

Best-approximation error for parametric quantum circuits

Lena Funcke

Perimeter Institute for Theoretical Physics
Waterloo, ON, Canada
lfuncke@perimeterinstitute.ca

Tobias Hartung

Department of Mathematical Sciences
University of Bath
Bath, United Kingdom
th2040@bath.ac.uk

Karl Jansen

NIC, DESY Zeuthen
Zeuthen, Germany
Karl.Jansen@desy.de

Stefan Kühn

Computation-Based Science and Technology Research Center
The Cyprus Institute
Nicosia, Cyprus
s.kuehn@cyi.ac.cy

Manuel Schneider

NIC, DESY Zeuthen
Zeuthen, Germany
manuel.schneider@desy.de

Paolo Stornati

NIC, DESY Zeuthen
Zeuthen, Germany
paolo.stornati@desy.de

Abstract—In Variational Quantum Simulations, the construction of a suitable parametric quantum circuit is subject to two counteracting effects. The number of parameters should be small for the device noise to be manageable, but also large enough for the circuit to be able to represent the solution. Dimensional expressivity analysis can optimize a candidate circuit considering both aspects. In this article, we will first discuss an inductive construction for such candidate circuits. Furthermore, it is sometimes necessary to choose a circuit with fewer parameters than necessary to represent all relevant states. To characterize such circuits, we estimate the best-approximation error using Voronoi diagrams. Moreover, we discuss a hybrid quantum-classical algorithm to estimate the worst-case best-approximation error, its complexity, and its scaling in state space dimensionality. This allows us to identify some obstacles for variational quantum simulations with local optimizers and underparametrized circuits, and we discuss possible remedies.

Index Terms—parametric quantum circuits, dimensional expressivity analysis, best-approximation error, variational quantum simulations, Voronoi diagrams

I. INTRODUCTION

Noisy intermediate-scale quantum (NISQ) computers [1] are opening up a new avenue to address a large class of computational problems that cannot be solved efficiently with classical computers. Applications of quantum computing range from machine learning [2] to finance [3] to various optimization problems [4], [5]. In physics, quantum computers intrinsically circumvent the sign problem that prevents Monte Carlo simulations of strongly correlated quantum-many body problems in certain parameter regimes [6]. Although current hardware is of limited size and suffers from a considerable level of noise, the

ability for NISQ devices to outperform classical computers has already been demonstrated successfully [7], [8] and techniques for mitigating the effects of noise are rapidly developing [9]–[13].

Many algorithms designed for NISQ devices make use of parametric quantum circuits. Variational quantum simulations (VQSs) [14], [15], a class of hybrid quantum-classical algorithms for solving optimization problems, are a particularly important example. Using a parametric quantum circuit, i.e., a quantum circuit composed of parameter dependent gates, a cost function is evaluated efficiently for a given set of parameters using the quantum coprocessor. The cost function is then minimized on a classical computer in a feedback loop based on the measurement outcome obtained from the quantum device. Using cost functions related to the energy of the quantum state prepared on the quantum device, VQSs have been successfully applied to quantum many-body systems in quantum chemistry [10], [14], [16], [17] and even quantum mechanics and quantum field theory [18]–[22].

Since VQSs depend on the choice of parametric quantum circuits, there are many open questions related to finding a good or optimal quantum circuit. For example, in order to be able to find the solution – or at least find a good approximation – a parametric quantum circuit needs to have many parameters. However, many parameters means many gates and thus large noise. One measure for a “good” quantum circuit is therefore to have as many parameters as necessary while being able to parametrize the entire state space of the simulated model. An optimal circuit taking this point of view would be minimal (there are no redundant parameters) and maximally expressive (the circuit can generate all relevant states). Parametric quantum circuits can be analyzed from this point of view using dimensional expressivity analysis (DEA) [23] which we will review in section II. However, DEA can only tell us whether or not a given circuit is minimal and maximally expressive. The construction of a maximally

Research at Perimeter Institute is supported in part by the Government of Canada through the Department of Innovation, Science and Industry Canada and by the Province of Ontario through the Ministry of Colleges and Universities. S.K. acknowledges financial support from the Cyprus Research and Innovation Foundation under project “Future-proofing Scientific Applications for the Supercomputers of Tomorrow (FAST)”, contract no. COMPLEMENTARY/0916/0048. M.S. and P.S. thank the Helmholtz Einstein International Berlin Research School in Data Science (HEIBRiDS) for funding.

expressive circuit is still highly non-trivial even though DEA does provide some guiding information. Additionally, having a maximally expressive circuit may not always be necessary or prudent for a given application, and DEA as described in [23] does not provide any information on how useful a non-maximally expressive circuit is.

In this article, we will expand on the DEA by providing a method of constructing a minimal, maximally expressive circuit on $N+1$ qubits given a minimal, maximally expressive circuit on N qubits in section III.

From section IV onwards we will move to non-maximally expressive circuits. These, by definition, cannot represent arbitrary states and, thus, the best case scenario is for a VQS to reproduce the best-approximation of the solution. In section IV we will define the necessary notation to discuss the worst-case error of the best-approximation. In particular, we will discuss Voronoi diagrams [24], [25] to estimate the best-approximation error. Given a set of points p_i , a Voronoi diagram is a partition of space into regions. Each region is associated to one of the points p_i and contains all points x that are closer to p_i than any other p_j . In other words, the Voronoi region associated with p_i is the set of points whose best-approximation is p_i . Knowing the Voronoi diagrams therefore allows us to compute the worst-case best-approximation error which is the largest distance a point x can be from its corresponding best-approximation p_i , i.e., the largest possible distance between x and p_i where x is an arbitrary point in the Voronoi region of p_i . This leads to an upper bound for the best-approximation error of a parametric quantum circuit by choosing the points p_i to be a discrete set of sample states the given parametric quantum circuit can generate.

In section V we will consider a minimal, maximally expressive circuit on the Bloch sphere as an example for the behavior of the here proposed best-approximation error estimates if applied to a maximally expressive circuit. A first example of a non-maximally expressive circuit will be provided in section VI. This example also highlights a potential problem if the best-approximation error is computed naïvely. In section VII we will expand on this potential problem with some practical remarks. In particular, we will propose splitting the analysis into two steps. The first step will provide conditions that allow us to circumvent the obstacles identified in section VI and provide a large lower bound on the best-approximation error should these conditions not be fulfilled. In the second step, once these conditions are fulfilled, the detailed analysis described in section IV is applicable without running into the potential problems observed in section VI. In section VIII we will discuss the computational complexity of the best-approximation error estimate, in section IX we propose an efficient mapping procedure that allows for exponential speed-up using a quantum device, and in section X we will discuss the scaling with respect to dimensionality. In section XI we will describe a relationship between the best-approximation error and the internal volume of the circuit image. This relationship will be illustrated with an example in section XII, and we will discuss its impact on VQSS in section XIII.

Finally, we conclude the article in section XIV.

II. DIMENSIONAL EXPRESSIVITY ANALYSIS

Within the context of dimensional expressivity analysis [23], we consider a parametric quantum circuit to be a continuously differentiable map $C : \mathcal{P} \rightarrow \mathcal{S}$. Here, \mathcal{P} is the parameter space and assumed to be a (usually compact) manifold without boundary. For example, if only rotation gates $R_G(\vartheta) = \exp(-i\frac{\vartheta}{2}G)$ for gates G with $G^2 = 1$ are used, then \mathcal{P} is the flat torus $(\mathbb{R}/2\pi\mathbb{Z})^N$, where N is the number of parameters in the circuit. \mathcal{S} is the “relevant” state space, i.e., a submanifold of the unit sphere in the complex 2^Q -dimensional Hilbert space for a Q -qubit quantum device.

The primary objective of the DEA is to identify redundant parameters in a given circuit C . Once these redundant parameters are identified, the parameter space \mathcal{P} can be restricted by setting all redundant parameters to constants, and the thus restricted circuit is locally surjective to the image \mathcal{M} of the original (unrestricted) circuit. This furthermore allows us to prove that \mathcal{M} is locally a manifold and that the restricted circuit is minimal, i.e., removal of any further parameters (by setting them constant) will result in a restriction of the image \mathcal{M} .

DEA identifies independent and redundant parameters inductively. The first parameter ϑ_1 is always independent (unless the circuit is invariant under change of ϑ_1 which is assumed to be not the case). Assuming that $\vartheta_1, \dots, \vartheta_k$ are already identified as independent, we need to check ϑ_{k+1} . The parameter ϑ_{k+1} is redundant if and only if the change in $C(\vartheta)$ given a perturbation of ϑ_{k+1} can also be achieved keeping ϑ_{k+1} fixed and varying $\vartheta_1, \dots, \vartheta_k$ appropriately. In other words, ϑ_{k+1} is redundant if and only if the tangent vector $\partial_{\vartheta_{k+1}}C(\vartheta)$ is a linear combination of the tangent vectors $\partial_{\vartheta_1}C(\vartheta), \dots, \partial_{\vartheta_k}C(\vartheta)$. Since our parameter space \mathcal{P} is a real manifold, this linear combination is to be taken with respect to real coefficients. To test this linear independence, we consider the matrix

$$S_{k+1} = (\Re\langle\partial_{\vartheta_m}C(\vartheta), \partial_{\vartheta_n}C(\vartheta)\rangle)_{1 \leq m, n \leq k+1}. \quad (1)$$

This matrix is invertible if and only if ϑ_{k+1} is independent. If ϑ_{k+1} is found to be redundant (i.e., S_{k+1} is not invertible), then ϑ_{k+1} is removed as a parameter (i.e., it is kept constant) and the remaining parameters are tested after possibly re-labeling $\vartheta_j \mapsto \vartheta_{j-1}$ for $j > k+1$ in the redundant case.

Furthermore, the matrix S_{k+1} can be measured efficiently using a quantum device. If all parametric gates are rotation gates $R_{G_n}(\vartheta_n) = \exp(-i\frac{\vartheta_n}{2}G_n)$ for some gate G_n (which can be non-trivial like CCNOT or $X_{q_1}X_{q_2}Y_{q_3}$, i.e., two X gates acting on qubits q_1 and q_2 and a Y gate acting on qubit q_3 simultaneously), then $\partial_{\vartheta_n}C(\vartheta) = -\frac{i}{2}\gamma_n$, where γ_n is a quantum circuit with the additional gate G_n inserted after the application of the gate $R_{G_n}(\vartheta_n)$. Hence, we obtain

$$\Re\langle\partial_{\vartheta_m}C(\vartheta), \partial_{\vartheta_n}C(\vartheta)\rangle = \frac{1}{4}\Re\langle\gamma_m, \gamma_n\rangle, \quad (2)$$

and $\Re\langle\gamma_m, \gamma_n\rangle$ can be measured on the quantum device.

Following the algorithm outlined in section 6 of [23], we add an ancilla qubit, which we initialize in $\frac{|0\rangle+|1\rangle}{\sqrt{2}}$. We then apply the gate sequence of our parametric quantum circuit but after application of R_{G_m} and R_{G_n} we insert the gates $X_{\text{anc}}CG_mX_{\text{anc}}$ and CG_n , respectively. Here, CG_j denotes a controlled G_j gate that uses the ancilla as control, and the pair of X gates on the ancilla (X_{anc}) ensures that CG_m is conditioned on the ancilla being in $|0\rangle$. Finally, we apply another Hadamard gate to the ancilla. The state of the quantum device after all these steps is

$$\frac{|0\rangle \otimes (|\gamma_m\rangle + |\gamma_n\rangle) + |1\rangle \otimes (|\gamma_m\rangle - |\gamma_n\rangle)}{2}, \quad (3)$$

and measuring the ancilla qubit yields the probability of finding the ancilla in $|0\rangle$

$$\text{prob}(\text{anc} = 0) = \frac{1 + \Re\langle\gamma_m, \gamma_n\rangle}{2}. \quad (4)$$

This algorithm of testing linear independence requires $\mathcal{O}((k+1)^2\epsilon^{-2})$ quantum device calls, where ϵ is the acceptable level of uncertainty for the measurement $\text{prob}(\text{anc} = 0)$ (a result of quantum device noise and finite statistics). The test for invertibility is classical and scales like $\mathcal{O}((k+1)^3)$, and the entire process needs to be repeated for $2 \leq k+1 \leq N$. Hence, this independence test algorithm scales like $\mathcal{O}(N^4\epsilon^{-2})$. On the quantum device it also requires an additional ancilla qubit and six additional gates.

It is important to note that this algorithm can be extended to compute any $\Re\langle C_1, C_2 \rangle$ for circuits C_1 and C_2 by controlling all gates in C_1 and C_2 on an ancilla and ensuring that C_1 is conditioned on the ancilla being in $|0\rangle$ and C_2 is conditioned on the ancilla being in $|1\rangle$. This will become important in section VIII, section IX, and section X.

III. CONSTRUCTION OF MINIMAL, MAXIMALLY EXPRESSIVE CIRCUITS ON $Q + 1$ QUBITS

DEA as described in section II requires a candidate circuit to be optimized. In many cases, finding such a candidate circuit is not an easy task. However, once a candidate is found, even if it is “too expressive” in the sense that it can generate states that are not physically meaningful, DEA provides a powerful tool. In particular, DEA can be used to remove symmetries such as invariance under a global phase transformation or even gauge invariance. In this section, we therefore want to discuss a means of generating candidate circuits. In order for this construction to be sufficiently general, we want the candidate circuits to be maximally expressive on Q qubits. Furthermore, for the method to be efficient, we want the candidate circuits to be minimal in the number of parameters. Thus, if we can construct a minimal, maximally expressive circuit with an image manifold that still contains symmetries we wish to remove, then DEA can be used for the symmetry removal, while not having to remove any additional redundant parameters, which keeps the entire circuit construction process streamlined.

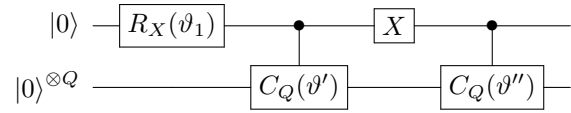
Such streamlining can be achieved using a minimal, maximally expressive circuit for the entire Q -qubit quantum device

state space. Of course, further streamlining can be achieved if some of the unwanted symmetries can already be removed at this stage, but in the most general case, in which all symmetries are to be removed via DEA, a minimal, maximally expressive circuit for the entire device state space is the optimal starting point. Since the Q -qubit state space is the unit sphere in the complex 2^Q -dimensional Hilbert space, any minimal, maximally expressive circuit must have $2^{Q+1} - 1$ real parameters. Starting with a single qubit, a minimal, maximally expressive circuit C_1 is given by

$$C_1(\vartheta) = R_Y(\vartheta_3)R_Z(\vartheta_2)R_X(\vartheta_1)|0\rangle. \quad (5)$$

The fact that this circuit is minimal and maximally expressive can be checked using DEA [23].

Given a Q -qubit minimal, maximally expressive circuit C_Q , we want to construct a minimal, maximally expressive circuit C_{Q+1} on $Q + 1$ qubits. The circuit $C_{Q+1}(\vartheta_1, \vartheta', \vartheta'')$ can be constructed by controlling the circuit C_Q on the newly added qubit and considering



where ϑ' and ϑ'' are a set of $2^{Q+1} - 1$ parameters each. The new circuit C_{Q+1} therefore has $2^{Q+2} - 1$ parameters, i.e., it is minimal if and only if it is maximally expressive.

To show maximal expressivity of C_{Q+1} , we recall that any $(Q + 1)$ -qubit state $|\psi\rangle$ can be expressed as

$$|\psi\rangle = \cos \frac{\vartheta_0}{2} |0\rangle \otimes |\psi_0\rangle + \sin \frac{\vartheta_0}{2} |1\rangle \otimes |\psi_1\rangle \quad (6)$$

with Q -qubit states $|\psi_0\rangle$ and $|\psi_1\rangle$. Through direct computation, we obtain

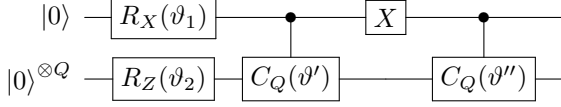
$$\begin{aligned} C_{Q+1}(\vartheta_1, \vartheta', \vartheta'') \\ = \cos \frac{\vartheta_1}{2} |1\rangle \otimes C_Q(\vartheta'') - i \sin \frac{\vartheta_1}{2} |0\rangle \otimes C_Q(\vartheta'), \end{aligned} \quad (7)$$

and choosing $\vartheta_1 = \pi - \vartheta_0$, $-iC_Q(\vartheta') = |\psi_0\rangle$, and $C_Q(\vartheta'') = |\psi_1\rangle$ shows that C_{Q+1} is maximally expressive on $Q+1$ qubits if C_Q is maximally expressive on Q qubits (which we assumed by induction).

Finally, we need to consider the controlled circuit CC_Q . By induction, we assume that C_Q contains only single-qubit gates and CNOT gates. Then, CC_Q contains only controlled single-qubit gates and Toffoli gates. Controlled single-qubit gates can be constructed from CNOT and single-qubit gates, and the Toffoli gate can be constructed from R_Y and CNOT gates [26]. Hence, C_{Q+1} can be constructed using only CNOT and single-qubit gates.

It should also be noted that this construction can be extended to automatically remove some unwanted symmetries. For example, two states on the quantum device differing only by a phase factor are equivalent. Therefore, we may wish to remove global phase symmetry from the set of states the proposed circuit can generate. On a single qubit, such a circuit without global phase symmetry is

given by $C_1(\vartheta) = R_Z(\vartheta_2)R_X(\vartheta_1)|0\rangle$. Similarly, this global phase symmetry can be incorporated into a circuit if we need it. To illustrate this, let us consider a Q -qubit circuit $C_Q(\vartheta) = U(\vartheta)|\text{init}\rangle$ and consider a gate set U_{init} such that $|\text{init}\rangle = U_{\text{init}}|0\rangle$. Then, we can incorporate global phase symmetry into C_Q by adding an additional parameter φ and considering the circuit $\tilde{C}_Q(\varphi, \vartheta) = U(\vartheta)U_{\text{init}}R_Z(\varphi)U_{\text{init}}^*|\text{init}\rangle$ where the R_Z gate may act on any qubit. Choosing the Q -qubit initial state $|\text{init}\rangle$ to be $|0\rangle$, we can adapt the construction described above by considering the $(Q+1)$ -qubit circuit



which generates the state

$$C_{Q+1}(\vartheta_1, \vartheta_2, \vartheta', \vartheta'') = \cos \frac{\vartheta_1}{2} |1\rangle \otimes C_Q(\vartheta'') - i \sin \frac{\vartheta_1}{2} |0\rangle \otimes \tilde{C}_Q(\vartheta_2, \vartheta'). \quad (8)$$

This circuit C_{Q+1} does not have a global phase symmetry as the choice of phase factor is fixed by $\cos \frac{\vartheta_1}{2} |1\rangle \otimes C_Q(\vartheta'')$.

IV. BEST-APPROXIMATION ERROR

In section III we have shown how to generate minimal, maximally expressive circuits on arbitrarily many qubits. Should we wish to remove symmetries from the image \mathcal{M} , we can achieve this using DEA as well. This leaves us with the problem that sometimes maximally expressive circuits may not be feasible experimentally. In such circumstances, we want to use a circuit that is no longer maximally expressive and therefore may not be able to represent the solution of a given VQS. In this case, we want to estimate the worst-case error of a best-approximation for a given state using a given non-maximally expressive circuit.

Let us consider a parametric quantum circuit $C : \mathcal{P} \rightarrow \mathcal{S}$ mapping a parameter space \mathcal{P} into a state space \mathcal{S} that is no longer necessarily the entire unit sphere as we discussed in section III. Let $d : \mathcal{S}^2 \rightarrow \mathbb{R}_{\geq 0}$ be a suitable distance in \mathcal{S} , e.g., the orthodromic (great-circle) distance as induced by the geometry of the quantum device state space, and define the distance of a point in \mathcal{S} to \mathcal{M} as

$$d_{\mathcal{M}} : \mathcal{S} \rightarrow \mathbb{R}_{\geq 0}; x \mapsto \inf_{\vartheta \in \mathcal{P}} d(x, C(\vartheta)). \quad (9)$$

Thus, the quantum circuit C is dense, i.e., it has a dense image, if and only if $\sup_{x \in \mathcal{S}} d_{\mathcal{M}}(x) = 0$. We note that this is only possible if \mathcal{P} is non-compact or C is maximally expressive. As \mathcal{P} is commonly a compact manifold, we note in such cases

$$d_{\mathcal{M}}(x) = 0 \iff \exists \vartheta \in \mathcal{P} : x = C(\vartheta). \quad (10)$$

In other words, if \mathcal{P} is compact, then C being dense is equivalent to C being maximally expressive. Furthermore, if \mathcal{P} is compact and C is not maximally expressive, then

$$\alpha_C := \sup_{x \in \mathcal{S}} d_{\mathcal{M}}(x) \quad (11)$$

is strictly positive and denotes the largest distance that a point in \mathcal{S} can be separated from \mathcal{M} . α_C is therefore the best-approximation error of the circuit C in the sense that

- the best-approximation $C(\vartheta_x)$ of a given point $x \in \mathcal{S}$ satisfies $d(x, C(\vartheta_x)) \leq \alpha_C$ and
- for all α with $0 < \alpha < \alpha_C$ there exists an $x \in \mathcal{S}$ such that the best-approximation $C(\vartheta_x)$ of x satisfies $d(x, C(\vartheta_x)) > \alpha$.

Hence, α_C is a measure of how good a given quantum circuit C approximates the state space \mathcal{S} .

In order to find an upper bound on α_C , let us approximate \mathcal{M} with a discrete point set \mathcal{D} . For example, we can choose $k \in \mathbb{N}$ sufficiently large, uniformly sample values $\vartheta_1, \dots, \vartheta_k \in \mathcal{P}$, and set $\mathcal{D} := \{C(\vartheta_1), \dots, C(\vartheta_k)\}$. We say that \mathcal{D} is an ε -dense approximation of \mathcal{M} if and only if $\sup_{x \in \mathcal{M}} \inf_{y \in \mathcal{D}} d(x, y) < \varepsilon$. Choosing an ε -dense approximation \mathcal{D} , we obtain

$$\alpha_C \leq \sup_{x \in \mathcal{S}} \inf_{y \in \mathcal{D}} d(x, y) \leq \alpha_C + \varepsilon. \quad (12)$$

Our objective is therefore to compute $\sup_{x \in \mathcal{S}} \inf_{y \in \mathcal{D}} d(x, y)$ as an estimate for α_C .

Let δ be a point in \mathcal{D} . The Voronoi region R_δ corresponding to δ is defined as the set

$$R_\delta := \left\{ x \in \mathcal{S}; d(x, \delta) = \min_{\delta' \in \mathcal{D}} d(x, \delta') \right\}. \quad (13)$$

Furthermore, let V_δ be the set of vertices of R_δ . Then,

$$\sup_{x \in \mathcal{S}} \inf_{y \in \mathcal{D}} d(x, y) = \max_{\delta \in \mathcal{D}} \max_{v \in V_\delta} d(\delta, v) \quad (14)$$

can be checked in finitely many steps and provides the estimate for α_C we were looking for.

V. A FULL BLOCH SPHERE EXAMPLE

As an initial example, let us consider a circuit that is maximally expressive for the Bloch sphere. Given that the circuit in this section will be maximally expressive, this example should reproduce $\alpha_C = 0$ and highlight the optimal behavior of the analysis.

In this section, we will choose the representation of vectors $|\psi\rangle$ on the Bloch sphere

$$|\psi\rangle = \cos \frac{\Theta}{2} |0\rangle + e^{i\Phi} \sin \frac{\Theta}{2} |1\rangle \quad (15)$$

with $0 \leq \Theta \leq \pi$ and $0 \leq \Phi < 2\pi$, as well as, the circuit

$$C(\vartheta) = R_Z(\vartheta_2)R_Y(\vartheta_1)|0\rangle \cong \cos \frac{\vartheta_1}{2} |0\rangle + e^{i\vartheta_2} \sin \frac{\vartheta_1}{2} |1\rangle. \quad (16)$$

From Equation 15 and Equation 16, we immediately observe that the circuit is maximally expressive – in fact, a double cover for $0 \leq \vartheta_1 \leq 2\pi$ – and we may restrict ϑ_1 to the interval $[0, \pi]$. The mapping of $|\psi(\Theta, \Phi)\rangle$ into \mathbb{R}^3 is then given by

$$|\psi(\Theta, \Phi)\rangle \mapsto \begin{pmatrix} \sin \Theta \cos \Phi \\ \sin \Theta \sin \Phi \\ \cos \Theta \end{pmatrix}. \quad (17)$$

Using

$$\alpha_C(N) := \max_{\delta \in \mathcal{D}_N} \max_{v \in V_\delta} d(\delta, v) \quad (18)$$

as an estimate for α_C , where $\mathcal{D}_N = \{C(\vartheta_1), \dots, C(\vartheta_N)\}$, and generating the ϑ_k using a scrambled Sobol' sequence in $[0, \pi] \times [0, 2\pi]$, we expect to see $\alpha_C(N) \rightarrow 0$ as $N \rightarrow \infty$. Figure 1 shows that this is indeed the case and that the rate of convergence is comparable to Monte Carlo type convergence $\alpha_C(N) \approx 1/\sqrt{N}$. Furthermore, we will observe in section X that this rate of convergence (including the prefactor) is comparable to the theoretically optimal rate of convergence for the estimation of α_C based on Voronoi vertices.

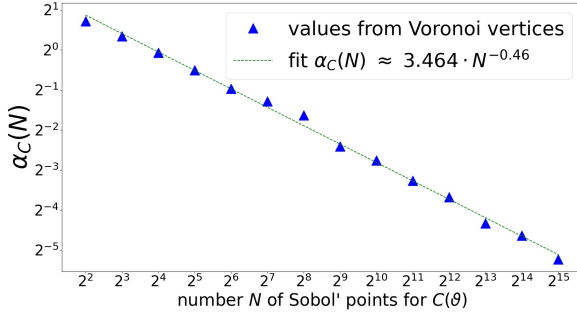


Fig. 1. Estimate of $\alpha_C(N)$ for the circuit C given in Equation 16 using N scrambled Sobol' points.

The blue triangles show $\alpha_C(N) = \max_{\delta \in \mathcal{D}_N} \max_{v \in V_\delta} d(\delta, v)$ as obtained from spherical Voronoi diagrams. The dashed green line is a convergence rate fit of the form $\alpha_C(N) \propto N^e$.

VI. A FIRST NON-MAXIMALLY EXPRESSIVE EXAMPLE

In contrast to the maximally expressive example of section V, we will now consider a non-maximally expressive circuit on the Bloch sphere. This circuit will represent a very bad choice of non-maximally expressive examples since it will only map to a great circle of the Bloch sphere. This has two important effects. Firstly, the worst-case best-approximation error will be very large; $\alpha_C = \frac{\pi}{2}$ in fact. Secondly, the computation of the Voronoi regions will encounter some non-trivial obstacle which we will discuss in more detail in section VII.

For ease of computation, we will consider the Bloch sphere embedded into \mathbb{R}^3 using the real basis $|0\rangle$, $|1\rangle$, and $i|1\rangle$. Furthermore, let us consider the circuit

$$C(\vartheta) = R_Y(\vartheta) |0\rangle = \cos \frac{\vartheta}{2} |0\rangle + \sin \frac{\vartheta}{2} |1\rangle. \quad (19)$$

Then \mathcal{M} is the great circle in the $|0\rangle$ - $|1\rangle$ -plane and consequently the worst-case best-approximation error is given by the distance of \mathcal{M} to the poles $\pm i|1\rangle$, i.e.,

$$\alpha_C = d_{\mathcal{M}}(i|1\rangle) = \frac{\pi}{2}. \quad (20)$$

To compute α_C from a set \mathcal{D} and its corresponding Voronoi vertices, we draw values

$$0 \leq \vartheta_1 < \vartheta_2 < \dots < \vartheta_k < 2\pi \quad (21)$$

and consider the points

$$C(\vartheta_k) = \cos \frac{\vartheta_k}{2} |0\rangle + \sin \frac{\vartheta_k}{2} |1\rangle. \quad (22)$$

We assume that k is large enough and that the ϑ_k are sufficiently well distributed such that

$$\mathcal{D} = \{C(\vartheta_j); 1 \leq j \leq k\} \quad (23)$$

is an ε -dense approximation with $\varepsilon < \pi$. For example, we may choose $k \geq 3$ equidistant points $\vartheta_j = \frac{2\pi(j-1)}{k}$.

For $1 < j < k$, the corresponding Voronoi region therefore is bounded by the great circles connecting $i|1\rangle$ and $-i|1\rangle$, which pass through $\cos \frac{\vartheta_j + \vartheta_{j-1}}{4} |0\rangle + \sin \frac{\vartheta_j + \vartheta_{j-1}}{4} |1\rangle$ and $\cos \frac{\vartheta_j + \vartheta_{j+1}}{4} |0\rangle + \sin \frac{\vartheta_j + \vartheta_{j+1}}{4} |1\rangle$. Hence, we can set $V_{C(\vartheta_j)}$ to these four points and obtain

$$\max_{v \in V_{C(\vartheta_j)}} d(C(\vartheta_j), v) = d(C(\vartheta_j), i|1\rangle) = \frac{\pi}{2}. \quad (24)$$

For $V_{C(\vartheta_1)}$ and $V_{C(\vartheta_k)}$, we note that $C(2\pi) = -|0\rangle$ is equivalent to $C(0) = |0\rangle$. Thus, $V_{C(\vartheta_1)}$ and $V_{C(\vartheta_k)}$ contain the already discussed points $i|1\rangle$ and $-i|1\rangle$, as well as, $\cos \frac{\vartheta_1 + \vartheta_2}{4} |0\rangle + \sin \frac{\vartheta_1 + \vartheta_2}{4} |1\rangle$ and $\cos \frac{\vartheta_k + \vartheta_{k-1}}{4} |0\rangle + \sin \frac{\vartheta_k + \vartheta_{k-1}}{4} |1\rangle$, respectively. Furthermore, both $V_{C(\vartheta_1)}$ and $V_{C(\vartheta_k)}$ contain the point between $C(\vartheta_1)$ and $C(\vartheta_k)$, i.e.,

$$C\left(\frac{2\pi + \vartheta_1 + \vartheta_k}{2} \bmod 2\pi\right). \quad (25)$$

Again, we conclude

$$\max_{v \in V_{C(\vartheta_1)}} d(C(\vartheta_1), v) = d(C(\vartheta_1), i|1\rangle) = \frac{\pi}{2} \quad (26)$$

and

$$\max_{v \in V_{C(\vartheta_k)}} d(C(\vartheta_k), v) = d(C(\vartheta_k), i|1\rangle) = \frac{\pi}{2}. \quad (27)$$

Hence, we have shown that

$$\alpha_C \leq \max_{\delta \in \mathcal{D}} \max_{v \in V_\delta} d(\delta, v) = \frac{\pi}{2} \leq \alpha_C + \varepsilon \quad (28)$$

holds for all $\varepsilon > 0$ and thus

$$\alpha_C = \frac{\pi}{2}. \quad (29)$$

VII. PRACTICAL REMARKS

The computation of α_C relies on the computation of Voronoi diagrams. This may be tricky due to the nature of a lower dimensional circuit manifold \mathcal{M} . As the example above shows, this is relatively easy to see if the state space \mathcal{S} is the entire unit sphere $\partial B_{\mathbb{C}^{2^Q}}$ of the Q -qubit Hilbert space \mathbb{C}^{2^Q} and if \mathcal{M} is contained in a lower dimensional subspace L of \mathbb{C}^{2^Q} . The Voronoi regions then only have two proper vertices ($i|1\rangle$ and $-i|1\rangle$ in the case above) and we artificially added the mid-points between samples to properly define the Voronoi regions. This is likely to be difficult for an automated process to handle. However, if we use spherical Voronoi diagrams with samples such that the matrix $(C(\vartheta_1) \dots C(\vartheta_k))$ has rank 2^{Q+1} , then the Voronoi regions are well-defined on the sphere $\partial B_{\mathbb{C}^{2^Q}}$. Should \mathcal{S} be a submanifold of the quantum device state space,

then this may require further analysis of Voronoi diagrams in \mathcal{S} . However, if \mathcal{M} is contained in a lower dimensional subspace L of the quantum device Hilbert space, then the poles normal to L will be separated from \mathcal{M} and therefore $\alpha_C \geq \frac{\pi}{2}$. This can be extended to state spaces \mathcal{S} that are not the unit sphere by considering points in \mathcal{S} that are not in L and taking their distance as a lower bound for α_C .

In many physical examples, we consider a state space \mathcal{S} that is an intersection of a lower dimensional subspace L of the quantum device Hilbert space and the unit sphere in the quantum device Hilbert space, e.g., if we are considering state spaces \mathcal{S} that are eigenspaces of an operator describing a symmetry, such as the momentum 1 sector for translational symmetry (cf., e.g., section 5 in [13]). Then, \mathcal{S} is the unit sphere in the lower dimensional subspace. In other words, all statements about spherical state spaces above hold. In particular, we need at least $\dim_{\mathbb{R}} L$ points for the matrix $(C(\vartheta_1) \cdots C(\vartheta_k))$ to have rank $\dim_{\mathbb{R}} L$. As long as $\text{rank}(C(\vartheta_1) \cdots C(\vartheta_k)) < \dim_{\mathbb{R}} L$, we know that all points lie in a lower dimensional subspace intersected with \mathcal{S} , which implies $\alpha_C \geq \frac{\pi}{2}$ considering the poles of \mathcal{S} with respect to that lower dimensional space.

With respect to the problem of $|0\rangle \cong -|0\rangle$ we noticed in the example, we can circumvent this by artificially introducing a phase symmetry to the circuit. In other words, if $C(\vartheta) = U(\vartheta)|\text{init}\rangle$ and $|\text{init}\rangle = U_{\text{init}}|0\rangle$, then we may analyze

$$\tilde{C}(\varphi, \vartheta) = U(\vartheta)U_{\text{init}}R_Z(\varphi)U_{\text{init}}^*|\text{init}\rangle \quad (30)$$

instead.

Alternatively, for rotation gates $R_G(\vartheta) = \exp(-i\frac{\vartheta}{2}G)$ with $G^2 = 1$, we could also consider $\vartheta \in [0, 4\pi]$ instead of $\vartheta \in [0, 2\pi]$.

VIII. COMPLEXITY ANALYSIS

The estimation of α_C on a spherical state space \mathcal{S} has two primary components. First, we need to ensure that the sample matrix

$$\Sigma_N = \begin{pmatrix} \Re C(\vartheta_1) & \Re C(\vartheta_2) & \cdots & \Re C(\vartheta_N) \\ \Im C(\vartheta_1) & \Im C(\vartheta_2) & \cdots & \Im C(\vartheta_N) \end{pmatrix} \quad (31)$$

has rank $\dim_{\mathbb{R}} \mathcal{S} + 1$. If this is not the case, then all $C(\vartheta_k)$ lie in the intersection of \mathcal{S} with a linear space L of dimension $\dim_{\mathbb{R}} L \leq \dim_{\mathbb{R}} \mathcal{S}$. This implies that there exist points in \mathcal{S} that are orthogonal to L and thus $\alpha_C(N) \geq \frac{\pi}{2}$. We can compute $\text{rank} \Sigma_N$ by computing the rank of

$$S_N = \Sigma_N^* \Sigma_N. \quad (32)$$

For this, we observe that the (j, k) -element of S_N is given by

$$(S_N)_{j,k} = \Re \langle C(\vartheta_j), C(\vartheta_k) \rangle \quad (33)$$

and note that the hybrid quantum-classical algorithm described in section II can be used to compute the $N \times N$ -matrix $S_N = \Sigma_N^* \Sigma_N$ with $\mathcal{O}(N^2 \epsilon^{-2})$ many QPU calls, where ϵ is the acceptable level of error on the QPU data. The rank of S_N can then be computed classically, e.g., using QR decomposition in $\mathcal{O}(N^3)$. Overall, this step therefore costs $\mathcal{O}(N^3 \epsilon^{-2})$.

Only once $\text{rank} S_N = \dim_{\mathbb{R}} \mathcal{S} + 1$, we need to consider the Voronoi diagrams. To compute the Voronoi diagrams, we need to embed \mathcal{S} into $\mathbb{R}^{\dim_{\mathbb{R}} \mathcal{S} + 1}$ and compute the vectors $C(\vartheta_1), \dots, C(\vartheta_N)$ using this $\mathbb{R}^{\dim_{\mathbb{R}} \mathcal{S} + 1}$ -representation. This step is discussed in section IX. Once this embedding is achieved, computing the Voronoi diagrams is efficient with worst-case cost in $\mathcal{O}\left(N^{\lceil \frac{\dim_{\mathbb{R}} \mathcal{S}}{2} \rceil}\right)$ via stereographic projection and Delaunay triangulation [27]. Since the Voronoi vertices are the circumcenters of Delaunay simplices, the number of Voronoi vertices coincides with the number of Delaunay simplices of N points in $D = \dim_{\mathbb{R}} \mathcal{S} + 1$ dimensions, i.e., $\mathcal{O}(N^{\lceil D/2 \rceil})$. Computing $\alpha_C(N)$ inefficiently via

$$\alpha_C(N) = \max_{v \in V} \min_{1 \leq j \leq N} d(v, C(\vartheta_j)), \quad (34)$$

where V is the entire set of Voronoi vertices – i.e., by comparing all Voronoi vertices to all sample points rather than restricting the comparison only to vertices of each sample's Voronoi region – we obtain a worst-case cost in $\mathcal{O}(N^{\lceil D/2 \rceil + 1})$ to compute $\alpha_C(N)$.

IX. MAPPING $C(\vartheta)$ EFFICIENTLY INTO $\mathbb{R}^{\dim_{\mathbb{R}} \mathcal{S} + 1}$

In order to efficiently map $C(\vartheta)$ into $\mathbb{R}^{\dim_{\mathbb{R}} \mathcal{S} + 1}$, we need to employ the quantum device again. But first, since we are working with real vector spaces, we cannot use $C(\vartheta)$ directly. Instead, we need to consider the real vector

$$C_{\mathbb{R}}(\vartheta) := \begin{pmatrix} \Re C(\vartheta) \\ \Im C(\vartheta) \end{pmatrix}. \quad (35)$$

Hence, we are looking to map the sequence of samples $C_{\mathbb{R}}(\vartheta_j) \in \mathbb{R}^{2^{\mathcal{Q}+1}}$ efficiently into $\mathbb{R}^{\dim_{\mathbb{R}} \mathcal{S} + 1}$. To make this mapping efficient, we will map the set $\{C_{\mathbb{R}}(\vartheta_j); j \leq k\}$ into the subspace $\text{lin}\{e_j; j \leq r\}$, where $\{e_j; j \leq \dim_{\mathbb{R}} \mathcal{S} + 1\}$ denotes the canonical basis of $\mathbb{R}^{\dim_{\mathbb{R}} \mathcal{S} + 1}$ and

$$r = \text{rank} \begin{pmatrix} C_{\mathbb{R}}(\vartheta_1) & C_{\mathbb{R}}(\vartheta_2) & \cdots & C_{\mathbb{R}}(\vartheta_k) \end{pmatrix}. \quad (36)$$

$C_{\mathbb{R}}(\vartheta_1)$ is automatically mapped to

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \quad (37)$$

Then, we map $C_{\mathbb{R}}(\vartheta_2)$ to

$$v_2 = \begin{pmatrix} v_{2,1} \\ v_{2,2} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad (38)$$

with

$$v_{2,1} = \langle v_2, v_1 \rangle = \Re \langle C(\vartheta_2), C(\vartheta_1) \rangle, \quad (39)$$

which we can compute efficiently on the quantum device, and

$$v_{2,2} = \sqrt{1 - v_{2,1}^2}. \quad (40)$$

At this point, we need to keep track of the current subspace basis B_k after mapping $C(\vartheta_1), \dots, C(\vartheta_k)$. The initial basis B_1 contains only the vector v_1 . The vector v_2 is an element of B_k with $k \geq 2$ if and only if $v_{2,2} \neq 0$.

If $v_2 \notin B_2$, then $B_2 = \{v_1\}$ and $C(\vartheta_3)$ is mapped to

$$v_3 = \begin{pmatrix} v_{3,1} \\ v_{3,2} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad (41)$$

with

$$v_{3,1} = \Re \langle C(\vartheta_3), C(\vartheta_1) \rangle \quad (42)$$

and

$$v_{3,2} = \sqrt{1 - v_{3,1}^2}. \quad (43)$$

Similarly, v_3 is added to B_3 if and only if $v_{3,2} \neq 0$.

On the other hand, if $v_2 \in B_2$, then $C(\vartheta_3)$ is mapped to

$$v_3 = \begin{pmatrix} v_{3,1} \\ v_{3,2} \\ v_{3,3} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad (44)$$

with

$$\begin{aligned} v_{3,1} &= \langle v_3, v_1 \rangle = \Re \langle C(\vartheta_3), C(\vartheta_1) \rangle, \\ v_{3,2} &= \frac{\Re \langle C(\vartheta_3), C(\vartheta_2) \rangle - v_{3,1} v_{2,1}}{v_{2,2}}, \\ v_{33} &= \sqrt{1 - v_{3,1}^2 - v_{3,2}^2}, \end{aligned} \quad (45)$$

and v_3 is added to B_3 if and only if $v_{3,3} \neq 0$.

In order to map $C(\vartheta_k)$, we have already mapped $C(\vartheta_1), \dots, C(\vartheta_{k-1})$ and identified a basis B_{k-1} with b_{k-1} elements. We denote the elements of B_{k-1} by $v_{\beta_1}, \dots, v_{\beta_{b_{k-1}}}$. Then $C(\vartheta_k)$ is mapped to

$$v_k = \begin{pmatrix} v_{k,1} \\ \vdots \\ v_{k,b_{k-1}+1} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \quad (46)$$

with

$$\begin{aligned} v_{k,1} &= \Re \langle C(\vartheta_k), C(\vartheta_{\beta_1}) \rangle, \\ v_{k,2} &= \frac{\Re \langle C(\vartheta_k), C(\vartheta_{\beta_2}) \rangle - v_{k,1} v_{\beta_2,1}}{v_{\beta_2,2}}, \\ &\vdots \\ v_{k,b_{k-1}} &= \frac{\Re \langle C(\vartheta_k), C(\vartheta_{\beta_{b_{k-1}}}) \rangle - \sum_{j=1}^{b_{k-1}-1} v_{k,j} v_{\beta_{b_{k-1}},j}}{v_{\beta_{b_{k-1}},b_{k-1}}}, \\ v_{k,b_{k-1}+1} &= \sqrt{1 - \sum_{j=1}^{b_{k-1}} v_{k,j}^2}. \end{aligned} \quad (47)$$

This is possible since by construction $v_{\beta_j,j} \neq 0$. Of course, v_k is added to B_k if and only if $v_{k,b_{k-1}+1} \neq 0$.

Combining this mapping of quantum states with the complexity analysis of section VIII shows that α_C can be estimated efficiently.

X. SCALING FOR SPHERICAL STATE SPACES AND MAXIMALLY EXPRESSIVE CIRCUITS

Now that we know that the Voronoi estimation of α_C is efficiently implementable, the question becomes how many points are necessary for the Voronoi estimation to yield a sufficiently reliable estimate of α_C .

The optimal scaling for $\alpha_C(N)$ on a $(D-1)$ -dimensional sphere would be given by equidistant points. As such a configuration is not possible for arbitrary values of N , we must approximate this type of scaling. Considering spherical coordinates in $D-1$ dimensions

$$\begin{aligned} x_1 &= \cos \varphi_1 \\ x_2 &= \sin \varphi_1 \cos \varphi_2 \\ x_3 &= \sin \varphi_1 \sin \varphi_2 \cos \varphi_3 \\ &\vdots \\ x_{D-1} &= \sin \varphi_1 \sin \varphi_2 \cdots \sin \varphi_{D-2} \cos \varphi_{D-1} \\ x_D &= \sin \varphi_1 \sin \varphi_2 \cdots \sin \varphi_{D-2} \sin \varphi_{D-1}, \end{aligned} \quad (48)$$

we choose circles of constant φ_1 which are spaced a distance d_1 apart on the surface of the sphere. Then, we discretize φ_2 on each of the φ_1 circles such that the φ_2 values are spaced a distance d_2 apart (d_2 depends on the specific circle of constant φ_1) and each value of d_2 is as close as possible to d_1 . Iteratively, we can space φ_k for each chosen $\varphi_1, \dots, \varphi_{k-1}$ such that the corresponding distances satisfy $d_1 \approx d_2 \approx \dots \approx d_k$ as closely as possible. As such, the volume of the sphere attributed to each point is approximately $\prod_{j=1}^{D-1} d_j$, i.e., approximately d_1^{D-1} . Since the surface area of a $(D-1)$ -dimensional unit sphere is $\frac{2\pi^{D/2}}{\Gamma(D/2)}$, this means we can place $N \approx \frac{2\pi^{D/2}}{\Gamma(D/2)d_1^{D-1}}$ points. Conversely, although we wish to place N points, we may not be perfectly able to, but we can place $N' \approx N$ points by choosing d_1 to be a fraction of π as close as possible to $\left(\frac{2\pi^{D/2}}{\Gamma(D/2)N}\right)^{1/(D-1)}$. Each sample can

then be interpreted as the approximate center of a cube with side length $\left(\frac{2\pi^{D/2}}{\Gamma(D/2)N}\right)^{1/(D-1)}$, which means the maximal distance that an arbitrary point can be away from any of these sample points is approximately $\left(\frac{2\pi^{D/2}}{\Gamma(D/2)N}\right)^{1/(D-1)}\frac{\sqrt{D}}{2}$. In other words, the best case scaling we can expect for $\alpha_C(N)$ with $D = \dim_{\mathbb{R}} \mathcal{S} + 1$ is

$$\alpha_C^{\text{opt}}(N) = \left(\frac{2\pi^{D/2}}{\Gamma(D/2)N}\right)^{1/(D-1)}\frac{\sqrt{D}}{2}. \quad (49)$$

Returning to the Bloch sphere example, we have $D = 3$, i.e.,

$$\alpha_C^{\text{opt}}(N) = \sqrt{3\pi}N^{-\frac{1}{2}} \approx 3.07N^{-\frac{1}{2}}. \quad (50)$$

A comparison with the observed $\alpha_C(N) \approx 3.46N^{-.46}$ using a Sobol' sequence to generate samples $C(\vartheta_k)$ in Figure 1 shows that the Sobol' point approximation is almost optimal.

XI. α_C AND THE VOLUME OF LOWER DIMENSIONAL \mathcal{M}

The scaling analysis of section X is optimized for spherical state spaces and maximally expressive circuits. If we wish to use a circuit C with lower dimensional image \mathcal{M} and guarantee a chosen best-approximation error α_C , then we need to compare α_C to $\text{vol}(\mathcal{M})$ and observe that \mathcal{M} has to have a large volume. In this section, we want to approximate a lower bound for $\text{vol}(\mathcal{M})$ in terms of α_C .

First, we note that $\dim \mathcal{M} = \dim \mathcal{S} - 1$ is the best case scenario (in terms of smallest possible volume of \mathcal{M}) since the projection onto a $\dim \mathcal{M} + 1$ dimensional subspace of \mathcal{S} reduces the distance between points and thus, reduces the volume of \mathcal{M} as well as α_C . After a suitable choice of coordinate transformation, we may therefore assume that $C(\vartheta)$ takes the form

$$C(\vartheta) = \begin{pmatrix} \cos(\gamma_1(\vartheta_1)) \\ \sin(\gamma_1(\vartheta_1)) \cos(\gamma_2(\vartheta_1)) \\ \sin(\gamma_1(\vartheta_1)) \sin(\gamma_2(\vartheta_1)) \sigma(\vartheta_2, \dots, \vartheta_n) \end{pmatrix}, \quad (51)$$

where σ is surjective onto the $(\dim \mathcal{M} - 1)$ -dimensional sphere, and (γ_1, γ_2) is the path C takes in the projection of \mathcal{S} onto the 2-dimensional subspace of \mathcal{S} orthogonal to the sphere σ maps to.

To estimate the minimal volume of \mathcal{M} , we can use a greedy algorithm to construct a path γ on the 2-dimensional sphere that is α_C close to all points. For example, we may start at the north pole and move south for a distance $2\alpha_C$. Then, we move east along the reached latitude stopping $2\alpha_C$ short of a full revolution. Now we move south again for $2\alpha_C$ and then move east stopping $2\alpha_C$ short of a full revolution. We continue this process until we reach the south pole. For simplicity, we assume that $2\alpha_C$ is an integer fraction of π , i.e., $2\alpha_C = \frac{\pi}{n}$. Then, this path contains n sections of going south for $2\alpha_C$ and $n - 1$ sections of moving east for a full revolution but stopping short $2\alpha_C$. The length of this path ℓ_1 is therefore

$$\ell_1 = 2\alpha_C + \sum_{j=1}^{n-1} 2\pi \sin \frac{\pi j}{n} = \frac{\pi}{n} + 2\pi \cot \frac{\pi}{2n}. \quad (52)$$

Finally, we obtain the corresponding volume

$$V_1 = \ell_1 \text{vol} S^{\dim \mathcal{M}-1} = \frac{4\pi^{\frac{\dim \mathcal{M}}{2}} (\alpha_C + \pi \cot \alpha_C)}{\Gamma\left(\frac{\dim \mathcal{M}}{2}\right)}. \quad (53)$$

Another possible greedy algorithm is to use a spiral with $\gamma_1(\vartheta_1) = \pi\vartheta_1$ and $\gamma_2(\vartheta_1) = 2\pi n\vartheta_1$ where $\vartheta_1 \in [0, 1]$. This spiral wraps n times around the sphere and therefore has $2\alpha_C \approx \frac{\pi}{n}$ for large n . We may compute the length of this curve as

$$\ell_2 = \int_0^1 \|\gamma'(t)\| dt = \sqrt{\pi^2 + 4\pi^2 n^2} \quad (54)$$

and thus the corresponding volume

$$V_2 = \ell_2 \text{vol} S^{\dim \mathcal{M}-1} \approx \frac{2\pi^{\frac{\dim \mathcal{M}}{2}+1} \sqrt{1 + \frac{\pi^2}{\alpha_C^2}}}{\Gamma\left(\frac{\dim \mathcal{M}}{2}\right)}. \quad (55)$$

For $\alpha_C \ll 1$, we can use $\cot(\alpha_C) = \frac{1}{\alpha_C} + \mathcal{O}(\alpha_C)$ to observe

$$\frac{V_1}{V_2} \approx \frac{2 \cot \alpha_C}{\sqrt{1 + \frac{\pi^2}{\alpha_C^2}}} \approx \frac{2}{\pi}, \quad (56)$$

i.e., both estimates are of the same order of magnitude, and given the volume of the circuit image \mathcal{M} , we can re-arrange the V_1 estimate to obtain

$$\alpha_C \gtrsim \frac{4\pi^{\frac{\dim \mathcal{M}}{2}+1}}{\Gamma\left(\frac{\dim \mathcal{M}}{2}\right) \text{vol}(\mathcal{M})}. \quad (57)$$

Of course, for this relation to be useful, we either need to compute α_C , e.g., from Voronoi diagrams, or $\text{vol}(\mathcal{M})$. If we wish to compute $\text{vol}(\mathcal{M})$, then we may choose a quadrature rule with points ϑ_j and weights ω_j and estimate

$$\begin{aligned} \text{vol}(\mathcal{M}) &= \int_{\mathcal{P}} \sqrt{\det g(\vartheta)} d\text{vol}_{\mathcal{P}}(\vartheta) \\ &\approx \sum_j \omega_j \sqrt{\det g(\vartheta_j)}, \end{aligned} \quad (58)$$

where

$$g_{jk}(\vartheta) = \Re\langle \partial_j C(\vartheta), \partial_k C(\vartheta) \rangle, \quad (59)$$

for which we already have an efficient algorithm in section II. It is important to note that g is the same matrix as S_N in section II. Thus, for $\det g$ to not be zero, the vectors $\partial_j C(\vartheta)$ need to be linearly independent. In other words, the volume element $\sqrt{\det g(\vartheta)} d\text{vol}_{\mathcal{P}}(\vartheta)$ needs to be computed with a minimal circuit.

XII. A BLOCH SPHERE SPIRAL EXAMPLE

To illustrate this relationship between α_C and $\text{vol}(\mathcal{M})$, let us consider the family of circuits

$$\begin{aligned} C_n(\vartheta) &= R_Z(n\vartheta) R_Y(\vartheta) |0\rangle \\ &= \cos \frac{\vartheta \bmod \pi}{2} |0\rangle + e^{2in\vartheta} \sin \frac{\vartheta \bmod \pi}{2} |1\rangle \end{aligned} \quad (60)$$

on the Bloch sphere with $\vartheta \in [0, 2\pi]$. Here the ‘‘mod π ’’ condition simply ensures that the coefficient in front of $|0\rangle$

is non-negative as per our representation of the Bloch sphere introduced in section V. In terms of the Bloch sphere representation

$$\cos \frac{\Theta}{2} |0\rangle + e^{i\Phi} \sin \frac{\Theta}{2} |1\rangle \mapsto \begin{pmatrix} \sin \Theta \cos \Phi \\ \sin \Theta \sin \Phi \\ \cos \Theta \end{pmatrix}, \quad (61)$$

$C_n(\vartheta)$ is represented by the vector

$$\gamma(\vartheta) = \begin{pmatrix} \sin(\vartheta \bmod \pi) \cos(2n\vartheta) \\ \sin(\vartheta \bmod \pi) \sin(2n\vartheta) \\ \cos(\vartheta \bmod \pi) \end{pmatrix}, \quad (62)$$

and we obtain

$$\det g(\vartheta) = 1 + 4n^2 \sin(\vartheta \bmod \pi)^2, \quad (63)$$

i.e.,

$$\text{vol}(\mathcal{M}_n) = \int_0^{2\pi} \sqrt{\det g(\vartheta)} d\vartheta = 4E(-4n^2) \quad (64)$$

using the elliptic integral

$$E(m) = \int_0^{\frac{\pi}{2}} \sqrt{1 - m \sin(\vartheta)^2} d\vartheta. \quad (65)$$

As in this case $\dim(\mathcal{M}_n) = 1$, the estimated lower bound for α_{C_n} becomes

$$\alpha_{C_n} \gtrsim \frac{4\pi^{\frac{3}{2}}}{\Gamma(\frac{1}{2}) \text{vol}(\mathcal{M}_n)} = \frac{4\pi}{\text{vol}(\mathcal{M}_n)} = \frac{\pi}{E(-4n^2)}. \quad (66)$$

Since C_n is a ‘‘spiral’’ circuit that is also taking into account that the parameter space \mathcal{P} should be a manifold without boundary (here $\mathbb{R}/2\pi\mathbb{Z}$), we expect to see strong agreement between α_{C_n} estimated using $\frac{\pi}{E(-4n^2)}$ and α_{C_n} extracted from Voronoi diagrams. Figure 2 confirms this expectation.

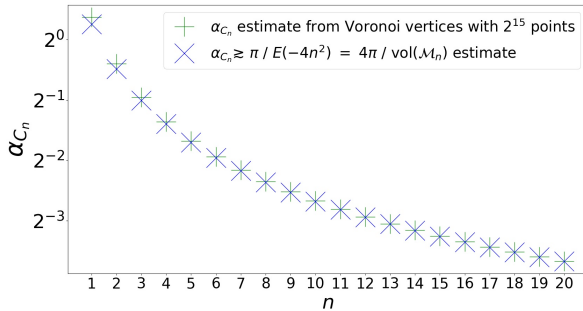


Fig. 2. Estimate of α_{C_n} for the circuit C given in Equation 60 using the volume estimate and the Voronoi diagram estimate. The Voronoi diagrams use 2^{15} samples generated from scrambled Sobol’ points.

XIII. VORONOI ANALYSIS FOR INITIALIZATIONS OF VQS WITH LOCAL OPTIMIZERS

The example in section XII and the estimate

$$\alpha_C \text{vol}(\mathcal{M}) \gtrsim \frac{4\pi^{\frac{\dim \mathcal{M}}{2} + 1}}{\Gamma(\frac{\dim \mathcal{M}}{2})} \quad (67)$$

show that quantum circuits that are not maximally expressive can be able to approximate points in the state space \mathcal{S} to a given precision α_C at the cost of having large $\text{vol}(\mathcal{M})$. In particular, the spiral example shows that points that are close in state space \mathcal{S} may have best-approximations with parameter values ϑ that are far apart in parameter space \mathcal{P} . This can be illustrated using the circuit

$$C_n(\vartheta) = R_Z(n\vartheta)R_Y(\vartheta) |0\rangle = \cos \frac{\vartheta \bmod \pi}{2} |0\rangle + e^{2in\vartheta} \sin \frac{\vartheta \bmod \pi}{2} |1\rangle \quad (68)$$

and a point p half-way between points $C_n(\vartheta)$ and $C_n(\vartheta')$ with $\vartheta' = \vartheta + \frac{2\pi}{n}$. Any arbitrarily small perturbation of p towards $C_n(\vartheta)$ will have best-approximation parameters close to ϑ , whereas a perturbation towards $C_n(\vartheta')$ has best-approximation parameters close to ϑ' , i.e., an entire revolution around the sphere later. If this point p is equatorial, then arbitrarily small perturbations in p yield best-approximations that are a distance approximately 2π apart as measured in \mathcal{M} . Considering an energy-type cost function

$$\text{Cost}(\vartheta) = \langle C_n(\vartheta) | H | C_n(\vartheta) \rangle \quad (69)$$

for some Hamiltonian H and supposing n large, \mathcal{M}_n is packed tightly around the solution state $|\psi\rangle$ that needs to be approximated by $C_n(\vartheta)$. However, this implies $\text{Cost}(\vartheta)$ has many local minima corresponding to states $C_n(\vartheta)$ that are close to $|\psi\rangle$ in \mathcal{S} , but these minima are far away from each other as measured in the metric of \mathcal{M} . Hence, having an initial guess $C_n(\vartheta)$ close to $|\psi\rangle$ in \mathcal{S} is not sufficient for a local optimizer to guarantee convergence to the best-approximation $C_n(\vartheta_\psi)$ of $|\psi\rangle$. Instead, the initial guess $C_n(\vartheta)$ must satisfy $\vartheta \approx \vartheta_\psi$ in parameter space as well.

The Voronoi analysis provides precisely such a set of initial guesses. Since every point of the state space \mathcal{S} is α_{C_n} close to one of the Voronoi sample points, we can use the Voronoi sample points as initial guesses and run a local optimizer VQS for each of them (or at least those that have near optimal cost function values). The best-approximation will then be selected as the $C(\vartheta_\psi)$ among the solution candidates $C(\vartheta_j)$ provided by each VQS for which $\text{Cost}(C(\vartheta_\psi)) = \min_j \text{Cost}(C(\vartheta_j))$.

XIV. CONCLUSION

In this article we have extended previous work on dimensional expressivity analysis (DEA) [23] in two important directions. First, we reviewed DEA and highlighted that it requires to have a candidate circuit to be optimized. However, finding a candidate circuit is not always straightforward. Hence, in section III we proposed an inductive process of constructing minimal, maximally expressive circuits on arbitrarily many qubits. This can be used directly, if the entire quantum device state space needs to be parameterized, or it may be optimized using DEA if certain symmetries are to be removed.

The second major open question surrounding DEA is related to non-maximally expressive circuits. Previously, DEA made no statements about non-maximally expressive circuits. Hence, we considered the worst-case best-approximation error and

proposed estimating it using Voronoi vertices in section IV. This was supplemented with examples highlighting the optimal behavior of the analysis in section V and some possible obstacles in section VI, a discussion of practical aspects surrounding these obstacles in section VII, and an example applying the Voronoi estimation of the best-approximation error on a maximally expressive circuit in section VII.

In section VIII and section IX we then discussed the computational complexity of the proposed Voronoi diagram based estimation of the best-approximation error using a hybrid quantum-classical algorithm that allows us to map the quantum states efficiently into classical memory and compute the Voronoi vertices. In particular, the quantum part of the algorithm scales quadratically in both the number of sample points N required for the Voronoi estimates and the inverse acceptable level of noise ϵ^{-1} coming from the quantum device. The algorithm without the Voronoi vertices computation is still in $\mathcal{O}(N^3\epsilon^{-2})$. The bottlenecks of the algorithm are therefore the computation of the Voronoi vertices which, in $D = \dim_{\mathbb{R}} S + 1$ dimensions, is worst case in $\mathcal{O}(N^{\lceil D/2 \rceil})$, and testing all $\mathcal{O}(N^{\lceil D/2 \rceil})$ Voronoi vertices against (worst case) all Voronoi sample points is in $\mathcal{O}(N^{\lceil D/2 \rceil + 1})$.

In section X we then discussed the necessary number of points N required to estimate the best-approximation error to a chosen level of accuracy. The optimal scaling is proportional to $N^{\frac{1}{D-1}}$ and achieved for maximally expressive circuits. For non-maximally expressive circuits, as discussed in section XI, the internal volume of the set of reachable states \mathcal{M} by the quantum circuit needs to be considered. In the same section, an efficient hybrid quantum-classical algorithm for the estimation of $\text{vol}(\mathcal{M})$ is also provided. Furthermore, we tested the relationship between the best-approximation error and $\text{vol}(\mathcal{M})$ using an example of a spiral circuit image \mathcal{M} on the Bloch sphere in section XII. Finally, we have discussed some of the expected obstacles when applying non-maximally expressive circuits to variational quantum simulations with local optimizers in section XIII.

The work presented here therefore extends DEA [23] by multiple practical aspects. We have provided minimal, maximally expressive candidate circuits that can be optimized using DEA on arbitrarily many qubits, which can further be optimized by removing unwanted symmetries via DEA. Should we choose to employ non-maximally expressive circuits, then we have provided the means of estimating the worst-case best-approximation error a priori to any VQS. In particular, if we want to ensure a certain best-approximation error α_C , then we can predict the minimal number of samples necessary to construct a set of VQS initial states, such that every element of the relevant state space is α_C close to at least one point in the VQS initialization set. The VQS can then be used to optimize any of these initial states to find the best-approximation of the VQS solution for non-maximally expressive circuits.

REFERENCES

[1] J. Preskill, “Quantum Computing in the NISQ era and beyond,” *Quantum*, 2:79, 2018.

[2] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, “Quantum machine learning,” *Nature*, 549 (7671):195–202, 2017.

[3] R. Orús, S. Mugel, and E. Lizaso, “Quantum computing for finance: Overview and prospects,” *Reviews in Physics*, 4:100028, 2019.

[4] A. Montanaro, “Quantum algorithms: an overview,” *npj Quantum Information*, 2(1), 2016.

[5] F. G. S. L. Brandao and K. M. Svore, “Quantum speed-ups for solving semidefinite programs,” in *2017 IEEE 58th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 415–426, 2017.

[6] M. Troyer and U.-J. Wiese, “Computational complexity and fundamental limitations to fermionic quantum monte carlo simulations,” *Phys. Rev. Lett.*, 94:170201, 2005.

[7] F. Arute et al., “Quantum supremacy using a programmable superconducting processor,” *Nature*, 574(7779):505–510, 2019.

[8] H.-S. Zhong et al., “Quantum computational advantage using photons,” *Science*, Vol. 370, Issue 6523, pp. 1460–1463, 2020.

[9] K. Temme, S. Bravyi, and J. M. Gambetta, “Error mitigation for short-depth quantum circuits,” *Phys. Rev. Lett.*, 119:180509, 2017.

[10] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, “Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets,” *Nature*, 549(7671):242–246, 2017.

[11] S. Endo, S. C. Benjamin, and Y. Li, “Practical quantum error mitigation for near-future applications,” *Phys. Rev. X*, 8:031027, 2018.

[12] K. Yeter-Aydeniz, R. C. Pooser, and G. Siopsis, “Practical quantum computation of chemical and nuclear energy levels using quantum imaginary time evolution and lanczos algorithms,” *npj Quantum Information*, 6(1), 2020.

[13] L. Funcke, T. Hartung, K. Jansen, S. Kühn, P. Stornati, and X. Wang, “Measurement error mitigation in quantum computers through classical bit-flip correction,” *arXiv:2007.03663*, 2020. URL <https://arxiv.org/abs/2007.03663>.

[14] A. Peruzzo, J. McClean, P. Shadbolt, M. Yung, X. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O’Brien, “A variational eigenvalue solver on a photonic quantum processor,” *Nat. Commun.*, 5:1, 2014.

[15] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, “The theory of variational hybrid quantum-classical algorithms,” *New J. Phys.*, 18(2):023023, 2016.

[16] Y. Wang et al., “Quantum simulation of helium hydride cation in a solid-state spin register,” *ACS Nano*, 9(8):7769–7774, 2015.

[17] C. Hempel et al., “Quantum chemistry calculations on a trapped-ion quantum simulator,” *Phys. Rev. X*, 8:031022, 2018.

[18] C. Kokail et al., “Self-verifying variational quantum simulation of the lattice schwinger model,” *Nature*, 569(7756):355, 2019.

[19] T. Hartung and K. Jansen, “Zeta-regularized vacuum expectation values,” *J. Math. Phys.*, 60(9):093504, 2019.

[20] K. Jansen and T. Hartung, “Zeta-regularized vacuum expectation values from quantum computing simulations,” *PoS(LATTICE2019)363*, 2020.

[21] D. Paulson, L. Dellantonio, J. F. Haase, A. Celi, A. Kan, A. Jena, C. Kokail, R. van Bijnen, K. Jansen, P. Zoller, and C. A. Muschik, “Towards simulating 2d effects in lattice gauge theories on a quantum computer,” *arXiv:2008.09252*, 2020. URL <http://arxiv.org/abs/2008.09252>.

[22] J. F. Haase, L. Dellantonio, A. Celi, D. Paulson, A. Kan, K. Jansen, and C. A. Muschik, “A resource efficient approach for quantum and classical simulations of gauge theories in particle physics,” *Quantum*, 5:393, 2021.

[23] L. Funcke, T. Hartung, K. Jansen, S. Kühn, and P. Stornati, “Dimensional Expressivity Analysis of Parametric Quantum Circuits,” *Quantum* 5, 422, 2021.

[24] G. Voronoi, “Nouvelles applications des paramètres continus à la théorie des formes quadratiques. Premier mémoire. Sur quelques propriétés des formes quadratiques positives parfaites” *J. Reine Angew. Math (133)*: 97–178, 1908.

[25] G. Voronoi, “Nouvelles applications des paramètres continus à la théorie des formes quadratiques. Deuxième mémoire. Recherches sur les paralléloèdres primitifs” *J. Reine Angew. Math (134)*: 198–287, 1908.

[26] M. A. Nielsen, I. L. Chuang, “Quantum Computation and Quantum Information, 10th Anniversary Edition,” Cambridge University Press, 2011.

[27] S. Fortune, “Voronoi Diagrams and Delaunay Triangulations,” in *Handbook of Discrete and Computational Geometry*, Edited by: J. E. Goodman, J. O’Rourke, C. D. Toth, CRC Press LLC, 2017.