MACHINE LEARNING
Science and Technology

Check for updates

PAPER

OPEN ACCESS

RECEIVED
12 March 2025

REVISED
28 January 2026

ACCEPTED FOR PUBLICATION
18 February 2026

PUBLISHED
26 February 2026

Original content from this work may be used under the terms of the Creative Commons Attribution 4.0 licence.

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Molecular representations of quantum circuits for quantum machine learning

Elham Torabian[1,2,*] **and Roman V Krems**[1,2]

[1] Department of Chemistry, University of British Columbia, Vancouver, B.C. V6T 1Z1, Canada
[2] Stewart Blusson Quantum Matter Institute, Vancouver, B.C. V6T 1Z4, Canada
[*] Author to whom any correspondence should be addressed.

E-mail: elhamtorabian@chem.ubc.ca

Keywords: quantum ansatz optimization, molecular descriptors, Gershgorin circles of Coulomb matrix, quantum machine learning

Abstract

We establish an isomorphism between quantum circuits (QCs) and a subspace of polyatomic molecules, which suggests that molecules can be used as descriptors of QCs for quantum machine learning. Our numerical results show that the performance of QCs for quantum support vector machines can be characterized by dimensionality-reduced molecular fingerprints as well as by the size of the largest and smallest Gershgorin circles derived from the Coulomb matrices of the corresponding molecules. This can be used to restrict the search space for the compositional optimization of QCs. We show that a high accuracy of a quantum algorithm can be achieved with high probability by sampling from a specific set of molecules. This work implies that quantum ansatz optimization can benefit from advances in cheminformatics and suggests an approach to identify key elements that enhance the accuracy of a quantum algorithm by mapping QCs onto molecules and exploring correlations between physical properties of molecules and circuit performance.

1. Introduction

Recent evidence suggests that general quantum learning algorithms are either efficiently classically simulable, or suffer from limitations due to vanishing of inner products of quantum states in exponentially large Hilbert spaces [1], which leads to barren plateaus (BP) in variational algorithms [2–8] or exponential concentration (EC) of quantum kernels [9–11]. BP and EC thus appear to preclude the quantum advantage of quantum machine learning (QML). On the other hand, it has been shown that quantum kernels can solve classification problems with NP [12] and PromiseBQP-complete complexity [13], indicating that the quantum advantage of QML can be achieved, in principle. These results suggest that, for practical applications, the search for the quantum advantage of QML should focus on general strategies of encoding inductive bias into the quantum algorithms (see also [1]).

For gate-based quantum computing, quantum states are generated by circuits of logic gates consisting of unitary operators acting on a given initial state of qubits. The most general strategy for constructing optimally biased quantum circuits (QCs) is the optimization of operator sequences in the space of gate permutations. However, the computational complexity of such optimization scales exponentially with the number of gates and qubits, even when greedy algorithms are employed [14]. To reduce this complexity, it is necessary to restrict the search space. This is usually achieved by either assuming a specific structure for the quantum ansatz or by building adaptive-structure ansatz from simple QCs. QC can also be chosen to respect relevant symmetries. For example, EC can be avoided by restricting the space of QC to covariant quantum kernels [15], or by quantum kernel bandwidth methods [16], quantum Fisher kernels [17], and projected quantum kernel methods [18]. However, these methods rely on some prior information about the data structure.

In the present work, we aim to restrict the search space of gate permutations, extending to deep circuits, by parameters determined from optimization of shallow circuits. More specifically, we consider the following problem: given a set of QC with depth $\mathcal{P} < \mathcal{T}$, identify the predictors of the performance

© 2026 The Author(s). Published by IOP Publishing Ltd

of QC with depth $\mathcal{P} > \mathcal{T}$. To achieve this, we require an algorithm for QC construction that satisfies the following conditions: (i) deep QC must inherit relevant properties of shallow circuits; (ii) the performance of QC of any depth can be discriminated by efficiently (polynomial in the number of gates and qubits) computable metrics. To develop an algorithm that meets these conditions, we establish and exploit an isomorphism between QC and polyatomic molecules.

We then demonstrate that the shape of the resulting molecules, encoded in the Gershgorin circles of Coulomb matrices for the spatial atomic arrangements, yields predictors of QC performance. With the proposed circuit—molecule mapping, the computation of the Gershgorin circles is, at most, quadratic in the number of qubits and gates. Our results demonstrate that these conclusions generalize both to different samples within a given data set and to different datasets. Our results imply that mapping QC onto physical objects, such as molecules, can be used to elucidate the important features of the quantum ansatz design. The correspondence between QCs and molecular structures is grounded in graph isomorphism principles, drawing parallels with methodologies in cheminformatics and quantum computing. In cheminformatics, graph-based representations, such as canonical atom numbering [19] and algorithms for cross-validating chemical and crystallographic data [20], leverage graph isomorphism to ensure consistent molecular descriptions. Similarly, QCs are often represented as directed acyclic graphs to capture gate dependencies and temporal ordering [21], providing a powerful framework for quantum computation. In addition, QML employs graph-theoretic mappings of QCs, as demonstrated in reinforcement learning-based optimization using ZX-calculus [22]. In this work, we explore an alternative approach by establishing an isomorphic mapping between molecular structures (as physical descriptors) and QCs, leveraging graph-theoretic principles to encode molecular connectivity. The bidirectional mapping-QCs to molecules and vice versa-remains consistent within each QC dataset, with QC matrix representations read row by row to ensure non-conflicting gate ordering. This consistency robustly captures substructures distinguishing performant from underperforming quantum kernels.

The proposed molecule—circuit mapping can be used to leverage numerical methods developed for optimization of molecular properties in chemical compound spaces for QC optimization. Approaches based on generative modeling, deep learning architectures, probabilistic search, and cheminformatics representations provide systematic strategies for navigating and optimizing high-dimensional design spaces under complex constraints. Generative adversarial networks (GANs) [23–25] and convolutional neural networks [26, 27] have been effectively used to explore and design molecules with specific properties by generating and optimizing chemical structures. Bayesian optimization (BO) techniques [28–30] have recently been extended to efficiently navigate molecular spaces, balancing exploration and exploitation. Cheminformatics approaches [31, 32] further enhance these methods by providing robust molecular representations, enabling precise optimization of desired properties for drug discovery. Given the proposed circuit—molecule mapping, these available methods can be repurposed to efficiently search QC ansatz spaces and optimize QC architectures for targeted quantum computational tasks.

2. Isomorphism between QC and molecules

We begin by establishing an isomorphism between universal QC and a specifically defined sub-space of polyatomic molecules. To prove isomorphism, we construct a mapping of QC with universal gate sets onto polyatomic molecules that ensures: (1) bijection; (2) structure preservation—the mapping must preserve the essential relationships such as adjacency and order of operations, and (3) reversibility—there must exist an inverse mapping that is also structure-preserving.

A set of quantum gates is universal if the gates in this set can approximate any desired unitary operation to arbitrary precision [33]. This implies that a universal set of gates must support arbitrary single-qubit rotations and, at least, two-qubit entanglement [34]. We consider, separately, three universal gate sets: set (a) including the rotation operators $R_x(\theta)$, $R_y(\theta)$, and $R_z(\theta)$, the phase shift gate $P(\phi)$ and CNOT [33, 35]; (b) the Clifford gates, including CNOT, the Hadamard gate H and the phase gate S, supplemented by the phase-shift T gate; (c) a combination of the Toffoli gates T and the Hadamard gates (computational universality) [36]. Though the Clifford set is not universal and can be simulated classically [37], the addition of the T gate makes the set universal. Similarly, the Toffoli gate alone can be used to simulate any classical reversible circuit, but is not sufficient for universal quantum computation. However, when combined with a quantum gate, such as H, the Toffoli gates can be used to build universal QCs.

We aim to obtain molecular representations that encode the architecture of universal QC applied to an arbitrary number n_Q of qubits. The representations adopted here encode the type and positions of gates in the QCs. We note that each QC can be uniquely represented by an $n_Q \times L$ matrix (shown in figure 1), where L is the number of gate layers. The matrix rows represent the qubits, and the columns

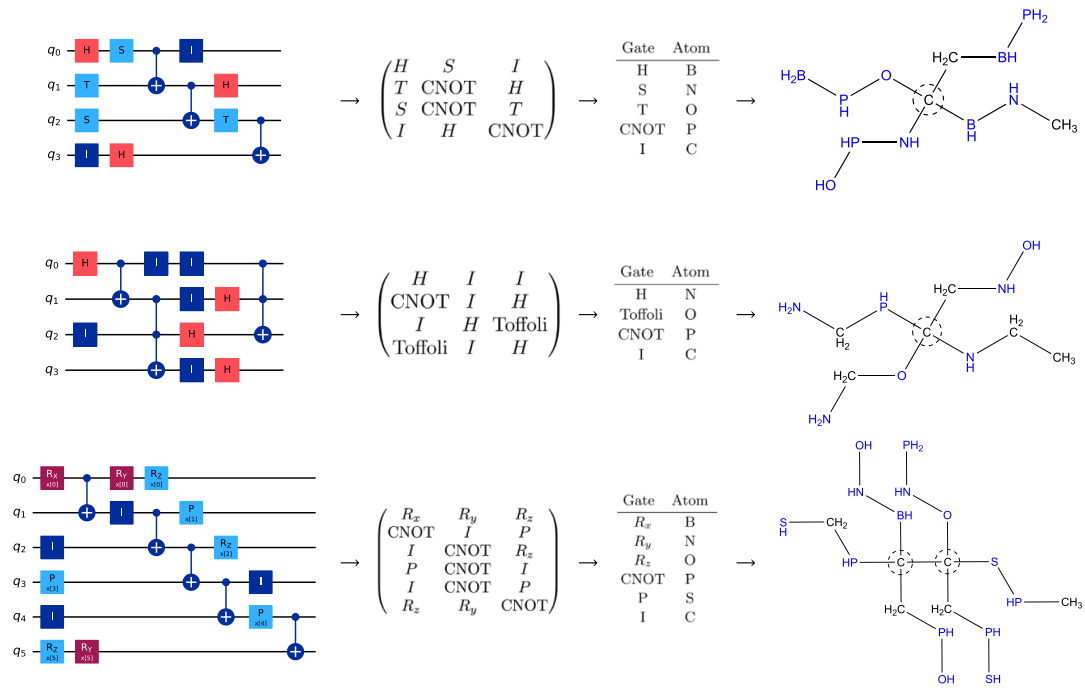


Figure 1. Three examples of mapping of quantum circuits with universal gate sets and the corresponding matrix representation of gates to molecules. Each qubit is represented by a molecular branch stemming from the backbone carbon atom/chain, encircled by the dashed line.

specify the gates or no-gate (identity I), which are applied to each qubit in a given circuit layer. We assign each row of the matrix to a branch of atoms stemming from a central carbon atom chain and identify each type of quantum gate with an atom that permits an appropriate number of covalent bonds. The identity operator is specified by the carbon atom, while the quantum gates are specified by non-carbon atoms placed in appropriate positions.

We encode the number of qubits into a polymeric carbon chain as illustrated in table 1. The goal is to ensure that the resulting molecules have $n_{\rm Q}$ branches. For example, a four-qubit circuit can be represented by a molecule with four branches of an sp^3-hybridized carbon atom, while a six-qubit circuit requires two covalently bonded sp-hybridized carbon atoms. A molecule with an arbitrary number of branches can be designed by inserting the appropriate number of carbon atoms in the central carbon atom chain. We choose the hybridization of the carbon atoms in the backbone chain to represent $n_{\rm Q}$ branches with the lowest number of carbon atoms. After assigning atoms to each molecular branch, we saturate the valency of all atoms by attaching hydrogen atoms to create stable molecules. A QC with an odd number of qubits is obtained by replacing one molecular branch with a hydrogen atom.

Each row of the matrix in figure 1 is represented by a molecular branch with quantum gates described by specific atoms that permit, at least, two covalent bonds. Because the pool of such atoms is limited, we restrict the entanglement gates between qubits with $|i-j| \leqslant \delta$, where i and j are the row indices of the matrices in figure 1 and δ is some finite integer. This does not restrict the universality of the resulting QCs, as nearest-neighbor entanglement is sufficient to generate universal QC [38, 39]. Thus, we need 3–5 atoms that permit at least two covalent bonds (bi-valent atoms) to identify all gates in any universal quantum gates set.

The mapping between QC gates and molecular structures, as shown in figure 1, may appear arbitrary due to the existence of multiple valid bijections. The variation across the three circuits arises from illustrating the feasibility of mapping different universal gate sets-each with distinct gate types and connectivity-onto molecular structures with varying atom choices, guided by atom valency, connectivity constraints, and entanglement range. In the results section, we examine alternative mappings and find that the results are largely insensitive to the specific choice of atoms. This insensitivity suggests that our conclusions depend primarily on the general structure of the QC—molecule correspondence rather than on a particular bijection.

3. Methodology for numerical calculations

3.1. Molecular descriptors

For machine learning and cheminformatics applications, molecules are encoded into molecular descriptors. Molecular descriptors must generally be numerical, universal, and capable of capturing essential features of molecules such as molecular shape, chemical bond arrangements, bond strengths and orders, vibrational frequencies, and conformational flexibility. While it is possible to design unique multi-dimensional descriptors, such as molecular fingerprints [40] or graph neural networks [41, 42] to explicitly represent the spatial arrangement and connectivity of atoms within a molecule, much recent work has focused on the development of physical descriptors, including Coulomb matrices [43], bag of bonds features [44], smooth overlap of atomic positions [45], and the Faber–Christensen–Huang–Lilienfeld representation [46]. Molecular fingerprints can be viewed as a digital signature of molecules, encapsulating the structural and chemical attributes through strings with 1000–4000 bits. Such descriptors are useful for cheminformatics applications, but impractical for interpreting the effects of individual atom groups on molecular properties or for optimization of molecular properties in chemical compound spaces. On the other hand, physical descriptors are more suitable for algorithms aiming to optimize the molecular properties, e.g. in spaces of molecular isomers [47].

To demonstrate the feasibility of molecular representations of QCs, we first consider molecular fingerprints as descriptors of QC. We use count-based molecular fingerprints implemented in RDKit and inspired by the Daylight fingerprinting method [48]. These fingerprints identify subgraphs of a molecule within a specified range of molecular groups, hash them into bit IDs, and use modulo operations to map these IDs to a fingerprint of fixed size. The count-based format records the frequency of each subgraph occurrence rather than a binary presence/absence. The hashing process incorporates atomic and bond properties, such as the atomic number (modulo 128), aromaticity, atom degree, and bond type [49].

In order to explore practical applications of molecular descriptors for structural optimization of QCs, we also employ the Coulomb matrix representations of molecules. As shown by Rupp *et al* [43], molecules can be effectively described by a symmetric matrix with the matrix elements given by

$$M_{ij} = \begin{cases} 0.5 Z_i^{2.4} & \text{for } i = j \\ Z_i Z_j / r_{ij} & \text{for } i \neq j, \end{cases} \tag{1}$$

where Z_i and Z_j are the atomic numbers of atoms i and j, respectively, r_{ij} is the distance between atoms i and j, with the indices $i, j \in [1, n_{\max}]$, and $n_{\max}$ is the number of atoms in the molecule. The diagonal elements of the Coulomb matrix represent a polynomial fit relating the atomic number to the total energy of free atoms [43], while the off-diagonal elements describe the electrostatic interaction between atoms i and j. The Coulomb matrix is invariant to molecular translations and rotations, though not invariant to atomic permutations.

It was previously shown that the eigenvalues of Coulomb matrices contain physical information about molecules [50]. However, within the QC-molecule mapping scheme proposed here, the number of atoms grows polynomially with the number of qubits and gates in the corresponding QC. Therefore, it is impractical to consider full Coulomb matrices or their eigenvalues as descriptors of QC. Instead, we propose and explore the application of the Gershgorin circle theorem. This theorem identifies a region in the complex plane that includes all eigenvalues of a complex square matrix [50]. More specifically, for an $m \times m$ matrix with complex entries, the eigenvalues of matrix M lie within the union of discs $D_1 \cup D_2 \cup \cdots \cup D_m$, where each disc D_i is defined as

$$D_i = \left\{ z \in C : |z - M_{ii}| < \sum_{j \neq i} |M_{ij}| \right\}. \tag{2}$$

We use the radii of the smallest and largest Gershgorin circles for the Coulomb matrices of the corresponding molecules to characterize the QC and to restrict the structural optimization of QC.

To generate molecular structures, we use RDKit [49]. More specifically, for the calculations presented here, we use molecular structures with the three-dimensional (3D) geometry optimized by the universal force field (UFF) method. The UFF optimization minimizes potential energy, including van der Waals interactions, and accounts for torsional interactions, in addition to energies of chemical bonds. This yields physical steric conformations [51]. We have also repeated the present calculations with molecules represented by two-dimensional (2D) arrangements of atoms. The 2D coordinates are computed using the Kamada–Kawai force-directed algorithm, which optimizes atomic positions by minimizing an energy

function based on linear springs [52]. This method preserves atomic connectivity but does not account for steric effects. Since UFF optimization is computationally expensive and scales significantly with the increasing number of atoms, bonded and non-bonded interactions [53], using 2D representations reduces the computational cost of generating molecules. Furthermore, a 2D molecular representation can be directly mapped to the coordinates of neutral atoms on quantum platforms with well-defined and well-controlled 2D atom arrays, such as Pasqal [54] and Quera [55], which allows the application of particular quantum algorithms for determining molecular properties on such quantum computers [56, 57].

3.2. Quantum support vector machines (SVMs)

In the present work, we consider applications of QC for quantum kernels of SVMs. The resulting SVMs are used for binary classification problems. The inputs $\boldsymbol{x}_i \in \mathbb{R}^n$ of the dataset $\mathcal{D} := \{\boldsymbol{x}_i, y_i \in \{0,1\}\}_{i=1}^N$ are encoded into quantum kernels

$$k(\boldsymbol{x},\boldsymbol{x}') = |\langle \Phi(\boldsymbol{x}')|\mathcal{U}^\dagger(\boldsymbol{x}')\mathcal{U}(\boldsymbol{x})|\Phi(\boldsymbol{x})\rangle|^2, \tag{3}$$

where

$$|\Phi(\boldsymbol{x})\rangle = \exp\left(\mathrm{i}\sum_k x_k Z_k\right) H^{\otimes n}|0\rangle^{\otimes n}, \tag{4}$$

x_k is the kth component of $\boldsymbol{x}$, Z_k is the Pauli Z-gate acting on kth qubit, and H is the Hadamard gate. The sequence of gates in $\mathcal{U}(\boldsymbol{x})$ determines the performance of the quantum SVM (QSVM) models [14]. We choose $\mathcal{U}$ to consist of L layers, each including one-qubit R_Z gates and/or two-qubit CNOT gates. Each of the R_Z gates for qubit k is parameterized by θ as follows:

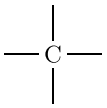
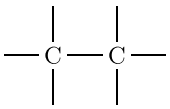
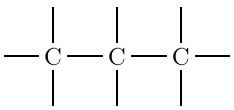
$$R_{Z,k}(\theta) = \begin{pmatrix} \mathrm{e}^{-\mathrm{i}\theta x_k} & 0 \\ 0 & \mathrm{e}^{\mathrm{i}\theta x_k} \end{pmatrix}. \tag{5}$$

This yields a QC with the number of free parameters θ equal to the number of R_Z gates. We optimize the parameters of all generated QCs using BO [58].

To illustrate that molecular descriptors can be used as representations of QCs, we consider four distinct classification problems: (I) given the ionic radii of A, B, B′ and X, classification of halide perovskites with the chemical formula $A_2BB'X_6$ into metals or non-metals [59] ; (II) classification of a four-dimensional (4D) synthetic data set implemented in [60]; (III) classification of a five-dimensional (5D) synthetic data set implemented in [60]; (IV) 5D- and ten-dimensional classification of digits 3 and 5 based on the MNIST dataset from [61]. Hereafter, we refer to these classification problems as perovskite (for problem I), hidden-manifold (for problems II and III), and MNIST (for problem IV). For each classification problem, we generate N QC $\mathcal{U}(\boldsymbol{x})$ including L layers of randomly sampled quantum gates, and train QSVM with the corresponding quantum kernels. The accuracy of the resulting QSVM is used as a metric of the corresponding quantum kernel performance. We use problems I and II to identify the molecular properties that can be used to optimize the search of performant QC. We then use problems III and IV to demonstrate the performance of the algorithm thus developed.

We use the following metric to quantify the performance of a quantum kernel (of equivalently a QC). The dataset for each classification problem is split into a training set and a test set. The training set for the perovskite classification problem includes 100 data points. For all other classification problems considered in this work, the training set includes 1000 data points. The test sets comprise 1442 randomly sampled materials for the perovskite classification problem and 1000 randomly selected data points for all other classification problems considered in this work. We label the two classes of each classification problem as positive and negative and define the true positive rate (TPR) and true negative rate (TNR) as the ratio of correct predictions to the total number of test points for a given class. The average accuracy of an SVM model is defined here as the balanced average of TPR and TNR computed over the test set. The boundary separating performant and underperforming circuits is defined as the midpoint between the maximum and minimum average classification accuracy over the entire set of QCs. We include a margin of $\pm 10\%$ to define a QC as performant, when the average accuracy of the corresponding SVM exceeds the boundary by more than 10%, and underperforming when the resulting average accuracy is more than 10% below the separating boundary. We discard the QC within the $\pm 10\%$ margin of the separating boundary.

Table 1. Central carbon atom chain designed to produce n_Q covalently bonded branches. As illustrated, the carbon polymer chain requires n_C carbon atoms for even n_Q. A QC with an odd number of qubits can be obtained by replacing one molecular branch with a hydrogen atom.

Number of qubits	Carbon backbone
4	
6	
8	
⋮	⋮
n_Q	$n_C = \frac{n_Q}{2} - 1$

3.3. Simplified QC—molecule mapping

As shown in section 2, the operators $\mathcal{U}(\boldsymbol{x})$ can be built with the universal set of quantum gates. In order to make computations more efficient, for numerical examples, we use two simplifications: (1) we allow $\mathcal{U}(\boldsymbol{x})$ to include only one-qubit R_Z gates and two-qubit CNOT gates; (2) we omit the identity operators from the QC $\rightarrow$ molecule mapping. This makes the inverse mapping from molecules to QC non-unique. However, our numerical tests show that this does not affect the results presented in this work.

The following section considers numerical examples with four-qubit and five-qubit QC. To map quantum gates to atoms for the numerical examples, we utilize the following scheme: $R_Z \rightarrow$ C, CNOT$(\delta = 1) \rightarrow$ N, CNOT$(\delta = 2) \rightarrow$ O, CNOT$(\delta = 3) \rightarrow$ S, and CNOT$(\delta = 4) \rightarrow$ P. The carbon atom backbone is constructed as illustrated in table 1. In particular, for the 5-qubit circuits, we use two sp^3-hybridized atoms with the valency of one of these atoms saturated by one hydrogen atom.

We consider two types of problems. First, we examine the performance of QC for QSVM in the space of molecular descriptors in order to show that molecular descriptors can serve as meaningful descriptors of QC. Second, we build molecules by constraining the algorithm to particular parts of a molecular space, and transform these molecules into QC to illustrate that molecular descriptors can reduce the complexity of compositional QC optimization.

4. Results

4.1. Molecular fingerprints as QC descriptors

We begin by illustrating the feasibility of applying widely used molecular fingerprints as variables to characterize the performance of QC for QSVM. For this calculation, we generate $N = 10\,000$ QC $\mathcal{U}(\boldsymbol{x})$ with $L = 5$ layers of randomly sampled quantum gates and train QSVM with the corresponding quantum kernels. Each QC is mapped onto a molecule, as described in section 3.3. High-dimensional fingerprints are constructed for the corresponding molecule using the method in [48] as described in section 3.1. The dimensionality of the molecular descriptors is then reduced by principal component analysis (PCA) to two principal components. Each of the principal components is a linear combination of fingerprint components.

Figure 2 shows the performance of QC for QSVM applied to two classification problems described in section 3.2 as a function of the two principal components thus obtained. Here and hereafter, the QSVM accuracy is represented by red circles for performant circuits and green triangles for underperforming kernels, yielding low accuracy of QSVM. The distinction between the performant and underperforming QC adopted here is described in section 3.2. Figure 2 demonstrates that there is a separation between performant QC and underperforming QC in the PCA-reduced space of molecular fingerprints. This separation is particularly clear for the 4D hidden-manifold classification problem, in which the data for two classes are generated by a classical artificial neural network. This suggests that PCA-reduced molecular fingerprints can be used to restrict the search of performant QC for QSVM through the QC $\leftrightarrow$ molecule mapping algorithm proposed here. The varying degree of separation between well-performing and poorly performing QCs in figure 2, as observed in the present molecular descriptor-based approach (e.g. clear separation right panel versus less distinct separation in the left panel), reflects

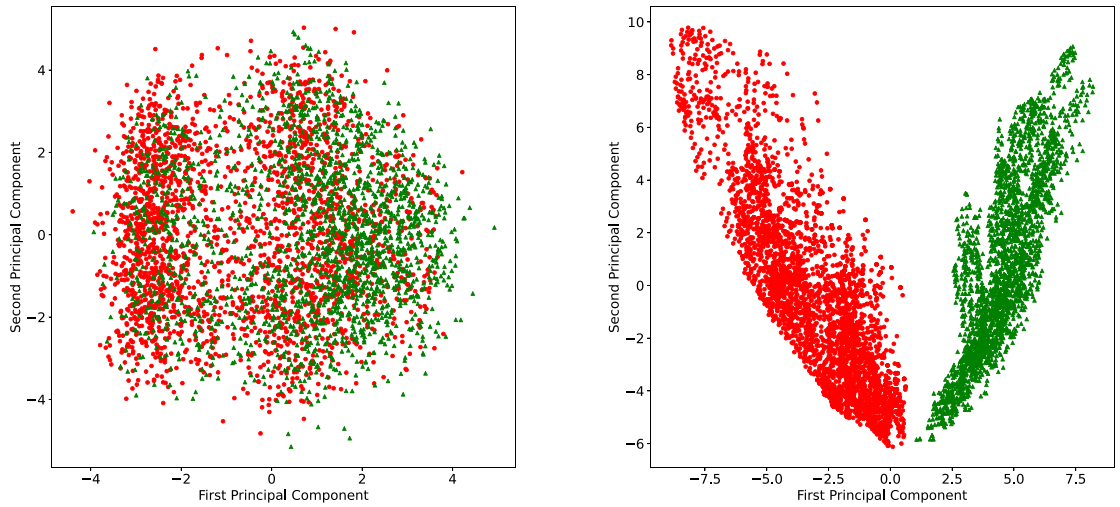


Figure 2. Distributions of performant QC (red circles), yielding high classification accuracy of QSVM, and underperforming QC (green triangles), leading to low classification accuracy of QSVM, in the space of two leading principal components after PCA dimensionality reduction of molecular fingerprints. The results are obtained with 10 000 randomly generated QC used to build kernels for QSVM to classify the 4D perovskite (left panel) and the 4D hidden-manifold (right panel) classification problems.

the problem-dependent nature of the classification task. We attribute this variability to the intrinsic complexity, patterns, and correlations within each dataset, which determine the architecture of performant and under-performing QCs and their corresponding molecular representations.

While figure 2 establishes a meaningful correlation between molecular descriptors and QC performance, reducing the dimensionality of molecular fingerprints by PCA requires diagonalization of high-dimensional matrices. The computational complexity of matrix diagonalization scales as $N_{\rm dim}^3$ with the matrix dimensionality $N_{\rm dim}$, making PCA-reduced molecular fingerprints computationally expensive for large QC. Furthermore, the complexity of molecular fingerprints complicates the interpretability of the basic elements that make a QC performant. Therefore, the following section explores alternative molecular representations of QC.

4.2. Gershgorin circles of Coulomb matrices as QC descriptors

A key goal of the present work is to identify molecular properties that are (i) either scalar or low-dimensional vectors; (ii) sensitive to the structural composition and shape of polyatomic molecules; (iii) efficient to compute, and (iv) provide a meaningful characterization of the corresponding QC performance. To achieve this, we explore the correlation between the performance of QC for QSVM and the values of the radii of the Gershgorin circles (2) derived from the Coulomb matrices for the corresponding molecules. The Gershgorin circle radii are represented by sums of the matrix elements and can thus be efficiently computed. For each molecule, we consider the radii of the largest and smallest Gershgorin circles, denoted for a given Coulomb matrix by $R_{\rm min}$ and $R_{\rm max}$ respectively.

The left panels of figure 3 illustrate the QC performance in the 2D space of $R_{\rm min}$ and $R_{\rm max}$. The right panels of figure 3 show the probability density of performant and low-accuracy QC computed from the distributions in the left panels by kernel density estimation (KDE) as functions of $R_{\rm min}$. The results are presented for the 4D perovskite (upper panels) and 4D hidden-manifold (lower panels) classification problems. Figure 3 shows that the magnitudes of the smallest and largest Gershgorin circles for a given molecule can be used to guide the search of performant QC for QSVM. To illustrate this directly, we generate two sets of 100 random molecules: one sampling from the quadrant with $R_{\rm min} > 220$ and $R_{\rm max} > 650$ (higher probability of finding performant QC), and another sampling from the quadrant with $R_{\rm min} < 220$ and $R_{\rm max} < 650$ (lower probability of finding performant QC).

All molecules in these random sets correspond to QC with 5 layers. We map each of these molecules onto a quantum kernel and apply the kernel thus obtained to the perovskite classification problem. The accuracy of the resulting models is shown in figure 4, with the results obtained from molecules with $R_{\rm min} > 220$ shown in the left panel and the results from molecules with $R_{\rm min} < 220$ in the right panel. The horizontal dashed line represents the average accuracy of 10 000 QSVM models illustrated in figure 3. Figure 5 presents the same results but for the 4D hidden manifold dataset and with the panels corresponding to sampling from the quadrant with $R_{\rm min} > 170$ and $R_{\rm max} > 400$ and the quadrant with $R_{\rm min} < 170$ and $R_{\rm max} < 400$.

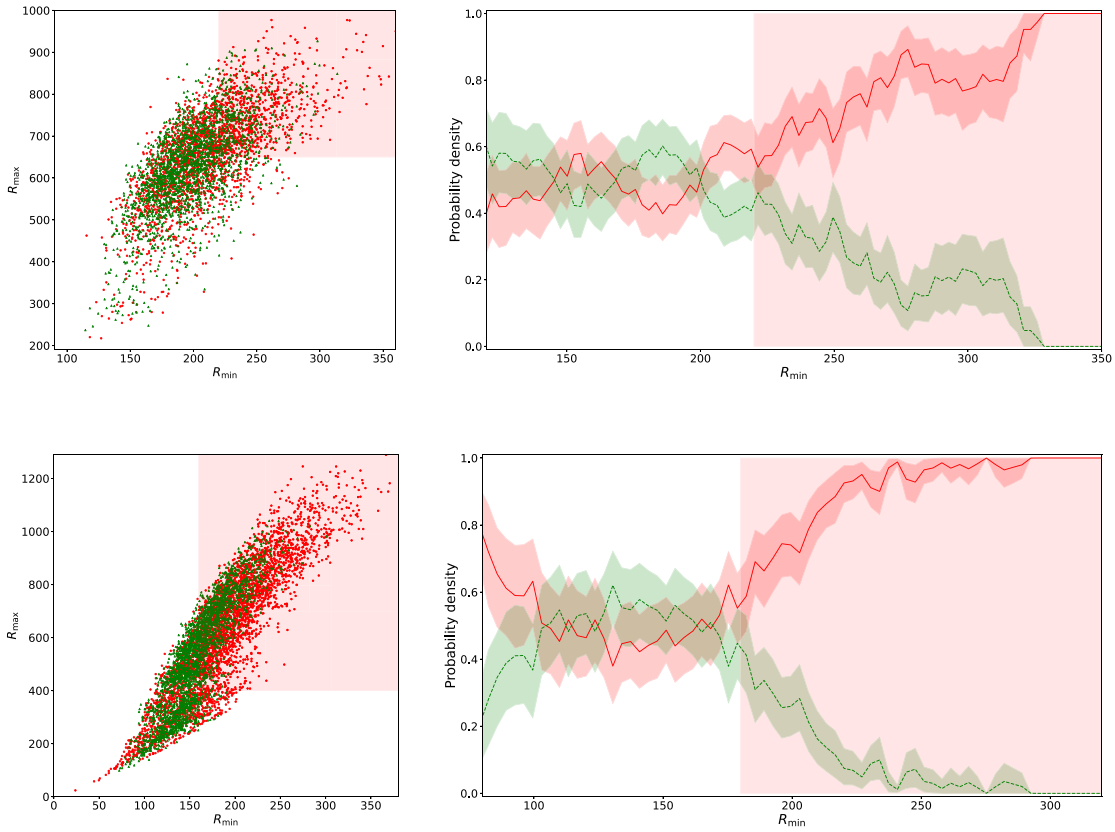


Figure 3. Left panels: distributions of performant QC (red circles) and underperforming QC (green triangles) in the space of the smallest ($R_{\min}$) and largest ($R_{\max}$) Gershgorin radii for the 4D perovskite (top) and 4D hidden-manifold (bottom) classification problems. Right panels: probability density of performant (solid red curves) and underperforming (broken green curves) as functions of $R_{\min}$ computed by KDE from the data in the left panels. The shaded region indicates the 95% confidence interval, reflecting the uncertainty in the density estimation due to the finite sample size.

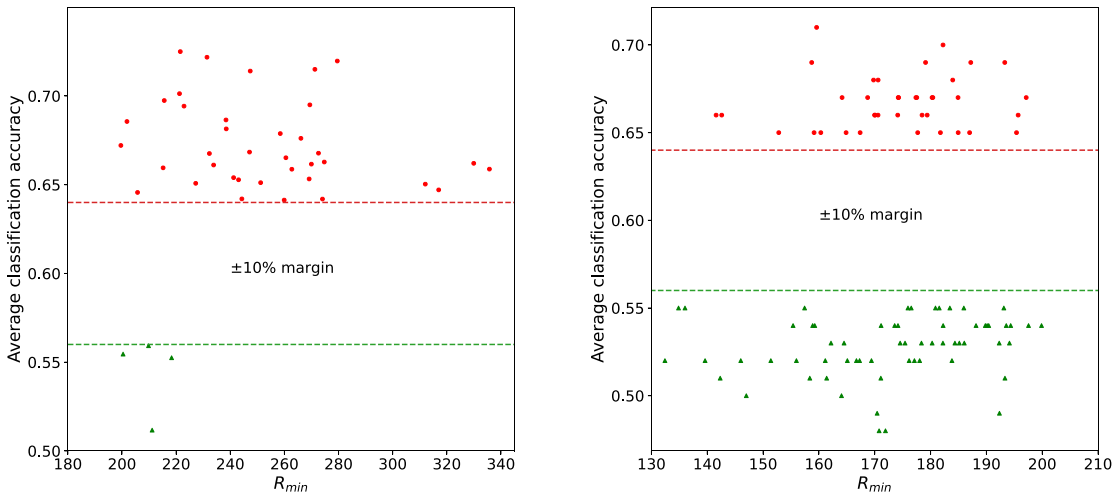


Figure 4. Distributions of average QSVM accuracy for the 4D perovskite classification problem with five-layer QC constructed from 100 randomly generated molecules with large radii ($R_{\min} > 220$) of the smallest Gershgorin circle (left panel) and low radii ($R_{\min} < 220$) of the smallest Gershgorin circle (right panel). The dashed lines show the $\pm 10\%$ margin centered at the average of the minimum and maximum classification accuracy in a sample of 10 000 random quantum circuits. The red circles and green triangles represent performant and underperforming QC designs, respectively.

Note that the random samples of molecules used in figures 4 and 5 are independent of the set in figure 3. The left panels of figures 4 and 5 show that randomly sampled molecules with a large value of $R_{\min}$ are much more likely to produce a performant quantum kernel for both datasets (e.g. the number of red circles is much larger than the number of green triangles). The right panels of figures 4 and 5

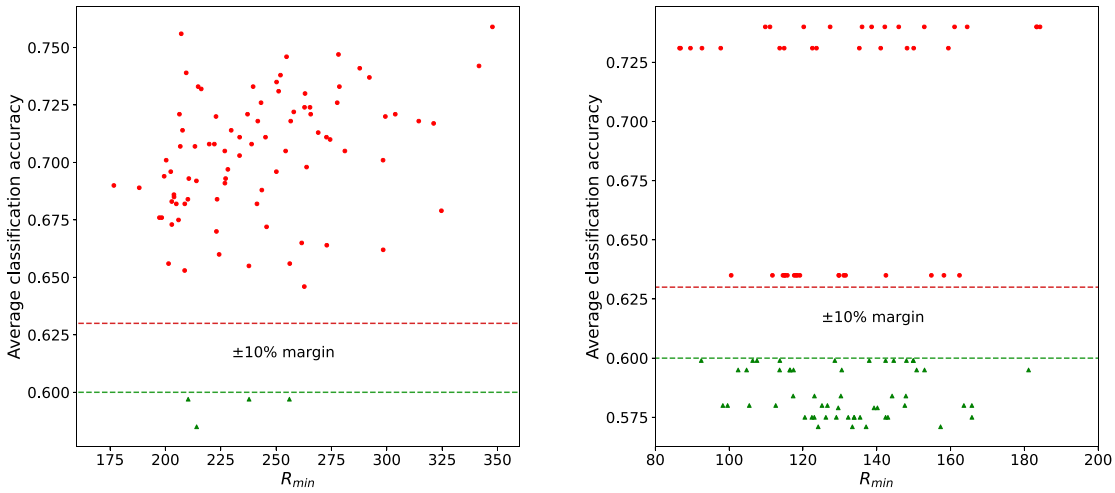


Figure 5. Distributions of average QSVM accuracy for the 4D hidden-manifold classification problem with five-layer QC constructed from 100 randomly generated molecules with large radii ($R_{\min} > 170$) of the smallest Gershgorin circle (left panel) and low radii ($R_{\min} < 170$) of the smallest Gershgorin circle (right panel). The dashed lines show the $\pm 10\%$ margin centered at the average of the minimum and maximum classification accuracy in a sample of 10 000 random quantum circuits. The red circles and green triangles represent performant and underperforming QC designs, respectively.

illustrate that the opposite is also true: randomly sampled molecules with a small value of $R_{\min}$ are much more likely to produce quantum kernels yielding low-accuracy models.

To extend the illustration of the present approach to a higher-dimensional problem, we analyze the ten-dimensional (10-feature) MNIST classification problem to assess the robustness of the suggested mapping approach. As shown in figure 6, the distributions of performant (red circles) and underperforming (green triangles) QCs in the space of the smallest $R_{\min}$ and largest $R_{\max}$ Gershgorin radii reveal distinct separation, with performant QCs exhibiting larger $R_{\min}$. The right panels of figure 6, displaying KDE of $R_{\min}$, confirm that performant QCs (solid red curves) are associated with higher $R_{\min}$ values compared to underperforming QCs (broken green curves), with 95% confidence intervals highlighting robust statistical separation. These results confirm the efficacy of the present method in higher-dimensional settings, demonstrating that balanced molecular connectivity enhances quantum kernel performance in complex classification tasks.

We next illustrate the proposed algorithm of designing performant QC by engineering polyatomic molecules for two independent classification problems: a 5D version of the hidden-manifold classification [60] and the 5D classification of digits 3 and 5 based on the MNIST dataset from [61]. We begin by generating a random sample of 10 000 molecules, each corresponding to QC with five layers of gates, for each classification problem. We then map 100 molecules with the highest values of the smallest Gershgorin circle radius and 100 molecules with the lowest values of the smallest Gershgorin circle radius onto the corresponding QCs. Using QC thus constructed, we build QSVM models for the 5D hidden-manifold and 5D MNIST classification problems. The results are presented in figure 7.

Designing quantum kernels by molecule $\rightarrow$ QC mapping provides insights into what determines the performance of quantum kernels for QSVM. Formally, a large radius of a Gershgorin circle indicates that, for the corresponding row of the matrix, the sum of the magnitudes of the off-diagonal elements is large compared to the magnitude of the diagonal element. For Hamiltonian matrices, a large Gershgorin circle may indicate strong coupling between different states, where the results are less determined by the local properties (diagonal elements) than by interactions with other states (off-diagonal elements). For molecules represented by Coulomb matrices, a large radius of the smallest Gershgorin circle suggests that the corresponding atom (associated with that particular row of the matrix) interacts strongly with many other atoms in the molecule. This implies a compact shape of tightly packed atoms. A large value of $R_{\min}$ also suggests that the central atom—one of the carbons in the backbone chain for the scheme adopted here—is in a highly symmetric or well-coordinated environment. For example, highly symmetrical molecules such as benzene or fullerenes include atoms that have strong and relatively uniform interactions with multiple neighboring atoms, leading to large radii of the Gershgorin circles. For quantum kernels produced by the molecule $\rightarrow$ QC mapping, this implies that the performance of QSVM is enhanced by increasing the density of the gate layout. These circuits, characterized by uniformly distributed gates, exhibit greater expressivity and improved performance in classification tasks due

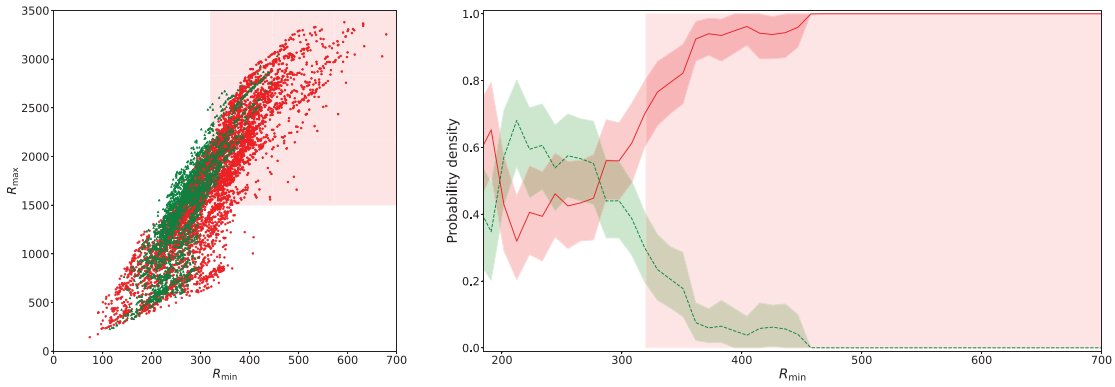


Figure 6. Left panels: distributions of performant quantum circuits (QCs, red circles) and underperforming QCs (green triangles) in the space of the smallest R_{min} and largest R_{max} Gershgorin radii for the 10D MNIST classification problem. Right panels: probability density of performant (solid red curves) and underperforming (broken green curves) QCs as functions of R_{min}, computed via kernel density estimation (KDE) from the data in the left panels. The shaded regions represent the 95% confidence intervals, accounting for uncertainty due to finite sample size.

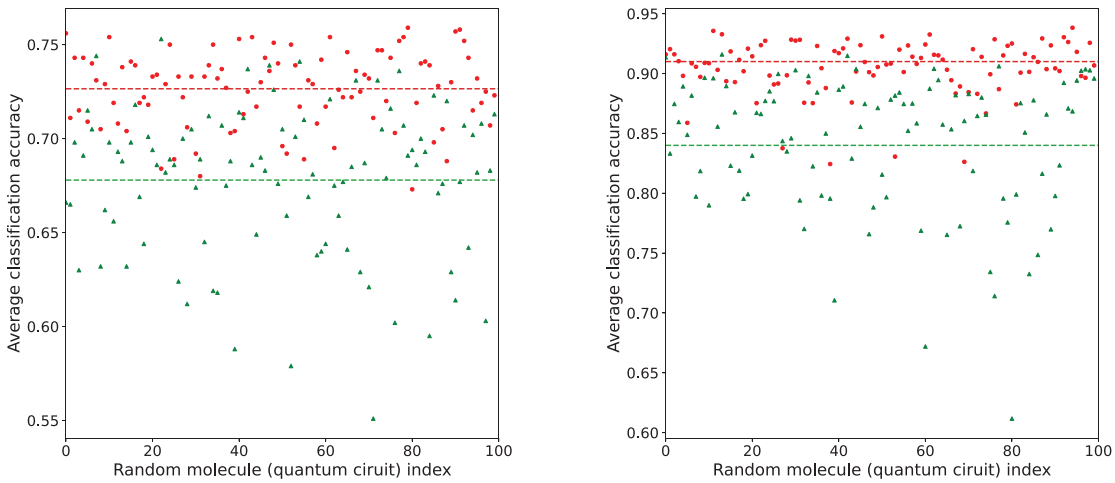


Figure 7. Distribution of average QSVM accuracy for the 5D hidden-manifold (left) and 5D MNIST (right) classification problems using five-layered QCs obtained from 10 000 randomly generated molecules. The red circles represent a subset of 100 molecules with the largest radius of the smallest Gershgorin circle, and the green triangles represent a subset of 100 molecules with the smallest radius of the smallest Gershgorin circle. The dashed lines show the $\pm 10\%$ margin centered at the average of the minimum and maximum classification accuracy in the entire sample of 10 000 quantum circuits.

to increased entanglement and gate interactions. This correlation suggests that the spherical, evenly connected architecture of such molecular graphs enables robust feature representation in the corresponding circuits, facilitating capture of complex patterns in quantum computation.

To investigate the impact of atom set selection on the representational power of QCs generated from molecular structures, we analyze the effect of varying atom types on the separation of performant and underperforming QCs in the space of molecular descriptors. As shown in figure 8, mapping the permutations of CNOT(i) gates to different sets of atoms—(1) {N, O, S}, (2) {N, P, O}, (3) {O, S, P}, (4) {N, S, P}—do not significantly alter the separation behavior. A large R_{min} indicates a central carbon atom in a highly symmetric, spherical, or well-coordinated molecular environment, with strong, uniform interactions among neighboring atoms, which maps to circuits with high gate density and robust feature-space coverage. This symmetry-driven property ensures that the representational power of the QCs remains consistent across different atom type choices.

Although in the simplified mapping scheme we only consider Z-rotation and CNOT gates, the molecular descriptor framework provides guidance at the level of circuit connectivity rather than for specific gate identities. Changing the ratio of single-qubit to multi-qubit gates impacts the effective connectivity of the molecular graph, which is directly reflected in the Gershgorin radii of the corresponding Coulomb matrices. In particular, increasing the density of entangling gates generally leads to larger values of the smallest Gershgorin radius R_{min}, while sparse or highly localized gate layouts reduce R_{min}. Extensions of the gate set, such as the inclusion of T gates, can be accommodated by extending the atom

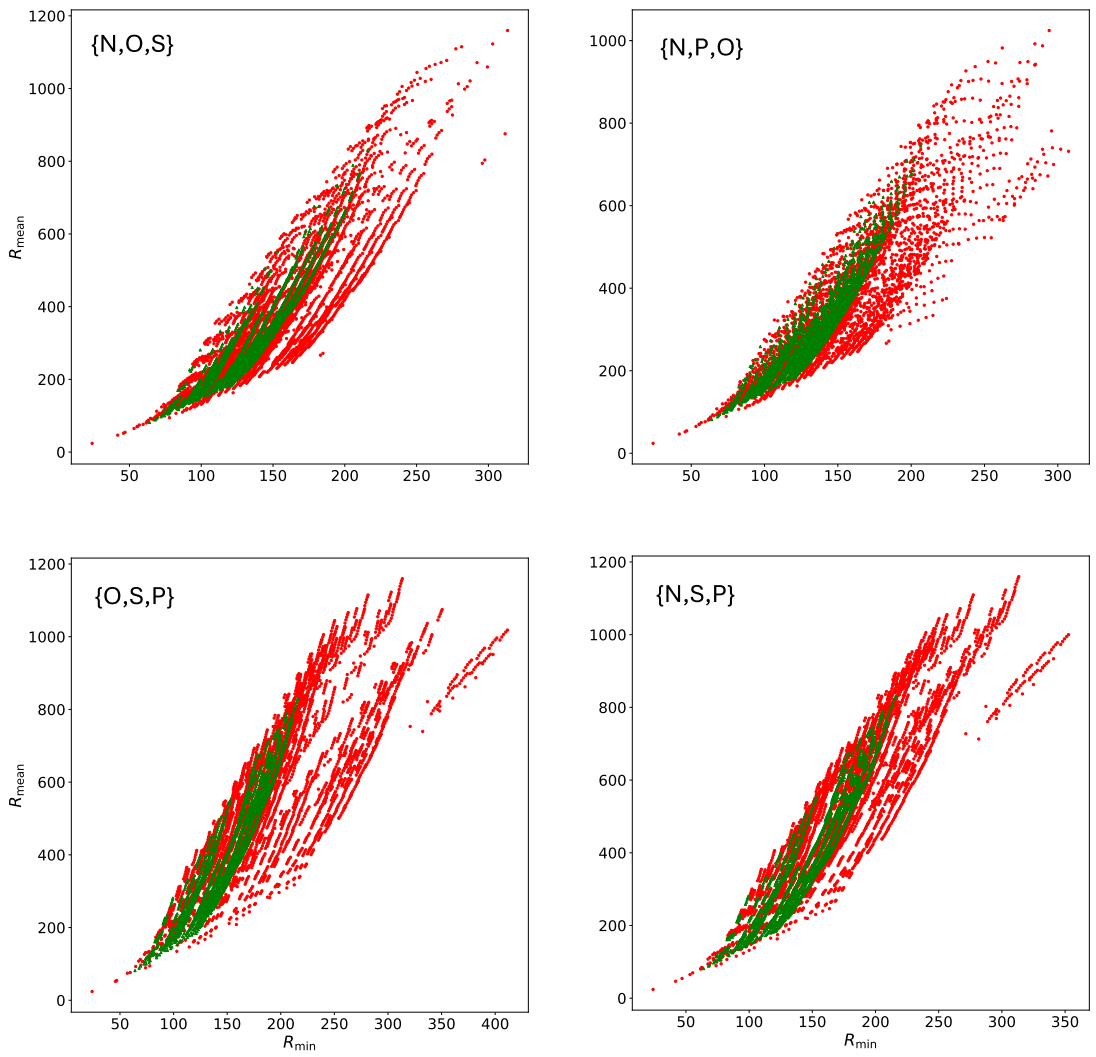


Figure 8. Distributions of performant QCs (red circles) and underperforming QCs (green triangles) in the space of the smallest R_{min} and mean R_{mean} of Gershgorin disks radii for the 4D hidden-manifold classification problem, corresponding to atom sets: (a) CNOT(i) → {N, O, S}, (b) CNOT(i) → {N, P, O}, (c) CNOT(i) → {O, S, P}, CNOT(i) →(d) {N, S, P}, from left to right panel.

set without altering the descriptor logic or the search-space restriction strategy. As a result, the proposed molecular descriptors provide gate-set-agnostic heuristics for designing expressive yet trainable QCs.

In the context of variational quantum algorithms, which are prone to BP, restricting the circuit search space through structured mappings provides an efficient way to circuit design. On one hand, by exploiting the molecule → QC isomorphism, we restrict the search space and consequently the distribution of QCs, avoiding random quantum gate selection and Haar random sampling, which are known to induce BP in highly expressive unstructured models [62]. This structured mapping promotes trainability by ensuring that gradients remain non-vanishing at moderate depths, as opposed to random unitaries that lead to exponential suppression. In addition, we observe that circuits with large R_{min} of molecular descriptors show better classification performance in QSVMs. For quantum kernels generated using this mapping, a large R_{min} indicates that QSVM performance is enhanced by denser gate layouts, allowing sufficient expressivity with moderate depth. Thus, this approach restricts the search space while balancing expressivity with shallow to moderate depth, ensuring that the trainable circuits are still expressive enough to capture relevant features.

These results indicate that molecular descriptors can provide meaningful structural guidance for QC design. Highly symmetric and well-coordinated molecular environments, characterized by large values of R_{min}, constrain the circuit search space toward architectures with balanced gate distributions, which are well suited for learning tasks requiring uniform coverage of the feature space. In contrast, molecular structures with increased branching map onto higher effective connectivity in the corresponding circuits, enabling stronger and more spatially distributed entanglement at shallow to moderate circuit depths.

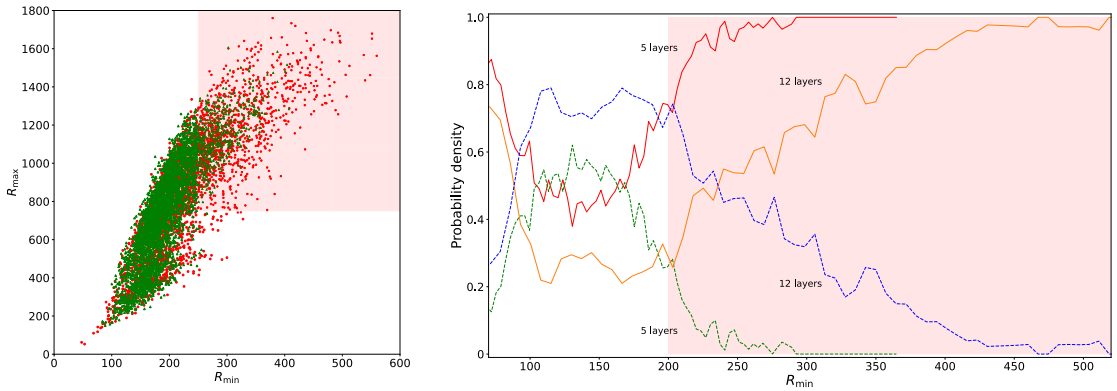


Figure 9. Left: same as in figure 3(lower left) but for QC with 12 layers. Right: probability density of performant (solid curves) and underperforming (broken curves) QC as functions of R_{min} computed for QC with 5 and 12 layers.

Taken together, molecular symmetry and branching provide complementary heuristics for limiting the circuit search space in a way that preserves both expressivity and trainability.

A second key goal of this work is to develop an algorithm for identifying QC performance metrics based on a set of shallow QC in order to design deep QC that yields high accuracy in the corresponding quantum model. The QC $\leftrightarrow$ molecule mapping is well suited for this purpose as molecules can be grown incrementally by increasing the length of the molecular branches. The shape of polyatomic molecules thus designed is determined by the shape of the parent molecules by construction. To confirm this by numerical examples, we extend the calculations in figure 3(lower panels) to QC with additional layers of quantum gates. The results of figure 3 are based on QC with five layers of gates, whereas figure 9 depicts the distribution of QC performance in the $(R_{\text{min}}, R_{\text{max}})$ space for QC with 12 layers. Figure 9 shows that the performance of deeper QC is enhanced in similar parts of the molecular parameter space. The comparison of figures 3 and 9 suggest the following approach: first, a distribution of efficient QC with a small number of gate layers is used to identify the range of R_{min} and R_{max} associated with a high performance of the quantum algorithm. This range of R_{min} and R_{max} can then be used to restrict the search of deeper QC that yield high performance.

The instability of QCs (e.g. vanishing gradients, decoherence, noise sensitivity) generally increases with circuit depth. Leveraging the molecule $\rightarrow$ QC isomorphism, molecular descriptors such as R_{min} can guide the design of moderately deep circuits. Circuits with 5–12 layers are considered shallow to moderately deep, providing a balance between mitigating noise accumulation and preserving sufficient expressivity for meaningful computations.

We have repeated the calculations in figures 1–9 using 2D molecular representations. The numerical tests indicate that using these representations produces similar distributions of performant and underperforming QCs in the space of Gershgorin circle descriptors for QCs with five and 12 layers, and for 4D- and 5D datasets. Since optimization of 3D molecular structures is computationally more time-consuming, these results indicate that 2D representations can be used as a less expensive yet accurate alternative for mapping QCs to molecules. Furthermore, these results suggest the possibility of determining molecular properties using quantum platforms better suited for 2D graphs [54, 55] as an alternative to experimental measurements or classical computations.

5. Conclusion

We have demonstrated the isomorphism between QCs and a subspace of polyatomic molecules, which suggests that molecules can be used as descriptors of QCs. This connection has several important implications. First, the numerical techniques for optimization of molecular properties in chemical compound spaces can be adopted for optimization of QCs for quantum algorithms. Thus, quantum ansatz optimization can benefit from advances in cheminformatics. In addition, a large number of descriptors developed for machine learning in chemistry can be applied to machine learning in QC spaces. In this work, we consider two qualitatively different molecular descriptors to characterize the performance of QCs.

We have shown that the performance of QC for QSVM can be characterized by PCA-reduced molecular fingerprints as well as by the size of the largest and smallest Gershgorin circles derived from the Coulomb matrices of the corresponding molecules. This can be used to restrict the search space for

the compositional optimization of QCs, as illustrated by figure 7 for two independent classification problems. We have shown that a high accuracy of a quantum algorithm can be achieved with high probability by sampling from a particular set of molecules.

The specific QC $\leftrightarrow$ molecule mapping proposed here provides an algorithm for increasing the depth of QCs, while improving the QC performance. Since qubits are encoded by different branches from a carbon polymer chain, large polyatomic molecules representing deep QCs inherit some properties of smaller molecules, corresponding to shallower QCs. This can be used to identify the range of molecular parameters yielding enhanced performance of a quantum algorithm with shallow circuits. The range of the molecular parameters thus determined can then be exploited to restrict the search of optimal QCs of larger depth.

The Gershgorin circles of Coulomb matrices effectively encode information about the distribution of atoms as determined by the molecular shape. Our results thus indicate that the shape of molecules used to construct QCs is important for the performance of the resulting quantum algorithm. More broadly, the present work suggests an approach to identify key elements that enhance the accuracy of a quantum algorithm by mapping QC onto molecules and exploring correlations between physical properties or atomic compositions of molecules and circuit performance.

Acknowledgment

This work was supported by NSERC of Canada.

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://github.com/ElhamTorabian/Molecular-representation-of-quantum-circuits [63].

ORCID iDs

Elham Torabian 0000-0002-3097-2519
Roman V Krems 0000-0002-9918-211X

References

[1] Kübler J, Buchholz S and Schölkopf B 2021 The inductive bias of quantum kernels *Advances in Neural Information Processing Systems* vol 34 p 12661
[2] McClean J R, Boixo S, Smelyanskiy V N, Babbush R and Neven H 2018 Barren plateaus in quantum neural network training landscapes *Nat. Commun.* **9** 4812
[3] Larocca M, Thanasilp S, Wang S, Sharma K, Biamonte J, Coles P J, Cincio L, McClean J R, Holmes Z and Cerezo M 2024 A review of barren plateaus in variational quantum computing (arXiv:2405.00781)
[4] Zhang H-K, Liu S and Zhang S-X 2024 Absence of barren plateaus in finite local-depth circuits with long-range entanglement *Phys. Rev. Lett.* **132** 150603
[5] Sack S H, Medina R A, Michailidis A A, Kueng R and Serbyn M 2022 Avoiding barren plateaus using classical shadows *PRX Quantum* **3** 020365
[6] Holmes Z, Arrasmith A, Yan B, Coles P J, Albrecht A and Sornborger A T 2021 Barren plateaus preclude learning scramblers *Phys. Rev. Lett.* **126** 190501
[7] Friedrich L and Maziero J 2022 Avoiding barren plateaus with classical deep neural networks *Phys. Rev. A* **106** 042433
[8] Kulshrestha A and Safro I 2022 Beinit: avoiding barren plateaus in variational quantum algorithms *Qce22* p 197
[9] Arrasmith A, Holmes Z, Cerezo M and Coles P J 2022 Equivalence of quantum barren plateaus to cost concentration and narrow gorges *Quantum Sci. Technol.* **7** 045015
[10] Thanasilp S, Wang S, Cerezo M and Holmes Z 2024 Exponential concentration in quantum kernel methods *Nat. Commun.* **15** 5200
[11] Kairon P, Jäger J and Krems R V 2025 Equivalence between exponential concentration in quantum machine learning kernels and barren plateaus in variational algorithms (arXiv:2501.07433)
[12] Liu Y, Arunachalam S and Temme K 2021 A rigorous and robust quantum speed-up in supervised machine learning *Nat. Phys.* **17** 1013
[13] Jäger J and Krems R V 2023 Universal expressiveness of variational quantum classifiers and quantum kernels for support vector machines *Nat. Commun.* **14** 576
[14] Torabian E and Krems R V 2023 Compositional optimization of quantum circuits for quantum kernels of support vector machines *Phys. Rev. Res.* **5** 013211
[15] Glick J R, Gujarati T P, Corcoles A D, Kim Y, Kandala A, Gambetta J M and Temme K 2024 Covariant quantum kernels for data with group structure *Nat. Phys.* **20** 479
[16] Shaydulin R and Wild S M 2022 Importance of kernel bandwidth in quantum machine learning *Phys. Rev. A* **106** 042407
[17] Suzuki Y, Kawaguchi H and Yamamoto N 2024 Quantum Fisher kernel for mitigating the vanishing similarity issue *Quantum Sci. Technol.* **9** 035050
[18] Huang H-Y, Broughton M, Mohseni M, Babbush R, Boixo S, Neven H and McClean J R 2021 Power of data in quantum machine learning *Nat. Commun.* **12** 2631

[19] Randić M 1977 On canonical numbering of atoms in a molecule and graph isomorphism *J. Chem. Inf. Comput. Sci.* **17** 171
[20] Merkys A, Vaitkus A, Grybauskas A, Konovalovas A, Quirós M and Gražulis S 2023 Graph isomorphism-based algorithm for cross-checking chemical and crystallographic descriptions *J. Cheminform.* **15** 25
[21] Qiskit Community 2025 Qiskit transpiler documentation (available at: https://docs.quantum.ibm.com/api/qiskit/transpiler)
[22] Riu J, Nogué J, Vilaplana G, Garcia-Saez A and Estarellas M P 2025 Reinforcement learning based quantum circuit optimization via ZX-calculus *Quantum* **9** 1758
[23] Putin E, Asadulaev A, Ivanenkov Y, Aladinskiy V, Sanchez-Lengeling B, Aspuru-Guzik A and Zhavoronkov A 2018 Reinforced adversarial neural computer for de novo molecular design *J. Chem. Inf. Model.* **58** 1194
[24] Putin E, Asadulaev A, Vanhaelen Q, Ivanenkov Y, Aladinskaya A V, Aliper A and Zhavoronkov A 2018 Adversarial threshold neural computer for molecular de Novo design *Mol. Pharmaceutics* **15** 4386
[25] Maziarka Ł, Pocha A, Kaczmarczyk J, Rataj K, Danel T and Warchoł M 2020 Mol-CycleGAN: a generative model for molecular optimization *J. Cheminform.* **12** 1
[26] Noguchi S and Inoue J 2022 Exploration of chemical space guided by PixelCNN for fragment-based de novo drug discovery *J. Chem. Inf. Model.* **62** 5988
[27] Peng S-P and Zhao Y 2019 Convolutional neural networks for the design and analysis of non-fullerene acceptors *J. Chem. Inf. Model.* **59** 4993
[28] Griffiths R-R and Hernández-Lobato J M 2020 Constrained Bayesian optimization for automatic chemical design using variational autoencoders *Chem. Sci.* **11** 577
[29] Kristiadi A, Strieth-Kalthoff F, Skreta M, Poupart P, Aspuru-Guzik A and Pleiss G 2024 A sober look at LLMs for material discovery: are they actually good for Bayesian optimization over molecules? (arXiv:2402.05015)
[30] Wang X, Jin Y, Schmitt S and Olhofer M 2023 Recent advances in Bayesian optimization *ACM Comput. Surv.* **55** 1–36
[31] Chen Y and Kirchmair J 2020 Cheminformatics in natural product-based drug discovery *Mol. Inf.* **39** 2000171
[32] David L, Thakkar A, Mercado R and Engkvist O 2020 Molecular representations in AI-driven drug discovery: a review and practical guide *J. Cheminform.* **12** 1
[33] Kitaev A Y, Shen A and Vyalyi M N 2002 *Classical and Quantum Computation* (American Mathematical Soc.) p 47
[34] Nielsen M A and Chuang I L 2010 *Quantum Computation and Quantum Information* (Cambridge University Press)
[35] Williams C P 2011 Explorations in quantum computing *Texts in Computer Science* (Springer) p 51
[36] Aharonov D 1999 Quantum computation *Annu. Rev. Comput. Phys. VI* 259–346
[37] Aaronson S and Gottesman D 2004 Improved simulation of stabilizer circuits *Phys. Rev. A* **70** 052328
[38] Fellner M, Messinger A, Ender K and Lechner W 2022 Universal parity quantum computing *Phys. Rev. Lett.* **129** 180503
[39] Klaver B, Rombouts S, Fellner M, Messinger A, Ender K, Ludwig K and Lechner W 2024 SWAP-less implementation of quantum algorithms (arXiv:2408.10907)
[40] Rogers D and Hahn M 2010 Extended-connectivity fingerprints *J. Chem. Inf. Model.* **50** 742
[41] Kipf T N and Welling M 2016 Semi-supervised classification with graph convolutional networks (arXiv:1609.02907)
[42] Wu Z, Pan S, Chen F, Long G, Zhang C and Yu P S 2020 A comprehensive survey on graph neural networks *IEEE Trans. Neural Netw. Learn. Syst.* **32** 4
[43] Rupp M, Tkatchenko A, Müller K-R and Von Lilienfeld O A 2012 Fast and accurate modeling of molecular atomization energies with machine learning *Phys. Rev. Lett.* **108** 058301
[44] Hansen K, Biegler F, Ramakrishnan R, Pronobis W, Von Lilienfeld O A, Muller K-R and Tkatchenko A 2015 Machine learning predictions of molecular properties: accurate many-body potentials and nonlocality in chemical space *J. Phys. Chem. Lett.* **6** 2326
[45] Bartók A P, Kondor R and Csányi G 2013 On representing chemical environments *Phys. Rev. B* **87** 184115
[46] Christensen A S, Bratholm L A, Faber F A and Anatole von Lilienfeld O 2020 FCHL revisited: faster and more accurate quantum machine learning *J. Chem. Phys.* **152** 044107
[47] Mao Y-W and Krems R V 2024 Efficient interpolation of molecular properties across chemical compound space with low-dimensional descriptors *Mach. Learn.: Sci. Technol.* **5** 015059
[48] I. Daylight Chemical Information Systems 2008 Daylight theory manual (available at: www.daylight.com)
[49] Landrum G 2023 RDKit documentation (available at: www.rdkit.org)
[50] Schrier J 2020 Can one hear the shape of a molecule (from its Coulomb matrix eigenvalues)? *J. Chem. Inf. Model.* **60** 3804
[51] Rappé A K, Casewit C J, Colwell K, Goddard W A III and Skiff W M 1992 UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations *J. Am. Chem. Soc.* **114** 10024
[52] Kamada T and Kawai S 1989 An algorithm for drawing general undirected graphs *Inf. Process. Lett.* **31** 7
[53] Jász A, Rák A, Ladjánszki I and Cserey G 2019 Optimized GPU implementation of Merck molecular force field and universal force field *J. Mol. Struct.* **1188** 227
[54] Scholl P *et al* 2021 Quantum simulation of 2D antiferromagnets with hundreds of Rydberg atoms *Nature* **595** 233
[55] Ebadi S *et al* 2022 Quantum optimization of maximum independent set using Rydberg atom arrays *Science* **376** 1209
[56] D'Arcangelo M, Henry L-P, Henriet L, Loco D, Gouraud N, Angebault S, Sueiro J, Forêt J, Monmarché P and Piquemal J-P 2024 Leveraging analog quantum computing with neutral atoms for solvent configuration prediction in drug discovery *Phys. Rev. Res.* **6** 043020
[57] Maskara N *et al* 2025 Programmable simulations of molecules and materials with reconfigurable quantum processors *Nat. Phys.* **21** 289–97
[58] Shahriari B, Swersky K, Wang Z, Adams R P and De Freitas N 2015 Taking the human out of the loop: a review of Bayesian optimization *Proc. IEEE* **104** 148
[59] Jain A *et al* 2013 Commentary: the materials project: a materials genome approach to accelerating materials innovation *APL Mater.* **1** 011002
[60] Bowles J, Ahmed S and Schuld M 2024 Better than classical? The subtle art of benchmarking quantum machine learning models (arXiv:2403.07059)
[61] LeCun Y 1998 The MNIST database of handwritten digits (available at: http://yann.lecun.com/exdb/mnist/)
[62] Cunningham J and Zhuang J 2025 Investigating and mitigating barren plateaus in variational quantum circuits: a survey *Quantum Inf. Process.* **24** 48
[63] Torabian E 2025 *Molecular representation of quantum circuits* (available at: https://github.com/ElhamTorabian/Molecular-representation-of-quantum-circuits)