gradient methods to optimize circuit parameters. Simplex methods are better at handling noise in cost function evaluations, which is crucial for experiments on real devices. Each cost function evaluation requires 20000 measurements, or "shots." With a circuit depth of 27, this process takes approximately 15 seconds. Using the random starting point shown in Figure 3.10, the algorithm required

an average of 260 cost function evaluations to complete 15 training epochs (one epoch corresponds to five Nelder-Mead iterations).

3.6.2 NISQ limitations and noise analysis

Figure 3.10 shows the learning process on a real quantum processor. The connected points represent the actual values measured during training. The shaded areas show the range of values obtained from 100 Monte Carlo simulations using the *FakeManila()* fake provider from the Qiskit library [41].

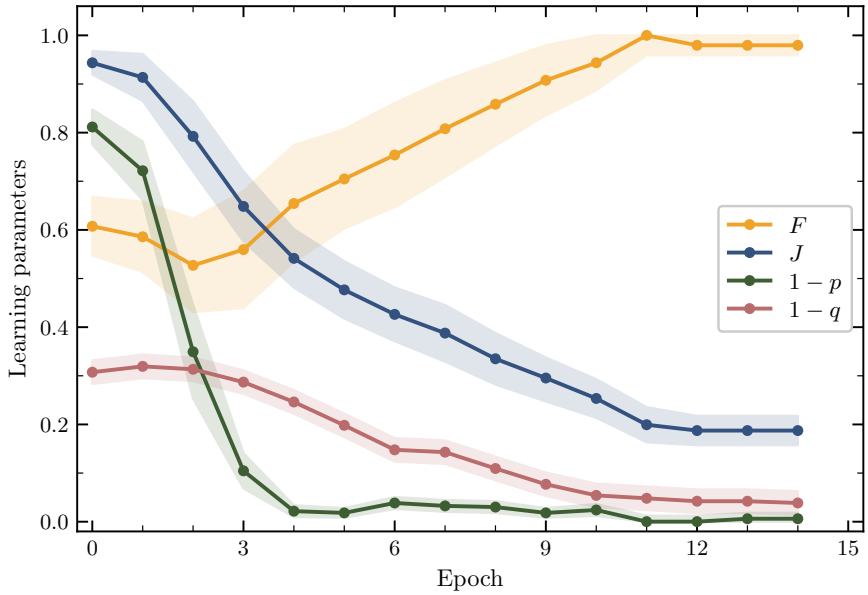


Fig. 3.10 **Experimental learning process performed on *ibmq_manila* quantum processor using three qubits.** Connected points represent actual values measured during training. Shaded areas depict the range of values obtained in 100 Monte Carlo simulations with *FakeManila()* fake provider, accounting for shot noise and transmon decoherence. Dark blue circles show cost function J, yellow circles show fidelity F, red circles show $1-q$, green circles show $1-p$. The shaded area does not include all points corresponding to J measurements on the real processor, suggesting the manufacturer's noise model may be inadequate for circuits of depth 27.

Monte Carlo simulations consider both Poisson noise resulting from a finite number of measurements and transmon decoherence according to the parameters in Table 3.2. However, the shaded area does not cover all the points corresponding to measurements of the cost function J on the actual processor. This suggests that the noise model provided by the manufacturer for *FakeManila()* may be inadequate for circuits with a depth of 27. The

probabilities p and q and the fidelity F are within the expected range, suggesting that the main deviations come from correlated errors not included in the model.

The algorithm successfully converges, the fidelity F increases from an initial value of approximately 0.5 to a final value of about 0.92 over 15 epochs. The probability of correct identification of a state generated as artificial by the discriminator $1 - q$ decreases from 0.8 to approximately 0.1. This indicates that the generator has learned to produce states that closely resemble real ones. Simultaneously, $1 - p$ decreases from 0.8 to approximately 0.2, suggesting that the discriminator correctly identifies about 80% of real states. The cost function J decreases from about 0.3 to 0.2, though it fluctuates more than the other quantities.

The main sources of error in the experiment fall into four categories. Phase decoherence arising from T_2 times in the range 24–109 μs limits the coherence window; for a circuit of depth 27, the execution time exceeds 1 μs, corresponding to 1–4% of T_2 depending on the qubit used, and resulting in a loss of approximately 2–8% in the amplitude of superposition states. Gate errors of 0.5–1.0% per CNOT gate accumulate across the roughly 10–15 two-qubit gates present after transpilation, yielding a cumulative error of 5–15%. Readout errors at the 2–3% level translate directly into inaccuracy in the cost function evaluation; for 20,000 measurements the statistical contribution is $\sqrt{0.03 \times 0.97/20{,}000} \approx 0.001$, although systematic calibration errors may be larger. Finally, automatic transpilation to the native gate set may introduce additional gates that increase circuit depth, making compiler-level optimization important for limiting this effect.

We applied a standard measurement error mitigation method corresponding to calibrating the detection part of a quantum computer. This method involves measuring the error matrix for each qubit by preparing the basis states $|0\rangle$ and $|1\rangle$ and observing the frequency of misclassification. The inverted matrix is then applied to correct the measured probability distributions. This simple technique improves measurement accuracy by 30-50% without requiring additional quantum resources.

The algorithm converged in approximately half of the experimental runs. If the algorithm did not converge, it stopped at a local minimum with a fidelity of $F \approx 0.6 - 0.7$. The user can restart the algorithm with a different starting point until convergence is achieved. In simulations without noise, the algorithm converged in each trial with a final cost function below 0.001, confirming that convergence problems are due solely to noise in real processors.

3.6.3 Single vs multi-qubit performance

The success of the single-qubit implementation contrasts with the difficulties encountered in larger systems. Attempts to extend the experiment to $n = 2$ qubits (preparing states of

the type $|00\rangle, |01\rangle, |10\rangle, |11\rangle$ or Bell states) encountered fundamental limitations related to circuit depth and coherence time.

For $n = 2$, Möttönen's ansatz requires $4^2 = 16$ parameters, which translates to a circuit depth of 92 layers (see Table 3.1). The execution time of such a circuit on *ibmq_manila* is approximately 10-15 μs, which corresponds to 10-60% of the T_2 time depending on the qubits used. For qubit 2, which has the shortest $T_2 = 24$ μs time, this results in a loss of approximately 40-50% of the amplitude of the superposition states, which makes the results practically useless.

We performed a series of experiments for $n = 2$ with 8192 measurements per cost function evaluation. The algorithm did not converge in any of the ten trials with different starting points. The fidelity oscillated around $F \approx 0.5 - 0.6$ without a clear upward trend. The probabilities p and q remained close to 0.5, indicating that the discriminator did not learn to distinguish between states. Implementing measurement error mitigation did not significantly improve the results, suggesting that decoherence is the main source of the problem rather than readout errors.

A key limitation is the exponential growth in the number of parameters with the number of qubits in Möttönen's ansatz. For $n = 3$, $4^3 = 64$ parameters and a depth of 317 layers are required, and for $n = 4$, 256 parameters and 976 layers are required. These values significantly exceed the capabilities of current NISQ devices, for which the practical limit of circuit depth while maintaining useful coherence is around 50-100 layers, depending on the processor.

For larger systems, the transition to a more efficient ansatz is inevitable. A hardware-efficient ansatz reduces the number of parameters to $\mathcal{O}(n)$ at the cost of losing the guarantee of universal approximation. For specific problems, such as preparing GHZ states, dedicated ansatze can be designed to exploit the target state's specific structure. For instance, the GHZ state $(|0\rangle^{\otimes n} + |1\rangle^{\otimes n})/\sqrt{2}$ can be prepared with a circuit of depth n using a Hadamard gate and a CNOT chain, without the need for variational optimization.

An alternative approach is to use dynamical decoupling techniques to reduce the effect of decoherence. This method involves inserting a sequence of fast pulses to compensate for magnetic field fluctuations, the main cause of depolarization. However, this approach necessitates modifications at the hardware control level, a capability that is often unavailable to users of cloud-based quantum platforms.

Experimental results confirm the feasibility of SQGEN on current NISQ devices for small systems ($n = 1$). However, extending it to larger dimensions requires hardware improvements, such as longer coherence times and smaller gate errors, as well as algorithmic improvements, such as more efficient ansatz and better error mitigation methods. The success of SQGEN

with $n = 1$ is a proof of concept, demonstrating that synergic generative learning is physically feasible. The challenges encountered with $n > 1$, however, point to areas for further research on adapting the method to the constraints of the NISQ era.

3.7 Synergic optimization and the structure of Nash equilibria

A recurring observation in the numerical benchmarks of Section 3.5 is that SQGEN reaches high fidelity reliably and without hyperparameter tuning, whereas QGAN requires careful balancing of the discriminator-to-generator training ratio. This empirical contrast has a precise theoretical explanation rooted in the geometry of the optimization landscape and the structure of quantum state space. The analysis presented in this section constitutes an original theoretical contribution of this dissertation.

3.7.1 Why quantum state space does not require adversarial training

The GAN framework was introduced to address a specific difficulty of classical generative modeling: high-dimensional data spaces such as images or audio waveforms possess no tractable, semantically meaningful similarity measure known a priori. The ℓ^2 norm between pixel arrays does not capture perceptual similarity; two face images may differ greatly in ℓ^2 distance while a face and random noise may differ little after spatial averaging. Computing distributional distances such as the Kullback-Leibler or Jensen-Shannon divergences requires density estimation, which is intractable in high dimensions. The GAN discriminator resolves this by learning a metric from data. Goodfellow et al. [70] showed that the optimal discriminator value function equals the Jensen-Shannon divergence, so the adversarial game implicitly provides the generator with a meaningful loss landscape.

Quantum state spaces do not share this deficiency. For pure states $|\psi\rangle, |\phi\rangle$ in an n-qubit Hilbert space, the fidelity $F(|\psi\rangle, |\phi\rangle) = |\langle\psi\rangle \phi|^2$ provides an axiomatic, efficiently computable, and semantically exact similarity measure. It satisfies $F = 1$ if and only if the states coincide up to a global phase; it is unitarily invariant; it obeys the data processing inequality under quantum channels; and it metrizes the state space through the Fuchs-van de Graaf inequalities,

$$1 - \sqrt{F(|\psi\rangle, |\phi\rangle)} \;\leq\; D_{\text{Tr}}(|\psi\rangle, |\phi\rangle) \;\leq\; \sqrt{1 - F(|\psi\rangle, |\phi\rangle)}, \tag{3.19}$$

Table 3.3 Comparison of generative learning paradigms. In the quantum setting, fidelity provides the exact metric that classical methods must learn adversarially.

Method	Loss function	Metric	Requires adversary	Stable
Classical autoencoder	$\|x - \hat{x}\|^2$	Poor (ℓ^2)	No	Yes
Variational AE	ELBO	Moderate	No	Yes
Classical GAN	Adversarial	Learned	Yes	No
QGAN	Adversarial	Learned†	Yes	No
SQGEN	$1 - F$	Exact (fidelity)	No	Yes
Quantum autoencoder	$1 - P(\|0\rangle)$	Exact (fidelity)	No	Yes

where D_{Tr} denotes the trace distance. On quantum hardware, F is accessible via the SWAP test using $O(\varepsilon^{-2})$ measurements, without any classical post-processing or density estimation. The discriminator's role as a learned metric is therefore structurally unnecessary for quantum state generation: the optimal loss function $J = 1 - F$ is available from the outset. The QGAN paradigm imports a solution to a problem that does not exist in Hilbert space.

Table 3.3 summarizes this contrast across generative learning paradigms. The progression in classical machine learning from autoencoders through variational autoencoders to GANs was driven by the need for progressively better metrics. In the quantum setting, the fidelity occupies the role that no classical measure can fill, rendering the adversarial step unnecessary.

3.7.2 Synergic minima as Nash equilibria

The relationship between the synergic cost function and the adversarial game can be made precise by examining the gradient structure of the two optimization objectives. Consider the cost function (3.14) for a fixed generator parameter θ_G with $\sigma \neq \rho$ (the generator has not yet converged). The partial derivative with respect to the discriminator rotation angle θ_{z_D} is:

$$\frac{\partial J}{\partial \theta_{z_D}} = 2 \cos\theta_{z_D} \sin\theta_{z_D} \cdot \mathcal{F}(\theta_G), \tag{3.20}$$

where $\mathcal{F}(\theta_G) = \sum_{z_G, z_R} p_\theta(z_G)\, q(z_R)\, \text{Tr}(\sigma_{z_G} \rho_{z_R})$ is the effective fidelity between the generator and source assemblages. The synergic discriminator minimizes J, which drives $\theta_{z_D} \to 0$ (the cooperative regime: the discriminator labels all states as real). The Nash discriminator of the adversarial game instead maximizes distinguishability, driving $\theta_{z_D} \to \pi/2$. Away from the fixed point $\sigma = \rho$, the synergic and Nash gradients with respect to θ_D are therefore antiparallel.

Despite this conflict away from convergence, the two objectives coincide at the solution. The following result establishes the precise relationship.

Theorem (Synergic minimum implies Nash equilibrium). *Let* (θ_G^*, θ_D^*) *be a synergic minimum of the cost function* (3.14) *with* $J(\theta_G^*, \theta_D^*) = 0$. *Then* (θ_G^*, θ_D^*) *is a Nash equilibrium of the quantum adversarial game.*

The proof proceeds in two steps. The condition $J = 0$ implies $F(\sigma, \rho) = 1$, hence $\sigma = \rho$ up to global phase. Generator stability then follows because any deviation from θ_G^* yields $F < 1$, increasing the cost, so the generator cannot improve unilaterally. Discriminator stability follows from an indifference argument: at $\sigma = \rho$, the trace $\mathrm{Tr}(M_D\sigma) - \mathrm{Tr}(M_D\rho)$ vanishes for every measurement operator $M_D(\theta_D)$, so the discriminator's payoff is constant in θ_D and no unilateral deviation can improve it.

A corollary of this result is that the set inclusion runs in one direction only. The set of synergic minima with $J = 0$ is a proper subset of the Nash equilibrium manifold. At $\sigma = \rho$, every discriminator configuration is Nash-optimal (the discriminator is indifferent), so the Nash equilibrium set has the full dimension $\dim \Theta_D$. The synergic minimum selects a unique point on this manifold, namely the cooperative configuration $\theta_{z_D} = 0$, which corresponds to complete cooperation between generator and discriminator. This selection mechanism explains the improved stability of SQGEN: rather than searching over the entire Nash manifold as the adversarial dynamics do, synergic optimization converges to a single well-defined cooperative solution.

3.7.3 Conditions for equivalence and failure modes

The implication of the theorem is one-directional: every synergic $J = 0$ minimum is Nash-optimal, but the converse does not hold in general. The question of when the two methods produce the same generator state in practice depends on four conditions: (C1) the ansatz is sufficiently expressive to achieve $J = 0$; (C2) the Nash equilibrium satisfying $F = 1$ is unique in generator state space, ruling out mode collapse; (C3) the optimizer reaches the global rather than a local minimum; and (C4) the target state $\rho = |\phi\rangle\langle\phi|$ is pure, so that fidelity $F = 1$ uniquely identifies the state. When all four conditions are met, the synergic and Nash solutions coincide in state space. All four conditions are naturally satisfied for pure-state generation with a universal ansatz, which is the regime studied in the GHZ benchmarks of Section 3.5.

The failure modes are instructive. When condition (C1) fails because the ansatz cannot represent the target state, the two methods diverge in a characteristic way: the synergic discriminator drives $\theta_{z_D} \to 0$, effectively hiding the generator's deficiency by treating

the mismatch as genuine, whereas the Nash discriminator drives $\theta_{z_D} \to \pi/2$, maximally exposing the discrepancy. This produces qualitatively different generator states: a synergic approximation σ_S that minimizes the cost subject to the ansatz constraint, and a Nash approximation σ_N that minimizes distinguishability from the target. The practical implication is that SQGEN with an under-expressive ansatz will converge to a stable but potentially misleading solution, whereas QGAN may oscillate or diverge. The observed susceptibility of SQGEN to degenerate minima for mixed-state targets (Section 3.5) is a manifestation of condition (C4) failing: when $\rho_{\text{target}} = \mathbb{I}/2^n$, many generator states achieve $J = 0$, and SQGEN may converge to any of them.

3.7.4 When adversarial training remains useful

The redundancy of adversarial training established above applies specifically to pure-state generation with coherent access to both the target preparation circuit and the generator. Three scenarios exist in which a discriminator provides a genuine benefit that direct fidelity optimization cannot replicate.

The first scenario concerns unknown mixed-state targets. If ρ is a mixed state accessible only through a preparation device with no classical description, computing $F(\sigma, \rho)$ directly requires full state tomography at cost $O(4^n)$. A discriminator trained on measurement outcomes can provide a useful gradient signal from polynomially many measurements, without reconstructing the full density matrix.

The second scenario arises when the target is a probability distribution $p(x)$ over measurement outcomes rather than a quantum state. This is the classical generative modeling problem embedded in a quantum circuit, and it inherits all the metric difficulties of the classical setting: fidelity is not defined, and a learned discriminator is again necessary.

The third scenario involves incoherent access: if the target preparation circuit $\mathcal{R}$ and the generator $\mathcal{G}$ cannot be coherently composed, for instance because they run on physically separate processors, the SWAP test required to evaluate F is not directly applicable. A discriminator operating locally on each processor can circumvent this constraint.

These boundary conditions delineate the operating regime of SQGEN and clarify that the choice between synergic and adversarial training is not universal but depends on the structure of the generative task.

3.7.5 SQGEN as a quantum autoencoder

The SQGEN circuit sequence,

$$|0\rangle^{\otimes(n+1)} \xrightarrow{\mathcal{R}} \xrightarrow{\mathcal{D}} \xrightarrow{X_{n+1}} \xrightarrow{\mathcal{D}^\dagger} \xrightarrow{\mathcal{G}^\dagger} \longrightarrow P(|0\rangle), \tag{3.21}$$

admits a natural interpretation as a quantum autoencoder [79]. The source circuit $\mathcal{R}$ prepares the target state from $|0\rangle$, playing the role of the encoder in an autoencoder. The discriminator $\mathcal{D}$ together with the ancilla qubit compresses the state information through a single-qubit channel, constituting the bottleneck. The inverse pair $\mathcal{D}^\dagger\mathcal{G}^\dagger$ then attempts to reverse the encoding and map the system back to $|0\rangle^{\otimes(n+1)}$, playing the role of the decoder. The cost function $J = 1 - P(|0\rangle^{\otimes(n+1)})$ is the probability of failing to recover the initial state, which is precisely the reconstruction loss of a quantum autoencoder.

In the synergic regime $\theta_{z_D} = 0$, the bottleneck becomes trivial: information passes through without compression and the circuit reduces to $\mathcal{G}^\dagger\mathcal{R}\,|0\rangle$. The cost function then simplifies to:

$$J = 1 - |\,\langle 0|^{\otimes n}\,\mathcal{G}^\dagger\mathcal{R}\,|0\rangle^{\otimes n}\,|^2 = 1 - F(\sigma, \rho), \tag{3.22}$$

which is structurally identical to the quantum autoencoder of Romero et al. [79], with $\mathcal{R}$ as the encoder and $\mathcal{G}$ as the decoder. The suggestion in Section 3.5 that the SQGEN architecture lends itself to quantum information compression is therefore more than an analogy: in the synergic regime ($\theta_{z_D} = 0$), SQGEN reduces to a quantum autoencoder by construction, with an adaptively learned decoder.

This identification has implications beyond the generative learning context. Given a coherent error channel $\mathcal{E}$ acting on a set of states $\{|\psi_i\rangle\}$, the recovery cost:

$$J_{\mathrm{EC}}(\theta_R) = 1 - \frac{1}{N}\sum_{i=1}^{N} F\Big(R(\theta_R)\,\mathcal{E}\,|\psi_i\rangle\,,\ |\psi_i\rangle\Big), \tag{3.23}$$

has a unique global minimum at $R^* = \mathcal{E}^\dagger$ with $J_{\mathrm{EC}} = 0$ whenever the ansatz is expressive enough to represent $\mathcal{E}^\dagger$. The adversarial approach is expected to fail in this setting because, at the solution $R^* = \mathcal{E}^\dagger$, the discriminator's gradient vanishes, and the generator-discriminator dynamics lack a restoring force—a well-known instability of adversarial training near equilibrium, as discussed in [70]. The synergic formulation does not have this instability, which makes SQGEN a natural framework for variational quantum error correction of coherent noise.

3.8 Discussion

3.8.1 SQGEN advantages and limitations

Compared to QGAN, SQGEN introduces a conceptually simpler approach to generative quantum learning. The main advantage stems from the synergic nature of optimization: the generator and discriminator are trained simultaneously by maximizing a single cost function, rather than alternately optimizing competing objectives. This eliminates the need to select the number of training iterations for each component. In QGAN, this is a key hyperparameter that requires experimental tuning by trial and error.

SQGEN requires fewer experiments per training epoch. For a single-qubit system, SQGEN requires an average of 33 experiments, whereas QGAN requires 150 experiments (the sum of the discriminator and generator components). This difference decreases with the number of qubits- for $n = 6$, SQGEN performs 177 experiments compared to 296 for QGAN- but the advantage of SQGEN remains clear. This is due to the ability to use the same sample from a real data source to optimize both components, which circumvents the prohibition on cloning quantum states without the need to duplicate source calls.

The design of the synergic circuit has deeper theoretical implications. When the measurement of the decision qubit is conditioned on the state of 0, only events in which the discriminator functions correctly are selected. Learning focuses on configurations in which the actual and generated states are close to converging, which should theoretically speed up the process. Using $D^\dagger$ to make the discriminator reversible allows for the simultaneous evaluation of generation and discrimination quality in a single experiment, which is impossible in the standard QGAN approach.

The main limitation of SQGEN is its exponential increase in circuit depth with the number of qubits. Möttönen's universal parameterization requires 4^n parameters for an n-qubit system, resulting in a depth of thousands of layers for $n = 6$. In comparison, QGAN distributes this complexity between the discriminator and generator, whose depths are additive rather than multiplicative with respect to total computation time. In practice, both approaches encounter the same fundamental limitations related to decoherence in NISQ devices.

The simulation times show the competitiveness of SQGEN for small systems ($n \leq 3$) but the growing advantage of QGAN for larger dimensions. For $n = 6$, SQGEN requires an average of 1559 seconds per epoch compared to 1000 seconds for QGAN. The significant variance in times (ranges differing by a factor of 5-10) in both methods is due to the sensitivity of optimization algorithms to the choice of starting parameters. In some configurations, SQGEN achieved comparable times to those of QGAN; in others, however, the convergence

process was significantly prolonged. This unpredictability poses a practical problem, requiring a strategy of multiple restarts from different starting points.

Both algorithms tend to converge to suboptimal solutions. For $n > 1$, we observed cases in which the fidelity F did not approach unity after 20 training epochs, even with continued optimization. This issue is more prevalent for SQGEN in the large n regime, suggesting that synergic optimization may be more susceptible to becoming trapped in local minima than the alternating QGAN strategy. Advanced initialization techniques are required, such as transfer learning from smaller systems or using classical neural networks to generate a starting point.

The susceptibility of SQGEN to degenerate minima has a precise theoretical origin, analyzed in Section 3.7: when the target state is mixed, condition (C4) fails and many generator states achieve $J = 0$ simultaneously, so the optimizer may converge to any of them. An effective mitigation strategy is to run multiple restarts from different initial points and select the solution achieving the highest fidelity F, measured independently at the end of training.

Implementation on the real quantum processor *ibmq_manila* confirmed the feasibility of SQGEN for $n = 1$ with a final fidelity of $F \approx 0.92$. The algorithm converged in about half of the trials with random starting points, which is acceptable given the level of noise in current devices. However, attempts to extend to $n = 2$ encountered fundamental limitations related to circuit depth exceeding the practical coherence threshold. Therefore, SQGEN applications on NISQ devices are currently limited to very small systems or require transitioning to more efficient variational approaches.

3.8.2 Limits of quantum advantage

Although SQGEN exploits the 2^n-dimensional structure of Hilbert space, it does not offer an advantage over classical methods in all cases. Three classes of problems can be identified for which the algorithm is efficiently simulable classically.

The first class consists of product states of the form $|\psi\rangle = \bigotimes_{i=1}^{n} |\psi_i\rangle$, whose classical description requires only $\mathcal{O}(n)$ parameters. Classical GANs can learn such distributions efficiently without any quantum resources.

The second class covers circuits composed entirely of Clifford gates: Hadamard, phase, and CNOT. By the Gottesman-Knill theorem [80], any such circuit can be simulated classically in $\mathcal{O}(\mathrm{poly}(n))$ time, ruling out exponential quantum advantage.

The third class comprises states satisfying the area law of entanglement, for which efficient tensor-network representations exist with bond dimension $\chi = \mathcal{O}(\mathrm{poly}(n))$, enabling classical simulation in polynomial time [81].

It follows that the advantage of SQGEN is conditional: it applies to distributions corresponding to states with long-range entanglement whose classical description requires exponentially many parameters, yet which admit an efficient quantum circuit representation of polynomial depth. GHZ states satisfy these conditions, since they can be prepared with a circuit of depth $\mathcal{O}(n)$ using one Hadamard gate and $n-1$ CNOT gates, whereas their classical simulation requires $\mathcal{O}(2^n)$ operations. For more complex quantum distributions arising in quantum chemistry or many-body physics, analogous complexity assessments remain an open research question.

Several theoretical questions remain open. The first concerns the scaling of gradient magnitudes in SQGEN. In randomly initialized parameterized circuits, the phenomenon known as barren plateaus causes the cost function gradient to vanish exponentially with the number of qubits, rendering optimization practically infeasible for large n [82]. There is a conjecture that the synergic cost function of SQGEN exhibits polynomially rather than exponentially vanishing gradients, which would account for the improved trainability observed empirically for systems up to $n = 6$; whether this persists at larger scales remains an open question. The intuition behind this conjecture is that post-selection on the $|0\rangle$ outcome in the decision qubit concentrates learning on circuit configurations where the discriminator performs correctly, effectively reducing the parameter space explored by the optimizer. A formal proof would require an analysis of gradient variance over the unitary group for the specific circuit structure of SQGEN and remains an active area of research.

A second open question concerns the mixed-state generalization of the theoretical results in Section 3.7. For mixed states, the fidelity $F(\rho,\sigma) = \left(\mathrm{Tr}\sqrt{\sqrt{\rho}\,\sigma\sqrt{\rho}}\right)^2$ is considerably more complex than its pure-state counterpart, and the cost landscape may have nontrivial structure with multiple local maxima. The conditions (C1)–(C4) established for pure-state generation do not transfer directly, and extending the equivalence theorem to mixed-state targets is an open problem.

A third question concerns the minimum circuit depth required for condition (C1) to hold for an arbitrary n-qubit target state. The Möttönen decomposition guarantees universality but at exponential depth; characterizing the minimum depth sufficient for a given class of target states would directly inform the design of hardware-efficient ansatze for SQGEN.

3.8.3 Comparison with kernel-based methods

An alternative perspective on how SQGEN works can be seen through an analogy to kernel methods in machine learning. These methods operate by mapping data to a high-dimensional feature space using a kernel function, $k(x, x')$, and then performing linear operations in that space. A key advantage is the ability to work in exponentially large feature spaces through the

kernel trick, which computes the scalar products $\langle \phi(x), \phi(x') \rangle$ without the need to explicitly construct the mapping ϕ.

The first part of the SQGEN circuit can be interpreted as a state preparation stage, and the second part (including the X gate) can be interpreted as a kernel evaluation circuit. From this perspective, SQGEN measures Gram matrix elements for a quantum kernel. Unlike standard kernel methods, however, we are not interested in evaluating the elements of the Gram matrix for a fixed feature mapping and specific pairs of points in feature space.

In SQGEN, the kernel is generated by the source state parameters, which are related to the generator circuit, and the discriminator parameters. The circuit parameters are variables that we optimize rather than fixed points in feature space. These points are provided by the generator and the source. In the variation circuit, we search for a kernel that minimizes the cost function J with respect to the circuit parameters.

These parameters appear with certain weights that must be determined by a classical algorithm, as in standard kernel method applications. Therefore, SQGEN can be considered a generative kernel learning method, which is currently being researched intensively as a promising tool for generative learning. The main difference is that SQGEN learns the optimal kernel for a specific generative task rather than using a predefined kernel.

This perspective sheds new light on the convergence problem [83]. Kernel methods are sensitive to the choice of kernel function; an inappropriate kernel can lead to poor performance, regardless of the quality of the training data. In SQGEN, kernel learning and generator learning occur simultaneously. This should theoretically lead to better adaptation. However, in practice, it increases the complexity of the optimization space, which may explain the observed problems with local minimums.

It is worth noting that quantum kernels offer an exponential advantage in feature space dimensionality: an n-qubit quantum state represents a point in a 2^n-dimensional Hilbert space. Classical kernel methods would require an exponential number of parameters to represent such a space. SQGEN exploits the natural exponential expressiveness of quantum systems, albeit at the cost of decoherence issues that limit practical scalability.

3.8.4 Application possibilities

SQGEN is a preliminary step toward understanding how a quantum computer can learn and demonstrate quantum entanglement. In the conducted experiments, the network learned to recognize and generate maximally entangled GHZ states. An interesting research direction would be to extend this learning approach to the general concept of entanglement, not just specific types of entangled states, but rather the abstract property that distinguishes entangled states from separable ones.

This would require more advanced classical learning to generate labels for multidimensional, multiparticle entanglement. The discriminator would need to learn a universal entanglement witness that works for any state in a given Hilbert space. The generator would need to produce a variety of classes of entangled states on demand. This problem exceeds the capabilities of current methods and presents a fascinating challenge that combines quantum information theory with machine learning.

Due to circuit depth requirements, practical applications of SQGEN in the current NISQ era are limited to small systems. In quantum chemistry, the algorithm could generate approximate ground states for small molecules as starting points for more accurate variational methods. In quantum system simulation, a natural use case is learning typical thermal states or ground states of small spin systems. The SQGEN architecture also lends itself to quantum information compression, where the generator acts as an encoder and the discriminator as a decoder in an autoencoder-like structure. Finally, the trained generator can serve as an initializer for quantum algorithms such as VQE or QAOA by providing a near-optimal starting point in parameter space.

A key area of development is adapting SQGEN to more efficient variational ansatze. A hardware-efficient ansatz reduces the number of parameters to linear scaling with the number of qubits, but it loses the guarantee of universality. For specific problem classes, dedicated ansatze can be designed based on the structure of target states. For example, for states with bounded entanglement depth, ansatze with polynomial depth instead of exponential depth are possible.

Developing quantum error mitigation and correction techniques is essential to expanding the range of SQGEN applications. Methods such as probabilistic error cancellation or zero-noise extrapolation can significantly improve results without requiring full error correction. Ultimately, the availability of quantum computers with error correction will eliminate the fundamental limitation of circuit depth, rendering universal ansatze feasible.

One interesting theoretical question is whether synergic learning offers a fundamental advantage over adversarial learning in terms of learning complexity. While our results show empirical differences in the number of experiments required and convergence time, there is no formal analysis determining in which classes of problems one method dominates the other. Developing such a theory would require combining tools from computational learning theory and quantum information theory.

Ultimately, SQGEN demonstrates that alternative approaches to quantum generative learning are possible and may offer advantages in specific scenarios. Synergic optimization is feasible on real quantum devices for small systems and offers conceptual simplicity by eliminating the need for hyperparameter selection. Further development of the method

requires algorithmic improvements, such as more efficient ansatze and better initialization strategies, as well as advances in hardware technology, including longer coherence times, smaller gate errors, and larger numbers of qubits. When combined with evolving mitigation and error correction techniques, SQGEN can become a useful tool in the arsenal of quantum machine learning algorithms.

The synergic training paradigm and the reduction in qubit overhead demonstrated in this chapter both depend critically on controlled manipulation of open-system dynamics: the unitarity of the discriminator exploited here is an idealization that holds only when decoherence rates are small compared to gate timescales. Understanding how dissipation shapes and, in particular, how its singularities can be turned into resources rather than nuisances is the subject of Chapter 4.

Chapter 4

Singularity in open quantum systems: Liouvillian exceptional points

The previous chapters addressed two aspects of working with nonlinear quantum systems: characterizing their correlations and generating states with prescribed properties. A third, equally important capability concerns the controlled exploitation of singularities that appear in the parameter space of these systems. Open quantum systems evolve under dissipation and decoherence, and their dynamics can become singular at specific points in parameter space. Rather than treating these singularities as pathological features to be avoided, one can engineer and observe them in a controlled setting, opening practical routes to enhanced sensing and signal amplification.

This chapter focuses on Liouvillian exceptional points (LEPs): degeneracies of the full Lindblad superoperator that have no counterpart in the semiclassical description based on effective non-Hermitian Hamiltonians. The central result is the first experimental demonstration, on an IBM Quantum processor, of a system that exhibits LEPs while possessing no corresponding Hamiltonian exceptional points (HEPs). The experiment uses quantum process tomography (QPT) rather than quantum state tomography, allowing the complete Liouvillian structure to be reconstructed even in the vicinity of singular points where noise is amplified. The interplay between theoretical model, circuit design, hardware optimization, and noise mitigation forms the main thread connecting the sections that follow.

4.1 Exceptional points in non-Hermitian systems: background and motivation

The concept of an exceptional point originates in the spectral theory of non-Hermitian matrices, where it denotes a parameter value at which two or more eigenvalues and their associated eigenvectors simultaneously degenerate [84]. Unlike an ordinary (diabolical) degeneracy, where coinciding eigenvalues leave the eigenvectors distinct and orthogonal, an exceptional point reduces the dimension of the eigenspace below its algebraic multiplicity: the operator becomes non-diagonalizable and admits only a Jordan normal form. This qualitative difference in spectral structure gives rise to a host of phenomena that have no analogue in Hermitian quantum mechanics.

Experimental interest in non-Hermitian physics accelerated significantly following the theoretical proposal [85] and subsequent realization that $\mathcal{PT}$-symmetric systems, invariant under combined parity and time-reversal operations, can exhibit entirely real spectra despite being non-Hermitian, with HEPs marking the transition between the exact and broken $\mathcal{PT}$-symmetric phases. Early demonstrations in the following decade employed optical waveguides and microwave cavities [86], and the field rapidly expanded to encompass electronic circuits, plasmonic structures, acoustic metamaterials, and optomechanical systems. This breadth of platforms established that non-Hermitian physics is not a curiosity of specific models but a universal feature of open systems with structured loss or gain [87].

Two classes of experimentally relevant phenomena concentrated the bulk of early work. The first concerns the strongly asymmetric response to perturbations near an EP: whereas a normal system responds linearly to a small parameter change, a system at an N-th order EP exhibits a response $\Delta\lambda \propto \epsilon^{1/N}$, where ϵ is the perturbation magnitude [88]. This sensitivity enhancement was demonstrated in optomechanical and microwave systems and immediately suggested applications in gyroscopy, mass sensing, and signal amplification [19]. The second class concerns the topological character of exceptional points: when a system parameter traces a closed loop encircling an EP in the complex plane, the eigenvalues are permuted in a direction-dependent way (a monodromy effect with no counterpart in Hermitian systems [89]). These two properties, sensitivity enhancement and non-trivial topology, remain the central motivation for studying EPs across all physical platforms.

The transition from classical and semiclassical descriptions to fully quantum systems introduced a conceptual gap that the effective non-Hermitian Hamiltonian framework could not bridge. The core difficulty is that an effective Hamiltonian $\hat{H}_e$ captures only the deterministic, continuously dissipative part of the dynamics, omitting the stochastic quantum jumps that arise when the environment is monitored or traced over. For small quantum systems evolving

under the Lindblad master equation, quantum jumps are not a correction but a constitutive part of the dynamics [90]: they modify the steady state, alter relaxation rates, and can create or destroy spectral degeneracies that $\hat{H}_e$ alone cannot predict. This recognition motivated the introduction of *Liouvillian exceptional points* by Minganti et al. [20] as degeneracies of the full Lindblad superoperator $\mathcal{L}$, which encompasses both the coherent drift and the incoherent jump contributions. The same work established that HEPs and LEPs are related but inequivalent: every HEP induces a corresponding LEP, but LEPs can exist in systems where $\hat{H}_e$ is diagonal and HEPs are permanently forbidden.

Subsequent theoretical work explored LEPs in linear bosonic systems, identified higher-order Liouvillian degeneracies [91], and established hybrid-Liouvillian formalisms connecting the two types of EP through quantum trajectory post-selection [92]. On the applications side, theoretical analyses of quantum heat engines [93] and quantum amplifiers [94] suggested that LEPs could improve thermodynamic performance by altering the relaxation spectrum in controlled ways. These results positioned LEPs not merely as theoretical refinements of HEPs but as genuinely new resources for quantum information processing and open-system engineering.

The experimental landscape, however, lagged considerably behind the theory. At the time of the work described in this chapter, only four experimental demonstrations of LEPs had been reported, all in single-qutrit platforms used to simulate qubit dissipation [95–97, 94, 93], and all relying on quantum state tomography (QST) rather than quantum process tomography. QST can track eigenvalue coalescence indirectly through the output state, but it cannot access the eigenmatrix overlap $\mathcal{O}_{12} = \langle\tilde{\sigma}_1|\tilde{\rho}_2\rangle$ that distinguishes a true EP from a diabolical degeneracy, nor can it reconstruct the full Liouvillian superoperator. The work presented here addresses both limitations simultaneously: by implementing QPT on multi-qubit quantum circuits, it provides the first direct reconstruction of a Liouvillian exhibiting LEPs and the first experimental verification that these LEPs exist in the complete absence of HEPs. The connection between this experimental strategy and the QPT machinery developed for the characterization toolkit of Chapter 2 is discussed in Section 4.3.1.

4.2 Hamiltonian and Liouvillian exceptional points

4.2.1 Effective non-Hermitian Hamiltonians

For an open quantum system governed by the Lindblad master equation [15, 16], a convenient decomposition separates the continuous non-unitary drift from the stochastic quantum jumps. Given a system Hamiltonian $\hat{H}$ and a set of jump operators $\hat{\Gamma}_\mu$, one defines the effective

non-Hermitian Hamiltonian:

$$\hat{H}_e = \hat{H} - \frac{i}{2}\sum_{\mu}\hat{\Gamma}_{\mu}^{\dagger}\hat{\Gamma}_{\mu}. \tag{4.1}$$

This operator governs the smooth, deterministic part of the evolution: the loss of energy and coherence to the environment between quantum jumps. Its eigenvalue problem,

$$\hat{H}_e|E_n\rangle = E_n|E_n\rangle, \tag{4.2}$$

yields complex eigenvalues whose imaginary parts encode decay rates. A distinguishing property of non-Hermitian operators is that their right eigenvectors $|E_n^R\rangle$ and left eigenvectors $\langle E_n^L|$, defined through $\hat{H}_e|E_n^R\rangle = E_n|E_n^R\rangle$ and $\langle E_n^L|\hat{H}_e = E_n\langle E_n^L|$ respectively, do not form an orthogonal set among themselves; orthogonality holds only between left and right eigenvectors corresponding to the same index n.

A HEP occurs when two or more eigenvalues E_n and their associated eigenvectors simultaneously coalesce as a function of some external parameter. At such a point the eigenspace collapses: the operator is no longer diagonalizable but can only be brought to Jordan form. The geometric degeneracy is strictly lower than the algebraic one, and this is the defining signature that distinguishes an exceptional point from an ordinary (diabolical) degeneracy, where eigenvalues coincide but eigenvectors remain distinct.

In a quantum trajectory picture, the jump operators $\hat{\Gamma}_{\mu}$ generate sudden stochastic changes in the wave function, corresponding to the emission or absorption of excitations through interaction with a monitored environment. The effective Hamiltonian captures the intervals between such events. This separation is physically meaningful: HEPs are entirely determined by $\hat{H}_e$ and therefore belong to the semiclassical regime in which quantum jumps are either absent or suppressed by trajectory post-selection.

4.2.2 Liouvillian superoperators

The complete dynamics of an open quantum system is described by the Lindblad master equation [15, 16], which can be written compactly using the Liouvillian superoperator $\mathcal{L}$ acting on the density matrix $\hat{\rho}(t)$:

$$\frac{\partial}{\partial t}\hat{\rho}(t) = \mathcal{L}\hat{\rho}(t) = -i[\hat{H}, \hat{\rho}(t)] + \sum_{\mu}\mathcal{D}[\hat{\Gamma}_{\mu}]\hat{\rho}(t), \tag{4.3}$$

where each dissipator takes the form:

$$\mathcal{D}[\hat{\Gamma}_{\mu}]\hat{\rho}(t) = \hat{\Gamma}_{\mu}\hat{\rho}(t)\hat{\Gamma}_{\mu}^{\dagger} - \frac{1}{2}\left(\hat{\Gamma}_{\mu}^{\dagger}\hat{\Gamma}_{\mu}\hat{\rho}(t) + \hat{\rho}(t)\hat{\Gamma}_{\mu}^{\dagger}\hat{\Gamma}_{\mu}\right). \tag{4.4}$$

An equivalent formulation separates the two contributions explicitly:

$$\mathcal{L}\hat{\rho}(t) = -i\left(\hat{H}_e\hat{\rho}(t) - \hat{\rho}(t)\hat{H}_e^\dagger\right) + \sum_\mu \hat{\Gamma}_\mu\hat{\rho}(t)\hat{\Gamma}_\mu^\dagger, \tag{4.5}$$

where the last term contains the quantum jump contributions absent from the effective Hamiltonian description.

In the vectorized representation [98], where density matrices are reshaped into column vectors $|\tilde{\rho}\rangle$, the superoperator $\mathcal{L}$ becomes an $N^2 \times N^2$ matrix:

$$\tilde{\mathcal{L}} = -i\left(\hat{H}_e \otimes \mathbb{I}_N - \mathbb{I}_N \otimes \hat{H}_e^T\right) + \sum_\mu \hat{\Gamma}_\mu \otimes \hat{\Gamma}_\mu^*. \tag{4.6}$$

The last term, involving the tensor product of jump operators, is the structural element that distinguishes the Liouvillian from its semiclassical counterpart: it cannot be obtained from the coherent part alone, and it is this term that enables the formation of LEPs in systems where HEPs are forbidden.

The eigenvalue problem for $\tilde{\mathcal{L}}$ reads:

$$\tilde{\mathcal{L}}|\tilde{\rho}_n\rangle = \lambda_n|\tilde{\rho}_n\rangle \qquad \text{and} \qquad \langle\tilde{\sigma}_n|\tilde{\mathcal{L}} = \lambda_n\langle\tilde{\sigma}_n|, \tag{4.7}$$

where $\hat{\rho}_n$ and $\hat{\sigma}_n$ are the right and left eigenmatrices respectively, and the real parts of all eigenvalues λ_n are non-positive, describing relaxation towards the steady state. The spectral gap $|\mathrm{Re}(\lambda_1)|$ (where $\lambda_0 = 0$ corresponds to the steady state) sets the rate of convergence. A LEP occurs when two eigenvalues λ_n and their associated eigenmatrices coalesce simultaneously, an event that requires a Jordan block structure and cannot be reduced to a diabolical degeneracy.

For short time steps dt, the evolution generated by $\mathcal{L}$ defines the quantum map:

$$\hat{\rho}(t + dt) = (\mathcal{L}\, dt + \mathbb{I})\,\hat{\rho}(t) \equiv \mathcal{S}\,\hat{\rho}(t). \tag{4.8}$$

The spectral properties of $\mathcal{S}$ and $\mathcal{L}$ are in one-to-one correspondence up to an affine shift, so LEPs of $\mathcal{L}$ appear as LEPs of $\mathcal{S}$. This observation is the foundation of the experimental strategy described in Section 4.4: rather than attempting to directly implement a continuous non-unitary evolution, one implements the discrete map $\mathcal{S}$ on a quantum processor and uses quantum process tomography to reconstruct $\mathcal{L}$.

4.2.3 Classification of exceptional points and their physical signatures

The distinction between HEPs and LEPs can be stated precisely in terms of which operator degenerates. HEPs are properties of $\hat{H}_e$ and capture all singular behavior accessible within a semiclassical or trajectory-selected description. LEPs are properties of the full $\tilde{\mathcal{L}}$ and account for both the drift and the jump contributions. Since $\hat{H}_e$ is embedded in $\tilde{\mathcal{L}}$ through the coherent term, every HEP gives rise to a corresponding LEP, but the converse is not true: there exist systems in which the jump term of $\tilde{\mathcal{L}}$ creates additional degeneracies that have no counterpart in $\hat{H}_e$.

The physical consequences of approaching an exceptional point are observable through the dynamics. Away from an LEP, the Liouvillian eigenvalues are either real (exponential decay) or come in complex conjugate pairs (decaying oscillations). At the LEP these two regimes meet: the imaginary part of $\lambda_{1,2}$ passes through zero and the eigenvalues coalesce. This transition is visible in the time evolution as a change from spiral to non-spiral trajectories on the Bloch sphere, and it serves as the primary experimental signature.

A second, more sensitive indicator is the overlap between the left and right eigenmatrices:

$$\mathcal{O}_{12} = \langle \tilde{\sigma}_1 | \tilde{\rho}_2 \rangle. \tag{4.9}$$

The eigenmatrices are normalized via the biorthogonality condition $\langle \tilde{\sigma}_n | \tilde{\rho}_m \rangle = \delta_{nm}$, so that $\mathcal{O}_{12}$ measures the extent to which the left eigenmatrix of mode 1 overlaps with the right eigenmatrix of mode 2. At a diabolical point, eigenvalues coalesce but eigenmatrices remain orthogonal, so $\mathcal{O}_{12} \to 0$. At a true exceptional point, the eigenmatrices merge and $\mathcal{O}_{12} \to 1$. Monitoring $\mathcal{O}_{12}$ therefore discriminates between these two types of degeneracy and provides direct confirmation that the observed coalescence is indeed an EP (Fig. 4.1).

4.2.4 Topological character of exceptional points and monodromy

Exceptional points are not merely isolated spectral degeneracies but carry a non-trivial topological structure in the parameter space of the system. To understand this, consider the eigenvalues $\lambda_{1,2}(\kappa)$ of $\tilde{\mathcal{L}}$ as analytic functions of a complex control parameter $\kappa = \kappa_{\mathrm{r}} + i\kappa_{\mathrm{i}}$. Away from an EP, these two functions are distinct branches of a multi-valued analytic function; the EP is a branch-point singularity at which the two branches meet. When κ traces a closed loop $\mathcal{C}$ in the complex parameter plane that encircles an EP, the two eigenvalues undergo a permutation upon return to the starting point:

$$\oint_{\mathcal{C}} d\kappa \, \frac{\partial}{\partial \kappa} \log(\lambda_1 - \lambda_2) = \pm 2\pi i, \tag{4.10}$$

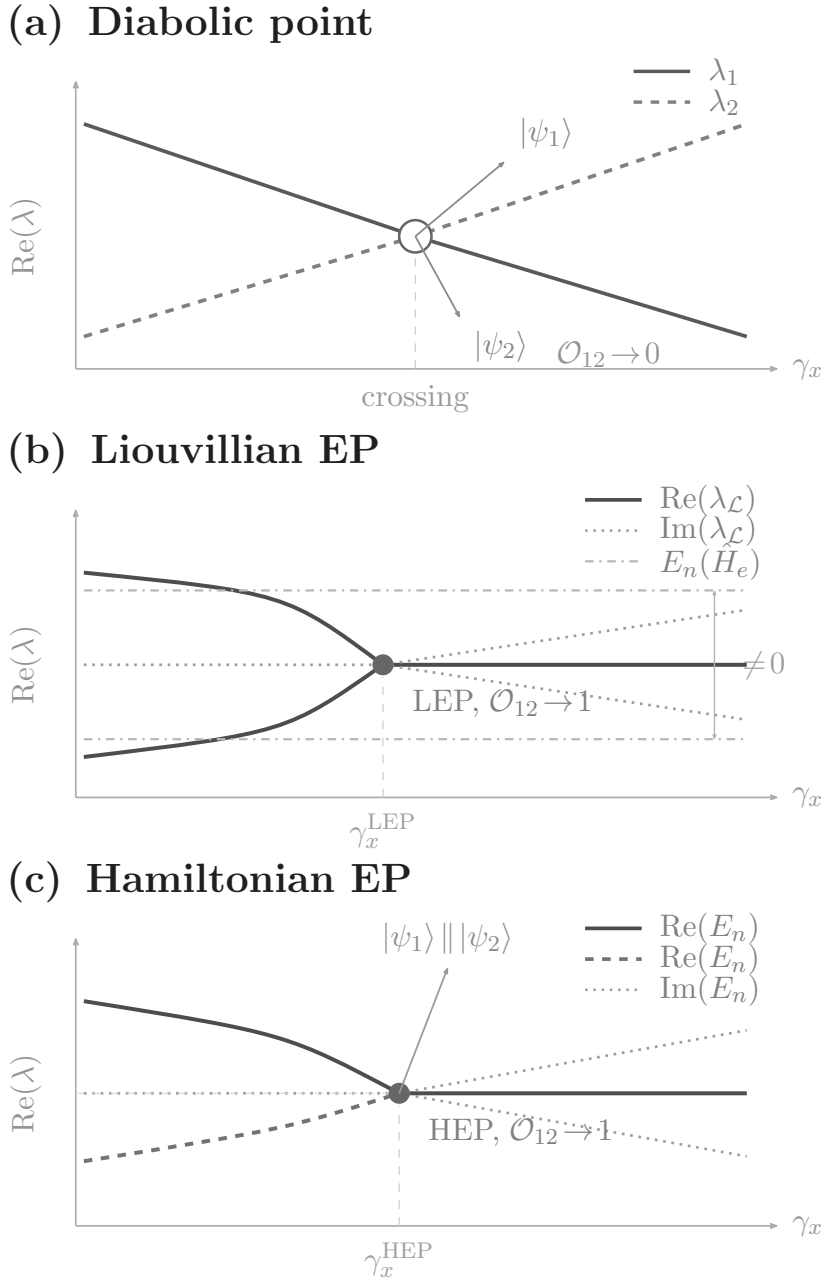


Fig. 4.1 **Schematic comparison of the three types of spectral degeneracy as a function of the dissipation parameter γ_x.** (a) Diabolical point: two eigenvalues cross while their eigenvectors remain orthogonal ($\mathcal{O}_{12} \to 0$). (b) Liouvillian exceptional point in the absence of a Hamiltonian exceptional point: the eigenvalues of $\mathcal{L}$ coalesce and the eigenmatrices merge ($\mathcal{O}_{12} \to 1$), while the eigenvalues of the effective Hamiltonian $\hat{H}_e$ remain distinct ($\omega \neq 0$). (c) Hamiltonian exceptional point: eigenvalues and eigenvectors of $\hat{H}_e$ coalesce simultaneously ($\mathcal{O}_{12} \to 1$). Solid and dashed curves show the two real parts; dotted curves show imaginary parts after coalescence.

where the sign depends on the orientation of $\mathcal{C}$. This permutation has no analogue for diabolical points, where both eigenvalues return to their original values after any closed loop, regardless of whether it encircles the degeneracy. The phase accumulated by each eigenvalue is π per encirclement (half the period of the branch cut), so two complete loops are needed to restore the original spectral labeling.

The associated eigenvectors undergo a corresponding exchange: after one encirclement of an EP, the eigenvector initially labeled $|\tilde{\rho}_1\rangle$ continuously deforms into $|\tilde{\rho}_2\rangle$ and vice versa, accumulating a geometric phase in the process. This behavior is the non-Hermitian analogue of the Berry phase and has been verified experimentally for HEPs in microwave cavities [99] and optical systems. For LEPs, the situation is richer because the parameter space is the space of dissipation rates $(\gamma_x, \gamma_y, \gamma_-)$ rather than a Hamiltonian coupling, and the branch-point structure is embedded in a real parameter manifold rather than a complex one. The sign of $z^2 = (\gamma_x - \gamma_y)^2 - \omega^2$ determines whether a given experimental trajectory in dissipation-rate space encircles an LEP: positive values (real z) correspond to the non-spiral regime, negative values (imaginary z) to the spiral regime, and the LEPs themselves sit at the boundary between the two.

For the model studied here, the two LEPs at $\gamma_x^{\pm} = \gamma_y \pm \omega$ are each second-order branch points, and any path connecting $\gamma_x < \gamma_x^-$ to $\gamma_x > \gamma_x^+$ must cross the spiral region. The QPT-based reconstruction of eigenvalues as a function of γ_x directly maps this branch structure: the imaginary parts of $\lambda_{1,2}$ emerge continuously from zero at each LEP and vanish again at the other, tracing the square-root branching expected from the analytic structure of Eq. (4.10). Observing this topology experimentally, for example by implementing closed parameter loops around individual LEPs [100] and tracking the resulting eigenvalue permutation via repeated QPT, represents a natural extension of the present work requiring process tomography at multiple parameter values along a two-dimensional trajectory in (γ_x, γ_y) space.

4.3 Quantum process tomography as a characterization tool

4.3.1 QPT as the natural language of Liouvillian physics

The choice of QPT as the primary experimental tool in this chapter reflects a structural correspondence between what QPT measures and what Liouvillian physics requires. Quantum state tomography (QST) [101] reconstructs the density matrix of a quantum system at a given moment by measuring a tomographically complete set of observables: it returns a single $\hat{\rho}_{\mathrm{out}}$ that encodes the response of the Liouvillian to exactly one input, in the same way

that measuring the output of a linear system at a single frequency characterizes only that frequency's gain and phase shift. The eigenspectrum of $\mathcal{L}$, by contrast, is a global property of the superoperator: it encodes how the system responds to any initial condition and describes the complete long-time behavior regardless of how the system was prepared.

QPT closes this gap [102, 103] by reconstructing the full input-output map $\mathcal{E} : \hat{\rho}_{\mathrm{in}} \mapsto \hat{\rho}_{\mathrm{out}}$ across an informationally complete set of input states. Once $\mathcal{E}$ is known, the Liouvillian is recovered by inversion of the short-time relation $\mathcal{S} = \mathcal{L}\,dt + \mathbb{I}$, and its eigenspectrum follows by numerical diagonalization. The full pipeline (prepare, evolve, measure, invert, diagonalize) is exactly the one implemented on the IBM Quantum processor for each value of γ_x in the experiments described in Section 4.4.

The distinction becomes sharper at exceptional points. At a diabolical degeneracy, two eigenvalues coincide but the associated eigenmatrices remain orthogonal; the degeneracy could in principle be detected from a time-series of QST measurements by resolving two contributions decaying at the same rate. At an LEP, the eigenmatrices themselves coalesce, which qualitatively changes the time evolution: instead of purely exponential decay, terms of the form $t\,e^{\lambda t}$ appear in the propagator. Separating this algebraic correction from purely exponential contributions in noisy data requires solving a numerically ill-conditioned problem, especially near the EP where the relevant decay rates are nearly equal. QPT bypasses this difficulty entirely: the Jordan block structure appears directly in the eigenspectrum of the reconstructed $\mathcal{L}$, and the eigenmatrix overlap $\mathcal{O}_{12}$ provides an unambiguous scalar diagnostic of exceptional-point character that no fixed-state measurement can furnish.

This structural advantage connects the LEP experiment to the broader characterization toolkit developed in Chapter 2. There, multicopy methods were shown to give access to collective properties of quantum states that no single-copy measurement can efficiently extract. The present chapter extends that philosophy to dynamical characterization: just as negativity requires multiple copies of a state to be directly measured, the exceptional-point structure of a Liouvillian requires the full process matrix, accessible only through QPT, to be unambiguously diagnosed. In both cases, the measurement overhead is justified by the mathematical structure of the quantity being measured, not merely by convenience.

The practical implementation of QPT on the *ibm_nairobi* processor is discussed in the following subsections, with particular attention to the choices that most affected the quality of the Liouvillian reconstruction: the selection of input states, the inversion strategy, and the error mitigation stack applied before and after circuit execution.

4.3.2 QPT methodology

For a single qubit, a tomographically complete set consists of the six eigenstates of the Pauli operators:

$$|\mathrm{in}_i\rangle,\ |\mathrm{out}_j\rangle\ \in\ \{|x_+\rangle, |x_-\rangle, |y_+\rangle, |y_-\rangle, |z_+\rangle, |z_-\rangle\}, \tag{4.11}$$

where $|x_\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$, $|y_\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \mp i|1\rangle)$, $|z_+\rangle \equiv |0\rangle$, and $|z_-\rangle \equiv |1\rangle$.

Three equivalent formulations of QPT were considered and implemented in the experiment. In Method 1, the 6×6 Liouvillian matrix is reconstructed by measuring all transition amplitudes:

$$\mathcal{L}'_{ij} = \langle \mathrm{out}_j | \mathcal{L}(|\mathrm{in}_i\rangle\langle \mathrm{in}_i|)\, |\mathrm{out}_j\rangle. \tag{4.12}$$

In Method 2, the 4×4 representation in the Pauli basis is obtained by evaluating:

$$\mathcal{L}''_{mn} = \frac{1}{2}\mathrm{Tr}[\mathcal{L}(\hat{\sigma}_m)\,\hat{\sigma}_n]\,, \quad m, n \in \{0, x, y, z\}, \tag{4.13}$$

where $\hat{\sigma}_0 = \mathbb{I}$. Method 3 uses projectors $|0\rangle\langle 0|$, $|0\rangle\langle 1|$, $|1\rangle\langle 0|$, $|1\rangle\langle 1|$ as input states; while formally simple, it requires non-physical input operators and is experimentally demanding. In practice, Method 1 yielded the least distorted Liouvillians and was adopted as the primary reconstruction approach [104, 105]. The formal equivalence of all three methods under ideal measurement conditions is established in Section 4.6.

4.3.3 Implementing non-unitary dynamics on a unitary processor

The central difficulty in realizing dissipative dynamics on a digital quantum processor is that all available gates are unitary, whereas the superoperator $\mathcal{S}$ representing a Lindblad step is generally not. The resolution lies in the Stinespring dilation [11]: any completely positive trace-preserving (CPTP) map $\mathcal{E}$ acting on a system of Hilbert space dimension N can be written as:

$$\hat{\rho}_{\mathrm{out}} = \mathrm{Tr}_{\mathrm{env}}\Big[\hat{U}\ (\hat{\rho}_{\mathrm{in}} \otimes |0\rangle\langle 0|_{\mathrm{env}})\ \hat{U}^\dagger\Big]\,, \tag{4.14}$$

where $\hat{U}$ is a unitary acting on the joint system-environment Hilbert space and the environment is initialized in a fixed pure state $|0\rangle_{\mathrm{env}}$.

The unitary $\hat{U}$ is obtained by Kraus decomposing $\mathcal{E}$. Writing:

$$\hat{\rho}_{\mathrm{out}} = \mathcal{E}(\hat{\rho}_{\mathrm{in}}) = \sum_l \hat{A}_l\, \hat{\rho}_{\mathrm{in}}\, \hat{A}_l^\dagger, \tag{4.15}$$

the Kraus operators $\hat{A}_l$ are related to the Choi matrix $\hat{\chi}$, defined via the Choi–Jamiołkowski isomorphism [106, 107] as $\hat{\chi} = \mathcal{E} \otimes \mathcal{I}(|\phi\rangle\langle\phi|)$ with $|\phi\rangle = \sum_j |j\rangle|j\rangle$ a maximally entangled

state, through:

$$A_l^{(ki)} = \sqrt{r_l}\,\langle k|\langle i|\pi_l\rangle, \tag{4.16}$$

where r_l and $|\pi_l\rangle$ are the eigenvalues and eigenstates of $\hat{\chi}$.

In the experimental implementation, the environment is represented by one or two auxiliary qubits, and the unitary $\hat{U}$ is compiled into sequences of the physically available gates $\{I, R_z(\theta), X, \sqrt{X}, \mathrm{CNOT}\}$. The number of auxiliary qubits controls the accuracy with which the target map is approximated: more qubits allow a richer Kraus decomposition but introduce additional error sources through the multi-qubit CNOT gates that entangle system and environment.

4.4 Experimental realization and observation of LEPs

4.4.1 Physical model and its exceptional point structure

The system studied is a single qubit with Hamiltonian $\hat{H} = \frac{\omega}{2}\hat{\sigma}_z$, subject to three dissipation channels: dephasing along x at rate γ_x, dephasing along y at rate γ_y, and amplitude damping at rate γ_-. The Liouvillian takes the form:

$$\mathcal{L}\hat{\rho}(t) = -i\Big[\hat{H}, \hat{\rho}(t)\Big] + \gamma_-\mathcal{D}[\hat{\sigma}_-]\hat{\rho}(t) + \gamma_x\mathcal{D}[\hat{\sigma}_x]\hat{\rho}(t) + \gamma_y\mathcal{D}[\hat{\sigma}_y]\hat{\rho}(t). \tag{4.17}$$

This model is $\mathbb{Z}_2$-symmetric under $\hat{\sigma}_- \to -\hat{\sigma}_-$ and can represent, among others, a spin-$\frac{1}{2}$ particle in a static magnetic field along z with transverse relaxation and spin-flip errors.

The model admits a clean separation between HEPs and LEPs, which is why it was selected as the experimental target. The effective non-Hermitian Hamiltonian is diagonal in the computational basis:

$$\hat{H}_e = \frac{1}{2}\,\mathrm{diag}(\omega - i\gamma_x - i\gamma_y - i\gamma_-,\ -\omega - i\gamma_x - i\gamma_y)\,, \tag{4.18}$$

with eigenvalues:

$$E_1 = \frac{\omega}{2} - \frac{i}{2}(\gamma_x + \gamma_y + \gamma_-), \tag{4.19}$$

$$E_2 = -\frac{\omega}{2} - \frac{i}{2}(\gamma_x + \gamma_y). \tag{4.20}$$

Their real parts differ by ω regardless of the dissipation rates; the degeneracy condition $E_1 = E_2$ can only be satisfied when $\omega = 0$ and $\gamma_- = 0$ simultaneously. For any oscillating

system with $\omega \neq 0$, the two eigenvalues of $\hat{H}_e$ remain permanently separated, and no HEPs exist in the model by construction.

The Liouvillian eigenvalues are:

$$\lambda_0 = 0, \tag{4.21}$$

$$\lambda_{1,2} = -\frac{\gamma_-}{2} - \gamma_x - \gamma_y \pm \Omega, \tag{4.22}$$

$$\lambda_3 = -\gamma_- - 2(\gamma_y + \gamma_x), \tag{4.23}$$

where $\Omega^2 = (\gamma_x - \gamma_y)^2 - \omega^2$. The corresponding right eigenmatrices are:

$$\hat{\rho}_0 \propto \mathrm{diag}(\gamma_x + \gamma_y,\ \gamma_x + \gamma_y + \gamma_-)\,, \tag{4.24}$$

$$\hat{\rho}_{1,2} \propto \mathrm{antidiag}(-i\omega \pm \Omega,\ \gamma_x - \gamma_y)\,, \tag{4.25}$$

$$\hat{\rho}_3 \propto \mathrm{diag}(-1, 1). \tag{4.26}$$

The eigenmatrix $\hat{\rho}_0$ is the steady state of the dynamics: its diagonal entries, proportional to $\gamma_x + \gamma_y$ and $\gamma_x + \gamma_y + \gamma_-$, reflect the populations set by the competition between the three dissipation channels.

LEPs occur when $\Omega = 0$, which gives the condition $(\gamma_x - \gamma_y)^2 = \omega^2$. For $\gamma_y > \omega$, this yields two exceptional points at:

$$\gamma_x^{\pm} = \gamma_y \pm \omega. \tag{4.27}$$

This condition arises from the nonlinear coupling between the coherent term $\mathcal{L}_{\mathrm{coh}}$ and the quantum jump term $\sum_\mu \gamma_\mu \hat{\sigma}_\mu \otimes \hat{\sigma}_\mu^*$ in the matrix representation of $\tilde{\mathcal{L}}$. It cannot be derived from $\hat{H}_e$ alone, confirming that the LEPs in this model are a genuinely quantum phenomenon with no semiclassical counterpart.

Experiments were conducted at $\gamma_- = 0$ and $\gamma_y = 2\omega$, placing the two LEPs theoretically at $\gamma_x/\omega = 1$ and $\gamma_x/\omega = 3$. The dissipation rate γ_x was varied systematically while the Liouvillian was reconstructed at each value.

4.4.2 Quantum circuit design

Three circuit configurations were designed to implement the same physical map $\mathcal{S}$, offering different trade-offs between representational accuracy and noise level (Fig. 4.2).

The single-qubit circuit applies a sequence of single-qubit gates directly to the system qubit, simulating the dissipative evolution without any auxiliary qubits. This implementation exploits the fact that, for this particular model and parameter regime, the CPTP map can be approximated to high accuracy within a single-qubit framework. The Qiskit compiler

optimized the gate sequence for the specific calibration parameters of the *ibm_nairobi* processor, resulting in circuits of 30–40 gate layers.

The two-qubit circuit adds one auxiliary qubit to represent the environment. The purification of $\mathcal{S}$ is implemented through controlled rotations and a CNOT gate connecting physical qubits #4 and #5. The larger system-environment Hilbert space provides a more faithful representation of the Kraus operators, at the cost of introducing CNOT gates whose error rate is roughly an order of magnitude higher than single-qubit gates. Circuit depth is typically 50–80 layers depending on γ_x.

The three-qubit circuit uses two auxiliary qubits, allowing the most complete Kraus decomposition of the target map. In principle this configuration most accurately represents the full non-Hermitian dynamics. In practice, the circuit depth exceeds 100 layers, and the accumulated error from CNOT gates and idle-time decoherence renders the results unreliable in the parameter regimes of interest.

The circuits were developed using Qiskit framework [41] and transpiled to the native gate set $\{I, R_z(\theta), X, \sqrt{X}, \mathrm{CNOT}\}$ using compiler optimization level 3, which performs gate merging, rotation cancellation, and layout optimization to minimize the total number of two-qubit gates.

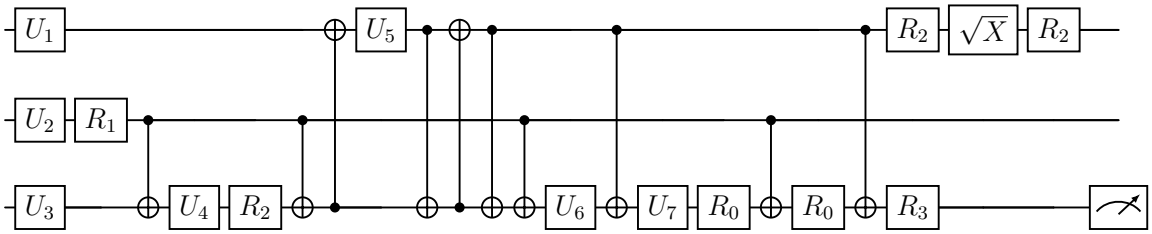
4.4.3 Hardware selection and noise characterization

The experiments were carried out on the *ibm_nairobi* processor [40], a seven-qubit device chosen on the basis of three criteria: the longest available T_2 coherence times across the IBM fleet at the time of the experiment, the lowest per-gate error rates for single-qubit operations, and the presence of a direct physical coupling between qubit #4 and qubit #5, eliminating the need for SWAP routing.

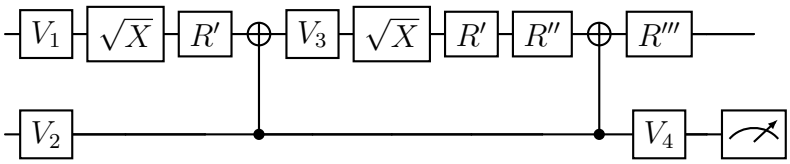
For the single-qubit experiment, qubit #0 was selected with relaxation time $T_1 = 126 \pm 15\,\mu$s, decoherence time $T_2 = 33 \pm 2\,\mu$s, and single-qubit gate error 2.27×10^{-4}. For the two-qubit experiment, qubits #4 and #5 were used: qubit #4 had $T_1 = 155 \pm 23\,\mu$s, $T_2 = 20 \pm 1\,\mu$s, and gate error 2.68×10^{-4}; qubit #5 had $T_1 = 109 \pm 14\,\mu$s, $T_2 = 76 \pm 7\,\mu$s, and gate error 2.98×10^{-4}. The $\mathrm{CNOT}_{4,5}$ error was 6.5×10^{-3}, an order of magnitude larger than the single-qubit errors and the dominant noise contribution in all multi-qubit circuits.

The main error sources fall into two categories. Statistical errors arise from the finite number of measurement shots (20 000 per input-output pair) and contribute uncertainties of order 0.7% for measurement probabilities near $\frac{1}{2}$. Systematic errors include imperfect gate calibration, residual ZZ coupling between neighboring qubits (crosstalk), and slow drift in qubit frequencies over the duration of the experiment. The latter was mitigated by

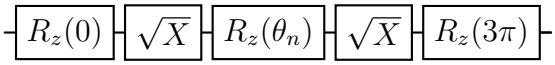
(a) Three-qubit circuit



(b) Two-qubit circuit



(c) $U_n \equiv U(\theta_n)$



(d) $V_n \equiv V(\varphi_n, \theta_n)$

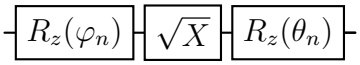


Fig. 4.2 **Quantum circuits optimized for the *ibm_nairobi* processor and applied in the experiments.** (a) Three-qubit circuit implementing the unitary evolution $\hat{U}^{(3)}$ of the system qubit coupled to a two-qubit environment (e_1, e_2). The system qubit (bottom wire) is measured after tracing out the two environment qubits. (b) Two-qubit circuit implementing $\hat{U}^{(2)}$ with a single ancillary environment qubit; the second qubit (bottom wire) is measured. Composite single-qubit gates $U_n \equiv U(\theta_n)$ and $V_n \equiv V(\varphi_n, \theta_n)$ are decomposed into sequences of native gates as shown in panels (c) and (d). Gate parameters: $U_1{=}U(0.5715)$, $U_2{=}U(0.4779)$, $U_3{=}U(0.5634)$, $U_4{=}U(0.2785)$, $U_5{=}U(0.3326)$, $U_6{=}U(0.3060)$, $U_7{=}U(0.3415)$; $V_1{=}V(0, 2\pi)$, $V_2{=}V(\pi/2,\, 0.0033)$, $V_3{=}V(0, \pi)$, $V_4{=}V(\pi/2, \pi/2)$; $R_0{=}R_z(0)$, $R_1{=}R_z(0.0460)$, $R_2{=}R_z(\pi/2)$, $R_3{=}R_z(1.5248)$, $R'{=}R_z(3\pi)$, $R''{=}R_z(1.5375)$, $R'''{=}R_z(1.6040)$.

interleaving calibration runs with data acquisition, but some residual drift remains in the reconstructed eigenvalues.

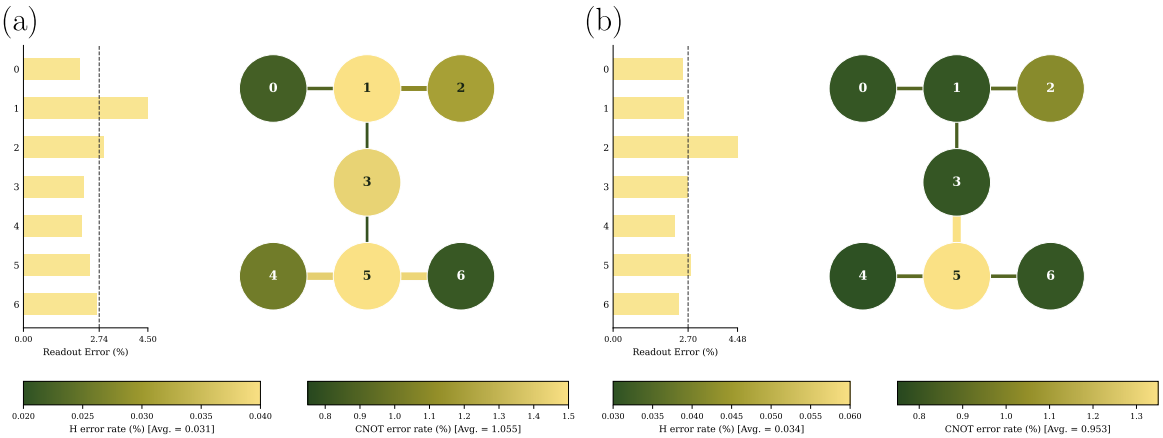


Fig. 4.3 **Calibration data of the *ibm_nairobi* seven-qubit processor** for (a) single-qubit experiment (qubit #0) and (b) two-qubit experiment (qubits #4 and #5), recorded on days of experiments. The Hadamard gate H appears explicitly in the initialization circuits. Full calibration parameters, including relaxation times T_1, decoherence times T_2, gate error rates, and qubit frequencies, are summarized in Table 4.1.

Table 4.1 Calibration parameters of the *ibm_nairobi* superconducting processor for the qubits used in the experiments. Relaxation times $\bar{T}_1$ and decoherence times $\bar{T}_2$ are averaged over 14–16 April 2023. The CNOT error rate between qubits #4 and #5 is 6.5×10^{-3}. Available native gates are $\{I, R_z(\theta), X, \sqrt{X}\}$ for all single-qubit operations and CNOT for the two-qubit experiments.

Qubit	$\bar{T}_1$ (μs)	$\bar{T}_2$ (μs)	Freq. (GHz)	Anharm. (GHz)	Gate error
#0	126 ± 15	33 ± 2	5.26	-0.34	2.27×10^{-4}
#4	155 ± 23	20 ± 1	5.29	-0.34	2.68×10^{-4}
#5	109 ± 14	76 ± 7	5.18	-0.34	2.98×10^{-4}

4.4.4 Noise mitigation strategy

A layered error mitigation approach was tailored to each circuit configuration. For the single-qubit experiment, compiler optimization level 1 was applied (combining successive single-qubit rotations and eliminating redundant identity gates), together with readout error mitigation at resilience level 1 [61], which calibrates the measurement matrix and inverts it statistically.

For the two-qubit experiment, compiler level 3 was used alongside dynamic decoupling (DD) [108, 109]. Dynamic decoupling inserts X-X pulse pairs symmetrically into idle

periods of each qubit, suppressing low-frequency dephasing noise that builds up while one qubit waits for the other to complete a gate sequence. Its effectiveness was quantified by comparing reconstructed Liouvillian spectra with and without DD: at $\gamma_x/\omega = 2$, the imaginary part $|\mathrm{Im}(\lambda_1)|$ changed from 1.42 ± 0.31 (without DD) to 1.04 ± 0.18 (with DD), against a theoretical value of 1.00. This corresponds to a reduction of approximately 40% in the effective decoherence noise contribution. Readout error mitigation applied on top of DD improved the accuracy of individual process matrix elements by a further $\sim 15\%$, and circuit optimization reduced the gate count by 20–30% relative to unoptimized compilation.

4.4.5 QPT protocol

For each value of γ_x, the Liouvillian was reconstructed by the following procedure. Six input states were prepared by applying the appropriate single-qubit rotations to $|0\rangle$: the eigenstates of $\hat{\sigma}_x$, $\hat{\sigma}_y$, and $\hat{\sigma}_z$. Each input state was then evolved according to the circuit implementing $\mathcal{S}$, with $\omega\, dt = 1/15$ chosen to satisfy the first-order approximation $\mathcal{S} \approx \mathcal{L}\, dt + \mathbb{I}$ while keeping the signal above the noise floor. After the circuit, six projection measurements were performed in the eigenbases of all three Pauli operators by prepending the appropriate rotation before the z-basis measurement. The full set of 36 output probabilities (six input states times six projections) was used to assemble the 6×6 process matrix. The Liouvillian matrix was then recovered by inverting $\mathcal{S} = \mathcal{L}\, dt + \mathbb{I}$, and its eigenspectrum was extracted by numerical diagonalization.

The theoretical curves shown in Fig. 4.4 incorporate a white-noise model for the input states: each nominally pure input $|\psi\rangle$ is replaced by the mixed state $\hat{\rho}_{\text{eff}} = (1-w)|\psi\rangle\langle\psi| + (w/d)\,\mathbb{I}_d$, where $d = 2$ is the Hilbert space dimension and w is a noise level fitted to the data. This model accounts for the observed systematic offset between ideal and experimental curves across the full parameter scan and provides the basis for the error-bar estimates in Section 4.6.

4.4.6 Results

The reconstructed Liouvillian spectra as a function of γ_x/ω constitute the main experimental dataset (Fig. 4.4). Prior to the hardware experiments, the three-qubit circuit was verified on the *AerSimulator* with a *FakeNairobi* noise model, confirming that the intended eigenvalue structure is correctly encoded in the circuits before hardware noise is introduced; the single- and two-qubit results from the actual processor then test its observability under realistic conditions. Both LEPs are expected theoretically at $\gamma_x/\omega = 1$ and $\gamma_x/\omega = 3$. In both experimental configurations the observed positions are slightly displaced from these values, an effect attributable to residual systematic errors and the finite density of the parameter scan.

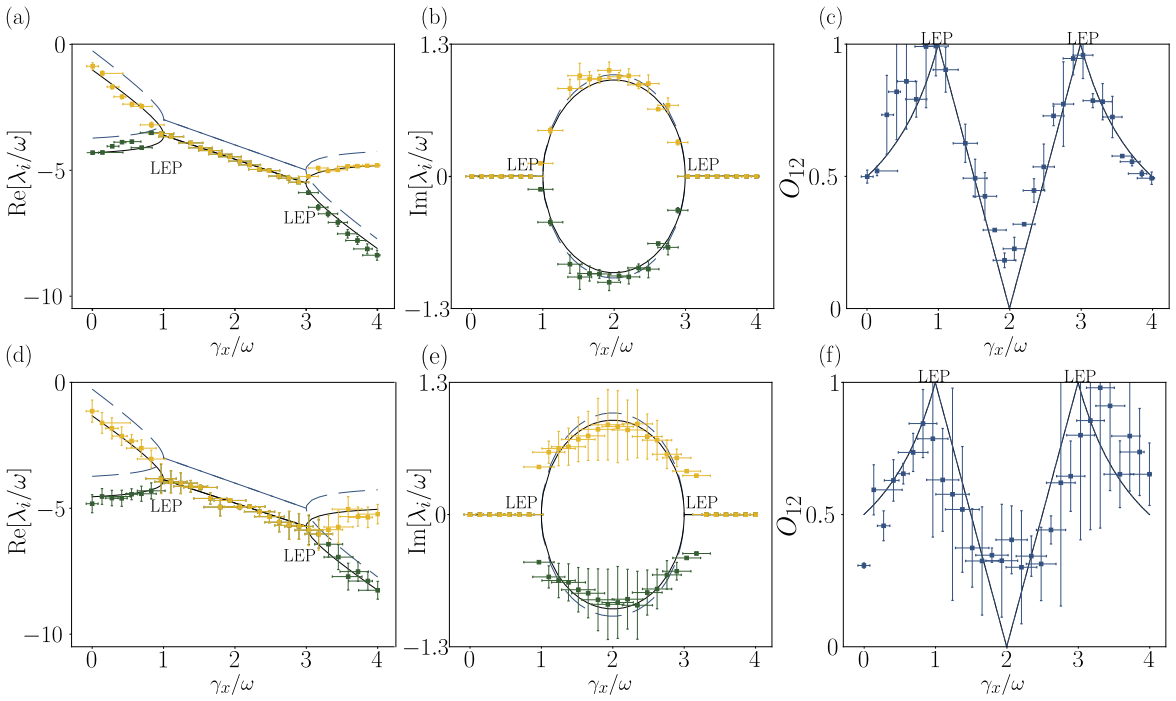


Fig. 4.4 Real (a),(d), imaginary (b),(e) parts of the Liouvillian eigenvalues λ_1 and λ_2, and eigenmatrix overlaps $\mathcal{O}_{12} = |\langle\tilde{\sigma}_1|\tilde{\rho}_2\rangle|$ (c),(f), reconstructed from single-qubit (a),(b),(c) and two-qubit (d),(e),(f) measurements on the *ibm_nairobi* processor. Experimental results are shown as squares (λ_1: red; λ_2: green when $\lambda_1 \neq \lambda_2$); theoretical predictions including white noise are shown as solid black curves, ideal predictions as dashed blue curves. Each measurement was performed with 20 000 shots at $\omega\, dt = 1/15$. The trivial eigenvalues $\lambda_0 = 0$ and λ_3 are omitted for clarity. Selected values are listed in Tables 4.2 and 4.3.

Table 4.2 Experimental nonzero eigenvalues $\lambda_n \equiv \lambda_n^{(m)}$ from single-qubit circuit measurements, for selected values of γ_x/ω. All quantities are given in units of ω. The Liouvillian spectral gap is defined as the nonzero eigenvalue with the smallest modulus of its real part. Error bars are indicated in Fig. 4.4.

m	γ_x/ω	$\mathrm{Re}\,\lambda_1$	$\mathrm{Im}\,\lambda_1$	$\mathrm{Re}\,\lambda_2$	$\mathrm{Im}\,\lambda_2$	λ_3	Gap
1	0	−0.88	0	−4.31	0	−4.95	0.88
2	0.83	−3.21	0	−3.52	0	−6.60	3.21
3	0.97	−3.55	0.13	−3.55	−0.13	−6.89	3.55
4	2.07	−4.64	1.04	−4.64	−1.04	−9.10	4.64
5	2.90	−5.48	0.33	−5.48	−0.33	−10.77	5.48
6	3.03	−5.25	0	−5.90	0	−11.01	5.25
7	4.00	−4.82	0	−8.38	0	−12.92	4.82

As γ_x increases from zero, the eigenvalues λ_1 and λ_2 are initially real and distinct, so the Bloch vector decays monotonically without oscillation. In the vicinity of $\gamma_x/\omega = 1$ the

Table 4.3 Same as Table 4.2, but for the two-qubit circuit measurements. The second LEP is shifted to the interval $(3.31, 3.45)$, compared to $(2.90, 3.03)$ for the single-qubit experiment, due to the higher effective noise level in the two-qubit configuration.

m	γ_x/ω	$\mathrm{Re}\,\lambda_1$	$\mathrm{Im}\,\lambda_1$	$\mathrm{Re}\,\lambda_2$	$\mathrm{Im}\,\lambda_2$	λ_3	Gap
1	0	-1.14	0	-4.82	0	-5.24	1.14
2	0.83	-3.04	0	-4.31	0	-6.67	3.04
3	0.97	-3.82	0.47	-3.82	-0.47	-7.11	3.82
4	2.48	-5.14	0.89	-5.14	-0.89	-9.66	5.14
5	3.31	-6.02	0.38	-6.02	-0.38	-11.40	6.02
6	3.45	-5.85	0	-6.43	0	-11.64	5.85

two eigenvalues merge and then re-emerge as a complex conjugate pair, acquiring non-zero imaginary parts that persist across the intermediate range. Throughout this spiral regime the Bloch vector traces decaying oscillations rather than monotone relaxation to the steady state. The process reverses near $\gamma_x/\omega = 3$, where the imaginary parts return to zero and the eigenvalues settle back on the real axis. This three-stage progression (pure exponential decay, oscillatory decay, pure exponential decay again) is the characteristic signature of two LEPs enclosing a region of spiral dynamics, and it is reproduced consistently in both the single-qubit and two-qubit reconstructions.

In the single-qubit experiment, the first LEP is localized to the interval $\gamma_x/\omega \in (0.83, 0.97)$ and the second to $(2.90, 3.03)$. In the two-qubit experiment, the second LEP shifts to $(3.31, 3.45)$, while the first remains close to the single-qubit result. The eigenmatrix overlap $\mathcal{O}_{12}$ reaches 0.95 near the first LEP and 0.92 near the second in the single-qubit experiment, confirming that the observed degeneracies are true exceptional points rather than diabolical ones.

The three configurations differ substantially in noise behavior. In the single-qubit experiment, the standard deviation of reconstructed eigenvalues is $\sigma_\lambda \approx 0.15\omega$ in the non-spiral regime and $\sigma_\lambda \approx 0.25\omega$ in the spiral regime. The two-qubit experiment shows $\sigma_\lambda \approx 0.30\omega$ and $\sigma_\lambda \approx 0.45\omega$ in the respective regimes, with CNOT errors accounting for the bulk of the increase. The three-qubit experiment yields $\sigma_\lambda \approx 1.2\omega$ in the middle of the spiral regime, large enough to obscure the qualitative structure of the spectrum and making this configuration impractical for the present hardware generation.

Taken together, these results establish that QPT can reliably reveal LEPs on NISQ hardware, provided that noise mitigation is applied carefully and the circuit complexity is kept within the coherence budget of the hardware. The two-qubit configuration achieves the best balance between model accuracy and noise level: it is more faithful to the physical Liouvillian

than the single-qubit approximation, while remaining well within the error threshold at which the LEP signatures remain unambiguous.

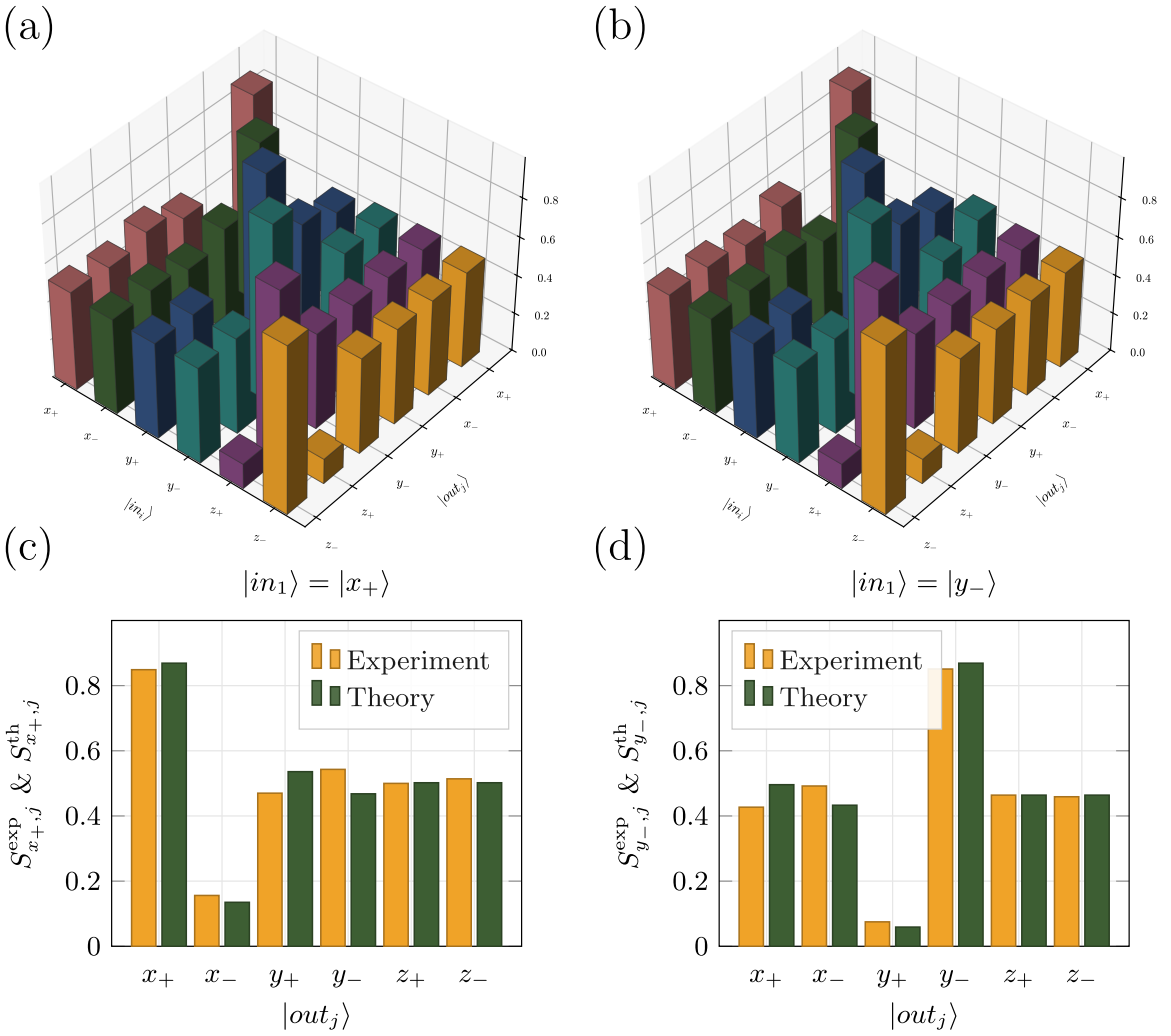


Fig. 4.5 Theoretical (a) and experimental (b) process matrix elements $S_{ij} = \mathcal{L}'_{ij}\, dt + \mathbb{I}$ for the single-qubit experiment at $\omega\, dt = 1/15$ and $\gamma_x/\omega = 0.96$, near the first LEP. Cross-sections for input states $|x_+\rangle$ (c) and $|y_-\rangle$ (d): theoretical predictions in blue, experimental results in red. Good agreement between theory and experiment confirms the reliability of QPT reconstruction near the singular point.

4.4.7 LEPs without Hamiltonian exceptional points

The most direct test of the theoretical prediction is to verify that no HEPs exist in the reconstructed data. The eigenvalues of $\hat{H}_e$, computed analytically from the same model parameters, are $E_1 = \omega/2 - i(\gamma_x + \gamma_y + \gamma_-)/2$ and $E_2 = -\omega/2 - i(\gamma_x + \gamma_y)/2$. For $\omega \neq 0$,

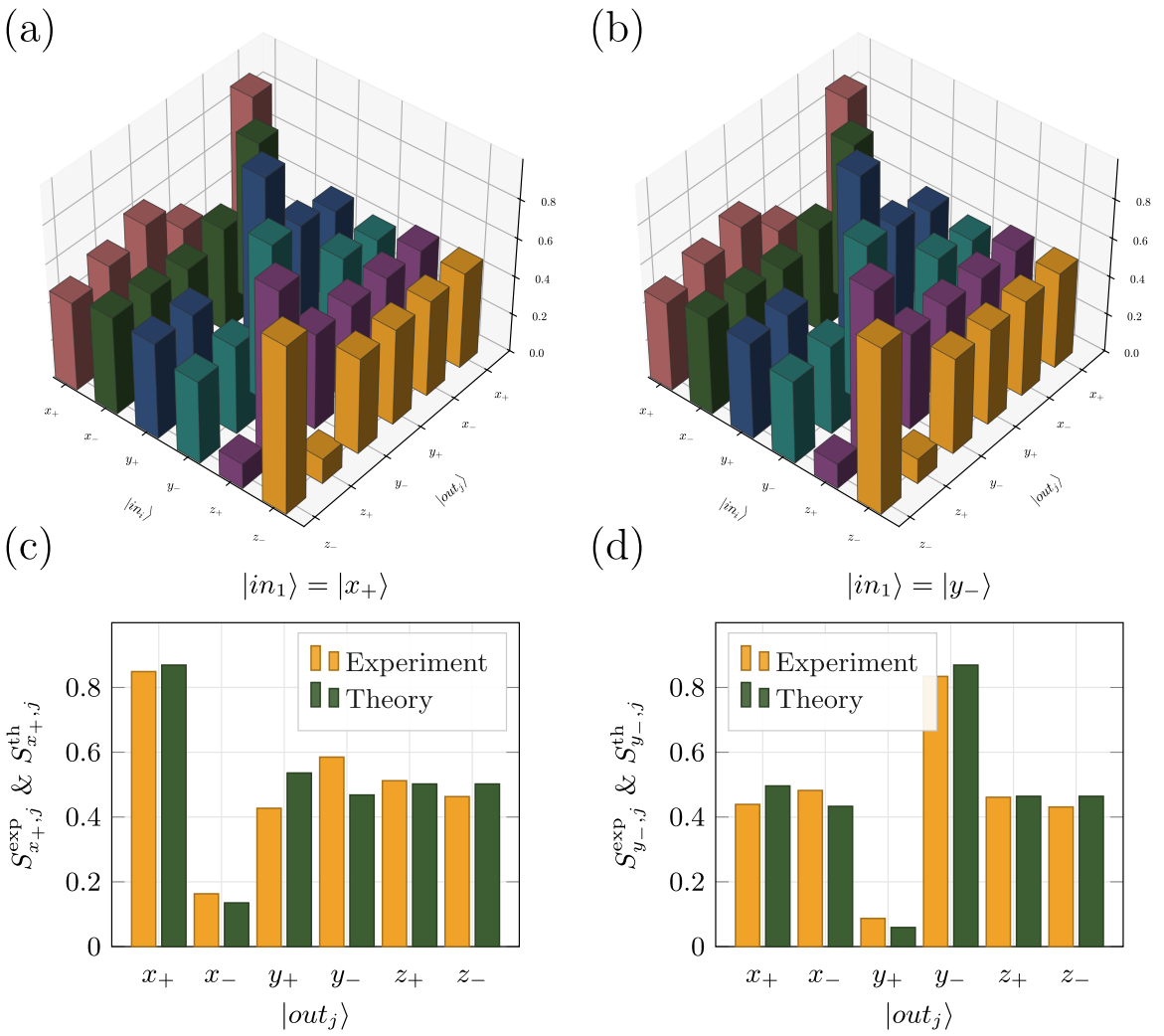


Fig. 4.6 Same as Fig. 4.5, but for the two-qubit experiment near the first LEP. The higher noise level in the two-qubit configuration is visible in the larger deviations between theoretical and experimental matrix elements, particularly in the off-diagonal blocks that are most sensitive to CNOT gate errors.

their real parts differ by ω regardless of the dissipation rates, so the eigenvalues can never coalesce. This was confirmed experimentally: across the full parameter range scanned, no crossing or avoided crossing of the $\hat{H}_e$ eigenvalues was observed, while the Liouvillian eigenvalues clearly exhibit two coalescing events.

This experimental confirmation completes the argument. The jump term $\sum_\mu \gamma_\mu \hat{\sigma}_\mu \otimes \hat{\sigma}^*_\mu$ in the Liouvillian matrix representation couples the coherent and dissipative contributions in a way that lies beyond the effective Hamiltonian description. Removing this term, which corresponds to projecting onto a single quantum trajectory or working in the semiclassical limit, eliminates the LEPs. Their presence therefore depends constitutively on the quantum

jump structure, not merely on the form of $\hat{H}_e$, and this is the clearest possible illustration of why LEPs are a quantum phenomenon that goes beyond semiclassical non-Hermitian physics.

As discussed in Section 4.1, all prior LEP experiments relied on QST in single-qutrit platforms [96, 97, 94, 93], which precludes direct observation of eigenmatrix coalescence. The present QPT-based reconstruction provides simultaneous access to the full Liouvillian spectrum and an unambiguous verification of exceptional-point character through $\mathcal{O}_{12}$ (Fig. 4.4).

4.5 Applications in sensing and signal amplification

4.5.1 Enhanced sensitivity near exceptional points

The coalescence of eigenvalues at an exceptional point has a well-known consequence for the system's response to perturbations. Near an N-th order EP, eigenvalue splitting under a perturbation of magnitude ϵ scales as $\epsilon^{1/N}$ rather than linearly in ϵ [88, 89], leading to an enhanced response that diverges as the EP is approached. This effect has been extensively studied for HEPs in classical and semiclassical systems and underlies proposals for EP-based sensors and signal amplifiers [19].

For the present model, the $\epsilon^{1/2}$ scaling follows directly from the analytic eigenvalue structure. Setting $\gamma_y = 2\omega$ and denoting $\gamma_x = \gamma_x^{\pm} + \epsilon$ for a small deviation from either LEP, the quantity $z = \sqrt{(\gamma_x - \gamma_y)^2 - \omega^2}$ expands as:

$$z \approx \sqrt{\pm 2\omega\,\epsilon + \epsilon^2} \approx \sqrt{\pm 2\omega}\,|\epsilon|^{1/2}, \tag{4.28}$$

to leading order in $|\epsilon|$. Since the eigenvalues $\lambda_{1,2} = -(\gamma_-/2 + \gamma_x + \gamma_y) \pm \Omega$ differ by $2\Omega = 2z\omega$, the spectral gap splitting scales as $|\epsilon|^{1/2}$ in the vicinity of each LEP. This is the square-root branch-point singularity characteristic of a second-order EP.

For LEPs, the situation is more nuanced because quantum jumps simultaneously enhance both the signal and the noise [110, 111]. Nevertheless, the experimental data demonstrate a measurable sensitivity enhancement in a neighborhood of the LEP. Near the first LEP, a change $\Delta\gamma_x = 0.1\omega$ in the control parameter produces a change $\Delta|\mathrm{Re}(\lambda_1)| \approx 0.35\omega$ in the spectral gap, compared to $\Delta|\mathrm{Re}(\lambda_1)| \approx 0.10\omega$ for the same parameter change away from the LEP. The roughly threefold enhancement is consistent with the square-root prediction of Eq. (4.28) evaluated at the operating point used in the single-qubit experiment.

The spectral gap reaches a local extremum at each LEP and can be monitored experimentally via QPT without prior knowledge of the Liouvillian structure, making it a practical observable for sensing applications. An unknown perturbation δ that shifts the LEP position

in parameter space, for example a weak magnetic field detuning the qubit frequency ω, displaces the extremum of the spectral gap and can be reconstructed from its new location. Near a second-order LEP, the shift of the spectral gap extremum scales as $|\delta|^{1/2}$, providing enhanced resolution for small δ compared to a system biased away from the EP.

4.5.2 QPT-based sensing protocol and resource analysis

A concrete sensing protocol based on the experimental infrastructure developed in this chapter proceeds in three stages. In the *calibration stage*, QPT is performed at a reference value $\gamma_x = \gamma_x^{(0)}$ near but not at an LEP, and the spectral gap $g_0 = |\mathrm{Re}(\lambda_1^{(0)})|$ is recorded. In the *perturbation stage*, the unknown signal δ (encoded, for example, as a shift $\omega \to \omega + \delta$) is applied and QPT is repeated at the same nominal circuit parameters, yielding a perturbed spectral gap $g(\delta)$. In the *readout stage*, the shift $g(\delta) - g_0$ is mapped to δ using the analytically computed sensitivity function $\partial g/\partial\delta$ at the operating point.

The resource cost of each QPT measurement is determined by the number of circuit executions N_{shots} required to achieve a target eigenvalue uncertainty σ_λ. For Method 1 (the 6×6 Liouvillian reconstruction used here), 36 input-output combinations are needed, each requiring N_{shots} independent measurements. With $N_{\mathrm{shots}} = 20\,000$ per combination, the total shot budget per QPT instance is $720\,000$. The resulting eigenvalue uncertainty is approximately $\sigma_\lambda \approx 0.15\omega$ in the non-spiral regime and $\sigma_\lambda \approx 0.25\omega$ near the LEP for the single-qubit circuit, with the two-qubit circuit approximately doubling both values due to CNOT-induced error.

The minimum detectable perturbation $\delta_{\min}$ at a given operating point $\gamma_x^{(0)}$ can be estimated from the condition that the spectral gap shift $|g(\delta_{\min}) - g_0|$ equals σ_λ:

$$\delta_{\min} \approx \frac{\sigma_\lambda}{|\partial g/\partial\delta|_{\gamma_x^{(0)}}}. \tag{4.29}$$

Near the first LEP at $\gamma_x^+ = \omega$ (with $\gamma_y = 2\omega$), the sensitivity $|\partial g/\partial\delta|$ is enhanced by the square-root singularity, reducing $\delta_{\min}$ by approximately the same factor of three observed in the eigenvalue response. Consequently, for the single-qubit circuit operating at $\gamma_x/\omega \approx 0.9$, the estimated minimum detectable frequency shift is $\delta_{\min} \approx 0.05\omega$, compared to $\delta_{\min} \approx 0.15\omega$ for the same circuit operated at $\gamma_x/\omega = 0$. This threefold improvement comes without any increase in measurement resources, solely from the operating point choice.

Two caveats qualify this improvement. First, the eigenvalue uncertainty σ_λ itself grows near the LEP due to noise amplification (Eq. (4.41)), partially offsetting the sensitivity gain; the net advantage depends on the ratio of sensitivity enhancement to noise amplification at each operating point. Second, the sensing protocol requires knowledge of the circuit

implementing $\mathcal{S}(\gamma_x^{(0)})$ and the ability to rerun it at the same parameters, which means any drift in the hardware (qubit frequency shifts, gate miscalibration) between the calibration and perturbation stages would appear as a false signal. Dynamical decoupling and interleaved calibration, as employed in the two-qubit experiment, mitigate but do not eliminate this systematic.

4.5.3 Signal-to-noise tradeoffs and platform considerations

The noise enhancement near an LEP is an inherent feature of the singular point and cannot be fully suppressed [110, 112]. In the two-qubit experiment, the eigenvalue uncertainty σ_λ increases by roughly a factor of 1.5 as γ_x/ω is swept from the non-spiral regime into the spiral regime, consistent with the expected noise amplification due to the near-degeneracy of the eigenmatrices. Dynamic decoupling and readout error mitigation reduce this increase but do not eliminate it.

A useful figure of merit for comparing platforms is the ratio:

$$\eta = \frac{|\partial g/\partial \delta|}{\sigma_\lambda}, \tag{4.30}$$

which measures the detectable signal per unit eigenvalue uncertainty. For the single-qubit experiment at the optimal operating point $\gamma_x/\omega \approx 0.9$, $\eta \approx 7\omega^{-1}$; for the two-qubit experiment, $\eta \approx 4\omega^{-1}$. The single-qubit circuit therefore achieves a better sensing figure of merit despite its less accurate representation of the full Liouvillian, because the noise penalty of the two-qubit CNOT gates outweighs the fidelity gain. This conclusion is hardware-specific: on a processor with CNOT errors below 10^{-3} (compared to 6.5×10^{-3} on *ibm_nairobi*), the two-qubit configuration would likely become the preferred choice.

The two-qubit configuration represents the practical optimum for current IBM Quantum hardware. Linear optical systems offer an alternative platform because damping channels can be applied directly to individual qubits encoded in photon polarization without requiring auxiliary modes to simulate the environment. Single- and two-qubit QPT methods transfer directly to such platforms, and the lower noise levels could enable more precise Liouvillian reconstruction, particularly in the vicinity of the LEP where the method is most sensitive to errors. Trapped-ion processors operate with gate fidelities one to two orders of magnitude higher than current superconducting processors, making them attractive candidates for experiments on larger systems or higher-order LEPs, where the scaling $\delta_{\min} \propto \sigma_\lambda^{N/(N+1)}$ at an N-th order EP would amplify both the gain from the EP and the penalty from gate noise.

4.5.4 Directions for further research

The experiments described in this chapter raise several questions for further investigation. Higher-order LEPs [91], where three or more eigenvalues coalesce, require either higher-dimensional systems (qutrits, multi-qubit registers) or more complex dissipation structures and would exhibit $\epsilon^{1/N}$ scaling for $N > 2$, potentially offering greater sensing advantage at the cost of increased circuit complexity and noise. The topological character of exceptional points in parameter space (Section 4.2.4), while described analytically here, has not yet been observed experimentally in the Liouvillian context: implementing closed parameter loops around individual LEPs via sequences of QPT measurements and tracking the resulting eigenvalue permutation would constitute the first direct topological experiment in Liouvillian physics.

Two further classes of exceptional points have been identified theoretically but remain experimentally unexplored. Hybrid Liouvillian exceptional points, introduced in Ref. [92], interpolate continuously between HEPs and LEPs by varying the rate at which quantum trajectories are post-selected: in the limit of perfect trajectory selection the Liouvillian reduces to the effective Hamiltonian and HEPs are recovered, while for zero post-selection the full Liouvillian governs and only LEPs remain. Accessing intermediate regimes requires partial measurement of the environment, a capability within reach of recent photonic experiments. Non-Markovian exceptional points arise when the system's memory of past states is non-negligible, as occurs in structured-bath or strong-coupling regimes. The pseudomode extension of the Lindblad formalism maps such non-Markovian dynamics onto an enlarged Markovian system, and the LEPs of this extended Liouvillian carry information about the spectral density of the original bath. The QPT infrastructure developed here, comprising Stinespring dilation, layered noise mitigation, and biorthogonal eigenmatrix overlap, transfers directly to these settings and provides a concrete starting point for their experimental investigation.

From an applications perspective, theoretical analyses predict that LEPs can enhance the efficiency of quantum heat engines by modifying the relaxation timescale at which the working medium exchanges energy with its reservoirs [93, 94]. These predictions have yet to be tested experimentally. The QPT infrastructure and circuit design strategies developed here provide a direct route to such experiments: the same Stinespring dilation approach used to implement the three-channel qubit dissipator can be adapted to more complex dissipation structures, and the calibrated noise mitigation stack (dynamic decoupling plus readout correction) provides a reproducible platform for quantitative comparison with theory.

4.6 Analytical structure of the Liouvillian and perturbation theory of errors

This section collects two sets of analytical results that underpin the experimental analysis. The first establishes the explicit matrix form of the Liouvillian and the formal equivalence of the three QPT representations used in the experiment. The second provides a perturbative framework for quantifying how measurement noise propagates into the reconstructed eigenvalues and eigenmatrix overlaps near an LEP, forming the theoretical basis for the error bars reported in Fig. 4.4.

4.6.1 Explicit Liouvillian matrices and eigenvalues

The general form of the Liouvillian superoperator for the three-channel qubit model can be written explicitly in the 6×6 representation corresponding to QPT Method 1 (see Section 4.3). In the case $\gamma_- = 0$, relevant to the main experiment, the process matrix takes the form:

$$\mathcal{L}' = \frac{\omega}{4} \begin{pmatrix} 4 & 4 & 1 & 1 & 0 & 0 \\ 4 & 4 & 1 & 1 & 0 & 0 \\ 1 & 1 & 2x & 2x & 0 & 0 \\ 1 & 1 & 2x & 2x & 0 & 0 \\ 0 & 0 & 0 & 0 & y_2 & y_2 \\ 0 & 0 & 0 & 0 & y_2 & y_2 \end{pmatrix}, \tag{4.31}$$

where $x = \gamma_x/\omega$ and $y_k = 2(x + k)$. The nonzero eigenvalues of $\mathcal{L}'$ are:

$$\lambda'_1 = -2\omega(2 + x), \tag{4.32}$$
$$\lambda'_2 = -\omega(2 + x - z), \tag{4.33}$$
$$\lambda'_3 = -\omega(2 + x + z), \tag{4.34}$$

where $z = \sqrt{x^2 - 4x + 3} = \sqrt{(x - 1)(x - 3)}$. The remaining three eigenvalues of $\mathcal{L}'$ are zero. The corresponding right eigenmatrices are:

$$|\tilde{\rho}'_1\rangle = [0,\ 0,\ 0,\ 0,\ 1,\ 1]^T, \tag{4.35}$$
$$|\tilde{\rho}'_2\rangle = [2 - x - z,\ 2 + x + z,\ 1,\ 1,\ 0,\ 0]^T, \tag{4.36}$$
$$|\tilde{\rho}'_3\rangle = [2 - x + z,\ 2 + x - z,\ 1,\ 1,\ 0,\ 0]^T. \tag{4.37}$$

The LEP condition $z = 0$ gives $x = 1$ or $x = 3$, corresponding to the two exceptional points at $\gamma_x/\omega = 1$ and $\gamma_x/\omega = 3$. At these values, $\lambda'_2 = \lambda'_3$ and the two eigenmatrices $|\tilde{\rho}'_2\rangle$ and $|\tilde{\rho}'_3\rangle$ become proportional, the defining signature of a non-diagonalizable operator.

In the 4×4 Pauli basis representation (Method 2), the same Liouvillian takes the form:

$$\mathcal{L}'' = \omega \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 4 & 1 & 0 \\ 0 & 1 & 2x & 0 \\ 0 & 0 & 0 & y_2 \end{pmatrix}, \tag{4.38}$$

with eigenvalues identical to those of $\mathcal{L}'$ (up to the three zero eigenvalues), confirming the spectral equivalence of the two QPT representations [113] under ideal measurement conditions. The transformation relating the two representations is given by the unitary matrix:

$$U'' = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & i & 0 \\ 0 & 1 & -i & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix}, \tag{4.39}$$

such that $\mathcal{L}'' = (U'')^\dagger \mathcal{L} U''$, where $\mathcal{L}$ is the 4×4 vectorized Liouvillian.

Let $\mathcal{E} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H})$ denote the quantum channel generated by the Liouvillian $\mathcal{L}$ over a fixed time step dt, so that $\mathcal{E}(\rho) = e^{\mathcal{L}\, dt}\rho$ to first order. Three distinct matrix representations of $\mathcal{E}$ arise naturally from three choices of operator basis.

Method 1: Input-output vectorization. Choosing the six input states $\rho^{(k)}$ to be the Hermitian generator set $\{|0\rangle\langle 0|,\ |1\rangle\langle 1|,\ |+\rangle\langle +|,\ |+i\rangle\langle +i|,\ |0\rangle\langle 1|,\ |1\rangle\langle 0|\}$ and recording the 6×6 input-output matrix $\mathcal{L}'_{jk} = \langle \phi_j | \mathcal{L}(\rho^{(k)}) | \phi_j \rangle$ yields a representation that includes three redundant zero-eigenvalue modes corresponding to the traceless non-Hermitian generators. The nonzero eigenvalues of $\mathcal{L}'$ coincide with those of the physical 4×4 Liouvillian by construction, since the redundant subspace is spanned by vectors annihilated by the physical trace constraint.

Method 2: Pauli basis vectorization. Expanding the density matrix in the orthonormal Pauli basis $\{\mathbb{I}/\sqrt{2},\ \sigma_x/\sqrt{2},\ \sigma_y/\sqrt{2},\ \sigma_z/\sqrt{2}\}$ and writing $\dot{\vec{r}} = \mathcal{L}''\,\vec{r}$ for the Bloch-vector-like coefficient vector yields the 4×4 matrix $\mathcal{L}''$ given in Eq. (4.38). Since the Pauli basis is a unitary rotation of the computational-basis vectorization used in Method 1 (restricted to the physical 4-dimensional subspace), the two representations are related by $\mathcal{L}'' = (U'')^\dagger\, \mathcal{L}\, U''$ with U'' unitary, and hence share the same eigenspectrum.

Method 3: Choi matrix. The Choi–Jamiłkowski isomorphism [106, 107] maps the channel $\mathcal{E}$ to the 4×4 matrix $C_\mathcal{E} = (\mathrm{id} \otimes \mathcal{E})(|\Phi^+\rangle\langle\Phi^+|)$, where $|\Phi^+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$. The eigenvalues of the process matrix $\hat{\chi}$ (defined by $C_\mathcal{E} = \sum_{mn} \chi_{mn} |\phi_m\rangle\langle\phi_n| \otimes \mathbb{I}$) are related to those of $\mathcal{L}$ by:

$$\mathrm{spec}(\hat{\chi}) = \exp\Big(\mathrm{spec}(\mathcal{L})\, dt\Big), \tag{4.40}$$

where the exponential acts componentwise. Since the map $\lambda \mapsto e^{\lambda\, dt}$ is injective on the relevant domain (all $\mathrm{Re}\,\lambda < 0$ for the physical Liouvillian), the eigenvalues of $\hat{\chi}$ determine those of $\mathcal{L}$ uniquely and vice versa. In particular, the coalescence condition $\lambda_2 = \lambda_3$ that marks the LEP is equivalent to a degeneracy $e^{\lambda_2 dt} = e^{\lambda_3 dt}$ in the Choi spectrum, so all three methods detect the same exceptional points.

The formal equivalence breaks down only in the presence of measurement noise: because $\mathcal{L}'$ is 6×6 while $\mathcal{L}''$ and $\hat{\chi}$ are 4×4, Method 1 requires the reconstruction of 12 additional matrix elements, which introduces a larger statistical uncertainty per eigenvalue for a fixed shot budget. The experimental comparison confirms this expectation: Methods 2 and 3 consistently yield lower variance in the reconstructed eigenvalues across the full parameter scan, while Method 1 produces systematically wider error bars near the LEPs where noise is already amplified by the near-degeneracy of the eigenmatrices. This comparison, together with the theoretical equivalence established above, provides the basis for using any of the three methods interchangeably when the goal is eigenvalue reconstruction, and for regarding their mutual consistency as an internal cross-check on the reliability of the experimental protocol.

4.6.2 Perturbative analysis of eigenvalue errors near LEPs

The enhanced sensitivity of eigenvalues to perturbations near an exceptional point has a precise quantitative formulation. For an experimental Liouvillian $\mathcal{L}_{\mathrm{exp}} = \mathcal{L}_0 + \delta\mathcal{L}$, where $\mathcal{L}_0$ is the ideal Liouvillian and $\delta\mathcal{L}$ a small perturbation (arising from gate errors, decoherence, and readout noise), first-order perturbation theory gives the eigenvalue shift:

$$\delta\lambda_n \equiv \lambda_n^{\mathrm{exp}} - \lambda_n^{(0)} \approx \langle\tilde{\sigma}_n^{(0)}|\delta\mathcal{L}|\tilde{\rho}_n^{(0)}\rangle, \tag{4.41}$$

where superscripts (0) denote ideal quantities. The corresponding first-order corrections to the right and left eigenmatrices are:

$$|\delta\tilde{\rho}_n\rangle \approx \sum_{i\neq n} \frac{\langle\tilde{\sigma}_i^{(0)}|\delta\mathcal{L}|\tilde{\rho}_n^{(0)}\rangle}{\lambda_i^{(0)}-\lambda_n^{(0)}}|\tilde{\rho}_i^{(0)}\rangle, \tag{4.42}$$

$$\langle\delta\tilde{\sigma}_n| \approx \sum_{i\neq n} \frac{\langle\tilde{\sigma}_n^{(0)}|\delta\mathcal{L}|\tilde{\rho}_i^{(0)}\rangle}{\lambda_i^{(0)}-\lambda_n^{(0)}}\langle\tilde{\sigma}_i^{(0)}|. \tag{4.43}$$

The denominators $\lambda_i^{(0)}-\lambda_n^{(0)}$ vanish at an LEP, which drives the divergence of eigenmatrix corrections and the accompanying noise amplification. The overlap $\mathcal{O}_{12}$ therefore accumulates a perturbative error:

$$\delta\mathcal{O}_{12}^{\text{exp}} = \langle\tilde{\sigma}_1^{\text{exp}}|\tilde{\rho}_2^{\text{exp}}\rangle - \langle\tilde{\sigma}_1^{(0)}|\tilde{\rho}_2^{(0)}\rangle \approx \langle\delta\tilde{\sigma}_1|\tilde{\rho}_2^{(0)}\rangle + \langle\tilde{\sigma}_1^{(0)}|\delta\tilde{\rho}_2\rangle + \langle\delta\tilde{\sigma}_1|\delta\tilde{\rho}_2\rangle, \tag{4.44}$$

where the last term is second order and can be neglected away from the LEP. To remove the phase ambiguity in $\delta\mathcal{O}_{12}^{\text{exp}}$, which arises because eigenvectors are defined only up to phase, the error bars in the figures use the phase-independent redefinition:

$$\overline{\delta\mathcal{O}}_{12}^{\text{exp}} = \left[\left|\langle\delta\tilde{\sigma}_1|\tilde{\rho}_2^{(0)}\rangle\right|^2 + \left|\langle\tilde{\sigma}_1^{(0)}|\delta\tilde{\rho}_2\rangle\right|^2 + |\langle\delta\tilde{\sigma}_1|\delta\tilde{\rho}_2\rangle|^2\right]^{1/2}. \tag{4.45}$$

This framework also provides an estimate of the effective uncertainty in γ_x arising from imperfect state preparation. At a given nominal parameter value γ_x, the effective input to the process is a mixture of the intended pure state $|\psi(\gamma_x)\rangle$ with white noise:

$$\hat{\rho}_{\text{eff}} = (1-w)|\psi(\gamma_x)\rangle\langle\psi(\gamma_x)| + \frac{w}{d}\,\mathbb{I}_d, \tag{4.46}$$

where d is the Hilbert space dimension and w the noise level. This state can equivalently be expressed as a distribution over pure input states, which maps to an uncertainty interval $[\gamma_x-\gamma_L, \gamma_x+\gamma_R]$ with asymmetric left and right error bars estimated as $P_L\,\gamma_L/\sqrt{6}$ and $P_R\,\gamma_R/\sqrt{6}$, respectively, for a triangular distribution. These asymmetric error bars account for the observation that LEP shifts in the two-qubit experiment are larger on one side of each EP than the other.

4.7 Summary

This chapter presented the first experimental evidence that a quantum system can exhibit Liouvillian exceptional points in the complete absence of Hamiltonian exceptional points. The

result rests on a theoretical model, a qubit with three competing dissipation channels, whose effective Hamiltonian has a diagonal structure that permanently separates its eigenvalues, while the full Liouvillian develops two exceptional points through the nonlinear coupling introduced by the quantum jump operators.

The experimental strategy is built around quantum process tomography rather than quantum state tomography. This choice is essential for two reasons: QPT reconstructs the full Liouvillian superoperator rather than a single output state, and it provides direct access to the eigenmatrix overlap $\mathcal{O}_{12}$ that distinguishes true exceptional points from diabolical degeneracies. Prior to this work it was unclear whether QPT could return reliable results near a singular point, where noise is amplified by the near-degeneracy of the eigenmatrices. The experiments demonstrate that it can, provided that appropriate noise mitigation is applied and the circuit complexity is kept within the coherence budget of the hardware.

Three circuit configurations were developed and compared on the *ibm_nairobi* processor. The single-qubit circuit offered the lowest noise but an approximate representation of the target dynamics; the two-qubit circuit provided better model fidelity at the cost of higher CNOT-induced errors; the three-qubit circuit, while theoretically most accurate, accumulated too much error in the deep circuits to yield interpretable results. Dynamic decoupling reduced decoherence noise by approximately 40% in the two-qubit experiment, and readout error mitigation improved process matrix accuracy by a further $\sim 15\%$. Compiler optimization at Qiskit level 3 reduced the gate count by 20–30% compared to unoptimized transpilation.

The measured sensitivity enhancement of roughly threefold near the LEP, together with the experimental localization of both exceptional points, demonstrates that Liouvillian singularities are observable, controllable, and potentially exploitable on current NISQ hardware. Together with the characterization methods of Chapter 2 and the state generation methods of Chapter 3, these results complete the set of tools for working with nonlinear quantum systems developed in this dissertation.

Chapter 5

Conclusion and Future Directions

The preceding chapters developed three interconnected tools for working with nonlinear quantum systems. Chapter 2 addressed the problem of characterizing quantum correlations using collective, multicopy measurement protocols. Chapter 3 introduced a generative machine learning approach to pure-state preparation on quantum processors. Chapter 4 demonstrated the first experimental observation of Liouvillian exceptional points in the complete absence of Hamiltonian exceptional points, using quantum process tomography as the primary diagnostic tool. Taken together, these contributions form a progression from measuring nonlinearity, through generating it, to deliberately engineering and exploiting it in open systems. This final chapter draws together the main results, identifies the limitations of each approach, and maps out the research directions that these findings suggest for the near-term and long-term development of quantum computing.

5.1 Summary of contributions

5.1.1 Characterization methods developed

The main challenge in characterizing quantum correlations with hardware available in the near future is that standard tools such as full-state quantum tomography and direct measurement of entanglement metrics tailored to specific families do not scale well. These tools require either exponential measurement resources or prior knowledge of the state's structure. Two complementary approaches to this problem were developed in Chapter 2.

The first is a resource-efficient protocol for measuring quantum correlations that combines multicopy state preparation with neural network processing. By measuring only five settings on two simultaneous copies of the state rather than the fifteen required by full two-qubit tomography, a 67% reduction in measurement overhead is achieved while retaining robustness

against readout noise through maximum likelihood estimation. The scaling advantage, from $\mathcal{O}(4^n)$ to $\mathcal{O}(2^n)$ in the number of measurement settings, becomes pronounced for systems of four or more qubits and was verified on IBM Quantum hardware.

The second contribution establishes multi-copy chirality as an entanglement witness with direct physical interpretation. The scalar spin chirality $\chi_{ijk} = \hat{S}_i \cdot (\hat{S}_j \times \hat{S}_k)$, imported from condensed-matter physics, translates through the identity $C_k = 8\,\mathrm{Tr}[\Omega_A \Omega_B \rho^{\otimes k}]$ into a multi-copy correlator that measures the handedness of joint quantum states. For pure states, C_4 is directly proportional to negativity and therefore constitutes a complete entanglement witness. For mixed states, negativity is reconstructed from the partial-transpose moments $\mu_k = \mathrm{Tr}[(\rho^{T_A})^k]$ via Newton–Girard identities, requiring only three joint measurements on two-copy circuits. The same circuit infrastructure, reinterpreted through the realignment map $R(\rho)$, gives access to the non-Hermiticity gap D_k that detects bound entanglement in all three known families of 3×3 PPT entangled states, including Horodecki, Tiles, and Chessboard states, across their full parameter ranges— a coverage that no previous single criterion had achieved. The Gröbner basis decision tree, a third algebraic contribution, reduces the average number of required measurement configurations by $5.3\times$ for two-qubit and $7.2\times$ for qubit–qutrit systems relative to full tomography by identifying the correct negativity formula from partial-transpose moment data alone.

Together, these methods address the measurement layer of the nonlinear quantum toolkit: they provide experimentally accessible, hardware-compatible routes to quantifying entanglement and correlation structure without assuming knowledge of the state.

5.1.2 Generation techniques introduced

Given methods for measuring quantum correlations, the question of how to efficiently prepare states with prescribed correlation structure becomes natural. Chapter 3 introduced the Synergic Quantum Generative Network (SQGEN), an alternative to the quantum generative adversarial network (QGAN) paradigm.

The distinguishing feature of SQGEN is that the generator and discriminator are trained jointly through a single synergic cost function J, rather than alternating between competing objectives as in standard adversarial training. This eliminates the QGAN model's main hyperparameter, the ratio of generator to discriminator training steps, as well as the need to track Nash equilibrium explicitly. The cost function is evaluated by a single circuit execution per training step, exploiting the reversibility of the discriminator via $D^\dagger$ to measure generation and discrimination quality simultaneously. For a single-qubit system, SQGEN requires an average of 33 experiments per epoch compared to 150 for QGAN, a more than fourfold reduction that persists, though with diminishing margin, up to six qubits.

The theoretical underpinning of SQGEN connects it to quantum kernel methods. The generator-discriminator circuit computes Gram matrix elements for a parameterized quantum kernel, with the circuit parameters playing the role of kernel hyperparameters. Therefore, training the SQGEN model is equivalent to generative kernel learning in a 2^n-dimensional Hilbert space. For target distributions characterized by long-range entanglement, quantum hardware provides an exponential representational advantage over classical methods. This connection suggests that the convergence properties of SQGEN should be analyzed within the framework of quantum kernel theory rather than game theory.

Experimental implementation on the *ibmq_manila* processor confirmed single-qubit operation at fidelity $F \approx 0.92$, demonstrating that synergic training is viable on current noisy hardware. The generation toolkit therefore complements the characterization toolkit: states prepared by SQGEN can be immediately characterized by the multicopy methods of Chapter 2, and the measured correlation properties provide a natural feedback signal for iterative refinement of the generator.

5.1.3 Singularity exploitation demonstrated

The third contribution moves beyond state preparation and measurement into the regime of deliberately engineered dynamics. Chapter 4 presented the first experimental demonstration that a quantum system can exhibit Liouvillian exceptional points (LEPs) while possessing no corresponding Hamiltonian exceptional points (HEPs).

The theoretical model of a single qubit with three competing dissipation channels was chosen precisely because its effective Hamiltonian is diagonal, permanently forbidding HEPs, while the quantum jump structure of the full Liouvillian creates two exceptional points at $\gamma_x^{\pm} = \gamma_y \pm \omega$. The full derivation is given in Section 4.4. This establishes that LEPs are a genuinely quantum phenomenon, not reducible to any semiclassical description.

The experimental strategy is built on quantum process tomography (QPT) rather than quantum state tomography. QPT reconstructs the full Liouvillian superoperator and gives direct access to the eigenmatrix overlap $\mathcal{O}_{12} = \langle\tilde{\sigma}_1|\tilde{\rho}_2\rangle$, which approaches unity at a true EP and zero at a diabolical degeneracy. Three circuit configurations were implemented and compared on the *ibm_nairobi* processor. The two-qubit configuration, combining dynamic decoupling, readout error mitigation, and compiler optimization at level 3, achieved the best balance between model fidelity and noise level, localizing both LEPs and measuring a threefold sensitivity enhancement in the spectral gap near each exceptional point.

The connection between the three chapters is structural. The QPT protocol used to reconstruct the Liouvillian is a dynamical generalization of the multicopy state characterization methods of Chapter 2: both access global properties of quantum objects (the superoperator

and the state, respectively) that are invisible to any single-shot measurement. The circuit design strategy of Chapter 4, based on Stinespring dilation and Kraus decomposition, parallels the parameterized circuit ansatz of SQGEN in Chapter 3: both translate a target quantum process into an experimentally implementable gate sequence. The three contributions thus form a coherent toolkit spanning measurement, preparation, and controlled exploitation of quantum nonlinearity. The three contributions were demonstrated independently on separate hardware platforms and distinct state families; no single experiment combines all three tools in an integrated pipeline. A natural direction for future work is an experiment that uses SQGEN-prepared states as direct input to the multicopy characterization protocol, or that engineers LEPs for dynamics generated by the SQGEN framework, thereby validating the toolkit as a unified experimental infrastructure rather than three independent demonstrations.

5.2 Current limitations

5.2.1 Hardware constraints

All three experimental contributions were demonstrated on IBM Quantum superconducting processors, and all three run into the same fundamental hardware bottleneck: the finite ratio of gate fidelity to circuit depth. For the multicopy entanglement measurement, this manifests as a floor on the detectable negativity: noise introduced by the controlled-SWAP gates used in two-copy circuits corrupts the signal for states with very low entanglement, making it difficult to distinguish weakly entangled states from separable ones. For SQGEN, the bottleneck is more severe: the Möttönen ansatz used for universal state preparation scales as 4^n parameters and $\mathcal{O}(4^n)$ circuit depth, which exceeds the practical coherence budget of current devices at six qubits and above. For the LEP experiment, the three-qubit circuit configuration, while theoretically the most accurate representation of the target Liouvillian, accumulated sufficient error to render the results uninterpretable, limiting the experiment to the two-qubit configuration.

A second hardware constraint is qubit connectivity. The IBM heavy-hex coupling map requires SWAP routing for operations between non-adjacent qubits, increasing effective circuit depth beyond what the logical circuit depth suggests. This penalty was partially mitigated in the LEP experiments by selecting qubits with direct physical coupling, but this strategy becomes increasingly restrictive as the system size grows.

Calibration drift over the duration of multi-hour experiments introduces a third class of hardware errors that is distinct from gate infidelity: slow drift in qubit frequencies and coupling strengths causes the Liouvillian reconstructed at the end of a parameter scan to

be slightly inconsistent with one reconstructed at the beginning. Interleaved calibration sequences mitigate but do not eliminate this effect. On current hardware generations, it represents an irreducible systematic that limits the precision of QPT-based protocols.

5.2.2 Theoretical gaps

Despite the empirical success of SQGEN, its convergence theory remains incomplete. The main hypotheses that the synergic cost function exhibits gradients that decay at a polynomial rather than exponential rate and that this avoids the phenomenon of barren plateaus have not been proven analytically. The argument from post-selection on the decision qubit is intuitive but not rigorous: it does not rule out the possibility that the effective parameter space explored by the optimizer remains exponentially large, merely concentrated near configurations where the discriminator performs correctly. A formal analysis would require computing the variance of the gradient over the Haar measure restricted to the SQGEN circuit architecture, a calculation that depends sensitively on the entanglement depth of the Möttönen ansatz.

For the characterization side, the connection between multi-copy chirality and bound entanglement remains partially unexplored. The realignment non-Hermiticity gap D_k detects all known PPT entangled families in 3×3 systems, but the mathematical condition under which $D_k > 0$ implies entanglement in arbitrary dimensions has not been established in full generality. The geometric interpretation of chirality as a multi-copy handedness correlator connects to topological spin physics, but a systematic classification of which entanglement structures produce non-zero C_k for each order k is not yet available.

For the LEP contribution, the perturbation theory of error propagation near the singular point is first-order. Near the LEP, the denominators $\lambda_i^{(0)} - \lambda_n^{(0)}$ in the eigenmatrix corrections vanish, and the perturbative expansion breaks down. In order to estimate the actual magnitude of measurement uncertainty as the system approaches the exceptional point (EP) asymptotically, a non-perturbative approach to the noise near that point would be necessary. This approach would be analogous to the resummation techniques used in degenerate perturbation theory for Hermitian systems.

5.2.3 Scalability challenges

The most pressing scalability limitation shared across all three contributions is the exponential growth of the state space with the number of qubits. For characterization, the multicopy approach reduces the scaling of measurement settings from $\mathcal{O}(4^n)$ to $\mathcal{O}(2^n)$, but $\mathcal{O}(2^n)$ is still exponential. For generation, the Möttönen universal ansatz requires $\mathcal{O}(4^n)$ parameters

and gates; even hardware-efficient ansatzes with linear parameter counts cannot guarantee faithful approximation of arbitrary target states. For LEP experiments, QPT of an n-qubit system requires $\mathcal{O}(4^{2n})$ measurements, making full process reconstruction intractable beyond a few qubits even in principle.

A structural obstacle shared by SQGEN and QPT is the classical post-processing bottleneck: once measurement data are collected, classical optimization or matrix reconstruction must be performed on objects of dimension 4^n. For systems of ten or more qubits, this classical computation may become the dominant cost even if the quantum measurement itself is efficient. Therefore, developing scalable methods for classical post-processing, whether through tensor network compression, stochastic linear algebra, or machine learning, is just as important as improving the quantum hardware itself for the long-term usefulness of these methods.

The sensing protocol based on LEPs faces an additional scalability challenge that is specific to the singular-point structure. Near the LEP, both the sensitivity and the noise amplification diverge as the EP is approached. The optimal operating point, where the figure of merit $\eta = |\partial g/\partial \delta|/\sigma_\lambda$ is maximized, sits at a finite distance from the EP and must be determined empirically for each hardware platform. For higher-order LEPs ($N > 2$), the $\epsilon^{1/N}$ scaling offers greater sensitivity gain but also amplifies noise more severely, and identifying the optimal operating point becomes a non-trivial optimization problem in itself.

5.3 Future research directions

5.3.1 Near-term improvements (NISQ era)

Several improvements to the methods developed in this dissertation are within reach using current or near-term NISQ hardware.

For the multicopy characterization approach, the most impactful near-term extension would be adaptive measurement protocols. Rather than measuring a fixed set of five settings, an adaptive protocol would select subsequent measurement bases based on the information already acquired, concentrating resources near the boundary between entangled and separable regions. Such protocols have been demonstrated for quantum state tomography and could in principle reduce the measurement overhead of the multicopy approach by a further factor of two to five for typical entangled states. A second direction is the extension of the chirality-based witness to multipartite systems. The scalar spin chirality χ_{ijk} is defined for triples of sites, so multi-copy circuits on three-qubit registers would give access to genuine tripartite entanglement beyond what two-qubit chirality can detect.

For SQGEN, the most pressing near-term improvement is the replacement of the Möttönen universal ansatz with hardware-efficient alternatives that trade universality for circuit depth. For specific target distributions with a known symmetry structure, assumptions that exploit this symmetry, equivalent circuits, or tensor product wave functions can achieve comparable accuracy with significantly fewer gates. Such approaches exploit the structure of the target Hamiltonian to design problem-specific circuit ansatze. Coupling SQGEN with zero-noise extrapolation or probabilistic error cancellation would further extend its practical reach on current hardware, potentially enabling reliable two-qubit operation at fidelity above 0.95.

For the LEP experiment, the most natural near-term extension is the topological verification described in Section 4.2.4. Implementing closed parameter loops in (γ_x, γ_y) space via a sequence of QPT measurements and tracking the eigenvalue permutation would constitute the first direct experimental observation of the monodromy structure of a Liouvillian. This experiment requires QPT simulations to be performed at approximately twenty parameter points along a two-dimensional trajectory. This is a challenging, yet feasible, extension of the current protocol that can be achieved with the existing hardware. A complementary direction is the exploration of hybrid LEPs, which interpolate continuously between HEPs and LEPs through quantum trajectory post-selection. The IBM Quantum mid-circuit measurement capability, available on more recent processor generations, provides a direct route to implementing the partial environment measurement required for hybrid LEP access.

5.3.2 Long-term vision (fault-tolerant era)

The three contributions of this dissertation were designed for NISQ hardware, but each points toward research directions that become fully accessible only in the fault-tolerant regime.

The multicopy measurement infrastructure scales polynomially in circuit depth with the copy number k. In the fault-tolerant era, where arbitrarily deep circuits are executable, this means that moments μ_k of the partial transpose can be accessed for any k, giving in principle complete information about the negativity spectrum and therefore about bound entanglement in arbitrary dimensions. The Newton–Girard reconstruction used here for $k \leq 3$ generalizes to a full spectral reconstruction for any finite-dimensional system, providing a route to complete entanglement characterization without full state tomography.

For SQGEN, fault-tolerant hardware removes the circuit depth constraint that currently limits the approach to small systems. The universal Möttönen ansatz, which requires depth $\mathcal{O}(4^n)$ in the worst case, becomes feasible for tens of qubits on fault-tolerant processors. The synergic training paradigm can also be extended to systems with mixed states described by Liouville dynamics. This would connect the generation framework of Chapter 3 with the dynamical infrastructure of Chapter 4. A generative network trained to reproduce a target

mixed state defined by its Liouvillian, rather than a pure-state density matrix, would combine the generation framework of Chapter 3 with the dynamical characterization infrastructure of Chapter 4, closing the loop of the nonlinear quantum toolkit.

For LEP-based sensing, fault-tolerant hardware opens access to higher-order exceptional points. A third-order LEP requires a three-level system (qutrit) or a carefully engineered three-channel qubit register, and exhibits $\epsilon^{1/3}$ sensitivity scaling in the eigenvalue splitting. The practical advantage over second-order EPs depends on the ratio of sensitivity gain to noise amplification, but for fault-tolerant hardware where gate errors are suppressed below 10^{-6}, the balance should shift decisively toward higher-order EPs. Theoretical predictions for LEP-enabled quantum heat engines suggest that modifying the relaxation spectrum at the EP can improve thermodynamic efficiency beyond the Carnot limit for specific non-equilibrium operating cycles [93, 94]; testing these predictions experimentally requires the combination of precise Liouvillian control and long coherence times that fault-tolerant hardware would provide.

A longer-horizon direction concerns the unification of the three toolkits. The chirality-based characterization of Chapter 2 detects entanglement through multi-copy handedness correlators. The SQGEN framework of Chapter 3 generates target states by optimizing a synergic circuit. The LEP framework of Chapter 4 controls the relaxation dynamics of open systems through dissipation engineering. A unified theory would describe how these three operations: characterize, generate, and control, interact when the target object is not a pure state but a quantum process: a channel or a Liouvillian. Process characterization via QPT, process generation via a Liouvillian analogue of SQGEN, and process singularity engineering via LEP design form a natural triplet of operations on the space of quantum channels, with potential applications to quantum error correction (characterizing noise channels), quantum simulation (generating target dissipative dynamics), and quantum sensing (exploiting channel singularities).

5.3.3 Open problems and conjectures

Several concrete open problems emerge from the results of this dissertation.

The first concerns the gradient landscape of SQGEN. The conjecture that the synergic cost function avoids barren plateaus due to the post-selection mechanism has been supported empirically but not proved. A proof or disproof would require bounding the variance of $\partial J/\partial\theta_k$ over random initializations of the Möttönen circuit, likely using concentration inequalities on the unitary group. If the conjecture holds, SQGEN would be provably trainable in regimes where standard QGAN is not, establishing a genuine theoretical advantage.

The second open problem concerns the completeness of the chirality witness. The identity $C_4 \propto \mathcal{N}$ holds for pure states and makes C_4 a complete entanglement witness in that case. For mixed states, the connection between C_k and negativity is mediated by the partial-transpose moments μ_k, and the question of which mixed-state entanglement structures produce $C_k \neq 0$ for a given k is open. In particular, it is not known whether there exist entangled states with zero chirality of all orders, which would constitute a blind spot of the multi-copy chirality framework.

The third open problem concerns the quantum-classical boundary in the LEP context. The third chapter demonstrates that LEPs can exist without HEPs. The converse, however, remains unresolved: can HEPs always be associated with at least one LEP in a physical open system? Theory has resolved this question in the affirmative (every HEP of the form $\hat{H}_e$ induces a LEP in $\mathcal{L}$)), but this has not yet been verified experimentally for a system in which both types of EPs occur simultaneously. Designing and implementing a two-channel qubit system with both HEPs and LEPs, and verifying their coexistence via QPT, would complete the experimental picture.

The fourth concerns the sensing figure of merit near higher-order LEPs. For a second-order EP, the optimal operating point satisfies a balance condition between $\epsilon^{1/2}$ sensitivity gain and $\epsilon^{-1/2}$ noise amplification. For an N-th order EP, the analogous balance gives an optimal offset $\epsilon^* \propto (\sigma_\lambda/\omega)^{2N/(N+1)}$ from the EP, suggesting that the net sensing advantage grows as N increases for fixed noise level. Whether this scaling is achievable in practice on superconducting hardware, where gate errors introduce a noise floor that may saturate before the enhancement regime is reached, remains an open experimental question.

Finally, the integration program outlined above suggests a broader conjecture: that the three nonlinear quantum resources studied in this dissertation, entanglement (Chapter 2), state generation capacity (Chapter 3), and Liouvillian singularity (Chapter 4), are not independent but are related by a deeper structural principle. A candidate for such a principle is the quantum resource theory of coherence in the space of quantum channels [114], which unifies state coherence, entanglement, and process non-classicality under a single framework. Whether LEPs correspond to a specific resource-theoretic object within this framework, and whether SQGEN can be interpreted as a coherence amplification protocol, are speculative but concrete questions that can be addressed with existing mathematical tools.

The results presented in this dissertation demonstrate that nonlinear quantum systems, understood as systems whose operationally relevant properties require nonlinear functions of observables or nonlinear dynamics to describe, are accessible, measurable, and controllable on current hardware. The three-part structure of characterization, generation, and singularity exploitation provides a modular framework that can be extended as hardware improves,

and the open problems identified above define a research agenda that connects near-term experimental work on NISQ processors to the long-term vision of fully programmable quantum systems.

References

[1] J. S. Bell. On the Einstein Podolsky Rosen paradox. *Physics Physique Fizika*, 1:195–200, Nov 1964.

[2] Peter W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26(5):1484–1509, 1997.

[3] Lov K. Grover. Quantum mechanics helps in searching for a needle in a haystack. *Phys. Rev. Lett.*, 79:325–328, Jul 1997.

[4] Charles H. Bennett and Gilles Brassard. Quantum cryptography: Public key distribution and coin tossing. volume 560, pages 7–11, 2014. Theoretical Aspects of Quantum Cryptography – celebrating 30 years of BB84.

[5] Artur K. Ekert. Quantum cryptography based on bell's theorem. *Phys. Rev. Lett.*, 67:661–663, Aug 1991.

[6] John Preskill. Quantum Computing in the NISQ era and beyond. *Quantum*, 2:79, August 2018.

[7] Asher Peres. Separability criterion for density matrices. *Phys. Rev. Lett.*, 77:1413–1415, Aug 1996.

[8] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. Separability of mixed states: necessary and sufficient conditions. *Physics Letters A*, 223(1):1–8, 1996.

[9] Seth Lloyd and Christian Weedbrook. Quantum generative adversarial learning. *Phys. Rev. Lett.*, 121:040502, Jul 2018.

[10] Pierre-Luc Dallaire-Demers and Nathan Killoran. Quantum generative adversarial networks. *Phys. Rev. A*, 98:012324, Jul 2018.

[11] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[12] Kai Chen and Ling-An Wu. A matrix realignment method for recognizing entanglement. *Quantum Inf. Comput.*, 3(3):193–202, 2003.

[13] Oliver Rudolph. Further results on the cross norm criterion for separability. *Quantum Information Processing*, 4(3):219–239, Aug 2005.

[14] Yuriy Makhlin. Nonlocal properties of two-qubit gates and mixed states, and the optimization of quantum computations. *Quantum Information Processing*, 1(4):243–252, Aug 2002.

[15] G. Lindblad. On the generators of quantum dynamical semigroups. *Communications in Mathematical Physics*, 48(2):119–130, Jun 1976.

[16] Vittorio Gorini, Andrzej Kossakowski, and E. C. G. Sudarshan. Completely positive dynamical semigroups of n-level systems. *Journal of Mathematical Physics*, 17(5):821–825, 05 1976.

[17] Jean Dalibard, Yvan Castin, and Klaus Mølmer. Wave-function approach to dissipative processes in quantum optics. *Phys. Rev. Lett.*, 68:580–583, Feb 1992.

[18] Klaus Mølmer, Yvan Castin, and Jean Dalibard. Monte carlo wave-function method in quantum optics. *J. Opt. Soc. Am. B*, 10(3):524–538, Mar 1993.

[19] Weijian Chen, Şahin Kaya Özdemir, Guangming Zhao, Jan Wiersig, and Lan Yang. Exceptional points enhance sensing in an optical microcavity. *Nature*, 548(7666):192–196, August 2017.

[20] Fabrizio Minganti, Adam Miranowicz, Ravindra W. Chhajlany, and Franco Nori. Quantum exceptional points of non-hermitian hamiltonians and liouvillians: The effects of quantum jumps. *Phys. Rev. A*, 100:062131, Dec 2019.

[21] W. K. Wootters and W. H. Żurek. A single quantum cannot be cloned. *Nature*, 299(5886):802–803, Oct 1982.

[22] Karol Bartkiewicz, Patrycja Tulewicz, Jan Roik, and Karel Lemr. Synergic quantum generative machine learning. *Scientific Reports*, 13(1):12893, 2023.

[23] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki. Quantum entanglement. *Rev. Mod. Phys.*, 81:865–942, Jun 2009.

[24] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 47:777–780, May 1935.

[25] John F. Clauser, Michael A. Horne, Abner Shimony, and Richard A. Holt. Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.*, 23:880–884, Oct 1969.

[26] B. S. Cirel'son. Quantum generalizations of Bell's inequality. *Letters in Mathematical Physics*, 4(2):93–100, 1980.

[27] Alain Aspect, Philippe Grangier, and Gérard Roger. Experimental tests of realistic local theories via bell's theorem. *Phys. Rev. Lett.*, 47:460–463, Aug 1981.

[28] Alain Aspect, Philippe Grangier, and Gérard Roger. Experimental realization of Einstein-Podolsky-Rosen-Bohm gedankenexperiment: A new violation of bell's inequalities. *Phys. Rev. Lett.*, 49:91–94, Jul 1982.

[29] H. M. Wiseman, S. J. Jones, and A. C. Doherty. Steering, entanglement, nonlocality, and the Einstein-Podolsky-Rosen paradox. *Phys. Rev. Lett.*, 98:140402, Apr 2007.

[30] G. Vidal and R. F. Werner. Computable measure of entanglement. *Phys. Rev. A*, 65:032314, Feb 2002.

[31] Paweł Horodecki and Artur Ekert. Method for direct detection of quantum entanglement. *Phys. Rev. Lett.*, 89:127902, Aug 2002.

[32] Paweł Horodecki. Measuring quantum entanglement without prior state reconstruction. *Phys. Rev. Lett.*, 90:167901, Apr 2003.

[33] Fabio Antonio Bovino, Giuseppe Castagnoli, Artur Ekert, Paweł Horodecki, Carolina Moura Alves, and Alexander Vladimir Sergienko. Direct measurement of nonlinear properties of bipartite quantum states. *Phys. Rev. Lett.*, 95:240407, Dec 2005.

[34] Karol Bartkiewicz and Grzegorz Chimczak. Two methods for measuring bell nonlocality via local unitary invariants of two-qubit systems in hong-ou-mandel interferometers. *Phys. Rev. A*, 97:012107, Jan 2018.

[35] Karol Bartkiewicz, Grzegorz Chimczak, and Karel Lemr. Direct method for measuring and witnessing quantum entanglement of arbitrary two-qubit states through hong-ou-mandel interference. *Phys. Rev. A*, 95:022331, Feb 2017.

[36] K. Banaszek, G. M. D'Ariano, M. G. A. Paris, and M. F. Sacchi. Maximum-likelihood estimation of the density matrix. *Phys. Rev. A*, 61:010304, Dec 1999.

[37] Dominik Koutný, Laia Ginés, Magdalena Moczała-Dusanowska, Sven Höfling, Christian Schneider, Ana Predojević, and Miroslav Ježek. Deep learning of quantum entanglement from incomplete measurements. *Science Advances*, 9(29):eadd7131, 2023.

[38] Scott Lundberg and Su-In Lee. A unified approach to interpreting model predictions, 2017.

[39] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, and Édouard Duchesnay. Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, 12(85):2825–2830, 2011.

[40] IBM Quantum. https://quantum.cloud.ibm.com/, 2026.

[41] Ali Javadi-Abhari, Matthew Treinish, Kevin Krsulich, Christopher J. Wood, Jake Lishman, Julien Gacon, Simon Martiel, Paul D. Nation, Lev S. Bishop, Andrew W. Cross, Blake R. Johnson, and Jay M. Gambetta. Quantum computing with Qiskit, 2024.

[42] Jens Koch, Terri M. Yu, Jay Gambetta, A. A. Houck, D. I. Schuster, J. Majer, Alexandre Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf. Charge-insensitive qubit design derived from the cooper pair box. *Phys. Rev. A*, 76:042319, Oct 2007.

[43] Łukasz Rudnicki, Paweł Horodecki, and Karol Życzkowski. Collective uncertainty entanglement test. *Phys. Rev. Lett.*, 107:150502, Oct 2011.

[44] Karel Lemr, Karol Bartkiewicz, and Antonín Černoch. Experimental measurement of collective nonlinear entanglement witness for two qubits. *Phys. Rev. A*, 94:052334, Nov 2016.

[45] V. Kalmeyer and R. B. Laughlin. Equivalence of the resonating-valence-bond and fractional quantum hall states. *Phys. Rev. Lett.*, 59:2095–2098, Nov 1987.

[46] Y. Taguchi, Y. Oohara, H. Yoshizawa, N. Nagaosa, and Y. Tokura. Spin chirality, berry phase, and anomalous hall effect in a frustrated ferromagnet. *Science*, 291(5513):2573–2576, 2001.

[47] Michał Horodecki, Paweł Horodecki, and Ryszard Horodecki. Mixed-state entanglement and distillation: Is there a "bound" entanglement in nature? *Phys. Rev. Lett.*, 80:5239–5242, Jun 1998.

[48] Xiao-Dong Yu, Satoya Imai, and Otfried Gühne. Optimal entanglement certification from moments of the partial transpose. *Phys. Rev. Lett.*, 127:060504, Aug 2021.

[49] Antoine Neven, Jose Carrasco, Vittorio Vitale, Christian Kokail, Andreas Elben, Marcello Dalmonte, Pasquale Calabrese, Peter Zoller, Benoît Vermersch, Richard Kueng, and Barbara Kraus. Symmetry-resolved entanglement detection using partial transpose moments. *npj Quantum Information*, 7(1):152, Oct 2021.

[50] Vojt ěch Trávníček, Karol Bartkiewicz, Antonín Černoch, and Karel Lemr. Experimental measurement of the hilbert-schmidt distance between two-qubit states as a means for reducing the complexity of machine learning. *Phys. Rev. Lett.*, 123:260501, Dec 2019.

[51] Patrycja Tulewicz, Karol Bartkiewicz, Adam Miranowicz, and Franco Nori. Resource-efficient quantum correlation measurements via multicopy neural network methods. *Scientific Reports*, 15(1):40868, Nov 2025.

[52] Andreas Elben, Richard Kueng, Hsin-Yuan (Robert) Huang, Rick van Bijnen, Christian Kokail, Marcello Dalmonte, Pasquale Calabrese, Barbara Kraus, John Preskill, Peter Zoller, and Benoît Vermersch. Mixed-state entanglement from local randomized measurements. *Phys. Rev. Lett.*, 125:200501, Nov 2020.

[53] Hsin-Yuan Huang, Richard Kueng, and John Preskill. Predicting many properties of a quantum system from very few measurements. *Nature Physics*, 16(10):1050–1057, Oct 2020.

[54] Karol Bartkiewicz, Ji ří Beran, Karel Lemr, Michał Norek, and Adam Miranowicz. Quantifying entanglement of a two-qubit system via measurable and invariant moments of its partially transposed density matrix. *Phys. Rev. A*, 91:022323, Feb 2015.

[55] L. I. Reascos, Bruno Murta, E. F. Galvão, and J. Fernández-Rossier. Quantum circuits to measure scalar spin chirality. *Phys. Rev. Res.*, 5:043087, Oct 2023.

[56] William K. Wootters. Entanglement of formation of an arbitrary state of two qubits. *Phys. Rev. Lett.*, 80:2245–2248, Mar 1998.

[57] Pawel Horodecki. Separability criterion and inseparable mixed states with positive partial transposition. *Physics Letters A*, 232(5):333–339, 1997.

[58] Charles H. Bennett, David P. DiVincenzo, Tal Mor, Peter W. Shor, John A. Smolin, and Barbara M. Terhal. Unextendible product bases and bound entanglement. *Phys. Rev. Lett.*, 82:5385–5388, Jun 1999.

[59] Dagmar Bruß and Asher Peres. Construction of quantum states with bound entanglement. *Phys. Rev. A*, 61:030301, Feb 2000.

[60] Andrew C. Doherty, Pablo A. Parrilo, and Federico M. Spedalieri. Complete family of separability criteria. *Phys. Rev. A*, 69:022308, Feb 2004.

[61] Paul D. Nation, Hwajung Kang, Neereja Sundaresan, and Jay M. Gambetta. Scalable mitigation of measurement errors on quantum computers. *PRX Quantum*, 2:040326, Nov 2021.

[62] Elias Amselem and Mohamed Bourennane. Experimental four-qubit bound entanglement. *Nature Physics*, 5(10):748–752, Oct 2009.

[63] Chao Zhang, Yuan-Yuan Zhao, Nikolai Wyderka, Satoya Imai, Andreas Ketterer, Ning-Ning Wang, Kai Xu, Keren Li, Bi-Heng Liu, Yun-Feng Huang, Chuan-Feng Li, Guang-Can Guo, and Otfried Gühne. Experimental verification of bound and multiparticle entanglement with the randomized measurement toolbox. 2023.

[64] Vaishali Gulati, Gayatri Singh, and Kavita Dorai. Using linear and nonlinear entanglement witnesses to generate and detect bound entangled states on an ibm quantum processor. *Physica Scripta*, 99(11):115122, oct 2024.

[65] Poetri Sonya Tarabunga and Tobias Haug. Quantifying mixed-state entanglement via partial transpose and realignment moments. 2025.

[66] Xiaofen Huang, Xishun Zhu, Bin Chen, Naihuan Jing, and Shao-Ming Fei. A note on entanglement detection via the generalized realignment moments. *Advanced Quantum Technologies*, 9(1):e00794, 2026.

[67] Jacob Biamonte, Peter Wittek, Nicola Pancotti, Patrick Rebentrost, Nathan Wiebe, and Seth Lloyd. Quantum machine learning. *Nature*, 549(7671):195–202, Sep 2017.

[68] Vojtěch Havlíček, Antonio D. Córcoles, Kristan Temme, Aram W. Harrow, Abhinav Kandala, Jerry M. Chow, and Jay M. Gambetta. Supervised learning with quantum-enhanced feature spaces. *Nature*, 567(7747):209–212, Mar 2019.

[69] K. Mitarai, M. Negoro, M. Kitagawa, and K. Fujii. Quantum circuit learning. *Phys. Rev. A*, 98:032309, Sep 2018.

[70] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. 27, 2014.

[71] Vedran Dunjko, Yimin Ge, and J. Ignacio Cirac. Computational speedups using small quantum devices. *Phys. Rev. Lett.*, 121:250501, Dec 2018.

[72] Mikko Möttönen, Juha J. Vartiainen, Ville Bergholm, and Martti M. Salomaa. Transformation of quantum states using uniformly controlled rotations. *Quantum Information & Computation*, 5:467–473, 2005.

[73] Christa Zoufal, Aurélien Lucchi, and Stefan Woerner. Quantum generative adversarial networks for learning and loading random distributions. *npj Quantum Information*, 5(1):103, Nov 2019.

[74] Maria Schuld, Ville Bergholm, Christian Gogolin, Josh Izaac, and Nathan Killoran. Evaluating analytic gradients on quantum hardware. *Phys. Rev. A*, 99:032331, Mar 2019.

[75] Abhinav Kandala, Kristan Temme, Antonio D. Córcoles, Antonio Mezzacapo, Jerry M. Chow, and Jay M. Gambetta. Error mitigation extends the computational reach of a noisy quantum processor. *Nature*, 567(7749):491–495, Mar 2019.

[76] Mark M. Wilde. *Quantum Information Theory*. Cambridge University Press, 2 edition, 2017.

[77] Stephen M. Barnett and Sarah Croke. Quantum state discrimination. *Adv. Opt. Photon.*, 1(2):238–278, Apr 2009.

[78] Abhinav Kandala, Antonio Mezzacapo, Kristan Temme, Maika Takita, Markus Brink, Jerry M. Chow, and Jay M. Gambetta. Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. *Nature*, 549(7671):242–246, Sep 2017.

[79] Jonathan Romero, Jonathan P Olson, and Alan Aspuru-Guzik. Quantum autoencoders for efficient compression of quantum data. *Quantum Science and Technology*, 2(4):045001, aug 2017.

[80] Daniel Gottesman. The Heisenberg representation of quantum computers. 1998.

[81] F. Verstraete, V. Murg, and J.I. Cirac. Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems. *Advances in Physics*, 57(2):143–224, 2008.

[82] Jarrod R. McClean, Sergio Boixo, Vadim N. Smelyanskiy, Ryan Babbush, and Hartmut Neven. Barren plateaus in quantum neural network training landscapes. *Nature Communications*, 9(1):4812, Nov 2018.

[83] Szymon Knop, Marcin Mazur, Przemysław Spurek, Jacek Tabor, and Igor Podolak. Generative models with kernel distance in data space. *Neurocomputing*, 487:119–129, 2022.

[84] Ş K. Özdemir, S. Rotter, F. Nori, and L. Yang. Parity–time symmetry and exceptional points in photonics. *Nature Materials*, 18(8):783–798, Aug 2019.

[85] Carl M. Bender and Stefan Boettcher. Real spectra in non-hermitian hamiltonians having $\mathcal{PT}$ symmetry. *Phys. Rev. Lett.*, 80:5243–5246, Jun 1998.

[86] Christian E. Rüter, Konstantinos G. Makris, Ramy El-Ganainy, Demetrios N. Christodoulides, Mordechai Segev, and Detlef Kip. Observation of parity–time symmetry in optics. *Nature Physics*, 6(3):192–195, Mar 2010.

[87] Emil J. Bergholtz, Jan Carl Budich, and Flore K. Kunst. Exceptional topology of non-hermitian systems. *Rev. Mod. Phys.*, 93:015005, Feb 2021.

[88] M. V. Berry. Physics of nonhermitian degeneracies. *Czechoslovak Journal of Physics*, 54(10):1039–1047, Oct 2004.

[89] W D Heiss. Exceptional points of non-hermitian operators. *Journal of Physics A: Mathematical and General*, 37(6):2455, jan 2004.

[90] Heinz-Peter Breuer and Francesco Petruccione. *The Theory of Open Quantum Systems*. Oxford University Press, 01 2007.

[91] Ievgen I. Arkhipov, Fabrizio Minganti, Adam Miranowicz, and Franco Nori. Generating high-order quantum exceptional points in synthetic dimensions. *Phys. Rev. A*, 104:012205, Jul 2021.

[92] Fabrizio Minganti, Adam Miranowicz, Ravindra W. Chhajlany, Ievgen I. Arkhipov, and Franco Nori. Hybrid-liouvillian formalism connecting exceptional points of non-hermitian hamiltonians and liouvillians via postselection of quantum trajectories. *Phys. Rev. A*, 101:062112, Jun 2020.

[93] J.-T. Bu, J.-Q. Zhang, G.-Y. Ding, J.-C. Li, J.-W. Zhang, B. Wang, W.-Q. Ding, W.-F. Yuan, L. Chen, Ş. K. Özdemir, F. Zhou, H. Jing, and M. Feng. Enhancement of quantum heat engine by encircling a liouvillian exceptional point. *Phys. Rev. Lett.*, 130:110402, Mar 2023.

[94] J.-W. Zhang, J.-Q. Zhang, G.-Y. Ding, J.-C. Li, J.-T. Bu, B. Wang, L.-L. Yan, S.-L. Su, L. Chen, F. Nori, Ş K. Özdemir, F. Zhou, H. Jing, and M. Feng. Dynamical control of quantum heat engines using exceptional points. *Nature Communications*, 13(1):6225, Oct 2022.

[95] M. Naghiloo, M. Abbasi, Yogesh N. Joglekar, and K. W. Murch. Quantum state tomography across the exceptional point in a single dissipative qubit. *Nature Physics*, 15(12):1232–1236, Dec 2019.

[96] Weijian Chen, Maryam Abbasi, Yogesh N. Joglekar, and Kater W. Murch. Quantum jumps in the non-hermitian dynamics of a superconducting qubit. *Phys. Rev. Lett.*, 127:140504, Sep 2021.

[97] Weijian Chen, Maryam Abbasi, Byung Ha, Serra Erdamar, Yogesh N. Joglekar, and Kater W. Murch. Decoherence-induced exceptional points in a dissipative superconducting qubit. *Phys. Rev. Lett.*, 128:110402, Mar 2022.

[98] Victor V. Albert and Liang Jiang. Symmetries and conserved quantities in lindblad master equations. *Phys. Rev. A*, 89:022118, Feb 2014.

[99] Jörg Doppler, Alexei A. Mailybaev, Julian Böhm, Ulrich Kuhl, Adrian Girschik, Florian Libisch, Thomas J. Milburn, Peter Rabl, Nimrod Moiseyev, and Stefan Rotter. Dynamically encircling an exceptional point for asymmetric mode switching. *Nature*, 537(7618):76–79, Sep 2016.

[100] Ievgen I. Arkhipov, Adam Miranowicz, Fabrizio Minganti, Şahin K. Özdemir, and Franco Nori. Dynamically crossing diabolic points while encircling exceptional curves: A programmable symmetric-asymmetric multimode switch. *Nature Communications*, 14(1):2076, Apr 2023.

[101] Daniel F. V. James, Paul G. Kwiat, William J. Munro, and Andrew G. White. Measurement of qubits. *Phys. Rev. A*, 64:052312, Oct 2001.

[102] Isaac L. Chuang and M. A. Nielsen. Prescription for experimental determination of the dynamics of a quantum black box. *Journal of Modern Optics*, 44(11-12):2455–2467, 1997.

[103] J. F. Poyatos, J. I. Cirac, and P. Zoller. Complete characterization of a quantum process: The two-bit quantum gate. *Phys. Rev. Lett.*, 78:390–393, Jan 1997.

[104] M. Mohseni, A. T. Rezakhani, and D. A. Lidar. Quantum-process tomography: Resource analysis of different strategies. *Phys. Rev. A*, 77:032322, Mar 2008.

[105] Adam Miranowicz, Karol Bartkiewicz, Jan Peřina, Masato Koashi, Nobuyuki Imoto, and Franco Nori. Optimal two-qubit tomography based on local and global measurements: Maximal robustness against errors as described by condition numbers. *Phys. Rev. A*, 90:062123, Dec 2014.

[106] Man-Duen Choi. Completely positive linear maps on complex matrices. *Linear Algebra and its Applications*, 10(3):285–290, 1975.

[107] Andrzej Jamiołkowski. Linear transformations which preserve trace and positive semidefiniteness of operators. *Reports on Mathematical Physics*, 3:275–278, 1972.

[108] Lorenza Viola and Seth Lloyd. Dynamical suppression of decoherence in two-state quantum systems. *Phys. Rev. A*, 58:2733–2744, Oct 1998.

[109] Michael J. Biercuk, Hermann Uys, Aaron P. VanDevender, Nobuyasu Shiga, Wayne M. Itano, and John J. Bollinger. Optimized dynamical decoupling in a model quantum memory. *Nature*, 458(7241):996–1000, Apr 2009.

[110] Hoi-Kwan Lau and Aashish A. Clerk. Fundamental limits and non-reciprocal approaches in non-hermitian quantum sensing. *Nature Communications*, 9(1):4320, Oct 2018.

[111] Mengzhen Zhang, William Sweeney, Chia Wei Hsu, Lan Yang, A. D. Stone, and Liang Jiang. Quantum noise theory of exceptional point amplifying sensors. *Phys. Rev. Lett.*, 123:180501, Oct 2019.

[112] Hudson Loughlin and Vivishek Sudhir. Exceptional-point sensors offer no fundamental signal-to-noise ratio enhancement. *Phys. Rev. Lett.*, 132:243601, Jun 2024.

[113] Jaromír Fiurášek and Zden ěk Hradil. Maximum-likelihood estimation of quantum processes. *Phys. Rev. A*, 63:020101, Jan 2001.

[114] Thomas Theurer, Dario Egloff, Lijian Zhang, and Martin B. Plenio. Quantifying operations with an application to coherence. *Phys. Rev. Lett.*, 122:190405, May 2019.

Summary

Quantum mechanics is a fundamentally linear theory. The evolution of quantum states, described by Schrödinger's equation, is based on the principle of superposition, and observables act linearly on Hilbert space. However, the quantities that are most significant in quantum information processing are, almost without exception, nonlinear functions of the state. Entanglement cannot be captured by a single expectation value. Measures such as negativity require spectral information about the partially transposed density matrix, which makes them inherently nonlinear functionals. Measurement is an irreversible, stochastic transformation that maps a density matrix to a probability distribution in a way that cannot be reduced to any linear observable. Furthermore, the dissipative dynamics that govern realistic quantum hardware produce singularities in parameter space where the spectral structure of the evolution operator changes qualitatively. This phenomenon is entirely absent from the unitary dynamics of closed systems.

These three sources of nonlinearity give rise to the central questions of this dissertation. On near-term superconducting processors, uncontrolled decoherence degrades entanglement on timescales comparable to gate durations. Qubit counts remain limited, and circuit depth is constrained by finite coherence. Standard quantum state tomography requires an exponential number of measurement settings proportional to the system size, making it impractical for even modest qubit registers. Preparing a state with prescribed entanglement properties requires navigating a highly nonconvex variational landscape where gradients can vanish and local minima are abundant. The singular behavior of open quantum systems near exceptional points had not been accessed experimentally through the full Lindblad description prior to the work presented here, despite being theoretically well-studied in effective non-Hermitian models. The three experimental contributions of this dissertation address these challenges in turn, progressing from characterization of nonlinear properties through their efficient generation to the deliberate exploitation of open-system singularities. All experiments were carried out on IBM Quantum superconducting processors.

The first chapter establishes the theoretical foundation for what follows. It reviews the postulates of quantum mechanics using the density-matrix formalism and introduces the

partial trace, the Bloch sphere picture for qubits, and the Pauli expansion of multi-qubit states. The Peres-Horodecki positive partial transposition criterion and the realignment criterion are presented as the two main tools for detecting entanglement beyond single-copy expectation values. Alongside their representation as singlet projections on multi-copy systems, the Makhlin invariants, which classify two-qubit states completely up to local unitary equivalence, are introduced. This connection forms the algebraic basis of the characterization methods developed in Chapter 2. The theory of open quantum systems is developed through the Markovian approximation to the Gorini-Kossakowski-Sudarshan-Lindblad master equation. The Liouvillian superoperator is treated in its vectorized form, and its eigenvalue problem is analyzed in detail. The quantum trajectory picture decomposes smooth Lindblad evolution into a continuous, non-unitary drift under an effective, non-Hermitian Hamiltonian, as well as discrete, stochastic quantum jumps. This makes the distinction between Hamiltonian exceptional points (HEPs) and Liouvillian exceptional points (LEPs) precise. This distinction is central to Chapter 4. HEPs are degeneracies of the effective Hamiltonian alone. Therefore, they belong to the semiclassical regime. LEPs are degeneracies of the full Liouvillian superoperator. They depend on quantum-jump terms that have no classical counterpart. This is why every HEP induces a corresponding LEP, but not vice versa. The chapter concludes with the no-go theorems of quantum information: no cloning, no deleting, no broadcasting, and no signaling, whose operational consequences influence the experimental strategies throughout the dissertation. The impossibility of copying quantum states motivates collective multi-copy measurements over independent single-copy tomography and prevents the naive transfer of classical generative learning algorithms to the quantum domain. Furthermore, it underlies the necessity of synergic circuit design in Chapter 3.

Chapter 2 addresses the problem of characterizing quantum correlations on near-term hardware without full state reconstruction. Three complementary approaches were developed and validated experimentally. The first approach is a multicopy estimation protocol that combines singlet-projection measurements with artificial neural networks. The key observation is that Makhlin invariants, from which negativity and Bell nonlocality can be recovered, can be expressed as expectation values of singlet projectors measured simultaneously on two copies of the state. This reduces the required number of measurement settings from the 4^n characteristic of full tomography to 2^n. A neural network trained on randomly generated two-qubit states can map these projections to entanglement measures without solving the associated polynomial equations. This approach achieves substantially better robustness to measurement noise than direct analytical inversion. The advantage over the analytical approach is structural: direct inversion of the polynomial root-finding problem amplifies noise through nonlinear expressions. In contrast, the network learns smooth mappings

that are less sensitive to small perturbations of the input. A systematic SHAP (Shapley additive explanations) analysis of feature importance reveals that only five of the twelve available projection configurations carry most of the predictive signal for negativity and Bell nonlocality. This reduces experimental overhead by 67% relative to the full projection set. The five configurations relevant for negativity are dominated by loop and inter-subsystem projections that capture local coherence and cross-subsystem correlations. Those relevant for Bell nonlocality, on the other hand, reflect the local measurement statistics of the CHSH operator. The protocol was implemented on *ibm_hanoi*, with qubit mappings averaged across multiple configurations to suppress systematic bias. The result was copy fidelity above 98%.

The second approach uses Gröbner basis decision trees to reconstruct adaptive negativity from partial transpose moments. Negativity is determined by the eigenvalues of the partially transposed density matrix. These eigenvalues are determined by the spectral moments $\mu_k = \mathrm{Tr}[(\rho^{T_A})^k]$ through the Newton-Girard identities and Ferrari's quartic formula. Each moment is measured by a single Hadamard test circuit acting on multiple copies of the state with a controlled cyclic permutation. A crucial algebraic identity holds universally: the second-order partial transpose moment is equal to the purity, $\mu_2 = \mathcal{I}_2 = \mathrm{Tr}[\rho^2]$, regardless of system dimension. This means that a single purity circuit provides the second-order invariant for both the partial transpose and purity moment series simultaneously, reducing the circuit count by one. The key observation that informs the decision-tree structure is that most two-qubit states belong to degeneracy classes of eigenvalues that can be identified from the first two moments via Gröbner basis polynomials, $\mathcal{G}_1$ and $\mathcal{G}_2$. A triple-plus degeneracy, where three eigenvalues coincide (covering Bell states, Werner states, and the maximally mixed state), requires only a single circuit. A two-pair degeneracy requires two circuits, and the generic case, with four distinct eigenvalues, requires all three moments $\{\mu_2, \mu_3, \mu_4\}$. Monte Carlo sampling of 10^5 Haar-random two-qubit states confirms that 95.3% fall into the generic branch and 4.7% into degenerate branches. On average 2.15 circuit configurations suffice, which corresponds to a $5.3\times$ reduction in measurement settings compared to full two-qubit tomography and a $7.2\times$ reduction for qubit-qutrit systems. Under depolarizing noise the decision nodes apply a tolerance $|G_i| < 3\sigma_G$ rather than an exact zero. Misclassification routes a state into the wrong branch, but it never produces an error in the final negativity since the generic quartic solver always serves as a reliable backup. Classification accuracy above 95% is maintained up to depolarizing strengths of approximately $p = 0.08$, which covers the typical operating range of current superconducting processors.

The third contribution identifies the physical content of the partial transpose moment corrections, $C_k = \mu_k - \mathcal{I}_k$ and uses this identification to expand the measurement framework for detecting bound entanglement. The central algebraic result is that this difference

decomposes exactly as a chirality-chirality correlator. There, the scalar spin chirality appears among qubits drawn from different copies of the state rather than from different lattice sites. This operator is well-known in condensed matter physics as the order parameter for chiral spin liquids and the driver of the topological Hall effect. The fourth-order result is $C_4 = 8\,\mathrm{Tr}[\Omega_A \Omega_B \rho^{\otimes 4}]$, where $\Omega = \frac{1}{2}\sum_{i<j<k} \chi_{ijk}$, and it is derived from the SWAP operator decomposition and Pauli commutator identities. This connection is not just an analogy. The same algebraic structure governs both phenomena, and the geometric interpretation is identical. A product state generates multi-copy spin configurations that remain coplanar across copies. This gives zero chirality. An entangled state, however, breaks this coplanarity simultaneously on both subsystems. This produces a nonzero correlator. At the second order, $C_2 = 0$ universally, which is a consequence of the fact that $\sigma = \sigma^{-1} = S_{12}$ for two copies. Thus, the anti-Hermitian operator $\Delta = \sigma^{-1} - \sigma$ vanishes identically. At the third order, the correction, C_3 is proportional to the commutator of two SWAP operators that share a copy index. This commutator is proportional to the scalar spin chirality of three copies. At the fourth order all four chirality triples of four copies contribute equally. For pure two-qubit states the fourth-order correction satisfies $C_4 = -4\mathcal{N}^2(1-\mathcal{N}^2)$, making chirality a complete entanglement witness. For Bell-diagonal states $C_4 = \frac{3}{4}\det(T)$, where T is the correlation tensor, giving chirality the geometric meaning of the signed volume of the correlation hypercube. The separable bounds $|C_3| \leq 1/36$ and $|C_4| \leq 1/27$ are achieved by rank-three equal-weight mixtures of three mutually unbiased basis product states, and they are tight. Multi-copy chirality is an intrinsically collective quantity. For a known state, it can be computed from the density matrix. However, no single-copy measurement on an unknown state can extract it without first performing complete state tomography.

Using the same controlled-SWAP infrastructure and reinterpreting it through the realignment map $R(\rho)$, one can access the non-Hermitian gap $D_k = \Sigma_k - G_k$ between the singular-value moments Σ_k and eigenvalue moments G_k of the realigned density matrix. This gap is directly related to spin physics as the antisymmetric part of the A–B spin correlation tensor and is connected to the cross-product, $\langle \mathbf{J}_A \times \mathbf{J}_B \rangle$. It is also sensitive to a type of entanglement that the PPT criterion cannot detect. For real-entry states the odd-order cross-restructuring difference $\delta G_k = G_k - G_k^{PT}$ provides a complementary signal. The algebraic reason that chessboard states evade this witness is traced to their checkerboard parity structure, which forces $R_- = 0$ and therefore $\delta G_k = 0$ identically, despite genuine entanglement. This requires rank-structure features to complete the detection. A Lasso logistic regression classifier trained on degree-2 polynomial combinations of the six spectral observables $\{\Sigma_1, G_1, \Sigma_2, G_2, D_1, D_2\}$ achieves 94% cross-validated recall of bound entanglement, with zero false positives. An ExtraTrees ensemble trained on 522 nonlinear

features raises this to 99.6%. Both classifiers cover all three known families of 3×3 PPT entangled states – Horodecki, Tiles UPB, and Chessboard – across their full parameter ranges, coverage not previously achieved by any single criterion. The work also identifies and experimentally demonstrates a new class of CCNR-invisible bound entangled states. These are marginal-noise mixtures, defined as follows: $\rho_{\mathrm{mn}}(t) = (1-t)\rho_0 + t\,\rho_A \otimes \rho_B$. For these states, the standard realignment trace-norm criterion, defined as $\|\rho^R\|_1 > 1$ fails for all $t \gtrsim 0.01$, while $D_2 > 0$ persists and enables detection with noise tolerance 10–20$\times$ greater than the CCNR threshold. The physical mechanism is that mixing with the marginal product suppresses the dominant singular value of the realignment matrix, which is responsible for the criterion failing, without eliminating the spectral asymmetry $D_2 > 0$.

All experimental results were obtained using the Heron architecture with a native gate set of $\{\mathrm{CZ}, \sqrt{X}, R_Z, X\}$, readout error mitigation, and Qiskit optimization level 3 transpilation on *ibm_kingston*, *ibm_torino*, and *ibm_fez*. The negativity mean errors are 0.002 (Kingston, 2×2), 0.012 (Torino, 2×2), and 0.027 (Torino, 2×3). The chirality mean errors are 0.022 (2×2) and 0.039 (2×3), with the larger scatter in the qubit-qutrit case being due to the roughly twofold increase in circuit depth for four-copy circuits at that dimension. The closed-form curve $-C_4 = 4\mathcal{N}^2(1-\mathcal{N}^2)$ is confirmed directly from the hardware data across the full entanglement range. This includes Werner state mixtures $\rho_W(p) = p|\Psi^-\rangle\langle\Psi^-| + (1-p)\mathbb{I}/4$ where the theoretical curve $-C_4 = \frac{3}{4}p^3$ is verified with chirality detection above the separable bound for $p \gtrsim 0.37$. All five bound entangled test states on *ibm_fez* are detected with positive D_2, while both separable control states yield D_2 consistent with zero.

Chapter 3 addresses the problem of efficiently preparing quantum states with a prescribed correlation structure. It introduces the Synergic Quantum Generative Network (SQGEN) as an alternative to the Quantum Generative Adversarial Network (QGAN) paradigm. A fundamentally different approach is motivated by a structural observation: in classical generative modeling, an adversarial discriminator is necessary because no tractable, semantically meaningful similarity measure is defined a priori on high-dimensional data spaces. In quantum state space, however, this deficiency does not exist. The fidelity, $F = |\langle\psi|\phi\rangle|^2$ between pure states, provides an axiomatic, efficiently computable, and unitarily invariant similarity measure that satisfies $F = 1$ if and only if the states coincide up to a global phase. This measure is accessible via the SWAP test without any classical post-processing or density estimation. Therefore, the role of the discriminator as a learned metric is structurally unnecessary for quantum state generation. Importing the adversarial paradigm from classical machine learning introduces a problem that simply does not exist in Hilbert space.

Nevertheless, the limitations of standard QGANs are systematic and worth characterizing precisely. The no-cloning theorem requires the independent preparation of two copies of the

source state: one to compare with the discriminator and one for fidelity evaluation, which brings the total qubit count to $3n + 2$ for an n-qubit target. Training instability arises because the ratio of discriminator to generator training steps per epoch (k_D/k_G) is a hyperparameter for which there is no principled choice. If the ratio is too large, the discriminator becomes too accurate too quickly, driving the generator's gradient signal to zero (vanishing gradients). If the ratio is too small, the generator can fool the discriminator without actually learning the target distribution (mode collapse). The parameter space has a dimension of 4^n for the Möttönen universal ansatz, which is used for both the generator and the discriminator. This results in 4096 parameters for $n = 6$ and circuit depths that exceed 4000 layers. This is far beyond the coherence budget of current hardware.

SQGEN replaces adversarial competition with cooperative optimization through a single synergic cost function, $J \propto \cos^2\theta \cdot \mathrm{Tr}(\sigma\rho)$, where the first factor describes the discriminator and the second describes the generator fidelity. This cost function is evaluated by executing a single circuit: the real-state path passes through the discriminator $\mathcal{D}$, a CNOT on the ancilla qubit, and then the adjoint $\mathcal{D}^\dagger$; the generated state passes through the same structure. Using the unitarity of $\mathcal{D}$ in this manner eliminates the need for independent generator and discriminator circuits, reducing the qubit overhead from $3n+2$ to $n+1$. A formal equivalence theorem establishes four conditions (C1)–(C4) under which the unique global minimum of J corresponds to perfect state generation: (C1) The Möttönen ansatz is expressive enough to represent the target. Conditions (C2)–(C4) are algebraic conditions that ensure there are no spurious local minima. For pure target states, all four conditions are satisfied by construction. For mixed targets, however, condition (C4) fails when $\rho = \mathbb{I}/2^n$, and many generator configurations achieve $J = 0$ simultaneously. Thus, the optimizer may converge to any of them. This failure mode has a precise theoretical origin: the non-uniqueness of the synergic minimum on the Nash equilibrium manifold.

The relationship between SQGEN and the adversarial game is precisely characterized: every synergic minimum with $J = 0$ is a Nash equilibrium of the quantum adversarial game, but not vice versa. At the solution $\sigma = \rho$, the Nash equilibrium set has the full dimension of the discriminator parameter space (since the discriminator is indifferent among all configurations), whereas the synergic minimum selects a unique cooperative configuration. This selection mechanism, which involves converging to a single, well-defined cooperative solution rather than wandering on the full Nash manifold, explains SQGEN's improved training stability empirically. The connection to quantum kernel methods is also precise: the generator-discriminator circuit computes Gram matrix elements for a parameterized quantum kernel. Training SQGEN is thus equivalent to generative kernel learning in a 2^n-dimensional Hilbert space. An additional connection emerges when recognizing that, in

the synergic regime, $\theta_{z_D} = 0$, SQGEN precisely reduces to a quantum autoencoder. In this case, the encoder is represented by $\mathcal{R}$ and the decoder by $\mathcal{G}$, offering an alternative approach to variational quantum error correction of coherent noise.

In simulations, SQGEN requires an average of 33 experiments per epoch for a single-qubit target compared to 150 for a comparable QGAN, a fourfold reduction. For $n = 2$ this advantage narrows to a factor of roughly two (61 versus 172 experiments), and by $n = 6$ disappears. SQGEN requires approximately 1559 seconds per epoch, compared to 1000 seconds for QGAN, because the deeper single SQGEN circuit (5385-depth) cannot be offset by the reduction in number of circuits. The significant variance in simulation times (a factor of five to ten between the best and worst runs) reflects sensitivity to initial parameters and the presence of degenerate minima in the high-dimensional landscape. When GHZ states are the target, both algorithms converge to a fidelity greater than 0.99 for $n \leq 2$ within five epochs. For $n = 3$ QGAN reaches $F \approx 0.98$ after three epochs while SQGEN reaches $F \approx 0.95$ after twenty. For $n = 4$ SQGEN shows a markedly slower convergence rate. SQGEN's advantage is conditional and applies to distributions corresponding to states with long-range entanglement. These states have a classical description that requires exponentially many parameters, but they admit an efficient quantum circuit representation. GHZ states satisfy this condition because they can be prepared in depth $\mathcal{O}(n)$. However, product states, Clifford circuits, and states satisfying the entanglement area law do not. Implementation on *ibmq_manila* demonstrated single-qubit operation at a fidelity of approximately 0.92, converging in roughly half of the trials with random initialization. However, extension to $n = 2$ failed due to the circuit depth exceeding the practical coherence threshold of approximately 90 gate layers.

Chapter 4 presents the first experimental demonstration that a quantum system can exhibit Liouvillian exceptional points in the complete absence of Hamiltonian exceptional points. This demonstration uses quantum process tomography on *ibm_nairobi* to reconstruct the full Liouvillian superoperator. At the time this work was conducted, only four experimental demonstrations of LEPs had been reported. All of these demonstrations were conducted on single-qutrit platforms that simulated qubit dissipation and relied on quantum state tomography. QST can track eigenvalue coalescence indirectly through the output state but cannot access the eigenmatrix overlap $\mathcal{O}_{12} = \langle\tilde{\sigma}_1|\tilde{\rho}_2\rangle$ that distinguishes a true exceptional point from an ordinary diabolical degeneracy, in which the eigenvalues coincide but the eigenvectors remain distinct, nor can it reconstruct the full Liouvillian superoperator. QPT closes both gaps simultaneously.

The theoretical model consisting of a single qubit with three competing dissipation channels was chosen precisely because its effective Hamiltonian is diagonal, permanently

forbidding HEPs for any $\omega \neq 0$, while the quantum jump structure of the full Liouvillian creates two exceptional points at $\gamma_x^{\pm} = \gamma_y \pm \omega$. The full derivation is given in Section 4.4.

Non-unitary dynamics are implemented on the processor via Stinespring dilation. Any completely positive, trace-preserving map can be written as a unitary on a joint system-environment Hilbert space. The environment is initialized in a fixed pure state and represented by one or two ancilla qubits. Three configurations were developed: a single-qubit direct implementation with 30–40 gate layers, a two-qubit dilation with 50–80 layers, and a three-qubit Kraus decomposition with more than 100 layers. Quantum process tomography (QPT) was performed for each value of γ_x using six input states (the eigenstates of all three Pauli operators) and six measurement bases. This yielded 36 output probability distributions, from which the 6×6 Liouvillian matrix was reconstructed via inversion of the short-time relation $\mathcal{S} \approx \mathcal{L}\, dt + \mathbb{I}$. The spectral equivalence of the three QPT representations (the 6×6 Hermitian generator set, the 4×4 Pauli basis, and the Choi-Jamiołkowski process matrix) was analytically confirmed, establishing that they all detect the same exceptional points.

The *ibm_nairobi* processor was selected for its longest coherence times and lowest error rates among available devices. Qubits #4 and #5 were used for the two-qubit configuration with direct physical coupling, eliminating SWAP routing. These qubits have coherence times of $T_1 = 155\,\mu$s and $T_2 = 20\,\mu$s, and $T_1 = 109\,\mu$s and $T_2 = 76\,\mu$s, respectively. A layered noise mitigation stack was tailored to each configuration. The single-qubit experiment used compiler level 1 and readout error mitigation at resilience level 1, while the two-qubit experiment used compiler level 3, dynamic decoupling (symmetric X–X pulse pairs in idle periods), and readout error mitigation. Dynamic decoupling reduced the imaginary part of the first non-trivial eigenvalue from 1.42 ± 0.31 (without DD) to 1.04 ± 0.18 (with DD, theoretical value 1.00). The compiler's level 3 reduced the gate count by 20–30% relative to level 1, thereby cutting circuit execution time and accumulated decoherence. While the three-qubit configuration is theoretically the most accurate, it yielded eigenvalue standard deviations $\sigma_\lambda \approx 1.2\omega$ in the spiral regime, which is large enough to obscure the qualitative structure of the spectrum. Therefore, it was not used for the main results.

The two-qubit configuration localizes both LEPs and measures the eigenmatrix overlap $\mathcal{O}_{12}$ reaching 0.95 near the first LEP and 0.92 near the second, confirming that the observed degeneracies are genuine exceptional points rather than diabolical ones. A sensitivity enhancement of approximately threefold is measured near each LEP: a change $\Delta\gamma_x = 0.1\omega$ produces $\Delta|\mathrm{Re}(\lambda_1)| \approx 0.35\omega$ at the operating point $\gamma_x/\omega \approx 0.9$, compared to 0.10ω far from the LEP. This is consistent with the square-root prediction of second-order exceptional point (EP) perturbation theory. The minimum detectable perturbation is reduced by a factor of three for the single-qubit circuit at its optimal operating point at no additional measurement

cost. However, the increase in noise enhancement near the LEP partially offsets this gain. The eigenvalue uncertainty increases by about $1.5\times$ as the parameter sweeps from the non-spiral to the spiral regime. Thus, the optimal operating point is a finite distance from the LEP.

It is worth stating explicitly the structural connections between the three chapters. QPT, or quantum process tomography, is a dynamical generalization of the multi-copy state characterization of Chapter 2. Both methods access global properties of quantum objects that are invisible to any single-shot measurement of the output alone. The circuit design strategy of Chapter 4, which translates a target dissipative process into an implementable gate sequence via Stinespring dilation and Kraus decomposition, parallels the parameterized circuit ansatz of Chapter 3. Both methods convert a target quantum object into an experimentally realized gate sequence with a controlled ancilla budget. These three contributions form a coherent progression: first, measuring the nonlinear properties of states; second, generating states with prescribed nonlinear properties; and third, characterizing and exploiting the singularities of the open-system dynamics that produce those properties.

The final chapter outlines the limitations and open problems that will guide future research. In terms of hardware, all three experiments face the same fundamental constraint: the finite ratio of gate fidelity to circuit depth. For multicopy characterization, this constraint results in a noise floor on detectable negativity. For SQGEN, it limits reliable operation to $n = 1$ with the Möttönen ansatz. For the LEP experiment, it renders the three-qubit configuration unusable. Near-term improvements include adaptive measurement protocols for the characterization layer, hardware-efficient ansatze that exploit symmetry for SQGEN, and topological verification of LEPs through closed parameter loops using mid-circuit measurement capabilities available on recent IBM Quantum generations. In the fault-tolerant era, multi-copy moments will be accessible for any order, enabling complete spectral reconstruction of the partial transpose. The Möttönen ansatz will be feasible with tens of qubits and could be extended to mixed-state targets described by Liouvillian dynamics. This will close the loop between the generation and singularity chapters. Higher-order LEPs ($N > 2$) with their sensitivity scaling $\epsilon^{1/N}$ will be accessible for sensing applications. Four concrete open problems are identified: the barren plateau conjecture for SQGEN, which requires proving that the synergic cost function avoids exponentially vanishing gradients; the completeness of the chirality witness for mixed entangled states since it is unknown whether entangled states with zero chirality of all orders exist; the experimental coexistence of HEPs and LEPs in a single system verified by QPT; and the optimal operating point near higher-order LEPs in the presence of a hardware noise floor.

This dissertation demonstrates that the nonlinear properties of quantum systems, which are defined as properties that require nonlinear functions of the state or nonlinear dynamics for

description, can be measured, prepared, and controlled on current superconducting hardware. The three-part framework of characterization, generation, and singularity exploitation is modular and can be expanded as hardware improves. The open problems identified here establish a research agenda that links short-term experimental work on NISQ processors to the long-term objective of creating fully programmable quantum systems.

Streszczenie

Mechanika kwantowa jest zasadniczo teorią liniową. Ewolucja stanów kwantowych, opisywana równaniem Schrödingera, podlega zasadzie superpozycji, a obserwable działają liniowo na przestrzeni Hilberta. Jednak istotne w kwantowym przetwarzaniu informacji wielkości są niemal zawsze nieliniowymi funkcjonałami stanu. Przykładowo splątanie nie może być w pełni charakteryzowane przez wartość oczekiwaną jakiejkolwiek obserwabli. Dla takich miar jak negatywność potrzebne są informacje o spektrum częściowo transponowanej macierzy gęstości, co czyni je z natury nieliniowymi. Sam pomiar jest nieodwracalną, stochastyczną transformacją odwzorowującą macierz gęstości na rozkład prawdopodobieństwa w sposób zbyt złożony, aby można było go opisać jako liniową obserwablę. W przeciwieństwie do unitarnej dynamiki układów izolowanych, dysypatywna dynamika rządząca rzeczywistym sprzętem kwantowym wytwarza osobliwości w przestrzeni parametrów, gdzie struktura spektralna operatora ewolucji zmienia się jakościowo.

Te trzy źródła nieliniowości motywują centralne pytania niniejszej rozprawy. Na obecnie dostępnych nadprzewodzących procesorach NISQ niekontrolowana dekoherencja niszczy splątanie na skali czasowej porównywalnej z czasem trwania bramek, liczba kubitów pozostaje ograniczona, a głębokość obwodów jest związana skończonym czasem koherencji. Standardowa tomografia stanu kwantowego wymaga liczby ustawień pomiarowych rosnącej wykładniczo z rozmiarem układu i szybko staje się niepraktyczna, nawet dla niewielkich rejestrów kubitowych. Tworzenie stanów ze specyficznym rodzajem splątania wymaga eksploracji wysoce niewypukłego krajobrazu wariacyjnego, w którym gradienty mogą zanikać, a lokalne minima występują wyjątkowo gęsto. Osobliwe zachowanie otwartych układów kwantowych w pobliżu punktów wyjątkowych było dotąd badane wyłącznie w ramach efektywnych modeli niehermitowskich i nie zostało zweryfikowane eksperymentalnie na gruncie pełnego opisu Lindblada. Trzy eksperymentalne wkłady niniejszej rozprawy kolejno podejmują te wyzwania, przechodząc od charakterystyki właściwości nieliniowych, poprzez ich skuteczne generowanie, aż po celowe wykorzystanie osobliwości układów otwartych. Wszystkie eksperymenty zostały przeprowadzone na nadprzewodzących procesorach IBM Quantum.

W pierwszym rozdziale przedstawiono podstawy teoretyczne wymagane przez kolejne rozdziały. Omówiono w nim postulaty mechaniki kwantowej w formalizmie macierzy gęstości, wprowadzając pojęcie śladu częściowego, model sfery Blocha dla kubitów oraz rozwinięcie Pauliego stanów wielokubitowych. Jako dwa główne narzędzia wykrywania splątania, wykraczające poza wartości oczekiwane dla układów jednokopiowych, przedstawiono kryterium dodatniej transpozycji częściowej Peresa–Horodeckiego oraz kryterium realignacji. Wprowadzono niezmienniki Makhlina, które klasyfikują stany dwukubitowe w sposób pełny z dokładnością do lokalnej równoważności unitarnej, wraz z ich reprezentacją jako projekcji singletowych na układach wielokopiowych. Stanowi to algebraiczny fundament metod charakteryzacji rozwinięty w rozdziale 2. Następnie rozwinięto teorię otwartych układów kwantowych poprzez przybliżenie Markowa równania głównego Goriniego– Kossakowskiego– Sudarshana– Lindblada, traktując superoperator Liouvilla w jego postaci wektorowej i szczegółowo analizując jego problem wartości własnych. Model trajektorii kwantowej rozkłada ewolucję Lindblada na ciągłą nieunitarną ewolucję w ramach efektywnego hamiltonianu niehermitycznego oraz dyskretne stochastyczne skoki kwantowe, co pozwala precyzyjnie rozróżnić punkty wyjątkowe hamiltonianu (HEP) od punktów wyjątkowych Liouvillianu (LEP). Rozróżnienie to ma kluczowe znaczenie w rozdziale 4. HEP są degeneracjami samego efektywnego hamiltonianu i należą do reżimu półklasycznego, podczas gdy LEP są degeneracjami pełnego superoperatora Liouvillego i zależą zasadniczo od składników skoków kwantowych nieposiadających klasycznego odpowiednika. Dlatego każdy HEP indukuje odpowiedni LEP, lecz nie odwrotnie. Rozdział zamykają twierdzenia zakazu w informacji kwantowej: zakaz klonowania, usuwania, transmisji i sygnalizowania, których operacyjne konsekwencje kształtują strategie eksperymentalne w całej rozprawie.

Rozdział drugi podejmuje problem charakteryzacji korelacji kwantowych na sprzęcie NISQ bez pełnej rekonstrukcji stanu. Opracowane zostają trzy komplementarne podejścia. Pierwsze to protokół wielokopijnego szacowania (MCE) łączący pomiary projekcji singletowych ze sztucznymi sieciami neuronowymi. Kluczowa obserwacja polega na tym, że niezmienniki Makhlina, z których można odtworzyć zarówno negatywność, jak i miarę nielokalności Bella, dają się wyrazić jako wartości oczekiwane projekcji na stan singletowy mierzonych jednocześnie na dwóch kopiach stanu. W ten sposób wymagana liczba ustawień pomiarowych redukuje się z rzędu 4^n do rzędu 2^n. Wytrenowana na losowo wygenerowanych stanach dwukubitowych sieć neuronowa, uczy się odwzorowywać te projekcje na miary splątania, bez rozwiązywania związanych równań wielomianowych. Metoda ta osiąga znacznie lepszą odporność na szum niż bezpośrednia inwersja analityczna. Przewaga nad podejściem analitycznym ma charakter strukturalny: bezpośrednia inwersja problemu znajdowania pierwiastków wielomianu wzmacnia szum poprzez wyrażenia nieliniowe, podczas gdy sieć

uczy się płynnych odwzorowań, które są mniej wrażliwe na niewielkie zaburzenia danych wejściowych. Systematyczna analiza SHAP (Shapley Additive Explanations) ważności cech ujawnia, że tylko pięć spośród dwunastu konfiguracji rzutowań niesie większość sygnału predykcyjnego, redukując nakład eksperymentalny o 67%. W pięciu konfiguracjach istotnych dla negatywności dominują projekcje pętlowe i krzyżowe, które oddają lokalną koherencję oraz korelacje krzyżowe, natomiast konfiguracje istotne dla nielokalności Bella odzwierciedlają lokalną statystykę pomiarową operatora CHSH. Protokół zaimplementowany na *ibm_hanoi*, z uśrednianiem po wielu konfiguracjach mapowania kubitów, osiąga efektywną wierność kopii powyżej 98%.

Drugie podejście wprowadza drzewa decyzyjne oparte na bazach Gröbnera do adaptacyjnej rekonstrukcji negatywności na podstawie wartości własnych częściowo transponowanej macierzy gęstości, które z kolei są określane przez momenty spektralne $\mu_k = \mathrm{Tr}[(\rho^{T_A})^k]$, poprzez tożsamości Newtona-Girarda. Każdy moment jest mierzony jednym obwodem testu Hadamarda, działającym na wielu kopiach stanu z kontrolowaną permutacją cykliczną. Istnieje kluczowa tożsamość algebraiczna o uniwersalnym charakterze: moment częściowej transpozycji drugiego rzędu jest równy czystości, $\mu_2 = \mathcal{I}_2 = \mathrm{Tr}[\rho^2]$, niezależnie od wymiaru układu, co wynika z tożsamości $\sigma = \sigma^{-1} = S_{12}$ dla dwóch kopii. Pojedynczy obwód czystości mierzy zatem oba niezmienniki jednocześnie, zmniejszając liczbę obwodów o jeden. Struktura drzewa decyzyjnego wynika z obserwacji, że większość stanów dwukubitowych należy do klas degeneracji wartości własnych identyfikowalnych z pierwszych dwóch momentów za pomocą wielomianów baz Gröbnera $\mathcal{G}_1$ i $\mathcal{G}_2$. Degeneracja potrójna plus, w której trzy wartości własne są zbieżne (obejmująca stany Bella, stany Wernera i stan maksymalnie mieszany), wymaga tylko jednego obwodu; degeneracja dwuparowa wymaga dwóch; natomiast jedynie przypadek ogólny z czterema różnymi wartościami własnymi wymaga pełnego zestawu trzech momentów $\{\mu_2, \mu_3, \mu_4\}$. Próbkowanie metodą Monte Carlo dla 10^5 losowanych z miary Haara stanów dwukubitowych potwierdza, że $95,3\%$ z nich należy do gałęzi ogólnej, a $4,7\%$ do gałęzi zdegenerowanych. Średnio wystarcza $2{,}15$ konfiguracji obwodów, co odpowiada $5{,}3$-krotnej redukcji ustawień pomiarowych w porównaniu z pełną tomografią dwukubitową oraz 7,2-krotnej redukcji dla układów kubit-kutrit. W warunkach szumu depolaryzującego węzły decyzyjne stosują tolerancję $|G_i| < 3\sigma_G$ zamiast dokładnego zera. Błędna klasyfikacja kieruje stan do niewłaściwej gałęzi, ale nigdy nie powoduje błędu w ostatecznej negatywności, ponieważ ogólny wynik równania czwartego rzędu jest zawsze poprawny jako rozwiązanie awaryjne. Dokładność klasyfikacji powyżej 95% jest utrzymywana dla szumu depolaryzacyjnego około $p \approx 0{,}08$, co obejmuje typowy zakres działania obecnych procesorów transmonowych.

Trzecia metoda identyfikuje fizyczną interpretację korekt momentów częściowej transpozycji $C_k = \mu_k - \mathcal{I}_k$ i wykorzystuje ją do rozszerzenia formalizmu wykrywania splątania związanego. Głównym wynikiem algebraicznym jest obserwacja, że różnica ta rozkłada się dokładnie jako korelator chiralności-chiralności: skalarny spin chiralności, operator znany w fizyce materii skondensowanej jako parametr porządku chiralnych cieczy spinowych oraz jako czynnik wywołujący topologiczny efekt Halla, pojawia się tutaj dla kubitów pochodzących z różnych kopii stanu, a nie z różnych węzłów sieci. Istotny wynik czwartego rzędu ma postać $C_4 = 8\,\mathrm{Tr}[\Omega_A \Omega_B \rho^{\otimes 4}]$, gdzie $\Omega = \frac{1}{2}\sum_{i<j<k} \chi_{ijk}$, wyprowadzony z rozkładu operatora SWAP i tożsamości komutatorów Pauliego. Związek ten nie jest jedynie analogią: ta sama struktura algebraiczna rządzi obydwoma zjawiskami, a interpretacja geometryczna jest identyczna: stan produktowy generuje konfiguracje spinowe o wielu kopiach, które pozostają współpłaszczyznowe między kopiami, dając zerową chiralność, podczas gdy stan splątany narusza tę współpłaszczyznowość jednocześnie w obydwu podukładach, wytwarzając korelator niezerowy. W drugim rzędzie $C_2 = 0$ dla wszystkich przypadków, co wynika z faktu, że $\sigma = \sigma^{-1} = S_{12}$ dla dwóch kopii, a zatem operator antyhermitowski $\Delta = \sigma^{-1} - \sigma$ znika identycznie. W trzecim rzędzie poprawka C_3 jest proporcjonalna do komutatora dwóch operatorów SWAP o tym samym indeksie kopii, co jest proporcjonalne do skalarnej chiralności spinowej trzech kopii. W czwartym rzędzie wszystkie trzykrotności chiralności czterech kopii mają równy udział. Dla stanów czystych dwukubitowych poprawka czwartego rzędu spełnia warunek $C_4 = -4\mathcal{N}^2(1-\mathcal{N}^2)$, co czyni chiralność kompletnym świadkiem splątania, a dla stanów diagonalnych Bella $C_4 = \frac{3}{4}\det(T)$, gdzie T jest tensorem korelacji, co nadaje chiralności geometryczne znaczenie jako objętość ze znakiem równoległościanu korelacji. Rozdzielne granice $|C_3| \le 1/36$ i $|C_4| \le 1/27$ są osiągane przez mieszaniny rangi 3 o równych wagach trzech stanów produktowych we wzajemnie nieobciążonych bazach i są ścisłe. Chiralność wielokopiowa jest wielkością z natury zbiorową: dla znanego stanu można ją obliczyć na podstawie macierzy gęstości, ale żaden pomiar jednokopiowy na nieznanym stanie nie pozwala jej wyodrębnić bez pełnej tomografii stanu.

Ta sama infrastruktura obwodów z kontrolowanymi bramkami SWAP, zreinterpretowana przez mapę realokacji $R(\rho)$, umożliwia dostęp do luki niehermitowskości $D_k = \Sigma_k - G_k$ między momentami wartości osobliwych Σ_k a momentami wartości własnych G_k realokowanej macierzy gęstości. Luka ta ma bezpośrednią interpretację spinowo-fizyczną jako część antysymetryczna tensora korelacji spinowej A–B, związana z iloczynem wektorowym $\langle \mathbf{J}_A \times \mathbf{J}_B \rangle$, i jest strukturalnie czuła na splątanie niewidoczne dla kryterium PPT. Dla stanów o rzeczywistych elementach macierzowych różnica w restrukturyzacji krzyżowej o nieparzystym rzędzie $\delta G_k = G_k - G_k^{PT}$ stanowi sygnał uzupełniający; algebraiczny powód, dla którego stany chessboard unikają tego świadka, wynika z ich struktury parzystości szachownicy, która

wymusza $R_{-} = 0$, a zatem $\delta G_k = 0$ dla wszystkich przypadków, pomimo rzeczywistego splątania. Klasyfikator regresji logistycznej Lasso wytrenowany na wielomianowych kombinacjach stopnia 2 sześciu obserwabli spektralnych $\{\Sigma_1, G_1, \Sigma_2, G_2, D_1, D_2\}$ osiąga 94% wielokrotnie walidowanego wykrywania splątania związanego przy zerowej liczbie wyników fałszywie pozytywnych, a zespół ExtraTrees na 522 nieliniowych cechach podnosi ten wynik do $99{,}6\%$. Obydwa klasyfikatory obejmują wszystkie trzy znane rodziny: Horodeckiego, Tiles UPB i Chessboard, splątanych stanów PPT wymiaru 3×3 w pełnych zakresach parametrów. Praca identyfikuje też nową klasę stanów niewidocznych dla kryterium CCNR: mieszaniny z szumem marginalnym $\rho_{\mathrm{mn}}(t) = (1-t)\rho_0 + t\,\rho_A \otimes \rho_B$, dla których $\|\rho^R\|_1 < 1$ dla wszystkich $t \gtrsim 0{,}01$, podczas gdy $D_2 > 0$ utrzymuje się, umożliwiając detekcję z tolerancją na szum 10–20-krotnie większą niż próg CCNR. Fizyczny mechanizm polega na tym, że mieszanie z iloczynem marginalnym tłumi dominującą wartość szczególną macierzy realignacji, która jest odpowiedzialna za $\Sigma_1 > 1$, bez eliminowania asymetrii spektralnej $D_2 > 0$.

Wszystkie wyniki eksperymentalne zostały uzyskane na procesorach *ibm_kingston*, *ibm_torino* i *ibm_fez*, wszystkich z architekturą Heron i zestawem bramek natywnych $\{\mathrm{CZ}, \sqrt{X}, R_Z, X\}$, mitygacją błędów odczytu i transpilacją Qiskit na poziomie optymalizacji 3. Średnie błędy negatywności wynoszą $0{,}002$ (Kingston, 2×2), $0{,}012$ (Torino, 2×2) i $0{,}027$ (Torino, 2×3). Średnie błędy chiralności wynoszą $0{,}022$ (2×2) i $0{,}039$ (2×3), przy czym większe rozrzuty w przypadku kubit-kutrit wynikają z około dwukrotnie większej głębokości obwodów czterokopijnych w tej wymiarowości. Krzywa analityczna $-C_4 = 4\mathcal{N}^2(1-\mathcal{N}^2)$ jest potwierdzana bezpośrednio z danych sprzętowych w pełnym zakresie splątania, w tym dla mieszanin Wernera $\rho_W(p) = p|\Psi^-\rangle\langle\Psi^-| + (1-p)\mathbb{I}/4$, gdzie teoretyczna krzywa $-C_4 = \frac{3}{4}p^3$ jest weryfikowana z detekcją chiralności powyżej granicy separowalnej dla $p \gtrsim 0{,}37$. Wszystkie pięć testowanych stanów splątanych związanych na *ibm_fez* zostało wykrytych z dodatnim D_2, podczas gdy obydwa separowalne stany kontrolne dają D_2 zgodne z zerem.

Rozdział trzeci omawia problem efektywnego przygotowywania stanów kwantowych o określonej strukturze korelacji. Synergiczna Kwantowa Sieć Generatywna (SQGEN) została wprowadzona jako alternatywa dla paradygmatu Kwantowej Generatywnej Sieci Przeciwstawnej (QGAN). Motywacja do przyjęcia zasadniczo odmiennego podejścia wynika ze strukturalnej obserwacji: w klasycznym uczeniu generatywnym dyskryminator jest konieczny, ponieważ nie istnieje żadna łatwa do obliczenia, semantycznie znacząca miara podobieństwa zdefiniowana a priori w wielowymiarowych przestrzeniach danych. W przestrzeni stanów kwantowych ta luka nie istnieje. Wierność $F = |\langle\psi|\phi\rangle|^2$ między czystymi stanami dostarcza aksjomatycznej, efektywnie obliczalnej i unitarnie niezmienniczej miary podobieństwa, która spełnia warunek $F = 1$ wtedy i tylko wtedy, gdy stany pokrywają

się z dokładnością do fazy globalnej. Jest ona dostępna poprzez test SWAP bez żadnego klasycznego przetwarzania końcowego. Dlatego rola dyskryminatora czyli wyuczonej metryki jest z definicji niepotrzebna przy generacji stanów kwantowych.

Ograniczenia standardowych sieci QGAN mają charakter systematyczny i warto je dokładnie scharakteryzować. Zakaz klonowania wymaga niezależnego przygotowania dwóch kopii stanu źródłowego, jednej do porównania z dyskryminatorem, drugiej do oceny wierności, co łącznie daje $3n + 2$ kubitów dla n-kubitowego stanu docelowego. Niestabilność uczenia wynika z faktu, że stosunek kroków treningowych dyskryminatora do generatora (k_D/k_G) na epokę jest hiperparametrem, którego wyboru nie można opierać na zasadach: zbyt duża wartość sprawia, że dyskryminator zbyt szybko staje się zbyt dokładny, zerując sygnał gradientu generatora; zbyt mała wartość sprawia, że generator może oszukać dyskryminator, nie ucząc się faktycznie rozkładu docelowego. Przestrzeń parametrów ma wymiar 4^n dla uniwersalnego ansatzu Möttönena, osiągając 4096 parametrów dla $n = 6$ i głębokości obwodów powyżej 4000 warstw, co znacznie wykracza poza zasoby koherencji obecnego sprzętu.

SQGEN zastępuje rywalizację opartą na przeciwnikach optymalizacją opartą na współpracy poprzez zastosowanie jednej synergicznej funkcji kosztu $J \propto \cos^2\theta \cdot \mathrm{Tr}(\sigma\rho)$. Funkcja ta jest obliczana przez wykonanie pojedynczego obwodu: ścieżka stanu rzeczywistego przechodzi przez dyskryminator $\mathcal{D}$, bramkę CNOT na kubicie pomocniczym, a następnie przez sprzężenie $\mathcal{D}^\dagger$; wygenerowany stan przechodzi przez tę samą strukturę. Wykorzystanie unitarności $\mathcal{D}$ eliminuje potrzebę stosowania niezależnych obwodów generatora i dyskryminatora, zmniejszając obciążenie kubitowe z $3n+2$ do $n+1$. Twierdzenie o równoważności ustala cztery warunki (C1)–(C4), przy których spełnieniu jedyne globalne minimum J odpowiada idealnemu wygenerowaniu stanu. Warunek (C1) wymaga, by ansatz Möttönena był wystarczająco ekspresyjny, aby reprezentować cel; (C2)–(C4) są warunkami algebraicznymi wykluczającymi fałszywe lokalne minima. Dla czystych stanów docelowych wszystkie cztery warunki są spełnione z założenia, podczas gdy dla stanów mieszanych warunek (C4) nie jest spełniony. Dla $\rho = \mathbb{I}/2^n$ wiele konfiguracji generatora jednocześnie osiąga $J = 0$, a optymalizator może zbiegać się do dowolnej z nich.

Związek między SQGEN a grą przeciwstawną jest precyzyjnie scharakteryzowany: każde minimum synergiczne przy $J = 0$ jest równowagą Nasha kwantowej gry przeciwstawnej, lecz nie odwrotnie. W punkcie rozwiązania $\sigma = \rho$ zbiór równowag Nasha ma pełny wymiar przestrzeni parametrów dyskryminatora (ponieważ dyskryminator jest obojętny wobec wszystkich konfiguracji), podczas gdy minimum synergiczne wybiera unikalną konfigurację kooperacyjną. Ten mechanizm selekcji, a wiec zbieżność do jednego dobrze zdefiniowanego rozwiązania kooperatywnego zamiast błądzenia po całej rozmaitości Nasha, empirycznie

wyjaśnia poprawioną stabilność uczenia SQGEN. Dodatkowy związek wynika z uznania, że w reżimie synergicznym $\theta_{z_D} = 0$ SQGEN sprowadza się dokładnie do kwantowego autoenkodera, z $\mathcal{R}$ jako enkoderem i $\mathcal{G}$ jako dekoderem, co otwiera niezależną drogę do wariacyjnej kwantowej korekcji błędów koherentnych szumów.

W symulacji algorytm SQGEN wymaga średnio 33 eksperymentów na epokę dla celu obejmującego jeden kubit, wobec 150 dla porównywalnej sieci QGAN, co stanowi czterokrotną redukcję. Dla $n = 2$ przewaga spada do około dwukrotności (61 wobec 172 eksperymentów), a przy $n = 6$ zanika: SQGEN wymaga około 1559 sekund na epokę wobec 1000 dla QGAN, ponieważ głębszy pojedynczy obwód SQGEN (głębokość 5385) nie daje się zrównoważyć przez redukcję liczby obwodów. Znaczna zmienność czasów symulacji (czynnik 5–10 między najlepszym a najgorszym przebiegiem) odzwierciedla wrażliwość na parametry startowe i obecność zdegenerowanych minimów w przestrzeni parametrów wysokiego wymiaru. W przypadku stanów GHZ jako celu obydwa algorytmy zbiegają do wierności powyżej $0{,}99$ dla $n \leq 2$ w ciągu pięciu epok; dla $n = 3$ QGAN osiąga $F \approx 0{,}98$ po trzech epokach, a SQGEN $F \approx 0{,}95$ po dwudziestu; dla $n = 4$ SQGEN wykazuje znacznie wolniejszą zbieżność. Przewaga SQGEN jest warunkowa: dotyczy ona rozkładów odpowiadających stanom o splątaniu dalekiego zasięgu, których opis klasyczny wymaga wykładniczo wielu parametrów, ale które dopuszczają wydajną reprezentację obwodu kwantowego. Implementacja na *ibmq_manila* potwierdziła działanie na pojedynczym kubicie z wiernością około $0{,}92$, zbiegając w mniej więcej połowie prób z losową inicjalizacją, podczas gdy rozszerzenie na $n = 2$ zawiodło z powodu przekroczenia praktycznego progu koherencji przez głębokość obwodu wynoszącą około 90 warstw bramek.

W rozdziale 4 przedstawiono pierwsze eksperymentalne dowody na to, że układ kwantowy może wykazywać punkty wyjątkowe Liouville'a przy całkowitym braku punktów wyjątkowych Hamiltonianu, wykorzystując tomografię procesów kwantowych na platformie *ibm_nairobi* do odtworzenia pełnego superoperatora Liouville'a. Do tej pory odnotowano jedynie cztery eksperymentalne demonstracje LEP, wszystkie wykorzystujące platformy jedno-kutritowe, symulujące dysypację kubitu. Wszystkie te demonstracje oparte były na tomografii stanu kwantowego (QST). QST może śledzić koalescencję wartości własnych pośrednio poprzez stan wyjściowy, ale nie ma dostępu do nakładania się macierzy własnych $\mathcal{O}_{12} = \langle\tilde{\sigma}_1|\tilde{\rho}_2\rangle$, co jednoznacznie odróżnia prawdziwy punkt wyjątkowy od zwykłej degeneracji diabolicznej, gdzie wartości własne pokrywają się, ale wektory własne pozostają odrębne, ani nie może zrekonstruować pełnego superoperatora Liouville'a. Tomografia procesu kwantowego (QPT) wypełnia obie luki jednocześnie.

Badany model to pojedynczy kubit o Hamiltonianie $\hat{H} = \frac{\omega}{2}\hat{\sigma}_z$ oraz trzech konkurujących ze sobą kanałach dyssypacji o częstościach γ_x, γ_y, γ_-. Efektywny niehermitowski hamiltonian

jest diagonalny w bazie energii własnych, a jego wartości własne mają części rzeczywiste trwale odseparowane o ω dla dowolnego niezerowego ω. W związku z tym w przestrzeni parametrów nie występują żadne punkty wyjątkowe HEP. Pełny superoperator Liouville'a tworzy jednak dwa punkty wyjątkowe drugiego rzędu w $\gamma_x^{\pm} = \gamma_y \pm \omega$ poprzez człon sprzężenia skoku kwantowego $\sum_\mu \gamma_\mu \hat{\sigma}_\mu \otimes \hat{\sigma}_\mu^*$, który nie posiada odpowiednika w efektywnym hamiltonianie. Przyjmując $\gamma_- = 0$ i $\gamma_y = 2\omega$, dwa LEP leżą teoretycznie w $\gamma_x/\omega = 1$ i $\gamma_x/\omega = 3$. Pomiędzy tymi wartościami wartości własne Liouville'a tworzą sprzężoną parę zespoloną, a trajektorie wektorów Blocha rysują zanikające spirale. Poza tym oknem wartości własne są rzeczywiste i zanik jest czysto wykładniczy. LEP jest formalnie osobliwością punktu rozgałęzienia funkcji własnych: gdy parametr sterujący opisuje zamkniętą pętlę otaczającą LEP, dwie wartości własne ulegają permutacji, zamiast powracać do wartości początkowych, co jest cechą monodromii nieobecną w degeneracjach diabolicznych.

Dynamika nieunitarna jest realizowana na procesorze za pomocą rozszerzenia Stinespringa: dowolne całkowicie dodatnie odwzorowanie zachowujące ślad można zapisać jako odwzorowanie unitarne w wspólnej przestrzeni Hilberta układu i otoczenia, przy czym otoczenie jest inicjowane w ustalonym stanie czystym i reprezentowane przez jeden lub dwa kubity pomocnicze.

Opracowane zostały trzy konfiguracje: bezpośrednia implementacja z jednym kubitem (głębokość 30–40 warstw bramek), rozszerzenie z dwoma kubitami (głębokość 50–80 warstw) oraz dekompozycję Krausa z trzema kubitami (głębokość powyżej 100 warstw). QPT przeprowadzono dla każdej wartości γ_x używając sześciu stanów wejściowych (stanów własnych wszystkich trzech operatorów Pauliego) i sześciu baz pomiarowych, co dało 36 rozkładów prawdopodobieństwa wyjściowego, z których rekonstruowano macierz 6×6 superoperatora Liouville'a poprzez odwrócenie relacji krótkotrwałej $\mathcal{S} \approx \mathcal{L}\, dt + \mathbb{I}$. Spektralna równoważność wszystkich trzech reprezentacji QPT: zestaw generatorów hermitiańskich 6×6, baza Pauliego 4×4 i macierz procesu Choi-Jamiołkowskiego, została ustalona analitycznie, potwierdzając, że wszystkie trzy wykrywają te same punkty wyjątkowe.

Procesor *ibm_nairobi* został wybrany ze względu na najdłuższe czasy koherencji T_2 i najniższe błędy bramkowe spośród dostępnych urządzeń; w konfiguracji dwubitowej zastosowano kubity nr #4 ($T_1 = 155\,\mu$s, $T_2 = 20\,\mu$s) i nr #5 ($T_1 = 109\,\mu$s, $T_2 = 76\,\mu$s) z bezpośrednim sprzężeniem fizycznym, co pozwoliło wyeliminować routing SWAP. Dla każdej konfiguracji dostosowano wielowarstwowy zestaw środków ograniczających zakłócenia. W eksperymencie z jednym kubitem zastosowano poziom kompilatora 1 oraz ograniczanie błędów odczytu na poziomie odporności 1, podczas gdy w eksperymencie z dwoma kubitami zastosowano poziom kompilatora 3, dynamiczne odsprzęganie (symetryczne pary impulsów X–X w okresach bezczynności) oraz ograniczanie błędów odczytu. Dynamiczne

odsprzęganie zmniejszyło część urojona pierwszej niebanalnej wartości własnej z $1{,}42 \pm 0{,}31$ (bez DD) do $1{,}04 \pm 0{,}18$ (z DD, wartość teoretyczna $1{,}00$). Poziom kompilatora 3 zmniejszył liczbę bramek o 20–30% w stosunku do poziomu 1, skracając czas wykonania obwodów i zmniejszając akumulację dekoherencji. Konfiguracja trzech kubitów, choć teoretycznie najdokładniejsza, dała odchylenia standardowe wartości własnych wynoszące $\sigma_\lambda \approx 1{,}2\omega$ w reżimie spiralnym, a więc wystarczająco duże, aby zasłonić jakościową strukturę widma.

Konfiguracja dwukubitowa lokalizuje oba punkty LEP i mierzy nakładanie się macierzy własnych $\mathcal{O}_{12}$ osiągając wartość $0{,}95$ w pobliżu pierwszego LEP oraz $0{,}92$ w pobliżu drugiego, potwierdzając, że obserwowane degeneracje są prawdziwymi punktami wyjątkowymi, a nie diabolicznymi. W pobliżu każdego punktu LEP zmierzono około trzykrotny wzrost czułości: zmiana $\Delta\gamma_x = 0{,}1\omega$ daje $\Delta|\mathrm{Re}(\lambda_1)| \approx 0{,}35\omega$ w punkcie pracy $\gamma_x/\omega \approx 0{,}9$, w porównaniu z $0{,}10\omega$ daleko od LEP, co jest zgodne z przewidywaniem pierwiastka kwadratowego z teorii perturbacji EP drugiego rzędu. Minimalna wykrywalna perturbacja jest odpowiednio zmniejszona trzykrotnie dla obwodu jednokubitowego w jego optymalnym punkcie pracy, bez dodatkowych zasobów pomiarowych. Wzmocnienie szumu w pobliżu LEP, co powoduje, że niepewność wartości własnych rośnie o czynnik około $1{,}5$, gdy parametr jest zmieniany z reżimu niespiralnego do spiralnego, częściowo kompensuje ten zysk, a optymalny punkt pracy leży w skończonej odległości od EP, a nie w nim.

Warto wyraźnie podkreślić powiązania strukturalne między tymi trzema rozdziałami. Rekonstrukcja Liouvilliana za pomocą QPT stanowi dynamiczne uogólnienie charakterystyki stanu wielokrotnego z rozdziału 2: obie metody pozwalają uzyskać dostęp do globalnych właściwości obiektów kwantowych, niewidocznych w przypadku pojedynczego pomiaru samego sygnału wyjściowego. Strategia projektowania obwodów z rozdziału 4, polegająca na przekształceniu docelowego procesu dyssypatywnego w sekwencję bramek, którą można zaimplementować za pomocą rozszerzenia Stinespringa i dekompozycji Krausa, jest analogiczna do podejścia opartego na obwodach parametryzowanych z rozdziału 3: obie metody przekształcają docelowy obiekt kwantowy w sekwencję bramek zrealizowaną eksperymentalnie przy kontrolowanym budżecie kubitów pomocniczych. Te trzy wkłady tworzą progresję od pomiaru nieliniowych właściwości stanów, poprzez generowanie stanów o określonych właściwościach nieliniowych, aż po charakteryzowanie i wykorzystywanie osobliwości dynamiki układu otwartego, która te właściwości wytwarza.

W ostatnim rozdziale wskazano ograniczenia i nierozwiązane problemy, które wyznaczają kierunek dalszych badań. Jeśli chodzi o sprzęt, we wszystkich trzech eksperymentach występuje to samo zasadnicze ograniczenie: skończony stosunek wierności bramek do głębokości obwodu. W przypadku charakterystyki wielokopiowej przejawia się to jako poziom szumu w wykrywalnej negatywności; w przypadku SQGEN ogranicza to niezawodne działanie

do $n = 1$ przy zastosowaniu ansatzu Möttönena; w przypadku eksperymentu LEP sprawia to, że konfiguracja trzech kubitów staje się bezużyteczna. Krótkoterminowe ulepszenia obejmują adaptacyjne protokoły pomiarowe dla warstwy charakterystyki, efektywne sprzętowo podejścia wykorzystujące symetrię dla SQGEN oraz topologiczną weryfikację LEP poprzez zamknięte pętle parametrów z wykorzystaniem możliwości pomiarów w środkowej części obwodu dostępnych w najnowszych generacjach IBM Quantum. W erze odporności na błędy momenty wielu kopii stają się osiągalne dla dowolnego rzędu, umożliwiając pełną rekonstrukcję spektralną transpozycji częściowej; podejście Möttönena staje się wykonalne przy dziesiątkach kubitów i może zostać rozszerzone na cele o stanach mieszanych opisane dynamiką Liouvillian, zamykając pętlę między rozdziałami dotyczącymi generowania i osobliwości; a LEP wyższego rzędu ze skalowaniem czułości $\epsilon^{1/N}$ stają się dostępne dla zastosowań sensorycznych. Zidentyfikowano cztery konkretne otwarte problemy: hipoteza barren plateau dla SQGEN, która wymagałaby udowodnienia, że synergiczna funkcja kosztu unika gradientów zanikających wykładniczo; zupełność świadka chiralności dla mieszanych stanów splątanych, ponieważ nie wiadomo, czy istnieją stany splątane o zerowej chiralności wszystkich rzędów; eksperymentalne współistnienie HEP i LEP w jednym układzie zweryfikowane przez QPT; oraz optymalny punkt pracy w pobliżu LEP wyższego rzędu w obecności szumu sprzętu.

Wyniki niniejszej rozprawy pokazują, że właściwości nieliniowe układów kwantowych, rozumiane jako właściwości wymagające do opisu funkcji nieliniowych stanu lub dynamiki nieliniowej, są jednocześnie mierzalne, możliwe do przygotowania i sterowalne na obecnym sprzęcie nadprzewodnikowym. Trzyczęściowa struktura obejmująca charakterystykę, generowanie i wykorzystanie osobliwości zapewnia ramową strukturę, którą można rozszerzać wraz z udoskonalaniem sprzętu, a zidentyfikowane tu otwarte problemy definiują program badawczy łączący krótkoterminowe prace eksperymentalne nad procesorami NISQ z długoterminowym celem, jakim są w pełni programowalne układy kwantowe.

Chapter 6

Statement of Contributions

This dissertation is based on four research contributions, each resulting from collaborative work. The following statement delineates the candidate's personal contribution to each project.

Chapter 2, Part I — Resource-efficient quantum correlation measurements via multicopy neural network methods

Published as: P. Tulewicz, K. Bartkiewicz, A. Miranowicz, F. Nori, *Resource-efficient quantum correlation measurements via multicopy neural network methods*, Scientific Reports **15**, 40868 (2025).

The Ph.D. student's contributions were crucial to the project's experimental and computational implementation. She developed a comprehensive experimental protocol that included selecting measurement bases, designing singlet projection circuits, and creating a qubit mapping strategy tailored to the topology and error characteristics of the *ibm_hanoi* processor. The candidate implemented and ran all quantum circuits on IBM Quantum hardware and was also responsible for managing calibration data and applying error mitigation processes, such as noise-free extrapolation, Bayesian readout correction, and posteriori selection based on purity. She developed an artificial neural network model, including its architecture, a training procedure on randomly selected two-qubit states, and a SHAP-based analysis that identified five optimal measurement configurations. The doctoral student performed all numerical calculations, simulations, and statistical analyses of shot noise scaling.

Chapter 2, Part II: Spin chirality across quantum state copies detects hidden entanglement

Submitted as: P. Tulewicz, K. Bartkiewicz, F. Nori, *Spin chirality across quantum state copies detects hidden entanglement* (in preparation).

The doctoral candidate contributed to this work in both theoretical and experimental ways. Theoretically, the candidate developed a comprehensive framework that links partial transposition moments with multiple chirality correlators. The candidate also established Newton–Girard relations, which transform moment differences into conditions for a Gröbner basis decision tree. They implemented all the quantum circuits necessary to evaluate chirality witnesses and non-Hermitian gaps. The candidate performed a significant portion of the experimental measurements on IBM Quantum processors, covering the Werner, Horodecki, Tiles, and Chessboard state families across their full parameter ranges.

Chapter 3: Synergic quantum generative machine learning

Published as: K. Bartkiewicz, P. Tulewicz, J. Roik, K. Lemr, *Synergic quantum generative machine learning*, Scientific Reports **13**, 12893 (2023).

The candidate contributed to the experimental implementation and theoretical development of the SQGEN structure. For the experimental work, the candidate co-developed and prepared numerical simulations and hardware experiments. This included implementing circuits based on the Möttönen approach and the reversible discriminator architecture on IBM Quantum processors. The candidate conducted comparative performance tests between SQGEN and the quantum version of GAN for system sizes $n = 1, ..., 6$, and collected convergence trajectories, fidelity metrics, and execution time statistics, which are presented in the chapter. Additionally, the candidate conducted a single-qubit hardware demonstration on the *ibmq_manila* processor. Theoretically, the candidate contributed to the analysis of the Nash equilibrium structure of the synergistic cost function and to identifying conditions under which SQGEN reduces to a quantum autoencoder at convergence.

Chapter 4: Experimental Liouvillian exceptional points in a quantum system without Hamiltonian singularities

Published as: S. Abo, P. Tulewicz, K. Bartkiewicz, Ş. K. Özdemir, A. Miranowicz *Experimental Liouvillian exceptional points in a quantum system without Hamiltonian singularities* New Journal of Physics **26**, 123032 (2024).

The candidate contributed to this work by focusing on the experimental implementation and analysis of results obtained using IBM Quantum hardware. She was responsible for preparing and conducting experiments on IBM Quantum superconducting processors, including single-qubit measurements on the *ibm_nairobi* processor. The candidate implemented Stinespring's extended non-unitary dynamics as unitary circuits on available gate sets and designed a quantum process tomography (QPT) protocol to reconstruct the full Liouville superoperator from hardware data. The candidate also applied dynamic decoupling strategies and compiler optimization.

Institute of Spintronics and Quantum Information, Faculty of Physics and Astronomy
Adam Mickiewicz University
ul. Uniwersytetu Poznańskiego 2, 61-614 Poznań, Poland

Poznań, 27.04.26r.

Statement about the contribution to the publication

We hereby declare that we are aware that the work:

Resource-efficient quantum correlation measurements via multicopy neural network methods,
by Patrycja Tulewicz, Karol Bartkiewicz, Adam Miranowicz, and Franco Nori,
Scientific Reports **15**, 40868 (2025), DOI: 10.1038/s41598-025-24607-2,

of which we are co-authors, has been included in the doctoral thesis of Patrycja Tulewicz, entitled *"Quantum correlations and exceptional points in superconducting processors: collective measurements, generative learning, and dynamical control"* (Chapter 2, Part I).

Patrycja Tulewicz designed and executed the full experimental protocol, including the selection of measurement bases, the design of singlet-projection circuits, and a qubit-mapping strategy tailored to the IBM Quantum Hanoi processor. She implemented and ran all quantum circuits on IBM Quantum hardware, managed calibration data, and applied error-mitigation techniques including zero-noise extrapolation, Bayesian readout correction, and purity-based post-selection. She developed the artificial neural network model—its architecture, training procedure on randomly sampled two-qubit states, and a SHAP-based analysis identifying five optimal measurement configurations—and performed all numerical simulations and statistical analyses of shot-noise scaling. Karol Bartkiewicz conceived and supervised the project and contributed to the hardware data analysis. Adam Miranowicz contributed to the theoretical development and manuscript preparation. Franco Nori contributed to the theoretical interpretation and supervised the research. All authors contributed to writing and reviewing the manuscript.

Karol Bartkiewicz

A. Miranowicz

Franco Nori

Institute of Spintronics and Quantum Information, Faculty of Physics and Astronomy
Adam Mickiewicz University
ul. Uniwersytetu Poznańskiego 2, 61-614 Poznań, Poland

Poznań, 27.04.26r.

Statement about the contribution to the publication

We hereby declare that we are aware that the work:

Spin chirality across quantum state copies detects hidden entanglement,
by Patrycja Tulewicz, Karol Bartkiewicz, and Franco Nori,
manuscript in preparation for publication,

of which we are co-authors, has been included in the doctoral thesis of Patrycja Tulewicz, entitled *"Quantum correlations and exceptional points in superconducting processors: collective measurements, generative learning, and dynamical control"* (Chapter 2, Part II).

Patrycja Tulewicz contributed both theoretically and experimentally. She developed the framework linking partial-transposition moments to multi-copy spin-chirality correlators and established the Newton–Girard relations that transform moment differences into conditions for a Gröbner-basis decision tree. She implemented all quantum circuits required to evaluate chirality witnesses and non-Hermitian gaps, and performed a substantial portion of the experimental measurements on IBM Quantum processors, covering the Werner, Horodecki, Tiles, and Chessboard state families across their full parameter ranges. Karol Bartkiewicz conceived and supervised the project and contributed to the interpretation of results. Franco Nori contributed to the theoretical interpretation and supervised the research. All authors contributed to writing and reviewing the manuscript.

Karol Bartkiewicz

Franco Nori

Institute of Spintronics and Quantum Information, Faculty of Physics and Astronomy
Adam Mickiewicz University
ul. Uniwersytetu Poznańskiego 2, 61-614 Poznań, Poland

Poznań, 27.04.26r.

Statement about the contribution to the publication

We hereby declare that we are aware that the work:

Synergic quantum generative machine learning,
by Karol Bartkiewicz, Patrycja Tulewicz, Jan Roik, and Karel Lemr,
Scientific Reports **13**, 12893 (2023), DOI: 10.1038/s41598-023-40137-1,

of which we are co-authors, has been included in the doctoral thesis of Patrycja Tulewicz, entitled *"Quantum correlations and exceptional points in superconducting processors: collective measurements, generative learning, and dynamical control"* (Chapter 3).

Karol Bartkiewicz developed the theoretical framework of SQGEN and supervised the project. Patrycja Tulewicz co-developed the numerical simulations and hardware experiments, implementing circuits based on the Möttönen state-preparation approach and the reversible-discriminator architecture on IBM Quantum processors. She conducted comparative performance tests of SQGEN against the quantum GAN for system sizes $n = 1, \ldots, 6$, collected convergence trajectories, fidelity metrics, and execution-time statistics, and performed a single-qubit hardware demonstration on the IBMQ Manila platform. She also contributed to the analysis of the Nash-equilibrium structure of the synergistic cost function and to identifying conditions under which SQGEN reduces to a quantum autoencoder at convergence. Jan Roik and Karel Lemr participated in scientific discussions and in manuscript preparation and review. All authors contributed to writing and reviewing the manuscript.

Karol Bartkiewicz

Roik

Institute of Spintronics and Quantum Information, Faculty of Physics and Astronomy
Adam Mickiewicz University
ul. Uniwersytetu Poznańskiego 2, 61-614 Poznań, Poland

Poznań, 27.04.26r.

Statement about the contribution to the publication

We hereby declare that we are aware that the work:

Experimental Liouvillian exceptional points in a quantum system without Hamiltonian singularities, by Shilan Abo, Patrycja Tulewicz, Karol Bartkiewicz, Şahin K. Özdemir, and Adam Miranowicz, New Journal of Physics **26**, 123032 (2024), DOI: 10.1088/1367-2630/ad98b6,

of which we are co-authors, has been included in the doctoral thesis of Patrycja Tulewicz, entitled *"Quantum correlations and exceptional points in superconducting processors: collective measurements, generative learning, and dynamical control"* (Chapter 4).

Patrycja Tulewicz focused on the experimental implementation and analysis of results obtained on IBM Quantum hardware. She prepared and conducted experiments on IBM Quantum superconducting processors, including single-qubit measurements on and two-qubit measurements on the IBM Nairobi processor. She implemented the Stinespring-extended non-unitary dynamics as unitary circuits on the available gate sets and designed a quantum process tomography (QPT) protocol to reconstruct the full Liouville superoperator from the hardware data. She also applied dynamic decoupling strategies and compiler optimization. Shilan Abo contributed to the theoretical foundations and performed analytical and numerical calculations and conducted experiments on IBM Quantum superconducting processor. Karol Bartkiewicz supervised the experimental work and contributed to the interpretation of results. Adam Miranowicz originated and oversaw the initial concept of the project. Şahin K. Özdemir contributed to contextualizing the significance of the results. All authors contributed to writing and reviewing the manuscript.

Shilan Abo

Karol Bartkiewicz

A. Miranowicz

S. Özdemir

Chapter 7

Appendix

7.1 Academic achievements

7.1.1 List of publications

1. P. Tulewicz, K. Bartkiewicz, *Applications of machine learning to long-range quantum routing*, Proc. SPIE **PC12133**, PC1213313 (2022). doi:10.1117/12.2621977

2. K. Bartkiewicz, P. Tulewicz, J. Roik, K. Lemr, *Synergic quantum generative machine learning*, Sci. Rep. **13**, 12893 (2023). doi:10.1038/s41598-023-40137-1

3. S. Abo, P. Tulewicz, K. Bartkiewicz, Ş. K. Özdemir, A. Miranowicz, *Experimental Liouvillian exceptional points in a quantum system without Hamiltonian singularities*, New J. Phys. **26**, 123032 (2024). doi:10.1088/1367-2630/ad98b6

4. P. Tulewicz, K. Bartkiewicz, A. Miranowicz, F. Nori, *Resource-efficient quantum correlation measurements via multicopy neural network methods*, Sci. Rep. **15**, 40868 (2025). doi:10.1038/s41598-025-24607-2

5. P. Tulewicz, K. Bartkiewicz, F. Nori, *Spin chirality across quantum state copies detects hidden entanglement* (in preparation).

7.1.2 List of conferences and seminars

1. **SPIE Photonics Europe 2022**, Strasbourg, France, 3–7 April 2022.
   Poster: *Applications of machine learning to long-range quantum routing*.

2. **1st Conference of Young Physicists in Poznań**, Poznań, Poland, 20–21 May 2022.
   Invited talk: *Applications of machine learning to long-range quantum routing*.

3. **22nd Polish-Slovak-Czech Optical Conference**, Wojanów, Poland, 5–9 September 2022.
   Poster: *Applications of machine learning to long-range quantum routing*.

4. **14th International Conference on Parallel Processing and Applied Mathematics**, Gdańsk, Poland, 11–14 September 2022.
   Oral presentation: *Collaborative generative quantum machine learning*.

5. **Quantum Machine Learning: from Fundamentals to Applications**, Udine, Italy, 12–16 September 2022.
   Assistance in poster presentation: *Synergic quantum generative machine learning*.

6. **Quantum characterization and control of quantum complex systems**, Como, Italy, 9–23 September 2022.
   Poster: *Application of reinforcement learning in long-range routing protocol search*.

7. **Current Problems in Physics 2022**, Zielona Góra, Poland, 24–27 October 2022.
   Oral presentation: *Synergic quantum generative machine learning*.

8. **Symposium on Spintronics and Quantum Information 2022**, Poznań, Poland, 8–9 December 2022.
   Oral presentation: *A direct method for measuring quantum entanglement and Bell nonlocality of two-qubit states through Hong-Ou-Mandel interferometry*.

9. **QUANTUMatter 2023**, Madrid, Spain, 23–25 May 2023.
   Poster: *Synergic quantum generative machine learning*.

10. **48th Congress of Polish Physicists**, Gdańsk, Poland, 1–7 September 2023.
    Invited talk: *Machine learning with quantum optics*.

11. **48th Congress of Polish Physicists**, Gdańsk, Poland, 1–7 September 2023.
    Poster: *A direct method for measuring quantum entanglement and Bell nonlocality of two-qubit states through Hong-Ou-Mandel interferometry*.

12. **Symposium on Spintronics and Quantum Information**, Będlewo, Poland, 11–13 January 2024.
    Poster: *Quantum generative machine learning*.

13. **QUANTUMatter 2024**, San Sebastián, Spain, 7–10 May 2024.
    Poster: *Experimental quantification of measuring quantum entanglement and Bell's nonlocality of two-qubit states*.

14. **International Young Quantum Scientists' Meet 2024**, Noida, India, 16–20 September 2024.
    Oral presentation: *Synergic Generative Machine Learning with Quantum Computing*.

15. **16th KCIK-ICTQT Symposium and "Quantum Horizons" Conference**, Gdańsk, Poland, 7–10 May 2025.
    Invited oral presentation: *Resource-Efficient Quantum Correlation Measurement: A Multicopy Neural Network Approach for Practical Applications*.

16. **Seminar at the Institute of Physics, University of Zielona Góra**, Zielona Góra, Poland, 11 June 2025.
    Invited oral presentation: *Efficient neural network-enhanced multicopy estimation for quantum correlation measurements*.

17. **Central European Workshop on Quantum Optics (CEWQO 2025)**, Vilnius, Lithuania, 23–27 June 2025.
    Poster: *Resource-efficient quantum correlation measurement using a multi-copy estimation approach*.

18. **Quantum Optics XI**, Kraków, Poland, 1–5 September 2025.
    Poster: *Measurement of quantum negativity via moment operators for two-qubits and qubit-qutrit cases*.

19. **Symposium on Spintronics and Quantum Information 2026**, Poznań, Poland, 23–25 April 2026.
    Oral presentation: *Spin chirality across quantum state copies detects hidden entanglement*.

7.2 Scientific projects

The research underlying this dissertation was carried out within the project: **Fundamental problems and implementations of dissipative quantum engineering**
Grant No. DEC-2019/34/A/ST2/00081 (2021 – present).
Principal Investigator: prof. dr hab. Adam Miranowicz.

7.3 Other scientific activities

Reviewing. Reviewer for *Physical Review A*, *Physical Review Letters*, and *Physical Review Research* (2021 – present).

Awards and distinctions. Second place award at the QLFuture Hackathon (2023).

Certifications.

- IBM Certified Associate Developer (2023);
- Qiskit Global Summer School 2023 – Quantum Excellence;
- IBM Quantum Challenge: Spring 2023 Achievement.

Academic service.

- Member of several committees of the Doctoral School of AMU, Poznań (2022 – 2024).
- Member of the scientific committee of the 1st Conference of Young Physicists in Poznań (May 2022).
- Assistance in organizing several conferences, Poznań (2021 – 2024).

Science outreach.

- Popular science lectures on quantum computing delivered in secondary schools across Poland (April – May 2023).
- Recording of a popular science podcast within the project *Here Happens Science II*, ID-UB program, AMU, Poznań (June 2024).
- Popular science lectures on quantum computing delivered at the event *Dziewczyny do ścisłych - PFNiS*, Poznań Poland (21 April 2025).

Declaration

Ja, niżej podpisana Patrycja Tulewicz studentka Szkoły Doktorskiej Szkoły Nauk Ścisłych w dziedzinie Nauk Fizycznych na Wydziale Fizyki i Astronomii Uniwersytetu im. Adama Mickiewicza w Poznaniu oświadczam, że przedkładaną rozprawę doktorską pt.: *Quantum correlations and exceptional points in superconducting processors: collective measurements, generative learning, and dynamical control* napisałam samodzielnie. Oznacza to, że przy pisaniu pracy, poza niezbędnymi konsultacjami, nie korzystałam z pomocy innych osób, a w szczególności nie zlecałam opracowania rozprawy lub jej istotnych części innym osobom, ani nie odpisywałam tej rozprawy lub jej istotnych części od innych osób. Równocześnie wyrażam zgodę na to, że gdyby powyższe oświadczenie okazało się nieprawdziwe, decyzja o nadaniu mi stopnia naukowego doktora zostanie cofnięta.

Patrycja Tulewicz