Dimensional Expressivity Analysis of Parametric Quantum Circuits

Lena Funcke[1], Tobias Hartung[2,3], Karl Jansen[4], Stefan Kühn[2], and Paolo Stornati[4,5]

[1]Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, ON N2L 2Y5, Canada

[2]Computation-Based Science and Technology Research Center, The Cyprus Institute, 20 Kavafi Street, 2121 Nicosia, Cyprus

[3]Department of Mathematics, King's College London, Strand, London WC2R 2LS, United Kingdom

[4]NIC, DESY Zeuthen, Platanenallee 6, 15738 Zeuthen, Germany

[5]Institut für Physik, Humboldt-Universität zu Berlin, Zum Großen Windkanal 6, D-12489 Berlin, Germany

Parametric quantum circuits play a crucial role in the performance of many variational quantum algorithms. To successfully implement such algorithms, one must design efficient quantum circuits that sufficiently approximate the solution space while maintaining a low parameter count and circuit depth. In this paper, develop a method to analyze the dimensional expressivity of parametric quantum circuits. Our technique allows for identifying superfluous parameters in the circuit layout and for obtaining a maximally expressive ansatz with a minimum number of parameters. Using a hybrid quantum-classical approach, we show how to efficiently implement the expressivity analysis using quantum hardware, and we provide a proof of principle demonstration of this procedure on IBM's quantum hardware. We also discuss the effect of symmetries and demonstrate how to incorporate or remove symmetries from the parametrized ansatz.

1 Introduction

Current noisy intermediate-scale quantum (NISQ) computers [1] open up a new route to tackle a variety of computational problems that cannot be addressed efficiently with classical computers. Potential applications range from machine learning [2] to finance [3] to various optimization problems [4, 5]. In particular, quantum computers intrinsically evade the sign problem that hinders Monte Carlo simulations of strongly correlated quantum-many body problems in certain parameter regimes [6].

While current hardware is still of limited size and suffers from a considerable level of noise, it has already been successfully demonstrated that NISQ devices can outperform classical computers [7]. Moreover, techniques for mitigating the effects of noise are rapidly developing [8–12].

Parametrized quantum circuits are at the heart of many algorithms designed for NISQ devices. A prominent example is variational quantum simulation (VQS) [13, 14], a class of hybrid quantum-classical algorithms for solving optimization problems. Using a quantum circuit composed of parametric gates as an ansatz, the quantum coprocessor is used to efficiently evaluate the cost function for a given set of variational parameters. The minimization procedure is then performed on a classical computer in a feedback loop based on the measurement outcome obtained from the quantum device. VQS has been successfully applied to quantum many-body systems in quantum chemistry [9, 13, 15, 16] and even quantum mechanics and quantum field theory [17–21].

The design of the parametrized quantum circuit used as an ansatz is crucial for the success of a VQS approach. While it should be sufficiently expressive to be able to approximate the solution of the problem to good accuracy, NISQ devices only allow for executing circuits of short depths faithfully before noise starts to dominate. Thus, one should avoid superfluous parameters in the circuit resulting in unnecessary gate operations. At the same time, having an insufficient number of independent parameters in the circuit can lead the classical minimization algorithm to fall into false local minima. A careful design of the quantum circuit is therefore essential to make optimal

use of current NISQ devices.

This question of circuit expressivity is an active area of research, see Refs. [22–32] and references therein. In particular, it has been proposed to assess the expressivity of a parametric quantum circuit by quantifying the circuit's ability to uniformly reach the full Hilbert space [23], which was accomplished by computing statistical properties based on randomly sampling states from a given circuit template. Using this ansatz, it has been shown that the number of single-qubit rotations in certain parameterized quantum circuits can be decreased without compromising the expressivity of the circuit [26]. On the other hand, it has been shown that the inclusion of redundant parametrized gates can make the quantum circuits more resilient to noise [29, 32]. In particular, Ref. [32] showed that increasing the depth of the circuit above a certain threshold, which heavily overparameterizes the problem, leads to a good approximation of the ground state assuming a noise-free quantum device. Thus, in practice, these two effects of increasing the circuit depth and the resilience to noise have to be weighted against each other, especially for different amounts of noise in the given circuit. A practical approach to circuit design should therefore incorporate both aspects. For example, the here proposed dimensional expressivity analysis can be used to ensure sufficient expressivity of a minimal circuit which could then be overparametrized to harness the positive effects studied in Ref. [32].

For non-parametric gates, it is also possible to follow an algebraic approach, as demonstrated in Ref. [24] for the group generated by CNOT and SWAP gates. The reduction rules developed in Ref. [24] can then be employed to minimize the number of CNOT and SWAP operations in a given quantum circuit, while the action of the quantum circuit, i.e., the input-to-output relation, is unchanged.

In this paper, we will devise a general method to analyze the expressivity of a given parametric quantum circuit and to identify redundant parameters, which is independent of the statistical approach proposed in Ref. [23]. In fact, we will develop a geometric approach to minimize the number of parameters in a parametric circuit, while the set of states that can be generated by the circuit remains unchanged. In this sense, the analysis proposed in this paper complements the algebraic approach [24], which can minimize the number of non-parametric gates, while our proposed approach minimizes the number of parametric gates.

To this end, we consider quantum circuits as operators that act on a state within a certain Hilbert space, typically the state a given quantum device is initialized in. Thus, we can consider parametric quantum circuits as a map from the set of parameters to a subset of the state space of the quantum device. In many cases, this image set of reachable states by the quantum circuit forms a submanifold of the state space of the quantum device. As such, the (dimensional) expressivity of the quantum circuit can be understood as the dimension of this manifold of reachable states. We will also call this manifold the circuit manifold.

In the ideal case, the dimension of the circuit manifold coincides with the number of parameters of the quantum circuit. This implies that there are no redundant parameters. Similarly, in an ideal situation, the circuit should be able to reach all physical states. The set of physical states of course depends on the model simulated using the VQS and can be much more restrictive than the full state space of the quantum device. For example, if the physical model has a certain symmetry, e.g., $\mathbb{Z}_2$ symmetry for fermionic models or translational invariance in homogeneous materials, then any state in the state space of the quantum device that does not satisfy these symmetries would be considered unphysical. For efficient quantum simulations of such models, the quantum circuit should therefore be able to reach all physical states (to be maximally expressive) while simultaneously not generate unphysical states.

The main goal of efficient circuit design is therefore to generate all physical states without introducing unphysical states or redundant parameters. In most cases, designing a circuit that satisfies physical symmetries is much easier than ensuring maximal expressivity while avoiding redundancy. Both maximal expressivity and avoiding redundancy can locally be tested using dimensional expressivity analysis. If there are more parameters than the circuit manifold dimension, then superfluous parameters exist and can be removed using the dimensional analysis techniques developed in this paper.

Similarly, the codimension[1] of the circuit manifold with respect to the manifold of physical states provides a measure of expressivity deficiency. If the codimension vanishes, then both manifolds coincide (locally), which ensures maximal expressivity of the quantum circuit. This is because the circuit manifold must be a submanifold of the manifold of physical states (provided all states generated by the quantum circuit are physical) and having the same dimension thus locally implies that they coincide. Additionally, the codimension, as a measure of expressivity deficiency, is a predictor of the number of parameters missing for a maximally expressive circuit. Thus, the dimensional expressivity analysis developed in this paper will provide information that can guide the construction of gates to be added to the quantum circuit.

The dimensional expressivity analysis developed in this paper will (i) allow us to determine which parameters are superfluous in the quantum circuit and (ii) tell us how to modify a given quantum circuit to eliminate these parameters, while respecting the symmetries of the physical model. In addition, we demonstrate how to modify a given quantum circuit to remove unwanted symmetries from the states generated by the quantum circuit, and how to approach the construction of a maximally expressive quantum circuit respecting the physical symmetries. As an example of a common symmetry found in physical models, we discuss translational invariance and describe how the physical state space is divided into sectors defined by the eigenvalues of the momentum operator.

The rest of the paper is structured as follows. In Sec. 2, we provide a simple introductory example of how to analyze the dimensional expressivity of a quantum circuit. We chose to analyze QISKIT's `EfficientSU2` 2-local circuit as an example of a commonly chosen circuit layout.

After this initial example, we then provide the general mathematical setup and develop the dimensional expressivity analysis in Sec. 3. In particular, Sec. 3.4 contains the technical details for QISKIT's `EfficientSU2` 2-local circuit as discussed in Sec. 2. Although Sec. 3 is mathematically very technical, we embed further illustra-

tive examples for the mathematical statements obtained.

In Sec. 4, we apply the dimensional expressivity analysis to the case of a translationally invariant 4-qubit circuit. This circuit was adapted from a well-working non-translationally invariant quantum circuit. However, incorporation of translational invariance led to a number of problems, such that the VQS algorithm failed to find the ground state. The example discussed in Sec. 4 therefore highlights how to use the dimensional expressivity analysis to identify circuit design shortcomings and how to use the obtained information to custom design a maximally expressive circuit.

In Sec. 5, we generalize the ideas underlying the example of Sec. 4 and analyze translationally invariant quantum circuits and momentum sectors of the physical state space. This serves as an example analysis for any type of physical symmetry that is generated through spectral invariants. As such, the ideas in Sec. 5 can be generalized further to obtain dimensions of physical state spaces and the necessary information underlying the custom design of maximally expressive quantum circuits.

Reducing the complete state space of a quantum device to the physical state space of a given model furthermore introduces questions surrounding the ability of a VQS to find the solution. For VQSs using a local optimizer in parameter space, this effectively translates into the question whether or not the physical state space is path connected and whether or not a path from the initial parameter set to a parameter set for the solution can be found using only information available in the physical state space, that is, information that can be obtained from quantum device calls. These path connectedness questions are addressed in App. A, since they are highly technical and explore a direction that is not necessary to delve into deeply for the examples provided in this paper. However, they can be helpful in applications of the dimensional expressivity analysis because they address some non-trivial mathematical properties of the dimensional expressivity analysis that can lead to unexpected behaviors such as failure of the VQS to converge.

Since the dimensional expressivity analysis requires complete state information, it is unlikely that an efficient classical algorithm exists to per-

---

[1]The codimension of a sub-manifold M' of a finite-dimensional manifold M is defined as the difference between the dimensions: $\mathrm{codim}(M') = \dim(M) - \dim(M')$.

form this analysis. In particular, any classical algorithm is likely to require exponential memory in the number of qubits even if there are only two parameters in the given circuit. In Sec. 6, we will therefore discuss a hybrid quantum-classical algorithm performing the dimensional expressivity analysis. Under some assumptions on the type of parametric gates used, this algorithm performs at polynomial cost in the number of circuit parameters, for both the classical part of the algorithm and in the number of quantum device calls. The memory requirements are polynomial in the number of circuit parameters for the classical part of the dimensional expressivity analysis, while the quantum part of the algorithm requires one ancilla qubit and six additional gates (independent of the number of qubits and parameters, cf. Sec. 6.1). As such, the proposed dimensional expressivity analysis has the potential to be automated and could be used for on the fly circuit optimization. A hardware implementation on IBMQ is discussed in Sec. 6.3. The technical details of the hardware experiments are presented in App. B.

We summarize and discuss our results in Sec. 7.

2 QISKIT's Efficient SU(2) 2-local circuit

As a simple introductory example, let us analyze QISKIT's [33] hardware efficient SU(2) 2-local circuit `EfficientSU2(3, reps=N)`, which the QISKIT documentation [34] proposes as "a heuristic pattern that can be used to prepare trial wave functions for variational quantum algorithms or classification circuit for machine learning." This circuit consists of $N + 1$ blocks of R_Y and R_Z gates applied to every qubit. These blocks are interlaced with N blocks containing $\mathrm{CNOT}(q, q')$ gates for all $q < q'$. For example, setting $N = 2$, this circuit is given in the upper panel of Fig. 1 on the next page.

The method allows the user to input various parameters in order to change the layout of this circuit, including different gates in the $N + 1$ single-qubit gate blocks and different entanglers, such as a circular CNOT pattern $\mathrm{CNOT}(q, q+1)$,

i.e.,
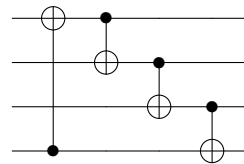
in the four-qubit case (to make the pattern more explicit). Leaving all optional parameters at their default value yields the circuit in the upper panel of Fig. 1.

Our aim for this section is to highlight some of the knowledge that can be gained using the dimensional expressivity analysis for the `EfficientSU2(3, reps=N)` circuit. As such, the analysis will be non-technical. For technical details we refer to Sec. 3, where we describe the method in general, in particular to Sec. 3.4 for specific details of the analysis regarding `EfficientSU2(3, reps=N)`. For a hands-on example that incorporates additional translational symmetry, we refer to Sec. 4. The technical details of incorporating such symmetries are given in Sec. 5.

2.1 Removing interdependent parameters

The state space of a three-qubit quantum device has real dimension 15, because it is the unit sphere of a 2^3-dimensional complex Hilbert space. Hence, a circuit without interdependent parameters has no more than 15 parameters, i.e., `EfficientSU2(3, reps=N)` with $N \geq 2$ must have at least three parameters that do not contribute to the circuits expressivity. Of course, in the $N = 2$ case, it is not possible to remove any three gates arbitrarily. For example, removing the three initial R_Y rotations renders the entire first layer ineffective. Thus, it effectively removes six gates and the resulting circuit cannot be maximally expressive. Similarly, there is no unique choice of parameter removal, and different applications might require different choices to be made. Dimensional expressivity analysis therefore aims to identify the parameters that need to be removed. In particular, gate prioritization will be discussed in Sec. 2.3. A more reasonable reduction to 15 parameters would be to remove the last three R_Z gates and consider the circuit given in the lower panel of Fig. 1. For most parameter sets, this will indeed provide a maximally expressive circuit. `EfficientSU2(3, reps=2)` there-
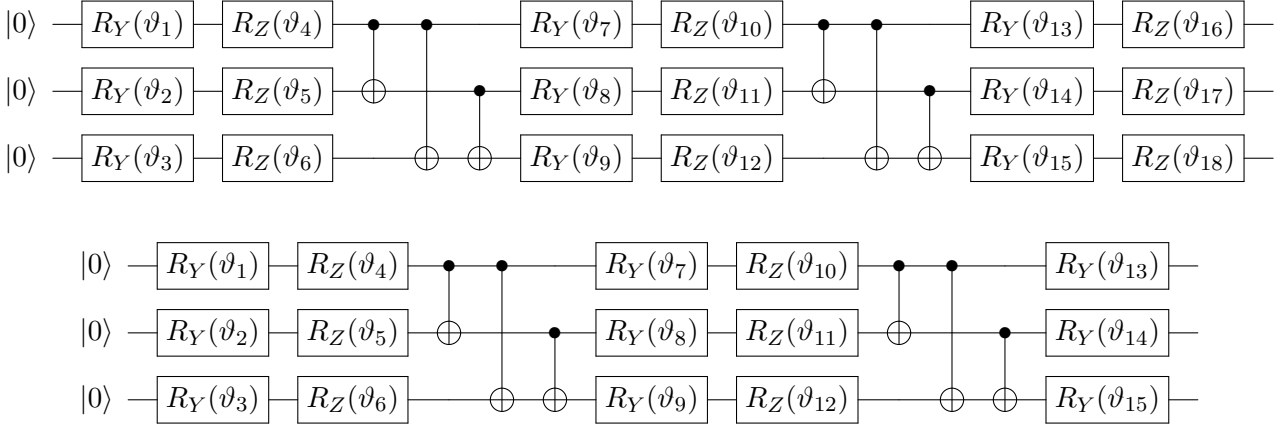
Figure 1: Upper panel: QISKIT's `EfficientSU2` 2-local circuit with $N = 2$ for three qubits. Lower panel: Reduction of QISKIT's `EfficientSU2` 2-local circuit with $N \geq 2$ to a maximally expressive circuit for three qubits using the minimal number of parameters.

fore provides an interesting test bed of the dimensional expressivity analysis, because it is simple enough to verify the analysis results, while being sufficiently complex to obtain interesting insights.

Dimensional expressivity analysis (Sec. 3.4) using random parameter values $\vartheta_1, \ldots, \vartheta_{18}$ shows that `EfficientSU2(3, reps=2)` is maximally expressive with 15 independent parameters (that is, the circuit is a locally surjective map into the state space). In fact, the circuit can be reduced to the circuit given in the lower panel of Fig. 1 which is still maximally expressive and only contains independent parameters. In this sense, `EfficientSU2(3, reps=2)` with global choice $\vartheta_{16} = \vartheta_{17} = \vartheta_{18} = 0$ is the most efficient, maximally expressive circuit in the `EfficientSU2(3, reps=N)`-family. In particular, it is not advisable to use `EfficientSU2(3, reps=N)` with $N > 2$ as it introduces additional parameters (and thus additional work for the optimizer) without increasing the expressivity of the circuit.

However, this maximal expressivity property is not true for all parameter sets (cf. Sec. 3.2). Some singular parameter sets exist. For example, a common starting position for a variational quantum simulation is likely to be $\vartheta_{\text{initial}} = 0$, i.e., all ϑ_j are initially set to 0. But starting with $\vartheta_{\text{initial}} = 0$, there are only seven independent parameters $\vartheta_1, \vartheta_2, \vartheta_3, \vartheta_4, \vartheta_7, \vartheta_8,$ and ϑ_{13}. This reduction is most easily seen with $\vartheta_4, \vartheta_5,$ and ϑ_6. If $\vartheta_1 = \vartheta_2 = \vartheta_3 = 0$, then $R_Z(\vartheta_4)$, $R_Z(\vartheta_5)$, and

$R_Z(\vartheta_6)$ all act on their respective qubit which is in the state $|0\rangle$. In other words, they all have the same effect of introducing a relative phase. As such, the three parameters are mutually dependent and we consider ϑ_4 to be the "independent" parameter because it is the first parameter (with respect to the chosen ordering of parameters) that generates arbitrary phases.

This effect can be detrimental to local optimizers. In particular, gradient based optimizers using `EfficientSU2(3, reps=2)` starting with $\vartheta = 0$ might be trapped in an artificial local minimum purely because a variation of the seven independent parameters remains inside a level set[2] of the energy landscape and hence the gradient is rendered zero. Replacing the starting value $\vartheta_{\text{initial}} = 0$ with a small random perturbation $\vartheta_{\text{initial}} = \varepsilon$ is unlikely to result in such an artificial minimum and provides the full dimensional expressivity of the circuit.

In the light of these singular sets, for which the dimensional expressivity is not maximal, it is interesting to note that performing the dimensional expressivity analysis on random parameters suffices to make general statements. This is a consequence of the generic structure of this type of circuit. The singular parameters are precisely the zeros of a function comprised of sums and products of sines and cosines (cf. Sec. 3.2). In other words, they are a lower dimensional subset of the

---

[2]A level set $L_c(f)$ of a function f of n variables is a set where the function takes on a given constant value c: $L_c(f) = \{(x_1, \cdots, x_n) \mid f(x_1, \cdots, x_n) = c\}$.
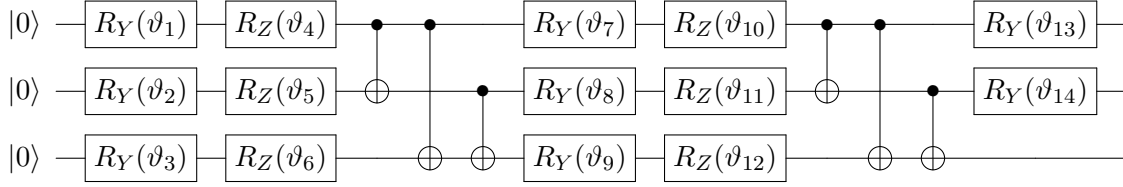
Figure 2: Reduction of QISKIT's `EfficientSU2` 2-local circuit similar to the lower panel of Fig. 1, but with removing the option to change a global phase.

parameter space. This implies that drawing[3] random parameters avoids this set of singular parameters with probability 1. Fundamentally, this is the reason why a random perturbation on the initial parameter set (like $\vartheta_{\text{initial}} = 0$ here) avoids artificial minima with probability 1 as well.

2.2 Removing global phases

The $\vartheta_{\text{initial}} = 0$ example also highlights a particular curiosity of the dimensional expressivity analysis. Above, we considered ϑ_4 to be an independent and "relevant" parameter because it introduces a global phase. Mathematically, ϑ_4 is necessary because it is impossible to generate an arbitrary vector in the unit sphere of the 2^3-dimensional complex Hilbert space if it is not possible to generate arbitrary phases. However, for quantum computing applications, such a global phase is irrelevant. We would therefore be interested in removing any gates or parameters whose only effect is a global phase change.

Assuming we wish to analyze a circuit $C(\vartheta)$
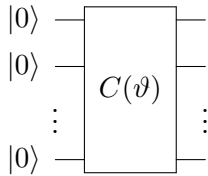
we may instead analyze the circuit $\tilde{C}(\varphi, \vartheta)$, which is given by $C(\vartheta)$ with an additional $R_Z(\varphi)$ gate:
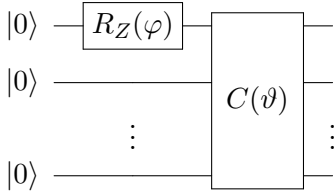
The additional gate in $\tilde{C}$ forces the parameter φ to be independent (being that it is the first parameter) and its effect is the introduction of a relative phase. Due to compactness of the parameter space and lower semi-continuity of parameter independence (cf., Sec. 3.2), checking this relative phase generation is sufficient to conclude about global phase symmetry of the circuit (cf., Sec. 3.5). Therefore, in the remaining part of the paper, we will call such phases "global" independently of whether they are applied to a single qubit or multiple qubits.

When analyzing $\tilde{C}$ instead of C, we may encounter a situation in which adding ϑ_j to the circuit C enlargens the set of reachable states (which is why it is considered an independent parameter) but the use of ϑ_j contributes at most a global phase factor to the set of reachable states (which is why we would want to consider ϑ_j as a dependent parameter). More precisely, dimensional expressivity analysis of C yields independence of the parameter ϑ_j but every state $|\psi\rangle$, that can be reached with C using only the parameters ϑ_1, ..., ϑ_j, is of the form $e^{i\alpha} |\psi'\rangle$ where $|\psi'\rangle$ can be reached with C using only the parameters ϑ_1, ..., ϑ_{j-1}. This precise situation presents itself when analyzing ϑ_4 in `EfficientSU2(3, reps=2)` at $\vartheta_{\text{initial}} = 0$. Analyzing $\tilde{C}$ instead of C, such global phases can already be generated using the parameters φ, ϑ_1, ..., ϑ_{j-1}. The addition of ϑ_j therefore no longer enlargens the set of reachable states and ϑ_j will be identified as a dependent parameter.

In application to `EfficientSU2(3, reps=2)` at $\vartheta_{\text{initial}} = 0$, we therefore obtain the six independent parameters ϑ_1, ϑ_2, ϑ_3, ϑ_7, ϑ_8, and ϑ_{13}. Application to `EfficientSU2(3, reps=2)` at random values of ϑ shows that for all tested values the parameter ϑ_{15} adds the global phase. Thus, if we globally set $\vartheta_{15} = 0$ as well, and only consider the circuit in Fig. 2, we still have a max-
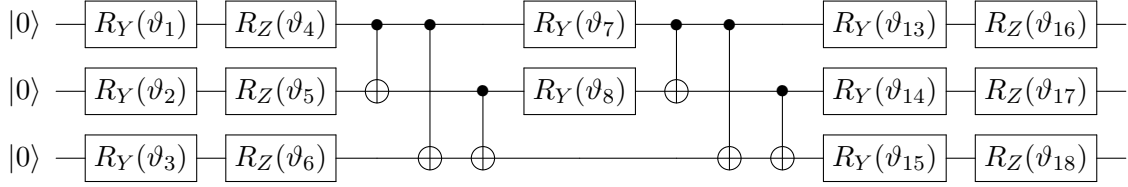
Figure 3: Reduction of QISKIT's `EfficientSU2` 2-local circuit similar to Fig. 2, but with different gate ordering.

imally expressive circuit – in the sense relevant to quantum computing – and we have optimized the circuit by removing the option to only change a global phase factor.

Of course, as an analysis tool, this idea can be generalized to other symmetries we wish to avoid. As long as we can add a one parameter block U
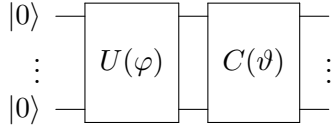
we can remove parameters from the quantum circuit which only add a symmetry that is to be avoided. We will discuss this procedure in more detail in Sec. 3.5.

2.3 Gate and parameter priority in dimensional expressivity analysis

The final circuit given in the lower panel of Fig. 1 is maximally expressive and has no redundant parameters. It is therefore optimal from the dimensional expressivity analysis point of view. However, it may not be the optimal reduced circuit for certain VQS applications. For example, it may be useful to have a circuit that can be easily initialized in an arbitrary product state. A possibility for such a circuit is the original `EfficientSU2(3, reps=N)`. Setting all parameters except the final layer to zero, reduces the circuit to
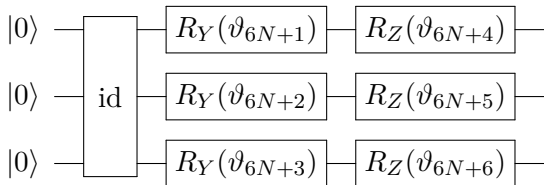

which is maximally expressive and free from redundant parameters for product states (this will be shown in Sec. 3.3).

However, generating arbitrary product states is at first sight not as easy for the circuit

in Fig. 2 as for `EfficientSU2(3, reps=N)`. If we want to ensure that the reduced circuit can generate product states just as easily as `EfficientSU2(3, reps=N)`, then we need to ensure the same $R_Y R_Z$ structure at the end of the reduced circuit. Luckily, we can adjust the dimensional expressivity analysis slightly in order to force this structure of the reduced circuit.

If ϑ_j and ϑ_k are mutually dependent and $j < k$, then the dimensional expressivity analysis will consider ϑ_j to be "independent" and ϑ_k "redundant". This is based on the implicit parameter priority order $j < k$, i.e, "j is more important than k". In general, let us consider a circuit with the structure
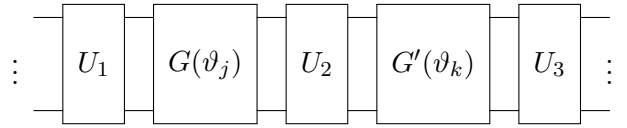
with mutually dependent ϑ_j and ϑ_k and $j < k$. If we wish to prioritize the gate block G' over the gate block G, then we can swap ϑ_j and ϑ_k to analyze the circuit
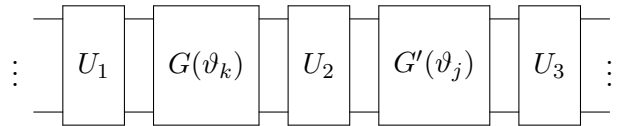
This change only affects the implicit gate priority as introduced by the ordering of parameters. The circuit itself is unchanged (provided the ranges for ϑ_j and ϑ_k are swapped as well). Dimensional expressivity analysis will still consider ϑ_j to be "independent" but now this means keeping G' rather than G.

Assuming we wish to prioritize the last layer in `EfficientSU2(3, reps=N)`, we can relabel the

parameters to give the structure
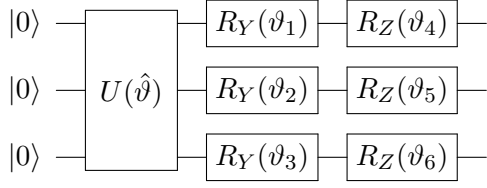
where $\hat{\vartheta} := (\vartheta_7, \ldots, \vartheta_{6N+6})$. Repeating the dimensional expressivity analysis (including the removal of a global phase) with this relabeling of the parameters (and returning to the initial parameter labeling) yields the maximally expressive redundancy free circuit in Fig. 3. Both circuits in Fig. 2 and Fig. 3 are maximally expressive and free from redundant parameters. From the dimensional expressivity analysis point of view they are therefore equivalent. For any given application this choice of gate and parameter prioritization may matter.

3 Circuit parameter reduction and manifold dimension

In order to quantify the expressivity of a given quantum circuit, we consider the dimension of the manifold of states that can be reached by the circuit. This choice is based on the fact that many quantum simulations use optimization loops to find optimal parameter sets for the quantum circuit. While it is possible for a low dimensional submanifold of reachable states to approximate every point of the state space (e.g., a dense spiral on the Bloch sphere), having a large codimension increases the likelihood of two states close in state space to be very far apart in parameter space. Hence, high codimension increases the likelihood of a local optimizer in a variational quantum simulation to fail.

Unfortunately, many quantum circuits have a large number of interdependent parameters. This means that each state can be reached through numerous choices of parameter values. In order to compute the dimension of the manifold of states reachable by a given quantum circuit, it is necessary to restrict the parameter space in such a way that the dimension can be extracted. This furthermore has the advantage that we can identify a minimal set of independent parameters. Optimizing only the independent parameters thus reduces computational effort and knowledge of the independent parameters can be used

as a guideline in circuit design. Similarly, knowing the codimension of the manifold of reachable states leads us to a natural measure of deficiency for a given quantum circuit. In circuit design, minimizing this deficiency will lead to maximal expressivity at a minimal number of parameters to optimize.

In Sec. 3.1 we will describe this parameter reduction in general terms. Of course, up to this point we have been assuming that the states reachable by a given quantum circuit form a manifold. This is not obvious a priori but we will show that this will at least be locally true, provided the conditions assumed in Sec. 3.1 are satisfied. In Sec. 3.2, we will discuss the local behavior of the parameter reduction and obtain estimates on the validity of reduced circuits, that is, how much variation of parameters is possible for any given reduction of the parameters. We will then consider a single-qubit example in Sec. 3.3 for explicit illustration and return to QISKIT's EfficientSU2 2-local circuit `EfficientSU2(3, reps=N)` in Sec. 3.4 to provide the details of the analysis in Sec. 2. Finally, in Sec. 3.5 we will discuss the removal of symmetries procedure that already allowed us to reduce `EfficientSU2(3, reps=N)` to a circuit that can generate arbitrary states up to a global phase.

3.1 General method

For the remainder of this paper, it is advantageous to consider a quantum circuit to be a map $C : P \to \partial B_{\mathcal{H}}$ where P is the parameter space and $\partial B_{\mathcal{H}}$ the unit sphere in the Hilbert space $\mathcal{H}$. For a Q-qubit system, $\mathcal{H} = \mathbb{C}^{2^Q}$. It is important to note that this definition of a quantum circuit now includes the evaluation on the initial state. For example, if the initial state of a single-qubit system is $|0\rangle$ and the gate sequence is $R_X(\vartheta)$, then

$$C(\vartheta) = R_X(\vartheta)|0\rangle = \cos\frac{\vartheta}{2}|0\rangle - i\sin\frac{\vartheta}{2}|1\rangle. \quad (1)$$

If P is a manifold (e.g., if the circuit contains N rotation gates and the entangling layers do not contain any parameters, P will be a flat torus $(\mathbb{R}/2\pi\mathbb{Z})^N$) and P locally contains a submanifold P_C such that $\check{C} := C|_{P_C} : P_C \to \partial B_{\mathcal{H}}$ is a diffeomorphism, then C defines a submanifold of $\partial B_{\mathcal{H}}$. To show that this is indeed the case, the important questions to ask are

(i) Does C locally define a submanifold of $\partial B_{\mathcal{H}}$?

(ii) If C locally defines a submanifold, what is its codimension?

(iii) If C locally defines a submanifold, can we restrict the parameter space to a minimal set?

Although the image of C is naturally described as a subset of a complex Hilbert space, because the parameter space is real, we have to treat the image of C as a real manifold. Furthermore, the questions (i), (ii), and (iii) are not independent of each other. In fact, the best way to show that we have a manifold already requires the restricted circuit $\check{C}$. Hence, it is advantageous to start with the construction of the parameter space restriction assuming that the image of C is a manifold and then show that the image of the restricted parameter set is already locally surjective on the image of C.

To restrict the parameter space locally, we can look at the tangent space of the image. If we have N parameters, then the tangent space is spanned by $\partial_1 C, \ldots, \partial_N C$. Thus, it is interesting to consider the Jacobian

$$C'(\vartheta) = \begin{pmatrix} | & | & & | \\ \partial_1 C(\vartheta) & \partial_2 C(\vartheta) & \ldots & \partial_N C(\vartheta) \\ | & | & & | \end{pmatrix}. \tag{2}$$

The rank of $C'(\vartheta)$ therefore gives us the number of independent parameters of C at $\vartheta \in P$. Furthermore, if we consider $J_\vartheta \subseteq \{1, \ldots, N\}$ such that $\{\partial_j C(\vartheta); j \in J_\vartheta\}$ is a maximal linearly independent subset of $\{\partial_j C(\vartheta); 1 \leq j \leq N\}$, then we can define the restriction $\check{C}_\vartheta$ in a neighborhood of ϑ by keeping ϑ_j constant whenever $j \notin J_\vartheta$. This implies that the columns of the Jacobian $\check{C}'_\vartheta(\vartheta)$ are precisely the $\partial_j C(\vartheta)$ with $j \in J_\vartheta$. In other words, $\check{C}_\vartheta$ is locally an immersion and any state reached by variation of parameters ϑ_j with $j \notin J_\vartheta$ can also be reached varying the ϑ_k with $k \in J_\vartheta$. To an abstract degree, this answers all three questions.

(i) The restricted circuit $\check{C}_\vartheta$ locally defines a submanifold of $\partial B_{\mathcal{H}}$ since it is an immersion and an immersion is locally an embedding.

(ii) The dimension of the manifold locally defined through $\check{C}_\vartheta$ has the dimension $\#J_\vartheta$.

Thus, the codimension of the circuit $\check{C}_\vartheta$ is $\dim \mathcal{H} - 1 - \#J_\vartheta$ (where $\mathcal{H}$ is the 2^Q-dimensional Hilbert space of a Q-qubit system).

(iii) The restriction to $\check{C}_\vartheta$ is minimal, that is, any further restriction of the circuit or parameter space will not be surjective on the image of the original circuit anymore and, therefore, reduce expressivity of the circuit.

3.2 Lower-semicontinuity and constancy of independent parameters

In this section, we will discuss the local behavior of the parameter reduction procedure. This will be of fundamental importance for estimating the set of parameters on which any given restricted quantum circuit is valid and for identification of singular points at which the dimensional expressivity of a given quantum circuit is reduced.

It is important to note that the circuit manifold dimension is lower semi-continuous, that is, if the restricted circuit $\check{C}_\vartheta$ defines a manifold of dimension n in a neighborhood of ϑ then the image of C has dimension *at least* n. It is possible that a small perturbation of ϑ increases expressivity but it cannot decrease.

This claim follows from lower-semicontinuity of $\mathrm{rank}C'(\vartheta)$ because the dimension of the manifold defined by $\check{C}_\vartheta$ is precisely the rank of $C'(\vartheta)$. This can be proven directly by assuming $\mathrm{rank}C'(\vartheta) =: r$. Then there exists a non-singular $r \times r$-submatrix C_0 of $C'(\vartheta)$. Thus $\det C_0 \neq 0$ and continuity of the determinant implies that for any sufficiently small perturbation $\tilde{C}$ of $C'(\vartheta)$ the $r \times r$-submatrix of $\tilde{C}$ corresponding to the same indices as C_0 must be non-singular as well. In other words, $\mathrm{rank}\tilde{C} \geq \mathrm{rank}C'(\vartheta)$ holds and lower-semicontinuity is proven.

Conversely, it is possible for the dimension to jump upwards under small perturbations. This behavior has already been observed in Sec. 2. An explicit example can be found in

`EfficientSU2(2, reps=0)`, i.e., the circuit

$$
\begin{aligned}
C(\vartheta) &:= R_{Z_2}(\vartheta_4) R_{Z_1}(\vartheta_3) R_{Y_2}(\vartheta_2) R_{Y_1}(\vartheta_1) \ket{00} \\
&= e^{-i\frac{\vartheta_3+\vartheta_4}{2}} \cos\frac{\vartheta_2}{2} \cos\frac{\vartheta_1}{2} \ket{00} \\
&\quad + e^{-i\frac{-\vartheta_3+\vartheta_4}{2}} \cos\frac{\vartheta_2}{2} \sin\frac{\vartheta_1}{2} \ket{01} \\
&\quad + e^{-i\frac{\vartheta_3-\vartheta_4}{2}} \sin\frac{\vartheta_2}{2} \cos\frac{\vartheta_1}{2} \ket{10} \\
&\quad + e^{i\frac{\vartheta_3+\vartheta_4}{2}} \sin\frac{\vartheta_2}{2} \sin\frac{\vartheta_1}{2} \ket{11} .
\end{aligned}
\tag{3}
$$

If $\vartheta_1 = \vartheta_2 = 0$, then $\partial_3 C(0,0,\vartheta_3,\vartheta_4)$ and $\partial_4 C(0,0,\vartheta_3,\vartheta_4)$ coincide, reducing the manifold dimension of $\check{C}_{(0,0,\vartheta_3,\vartheta_4)}$ to 3. However, this is not the case for any small perturbation in $(\vartheta_1, \vartheta_2)$.

This lower semi-continuity property of the circuit reduction indicates that it may be advantageous to introduce small random perturbations if a quantum simulation appears stuck in what seems to be a local minimum. In particular, if an initial parameter set is generated by some analytic method, an implementation might be improved through a small random perturbation of the initial parameter set.

The lower-semicontinuity property has shown that a small perturbation of the parameters can increase the number of independent parameters and that independent parameters stay independent under small perturbation. Crucially, it also allows us to estimate how large such a perturbation may be. Let $C_0(\vartheta_0)$ be the maximal invertible submatrix as determined in the lower-semicontinuity proof above. We will denote the $\text{rank}(C_0(\vartheta_0)) \times \text{rank}(C_0(\vartheta_0))$-submatrix of $C'(\vartheta)$ corresponding to the same indices as $C_0(\vartheta)$ for any parameter ϑ.[4]

Continuity of $\det \circ C_0$ implies the existence of

---

[4]To illustrate this submatrix construction, consider $C'(\vartheta) := \begin{pmatrix} a(\vartheta) & b(\vartheta) & c(\vartheta) \\ d(\vartheta) & e(\vartheta) & f(\vartheta) \\ g(\vartheta) & h(\vartheta) & i(\vartheta) \end{pmatrix}$. At some value ϑ_0, point evaluation may yield $C'(\vartheta_0) = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 2 & 1 \end{pmatrix}$. This matrix is of rank 2 and has the invertible 2×2-submatrix $C_0(\vartheta_0) = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} a(\vartheta_0) & b(\vartheta_0) \\ g(\vartheta_0) & h(\vartheta_0) \end{pmatrix}$. The submatrix of $C'(\vartheta)$ corresponding to the same indices as $C_0(\vartheta_0)$ is then given by $C_0(\vartheta) = \begin{pmatrix} a(\vartheta) & b(\vartheta) \\ g(\vartheta) & h(\vartheta) \end{pmatrix}$.

$R > 0$ such that

$$
\forall \vartheta \in B(\vartheta_0, R) : \ \det(C_0(\vartheta)) \neq 0
\tag{4}
$$

holds. This implies that the independent parameters identified for ϑ_0 remain independent for all ϑ in the open ball $B(\vartheta_0, R) := \{\vartheta; \ \text{dist}(\vartheta, \vartheta_0) < R\}$ where dist denotes the distance in parameter space. We can estimate such a value of R by choosing an initial $R_0 > 0$ and computing the derivative of $\det \circ C_0$ on the closed ball $B[\vartheta_0, R_0]$. Setting $\delta := \max_{\vartheta \in B[\vartheta_0, R_0]} \|(\det \circ C_0)'(\vartheta)\|_{\ell_2}$ where $\|(\det \circ C_0)'(\vartheta)\|_{\ell_2}$ denotes the ℓ_2-norm of the gradient $(\det \circ C_0)'(\vartheta)$, we obtain the estimate $R \geq \min\{R_0, \delta^{-1} |\det(C_0(\vartheta_0))|\}$.

We will call this R the radius of validity for a reduced circuit. It is of particular importance because it guarantees that no singular points are closer than R to the parameter set at which a reduced circuit was constructed. As such, can be used to anticipate problems arising from a loss of dimensional expressivity. We will therefore make heavy use of the radius of validity in Sec. 3.3, where we will construct reduced circuits for a single-qubit example, and in App. A, where it will be used to identify cases in which a quantum simulation can fail due to loss of dimensional expressivity. Fig. 8 and Fig. 9 in particular show how the radii of validity behave on paths with and without singular points.

To illustrate this estimation of the radius of validity, let us consider the single-qubit circuit

$$
C(\vartheta) := R_X(\vartheta) \ket{0} = \cos\frac{\vartheta}{2} \ket{0} - i \sin\frac{\vartheta}{2} \ket{1} .
\tag{5}
$$

The Jacobian is therefore given by

$$
C'(\vartheta) = \frac{1}{2} \begin{pmatrix} -\sin\frac{\vartheta}{2} \\ -i\cos\frac{\vartheta}{2} \end{pmatrix} .
\tag{6}
$$

However, this Jacobian is over $\mathbb{C}$ whereas our analysis is over $\mathbb{R}$. We therefore need to split real and imaginary parts, doubling the dimension, and yielding

$$
C'(\vartheta) = \frac{1}{2} \begin{pmatrix} -\sin\frac{\vartheta}{2} \\ 0 \\ 0 \\ -\cos\frac{\vartheta}{2} \end{pmatrix} .
\tag{7}
$$

For any point with $-\frac{1}{2}\sin\frac{\vartheta}{2} \neq 0$, the submatrix C_0 can be chosen as the 1×1-matrix $\left(-\frac{1}{2}\sin\frac{\vartheta}{2} \right)$. Hence, $\det(C_0(\vartheta)) = -\frac{1}{2}\sin\frac{\vartheta}{2}$ and

$(\det \circ C_0)'(\vartheta) = -\frac{1}{4}\cos\frac{\vartheta}{2}$ imply $\delta \leq \frac{1}{4}$ for any choice of R_0 as well as the estimate $R \geq 2\left|\sin\frac{\vartheta}{2}\right|$ provided $R_0 \geq 2$ is chosen.

Unfortunately, the estimate for the radius of validity R for the reduction to independent parameters converges to 0 as $\frac{\vartheta}{2}$ converges to a zero of the sine. This limitation is artificial since the one parameter in the circuit is always independent which we can also see switching to $-\frac{1}{2}\cos\frac{\vartheta}{2}$ near zeros of the sine. This problem can be addressed using the fact that $\mathrm{rank}(A) = \mathrm{rank}(A^*A)$ holds for all matrices A. Thus, if we choose $\tilde{C}_0$ to be the matrix of all independent columns of $C'(\vartheta)$, i.e.,

$$\tilde{C}_0 = \frac{1}{2}\begin{pmatrix} -\sin\frac{\vartheta}{2} \\ 0 \\ 0 \\ -\cos\frac{\vartheta}{2} \end{pmatrix}, \qquad (8)$$

then we obtain $\det(\tilde{C}_0^*\tilde{C}_0) = \frac{1}{4}$ for all ϑ. This shows directly that the circuit has only independent parameters and that this is valid for all parameters.

In practical terms, the lower-semicontinuity property and the corresponding radius of validity imply that the pointwise performed dimensional expressivity analysis is valid locally. More precisely, if we choose a point Θ in parameter space, identify the set independent parameters $\Theta_{i_1}, \ldots, \Theta_{i_k}$ at that point, and compute the radius of validity R at Θ, then all points ϑ with $\mathrm{dist}(\vartheta, \Theta) < R$ have independent parameters $\vartheta_{i_1}, \ldots, \vartheta_{i-k}$. In other words, independent parameters remain independent under small perturbations on point at which the dimensional expressivity analysis is performed.

Let us furthermore assume that the parameter space P is a compact manifold, as is the case for parametric quantum circuits containing only rotation gates and non-parametric gates. Then, we can use this perturbation resistance of independent parameters to find a finite set of points that need to be checked such that a set of independent parameters is known for every point in parameter space. To formalize this, let us define I_Θ as the set of independent parameters obtained from dimensional expressivity analysis at a point $\Theta \in P$, i.e., the parameters Θ_i with $i \in I_\Theta$ are independent. Let R_Θ be the corresponding radius of validity. Then, by lower-semicontinuity, the parameters ϑ_i with $i \in I_\Theta$ are indepen-

dent whenever $\mathrm{dist}(\vartheta, \Theta) < R_\Theta$. However, the open balls $B(\Theta, R_\Theta) = \{\vartheta \in P;\ \mathrm{dist}(\vartheta, \Theta) < R_\Theta\}$ form an open cover of P. By compactness, there exists a finite subcover, i.e., there exist finitely many $\Theta^{(1)}, \ldots, \Theta^{(n)}$ such that $P = \bigcup_{j=1}^{n} B(\Theta^{(j)}, R_{\Theta^{(j)}})$. Given any $\vartheta \in P$, we can therefore find a j_ϑ such that $\vartheta \in B(\Theta^{(j_\vartheta)}, R_{\Theta^{(j_\vartheta)}})$ and we conclude that the parameters ϑ_i with $i \in I_{\Theta^{(j_\vartheta)}}$ are independent. Hence, in order to find sets of independent parameters at every point in parameter space, it suffices to perform the dimensional expressivity analysis at finitely many points.

While this allows for the construction of finitely many minimal and maximally expressive quantum circuits that cover the entire parameter space, it should be noted that it may not always be prudent to construct a finite family of circuits, but to allow for an infinite family of circuits to exist on this finite cover. This leads to the notion of "seamless switching" as discussed in Sec. 3.3.3 and is, more generally, related to path connectivity questions (cf. App. A). It should also be noted that testing sufficiently many points in parameter space to obtain an a-priori atlas of independent parameters may not always be necessary. In some applications, it may be sufficient to choose a general method of constructing the circuit and performing the dimensional expressivity analysis for each update of a VQS. In other words, the idea would be to construct a set of independent parameters for the initial set of VQS parameters and to use the corresponding restricted circuit until the parameter update maps outside the current radius of validity. Then, a new circuit can be constructed and used until the VQS leaves its radius of validity. Of course, such an on the fly analysis requires an efficient and automated implementation which will be discussed in Sec. 6.

3.3 Single-qubit case example

In this section, we will illustrate the dimensional expressivity analysis on a single-qubit example. We will start our discussion of the single-qubit case analyzing a simple two-parameter circuit in Sec. 3.3.1. Since this first two-parameter circuit cannot be maximally expressive (that requires three parameters), we will then consider an artificial four-parameter circuit in Sec. 3.3.2 to highlight how to identify interdependent parameters and explicitly reduce the circuit. This will yield

multiple reduced circuits highlighting a behavior that is generically to be expected. In Sec. 3.3.3, we therefore discuss this generic behavior giving us multiple locally reduced circuits in more detail. We also prove that finitely many reduced circuits are sufficient to cover the entire parameter space provided the parameter space is compact. Finally, in Sec. 3.3.4 we will return the reduced four-parameter circuits of Sec. 3.3.2 and apply some symmetry removal techniques. These will be developed in Sec. 3.5 but at this point, they will already allow us to further simplify the reduced circuits of Sec. 3.3.2.

3.3.1 A two-parameter circuit

The first circuit we may consider is

$$C(\vartheta) := R_Z(\vartheta_2)R_X(\vartheta_1)\,|0\rangle \qquad (9)$$

with parameter space $(\mathbb{R}/2\pi\mathbb{Z})^2$. In the coordinates of the Hilbert space $\mathbb{C}^2$, $C(\vartheta)$ can thus be represented as

$$C(\vartheta) = \begin{pmatrix} \cos\frac{\vartheta_1}{2}\cos\frac{\vartheta_2}{2} - i\cos\frac{\vartheta_1}{2}\sin\frac{\vartheta_2}{2} \\ -i\sin\frac{\vartheta_1}{2}\cos\frac{\vartheta_2}{2} + \sin\frac{\vartheta_1}{2}\sin\frac{\vartheta_2}{2} \end{pmatrix}. \quad (10)$$

This implies

$$\partial_1 C(\vartheta) = \frac{1}{2}\begin{pmatrix} -\sin\frac{\vartheta_1}{2}\cos\frac{\vartheta_2}{2} + i\sin\frac{\vartheta_1}{2}\sin\frac{\vartheta_2}{2} \\ -i\cos\frac{\vartheta_1}{2}\cos\frac{\vartheta_2}{2} + \cos\frac{\vartheta_1}{2}\sin\frac{\vartheta_2}{2} \end{pmatrix} \qquad (11)$$

and

$$\partial_2 C(\vartheta) = \frac{1}{2}\begin{pmatrix} -\cos\frac{\vartheta_1}{2}\sin\frac{\vartheta_2}{2} - i\cos\frac{\vartheta_1}{2}\cos\frac{\vartheta_2}{2} \\ i\sin\frac{\vartheta_1}{2}\sin\frac{\vartheta_2}{2} + \sin\frac{\vartheta_1}{2}\cos\frac{\vartheta_2}{2} \end{pmatrix}. \qquad (12)$$

Since C is not constant, its Jacobian is either of rank 1 or 2. For the Jacobian to be of rank 1, there have to be non-zero *real* numbers α, β such that

$$0 = \alpha\partial_1 C(\vartheta) + \beta\partial_2 C(\vartheta). \qquad (13)$$

Checking all the possible cases of vanishing/non-vanishing sines/cosines, it is easy to see that this condition in unsatisfiable. Hence, both parameters ϑ_1 and ϑ_2 are independent at all points in parameter space, and the circuit defines a real two-dimensional submanifold of $\partial B_{\mathbb{C}^2}$.

Of course, in general, checking the solvability of such a set of equations is not efficient if ϑ is left a variable. Instead we can write down the Jacobian with real and imaginary parts split;

$$C'(\vartheta) = \frac{1}{2}\begin{pmatrix} -\sin\frac{\vartheta_1}{2}\cos\frac{\vartheta_2}{2} & -\cos\frac{\vartheta_1}{2}\sin\frac{\vartheta_2}{2} \\ \cos\frac{\vartheta_1}{2}\sin\frac{\vartheta_2}{2} & \sin\frac{\vartheta_1}{2}\cos\frac{\vartheta_2}{2} \\ \sin\frac{\vartheta_1}{2}\sin\frac{\vartheta_2}{2} & -\cos\frac{\vartheta_1}{2}\cos\frac{\vartheta_2}{2} \\ -\cos\frac{\vartheta_1}{2}\cos\frac{\vartheta_2}{2} & \sin\frac{\vartheta_1}{2}\sin\frac{\vartheta_2}{2} \end{pmatrix}. \qquad (14)$$

Checking this at a some point, e.g., $\vartheta = 0$ for simplicity, we obtain

$$C'(\vartheta) = \frac{1}{2}\begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 0 & -1 \\ -1 & 0 \end{pmatrix} \qquad (15)$$

which is of rank 2. Thus we conclude that at least in a neighborhood of $\vartheta = 0$ both parameters are independent. The matrix $\tilde{C}_0(\vartheta)$ of independent columns near $\vartheta = 0$ therefore coincides with $C'(\vartheta)$ and $\tilde{C}_0(\vartheta)^*\tilde{C}_0(\vartheta)$ simplifies to the $\frac{1}{4}\mathrm{id}_{2\times 2}$ where $\mathrm{id}_{2\times 2}$ is the 2×2-identity matrix. This proves that the radius of validity is infinite and therefore that the two parameters ϑ_1 and ϑ_2 are always independent.

3.3.2 Reduction of a four-parameter circuit

As a second example, let us consider the circuit

$$C(\vartheta) = R_Y(\vartheta_4)R_X(\vartheta_3)R_Z(\vartheta_2)R_X(\vartheta_1)\,|0\rangle. \quad (16)$$

Then

$$\partial_1 C(\vartheta) = \frac{-i}{2}R_Y(\vartheta_4)R_X(\vartheta_3)R_Z(\vartheta_2)XR_X(\vartheta_1)\,|0\rangle, \qquad (17)$$

$$\partial_2 C(\vartheta) = \frac{-i}{2}R_Y(\vartheta_4)R_X(\vartheta_3)ZR_Z(\vartheta_2)R_X(\vartheta_1)\,|0\rangle, \qquad (18)$$

$$\partial_3 C(\vartheta) = \frac{-i}{2}R_Y(\vartheta_4)XR_X(\vartheta_3)R_Z(\vartheta_2)R_X(\vartheta_1)\,|0\rangle, \qquad (19)$$

$$\partial_4 C(\vartheta) = \frac{-i}{2}YR_Y(\vartheta_4)R_X(\vartheta_3)R_Z(\vartheta_2)R_X(\vartheta_1)\,|0\rangle. \qquad (20)$$

If we assume we are using a local optimizer and the procedure starts with $\vartheta = 0$, we observe

$$\partial_1 C(0) = \frac{-i}{2}X\,|0\rangle = \frac{-i}{2}\,|1\rangle, \qquad (21)$$

$$\partial_2 C(0) = \frac{-i}{2}Z\,|0\rangle = \frac{-i}{2}\,|0\rangle, \qquad (22)$$

$$\partial_3 C(0) = \frac{-i}{2}X\,|0\rangle = \frac{-i}{2}\,|1\rangle, \qquad (23)$$

$$\partial_4 C(0) = \frac{-i}{2}Y\,|0\rangle = \frac{1}{2}\,|1\rangle. \qquad (24)$$

Of these vectors, the set $\{\partial_1 C(0), \partial_2 C(0), \partial_4 C(0)\}$ is linearly independent over $\mathbb{R}$. Therefore, we can ignore the parameter ϑ_3 for this step in the procedure. The submanifold locally defined by the reduced circuit $\check{C}(\check{\vartheta}) := C(\vartheta)|_{\vartheta_3=0}$ has dimension 3, which coincides with $\dim_{\mathbb{R}} \partial B_{\mathbb{C}^2}$. The reduced circuit $\check{C}(\check{\vartheta})$ is therefore maximally expressive at this point.

Considering

$$\check{C}_1(\check{\vartheta}) = R_Y(\check{\vartheta}_4)R_Z(\check{\vartheta}_2)R_X(\check{\vartheta}_1)\ket{0}, \qquad (25)$$

we know that the Jacobian has rank 3 near zero. Thus, we can compute $\check{C}_1'(\check{\vartheta})^*\check{C}_1'(\check{\vartheta})$ to estimate the area of validity of this reduced circuit. This computation yields

$$\check{C}_1'(\check{\vartheta})^*\check{C}_1'(\check{\vartheta}) = \begin{pmatrix} \frac{1}{4} & 0 & \frac{-\sin \check{\vartheta}_2}{4} \\ 0 & \frac{1}{4} & 0 \\ \frac{-\sin \check{\vartheta}_2}{4} & 0 & \frac{1}{4} \end{pmatrix} \qquad (26)$$

and therefore

$$\det\left(\check{C}_1'(\check{\vartheta})^*\check{C}_1'(\check{\vartheta})\right) = \frac{1 - \sin^2 \check{\vartheta}_2}{64}$$
$$= \frac{\cos^2 \check{\vartheta}_2}{64}. \qquad (27)$$

The reduced circuit is therefore valid as long as $\cos \check{\vartheta}_2 \neq 0$. Using the estimate on the radius of validity

$$R \geq \frac{\left|D(\check{\vartheta})\right|}{\max_\vartheta |D'(\vartheta)|} \qquad (28)$$

with $D(\check{\vartheta}) := \det\left(\check{C}_1'(\check{\vartheta})^*\check{C}_1'(\check{\vartheta})\right)$, we still obtain the estimate $R \geq 1$ for any $\check{\vartheta}$ with $\check{\vartheta}_2 \in \pi\mathbb{Z}$ since R_0 can be chosen arbitrarily.

At $\vartheta_2 = \frac{\pi}{2}$, we obtain that ϑ_4 is dependent on ϑ_1 and ϑ_2. Hence, we may choose the reduced circuit

$$\check{C}_2(\check{\vartheta}) = R_X(\check{\vartheta}_3)R_Z(\check{\vartheta}_2)R_X(\check{\vartheta}_1)\ket{0}, \qquad (29)$$

which now satisfies

$$\det\left(\check{C}_2'(\check{\vartheta})^*\check{C}_2'(\check{\vartheta})\right) = \frac{1 - \sin^2 \check{\vartheta}_2}{64}$$
$$= \frac{\sin^2 \check{\vartheta}_2}{64}. \qquad (30)$$

To ensure maximal expressivity, we may therefore consider the circuit

$$C^+(\vartheta) = R_Y(\vartheta_4)R_X(\text{const.})R_Z(\vartheta_2)R_X(\vartheta_1)\ket{0} \qquad (31)$$

whenever $\text{dist}(\vartheta_2, \pi\mathbb{Z}) \leq \frac{\pi}{4}$ and

$$C^-(\vartheta) = R_Y(\text{const.})R_X(\vartheta_3)R_Z(\vartheta_2)R_X(\vartheta_1)\ket{0} \qquad (32)$$

whenever $\text{dist}(\vartheta_2, \pi\mathbb{Z}) > \frac{\pi}{4}$. Repeating the analysis we performed above for const. $= 0$, we observe that this choice between C^+ and C^- always provides a valid maximally expressive circuit and the distance from any parameter set ϑ to the closest singular point of the restricted circuit $C^\pm$ is $\geq \frac{\pi}{4}$.

3.3.3 General reduction to finitely many circuits

For the particular example given in the previous subsection, none of this elaboration is necessary, but in general we can use this procedure to construct maximally expressive circuits. By starting with a redundant circuit C, we can choose a maximally expressive starting restriction $\check{C}_{\vartheta_0}$ at an initial ϑ_0 and compute/estimate its area of validity O_{ϑ_0}. Choosing a subset $\Omega_{\vartheta_0} \subseteq O_{\vartheta_0}$ (e.g., the interior of a compact) we can ensure that Ω_{ϑ_0} is uniformly bounded away from the set of singular points. If the starting circuit is sufficiently redundant, then each point ϑ outside Ω_{ϑ_0} has a maximally expressive restriction $\check{C}_\vartheta$ with area of validity Ω_ϑ, which, too, is uniformly bounded away from the singular points of $\check{C}_\vartheta$. Hence, if the parameter space is compact, there exists finitely many ϑ such that the corresponding finitely many Ω_ϑ still form an open cover of the parameter space. Each of these Ω_ϑ is associated with a maximally expressive restriction $\check{C}_\vartheta$, and restricting their validity to the Ω_ϑ implies that no singular point can be approximated in parameter space.

In the case above, although we could have restricted to two circuits by setting const. $= 0$ in both cases, it seemed more reasonable to still consider an infinite cover because this allows us to perform "seamless" switching, i.e., simply change which parameters are treated as constants. We could have chosen the finite cover

$$C_0^+(\vartheta) = R_Y(\vartheta_3)R_Z(\vartheta_2)R_X(\vartheta_1)\ket{0} \qquad (33)$$

whenever $\text{dist}(\vartheta_2, \pi\mathbb{Z}) \leq \frac{\pi}{4}$ and

$$C_0^-(\vartheta) = R_X(\vartheta_3)R_Z(\vartheta_2)R_X(\vartheta_1)\ket{0} \qquad (34)$$

whenever $\text{dist}(\vartheta_2, \pi\mathbb{Z}) > \frac{\pi}{4}$. However, that would require additional work whenever the algorithm crosses $\vartheta_2 = \frac{\pi}{4}$. In particular, if we are using C_0^+

crossing $\vartheta_2 = \frac{\pi}{4}$ to some point ϑ and want to switch to C_0^-, then we need to compute $\vartheta_{\text{new}} = (C_0^-)^{-1}(C_0^+(\vartheta))$ to find the new parameter set for the new circuit C_0^-. While this is possible (the $C_0^\pm$ are charts of the state space and thus diffeomorphisms), it is much more practical to consider the infinitely many restrictions $C^\pm$ as in the example above.

3.3.4 Global phase removal

Using the analysis techniques for removing symmetries, which we will develop in Sec. 3.5, we can further analyze $C^\pm$ and note that the common gates $R_Z(\vartheta_2)R_X(\vartheta_1)$ cannot generate an arbitrary phase. This means that the difference between C^+ and C^- is solely due to the ability of generating arbitrary phases. Since global phases

are irrelevant, we can consider the reduced circuit

$$\check{C}(\vartheta) := R_Z(\vartheta_2)R_X(\vartheta_1)\,|0\rangle\,. \qquad (35)$$

This circuit can generate an arbitrary single-qubit state up to a global phase. Using the analysis at the beginning of this current subsection, the circuit is free of singular points.

The same is true for the circuit $R_Z(\vartheta_2)R_Y(\vartheta_1)\,|0\rangle$. Hence, the first layer of QISKIT's EfficientSU2 2-local circuit `EfficientSU2(3, reps=N)` circuit (cf. Sec. 2) can generate arbitrary tensor products of single-qubit states (up to a global phase).

3.4 Return to QISKIT's EfficientSU2 2-local circuit

At this point, we can quickly return to the `EfficientSU2(3, reps=N)` and provide the details for the example of Sec. 2 as an illustration. Using the $N = 1$ circuit $C(\vartheta)$ (for simplicity of the output), we can compute the reduced row echelon form for $C'(\vartheta)$. At $\vartheta = 0$, we obtain

$$\begin{pmatrix} \vartheta_1 & \vartheta_2 & \vartheta_3 & \vartheta_4 & \vartheta_5 & \vartheta_6 & \vartheta_7 & \vartheta_8 & \vartheta_9 & \vartheta_{10} & \vartheta_{11} & \vartheta_{12} \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} \qquad (36)$$

with the dots indicating that the rest of each column is filled with zeros. Since the jth column represents the parameter ϑ_j, we can identify that parameter ϑ_9 depends on ϑ_3 and the parameters ϑ_5, ϑ_6, ϑ_{10}, ϑ_{11}, and ϑ_{12} all depend on ϑ_4. In comparison, the reduced row echelon form of $C'(\vartheta)$ evaluated at a random parameter set ϑ is (with probability 1) of the form $\begin{pmatrix} A \\ B \end{pmatrix}$ where A is

the 12×12 identity and B the 4×12 zero matrix; thus indicating that all twelve parameters are independent.

If, on the other hand, we consider the circuit $\tilde{C}(\varphi, \vartheta)$, which is leading $C(\vartheta)$ with $R_Z(\varphi)$ on the first qubit, and compute the reduced row echelon form of $\tilde{C}'(\varphi, \vartheta)$ at $\varphi = 0$ and $\vartheta = 0$, we obtain

$$\begin{pmatrix} \varphi & \vartheta_1 & \vartheta_2 & \vartheta_3 & \vartheta_4 & \vartheta_5 & \vartheta_6 & \vartheta_7 & \vartheta_8 & \vartheta_9 & \vartheta_{10} & \vartheta_{11} & \vartheta_{12} \\ 1 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix} \qquad (37)$$

with the dots again indicating that the rest of each column is filled with zeros. Once more we can identify ϑ_9 as dependent on ϑ_3 but also ϑ_4, ϑ_5, ϑ_6, ϑ_{10}, ϑ_{11}, and ϑ_{12} are all dependent on φ. Since, by construction, the impact of φ is only a global phase, we know that the gates corresponding to ϑ_4, ϑ_5, ϑ_6, ϑ_{10}, ϑ_{11}, and ϑ_{12} only change the phase if varied at $\vartheta = 0$.

In the `EfficientSU2(3, reps=2)` case, the reduced row echelon form of $C'(\vartheta)$ evaluated at random values of ϑ produces

$$\begin{pmatrix} \vartheta_1, \ldots, \vartheta_{15} & \vartheta_{16}, \vartheta_{17}, \vartheta_{18} \\ \text{id} & A \\ 0 & 0 \end{pmatrix} \qquad (38)$$

where A is some 15×3-matrix. This shows that $\vartheta_1, \ldots, \vartheta_{15}$ are independent. Similarly, if we introduce an initial $R_Z(\varphi)$ gate on the first qubit and compute the reduced row echelon form for $\tilde{C}'(\varphi, \vartheta)$ at random values of (φ, ϑ), we obtain

$$\begin{pmatrix} \varphi & \vartheta_1, \ldots, \vartheta_{14} & \vartheta_{15}, \ldots, \vartheta_{18} \\ 1 & 0 & A \\ 0 & \text{id} & B \\ 0 & 0 & 0 \end{pmatrix} \qquad (39)$$

which shows that the net contribution of ϑ_{15} is a global phase. Everything else can be generated by $\vartheta_1, \ldots, \vartheta_{14}$. It is important to note that although these are pointwise evaluations, sufficiently small perturbations cannot change the set of independent parameters. This is a corollary of the lower-semicontinuity property.

3.5 Removal of symmetries

As already alluded to at the end of Sec. 2 and Sec. 3.3, it can be important to construct a quantum circuit that avoids certain unwanted symmetries, like generating states that differ only by a global phase. Let us more generally consider an n-parameter circuit $C(\vartheta_1, \ldots, \vartheta_n)$ and suppose that the reduced row echelon form for $C'(\vartheta)$ is of the form

$$\begin{pmatrix} \text{id}_{n \times n} \\ 0_{(d-n) \times n} \end{pmatrix} \qquad (40)$$

where d is the real dimension of the Hilbert space, i.e., 2^{Q+1} if Q is the number of qubits. If we wish to avoid an unwanted symmetry that can be encoded using a parametric gate set $U(\varphi)$ (e.g., $U(\varphi) = R_Z(\varphi)$ on some qubit to generate a global phase as in Sec. 2 and Sec. 3.3), then instead of analyzing C, we may analyze $\tilde{C}(\varphi, \vartheta)$ given by
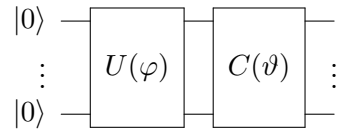
This encoding U should have at least one parameter value φ_0 for which $U(\varphi_0) = \text{id}$. This is necessary for U to "generate" the symmetry as opposed to "changing the circuit to contain the symmetry." Furthermore, the dimensionality of the parameter φ should coincide with the dimension of the unwanted symmetry, e.g., φ is one-dimensional for the global phase symmetry $U(1)$ or three-dimensional for an $SU(2)$ symmetry.

At this point, it is important to recall the connection between global, local, and pointwise testing of parameter dependence. A circuit is subject to a global symmetry if and only if the action of the symmetry maps the circuit manifold into itself. Given a compact symmetry group, it suffices to check this property locally for the symmetry group. Given a compact circuit manifold, it also suffices to check this property locally for the circuit manifold. The group properties

furthermore allow us to only check a neighborhood of the identity element of the symmetry group, as any other pair of open subsets U of the group and V of the circuit manifold can be replaced by $U' = g^{-1}U$ and $V' = g^{-1}V$ where $g \in U$. Then U' is a neighborhood of the identity. Finally, lower-semicontinuity of parameter independence implies that it suffices to test independence on a single point to obtain independence on a neighborhood of the point. Since we plan to use the dimensional expressivity analysis on an extended circuit $\tilde{C}(\varphi, \vartheta)$, this means that it suffices to check for parameter independence at φ_0 satisfying $U(\varphi_0) = \mathrm{id}$ and sufficiently many ϑ, as discussed in Sec. 3.2. In this sense, it is sufficient to consider global symmetries, such as the gate $R_Z(\varphi)$ which generates a global phase on the current state of the quantum circuit, in order to identify parameters that are responsible for global symmetries.

We can therefore identify the parameters generating a symmetry by looking at parameters that are shown as independent under dimensional expressivity analysis for the circuit C but dependent when analyzing $\tilde{C}$. For simplicity, let us assume a one-dimensional symmetry. One possible outcome for the reduced row echelon form of $\tilde{C}'(\varphi_0, \vartheta)$ would be

$$
\begin{pmatrix}
\varphi & \vartheta \\
1 & 0 \\
0 & \mathrm{id}_{n \times n} \\
0_{(d-n-1) \times 1} & 0_{(d-n-1) \times n}
\end{pmatrix}. \quad (41)
$$

If this is the case for almost all parameter sets ϑ, then all parameters $\vartheta_1, \ldots, \vartheta_n$ are independent and the symmetry encoded by U is not generated by the circuit $C(\vartheta)$.

The other option is that C does generate the symmetry encoded by U. If that is the case, then one[5] of the parameters $\vartheta_1, \ldots, \vartheta_n$ is dependent on φ and the remaining ϑ_j. Let this parameter be ϑ_m. The reduced row echelon form of $\tilde{C}'(\varphi_0, \vartheta)$ will then be

$$
\begin{pmatrix}
\varphi & \vartheta_1, \ldots, \vartheta_{m-1} & \vartheta_m & \vartheta_{m+1}, \ldots, \vartheta_n \\
1 & 0 & a & 0 \\
0 & \mathrm{id} & b & 0 \\
0 & 0 & 0 & \mathrm{id} \\
0 & 0 & 0 & 0
\end{pmatrix} \quad (42)
$$

---

[5]In general, the number of dependent parameters has to be equal to the dimension of the tested symmetry, i.e., the dimensionality of φ.

where b, both id, and each 0 are matrices/vectors of the appropriate dimension. Crucially, a is non-zero ($a = 0$ would violate Eq. 40). The vector b may or may not be zero. If $b = 0$, then ϑ_m has locally the same effect as U. The gate corresponding to ϑ_m thus locally only contributes the unwanted symmetry. In general, b is expected to be non-zero. In that case, the gate corresponding to ϑ_m is not locally equivalent to the unwanted symmetry but removal of the unwanted symmetry from the circuit manifold yields a submanifold that can locally be described by the reduced circuit after removal of the gate corresponding to ϑ_m. Thus, avoidance of the unwanted symmetry can locally be achieved by setting ϑ_m constant.[6]

While it is not clear a priori whether or not the same ϑ_m is always responsible for the additional symmetry, it is at least locally always true. In other words, if we find that the mth parameter (ϑ_m) generates an unwanted symmetry at a point ϑ, then there exists a neighborhood of ϑ such that the unwanted symmetry is always generated by the mth parameter.

The idea behind the proof is as follows. Since ϑ_m generates the unwanted symmetry at the point ϑ, its partial derivative of the circuit $\partial_m C(\vartheta)$ is a linear combination of the φ derivative and all other ϑ_j derivatives. In other words, we can write

$$
\partial_m C(\vartheta) = \alpha_0 \partial_\varphi \tilde{C}(\varphi_0, \vartheta) + \sum_{j \neq m} \alpha_j \partial_j C(\vartheta) \quad (43)
$$

for some values $\alpha_j \in \mathbb{R}$. In order to prove that ϑ_m is always responsible for the unwanted symmetry (at least in a neighborhood of ϑ), then we need to show that we can always find values α_j satisfying Eq. 43 if ϑ is subject to sufficiently small perturbations. This can be achieved using the implicit function theorem. We will work out the precise details of this in the remainder of this section.

Since we have n independent vectors $\partial_j C(\vartheta)$, we can assume without loss of generality that we are working in a representation of $\mathbb{R}^n$. Similarly, we may identify the parameter space (at least locally) with an open subset of $\mathbb{R}^n$. Furthermore, without loss of generality let $m = n$. Since

---

[6]We cannot simply remove the gate G corresponding to ϑ_m as this would replace $G(\vartheta_m)$ with the identity which changes the output of the circuit. If $G = R_X$ and $\vartheta_m = \pi$ then we would remove an X gate from the circuit and thus fundamentally change the quantum state.

$\partial_\varphi \tilde{C}(\varphi_0, \vartheta)$ is linearly dependent on the $\partial_j C(\vartheta)$, it too can be represented in $\mathbb{R}^n$. By Eq. 42, there now exist $\alpha_0, \ldots, \alpha_{n-1} \in \mathbb{R}$ such that

$$F(x,y) := \partial_n C(x) - y_0 \partial_\varphi \tilde{C}(\varphi_0, x) - \sum_{j=1}^{n-1} y_j \partial_j C(x) \tag{44}$$

satisfies $F(\vartheta, \alpha) = 0$. Using Eq. 42, we observe that $\partial_y F(\vartheta, \alpha)$ is invertible. The implicit function theorem therefore ensures the existence of a neighborhood U of ϑ, a neighborhood V of α, and a continuously differentiable $f : U \to V$ such that for all $(x, y) \in U \times V$

$$F(x, y) = 0 \iff y = f(x). \tag{45}$$

In particular, for all ϑ' in a neighborhood of ϑ there exists α' (in a neighborhood of α) such that

$$\partial_n C(\vartheta') = \alpha_0' \partial_\varphi \tilde{C}(\varphi_0, \vartheta') + \sum_{j=1}^{n-1} \alpha_j' \partial_j C(\vartheta'). \tag{46}$$

This proves that, at least locally, we can always keep the same parameter constant if we wish to remove an unwanted symmetry from the circuit.

4 Analysis and design of a translationally invariant 4-qubit circuit

In this section, we will apply the dimensional expressivity analysis to the case of a translationally invariant 4-qubit circuit. In addition to the question of maximal expressivity, which we focused on in Sec. 2, the aim of this analysis is to construct a quantum circuit that encodes a particular physical symmetry; in this case, translational invariance. As such, the analysis in this section will heavily depend on Sec. 5 in addition to Sec. 3.

The first circuit under consideration is a circuit that we adapted from a more general framework to study the translationally invariant Ising model, whose Hamiltonian is of the form

$$H_{\text{Ising}} = J \sum_{q=1}^4 X_q X_{q+1} + B \sum_{q=1}^4 Z_q,$$

where $X_5 = X_1$. However, upon integrating the translational invariance, we found the circuit failed to perform.

The implementation started with a known solution to an instance of the Ising model ($J = 0$ and $B = -1$) and then continuously deformed the Hamiltonian from the solved initial Hamiltonian $H_0 = -\sum_{q=1}^4 Z_q$ to the target Hamiltonian $H_1 = \sum_{q=1}^4 X_q X_{q+1} - \sum_{q=1}^4 Z_q$. Discretizing the path of the Hamiltonian $H(x) := (1-x)H_0 + xH_1$ into a large number of steps $H(x_i)$ and using the solution of $H(x_i)$ as initial guess for $H(x_{i+1})$, we aimed to ensure that a local (gradient based) optimizer would not incorrectly terminate due to a local minimum. However, the result is shown in Fig. 4(a). The ground state energy as obtained from a variational quantum simulation visibly deviates from the true ground state energy, although there is a sizable energy gap. This result was independent of the number of layers used for the circuit. Thus, the aim of this example is to use dimensional expressivity analysis to explain why something like this can happen and how to fix it.

Using a non-translationally invariant circuit that worked well for the more general Heisenberg model, the translational invariance was built into the circuit by replacing each single-qubit rotation gate $R_{O_q}(\vartheta) = e^{i\frac{\vartheta}{2} O_q}$, such as $R_{X_q}(\vartheta) = e^{i\frac{\vartheta}{2} X_q}$, with a rotation layer

$$L_O(\vartheta) := \prod_{q=1}^4 R_{O_q}(\vartheta) \tag{47}$$

and similarly for multi-qubit rotations, e.g.,

$$L_{X_q X_{q+1}}(\vartheta) := \prod_{q=1}^4 R_{X_q X_{q+1}}(\vartheta). \tag{48}$$

In other words, each gate is translated over all qubits and each translate has the same angle. The circuit in question then has an initial layer

$$L_0 := L_Z(\vartheta_2) L_X(\vartheta_1) \tag{49}$$

and arbitrarily many further layers of the form

$$L_j := L_Y(\vartheta_{4j+2}) L_Z(\vartheta_{4j+1}) L_X(\vartheta_{4j}) \\ \cdot L_{X_q X_{q+1}}(\vartheta_{4j-1}) \tag{50}$$

where $j \in \{1, \ldots, N\}$ is the layer index. The complete circuit with N layers for the momentum 1 sector (cf., Sec. 5) is then given by

$$C(\vartheta) := L_N \cdots L_1 L_0 |0\rangle. \tag{51}$$

In order to solve the translationally invariant Ising model, we started with an instance of the Ising model whose solution is known analytically (all parameters ϑ_k set to 0) and deformed the model continuously to the target model. This approach was chosen in order to avoid local minima in the energy landscape, which are likely to trap the state starting from a random configuration. Unfortunately, the solution failed to build up the entanglement we expected from the ground state of the final Hamiltonian. In fact, the circuit only produced product states as evidenced by the fact that the angles for the entangling layers $L_{X_qX_{q+1}}$ always remained zero. In other words, the direction of entangled states in the state space, which can be generated using this entangling layer, was orthogonal to the gradient of the energy landscape. Therefore, one of the reasons for the circuit's failure was that – up to product state optimization – we were stuck in a local minimum induced by the circuit.

Using the results of Sec. 3 and Sec. 5, we can analyze this circuit. The initial step of the analysis is to compute the tangent vectors of the circuit with respect to each parameter. These tangent vectors then span the tangent space of the manifold generated by the image of the circuit. Since we started in a known solution, we will exemplify the analysis using the initial parameter set $\vartheta_{\text{initial}} = 0$.

Computing the tangent space of the circuit manifold at $\vartheta_{\text{initial}} = 0$ according to Sec. 3 yields that the tangent space is spanned by

$$\partial_1 C(0) = \partial_{4j} C(0) = -\frac{i}{2} \sum_q X_q \,|0\rangle$$

$$\partial_2 C(0) = \partial_{4j+1} C(0) = -\frac{i}{2} \sum_q Z_q \,|0\rangle$$

$$\partial_{4j+2} C(0) = -\frac{i}{2} \sum_q Y_q \,|0\rangle \tag{52}$$

$$\partial_{4j-1} C(0) = -\frac{i}{2} \sum_q X_q X_{q+1} \,|0\rangle$$

for every value of N. The circuit therefore defines a 4-dimensional submanifold in a neighborhood of $\vartheta = 0$, and only the parameters ϑ_1, ϑ_2, ϑ_3, and ϑ_6 are relevant at the point the circuit failed. This explains why adding additional layers of the same structure has no impact in this case.

Naïvely, a reasonable idea to overcome this problem would be to replace the $L_{X_qX_{q+1}}$ en-
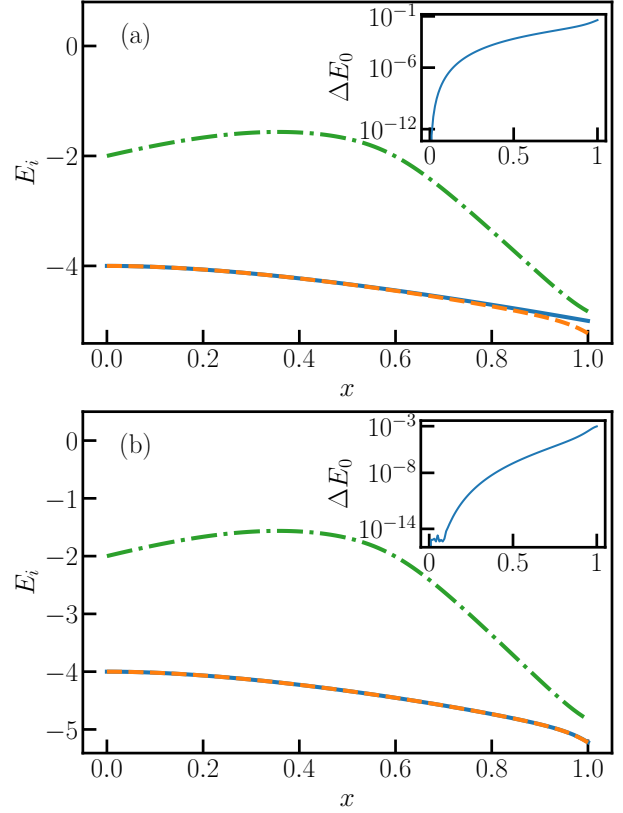

Figure 4: (a) Failure of gradient based variational quantum simulation to continually deform the known ground state of a Hamiltonian $H(0)$ to the target Hamitonian $H(1)$. The VQS ground state energy (blue solid line) deviates from the analytic ground state (orange dashed line). The first excited state energy (green dash-dotted line) is shown as well. The inset shows the relative deviation of the VQS ground-state energy from the exact result, which is of order $\mathcal{O}(0.1)$ for $x \to 1$. (b) Success of the same gradient based variational quantum simulation after adding a different entangler based on expressivity analysis, which reduces the deviation by two orders of magnitude (see inset in the lower panel).

tangler with a different translationally invariant entangler. However, the natural choices will fail, as well. An additional $L_{Y_qY_{q+1}}$ entangler would generate the direction $\partial_3 C(0)$ and an additional $L_{Z_qZ_{q+1}}$ would generate $\partial_2 C(0)$. The naïve choices for translationally invariant entanglers in combination with single-qubit rotations cannot generate a circuit manifold of dimension > 4. Yet the momentum 1 sector has dimension 11 (cf., Sec. 5).

Luckily, our analysis in Sec. 5 can guide us in overcoming the impasse. In Sec. 5, we will compute the dimension of the momentum eigensectors. In particular, we will construct a basis of eigenvectors of the momentum sectors by con-

structing equivalence classes. The construction of these basis states allows us to identify the missing directions from the circuit. Using an orthogonal set of these missing directions $|v_j\rangle$, we can find gates G_j that will generate the missing directions (here $|v_j\rangle = G_j |0\rangle$), and finally construct one-parameter layers $L_j(\vartheta_j) = \exp\left(-i\frac{\vartheta_j}{2}G_j\right)$. Adding these gates to the circuit adds the missing directions $|v_j\rangle$ to the tangent space.

Of course, there is a certain amount of freedom in the construction of these gates G_j. In this case, we chose the layers $L_{X_qY_{q+1}}$, $L_{X_qX_{q+2}}$, $L_{X_qY_{q+2}}$, $L_{X_qX_{q+1}X_{q+2}}$, $L_{Y_qY_{q+1}Y_{q+2}}$, $L_{X_qX_{q+1}X_{q+2}X_{q+3}}$, and $L_{X_qX_{q+1}X_{q+2}Y_{q+3}}$. However, we could have chosen $L_{X_qX_{q+1}Y_{q+2}}$ instead of $L_{Y_qY_{q+1}Y_{q+2}}$ to the same effect. In fact, any pair of X gates could be replaces by a pair of Y gates on the same qubits and vice versa.

This yields the (by construction minimal maximally expressive) circuit

$$
\begin{aligned}
C(\vartheta) := {} & L_{X_qX_{q+1}X_{q+2}Y_{q+3}}(\vartheta_{11}) \\
& \cdot L_{X_qX_{q+1}X_{q+2}X_{q+3}}(\vartheta_{10}) \\
& \cdot L_{Y_qY_{q+1}Y_{q+2}}(\vartheta_9) L_{X_qX_{q+1}X_{q+2}}(\vartheta_8) \\
& \cdot L_{X_qY_{q+2}}(\vartheta_7) L_{X_qX_{q+2}}(\vartheta_6) \\
& \cdot L_{X_qY_{q+1}}(\vartheta_5) L_{X_qX_{q+1}}(\vartheta_4) \\
& \cdot L_Y(\vartheta_3) L_Z(\vartheta_2) L_X(\vartheta_1) |0\rangle
\end{aligned}
\tag{53}
$$

and we can repeat the variational quantum simulation of Fig. 4(a) with this new circuit. The results are shown in Fig. 4(b). In particular, we can see that the ground state energy as obtained from the variational quantum simulation is now much closer to the analytic ground state energy, with a relative error that is two orders of magnitude smaller than the one in Fig. 4(a). Thus, this shows that the tailor-made circuit is not subject to the artificial restrictions which caused the original circuit to fail.

This construction can be generalized to larger numbers of qubits. For example, the two-qubit entanglers $L_{X_qX_{q+1}}$, $L_{X_qY_{q+1}}$, $L_{X_qX_{q+2}}$, $L_{X_qY_{q+2}}$ will no longer be sufficient to generate all two-qubit entangling directions. For more than four qubits, the entangling layers $L_{X_qX_{q+n}}$ and $L_{X_qY_{q+n}}$ for all $n \leq \frac{\#\text{qubits}}{2}$ are required.

Furthermore, it should be noted that the L_Y layer can be removed from the circuit because it only contributes the ability to generate arbitrary

phases (shown in Sec. 3.3), i.e., the circuit

$$
\begin{aligned}
C(\vartheta) := {} & L_{X_qX_{q+1}X_{q+2}Y_{q+3}}(\vartheta_{10}) \\
& \cdot L_{X_qX_{q+1}X_{q+2}X_{q+3}}(\vartheta_9) \\
& \cdot L_{Y_qY_{q+1}Y_{q+2}}(\vartheta_8) L_{X_qX_{q+1}X_{q+2}}(\vartheta_7) \\
& \cdot L_{X_qY_{q+2}}(\vartheta_6) L_{X_qX_{q+2}}(\vartheta_5) \\
& \cdot L_{X_qY_{q+1}}(\vartheta_4) L_{X_qX_{q+1}}(\vartheta_3) \\
& \cdot L_Z(\vartheta_2) L_X(\vartheta_1) |0\rangle
\end{aligned}
\tag{54}
$$

can generate arbitrary states with momentum 1 (cf., Sec. 5) up to a phase. By construction of the quantum simulation, we only needed to find a custom circuit for the momentum 1 sector. This is because having a non-vanishing energy gap throughout the entire quantum simulation implies a continuous deformation of the ground state from the known initial ground state to the target Hamiltonian ground state. In general, it is unlikely that prior knowledge of the system will allow us to know which momentum sector the ground state is in. The process highlighted in this section would therefore need to be repeated for each momentum sector. Applying the tools developed in Sec. 5 to the other momentum sectors, will yield custom circuits for each. In fact, the here considered case is the worst case scenario because all other momentum sectors need at most 8-parameter quantum circuits (Eq. 68 and removal of a global phase). The explicit computation in Sec. 5.2 indeed shows that the momentum -1 and momentum $\pm i$ sectors can be parametrized (up to a phase) with 6-parameter and 4-parameter circuits, respectively. Compared to the 30-parameter circuit that a naïve 4-qubit quantum simulation requires, this shows the power of custom circuit design based on dimensional expressivity analysis.

From a practical point of view, the three-qubit interactions in Eq. 54 are in general harder to realize than the two-qubit interactions in the original ansatz in Eq. 51. In particular, taking into account the restricted native gate set of current quantum hardware, synthesizing the three-qubit terms can lead to deeper circuits and, thus, to an increase in noise. As a result, on NISQ devices there might be a trade-off between maximal expressivity and simplicity of the ansatz. For example, using only single-qubit rotation layers and all two-qubit rotation layers is sufficient to overcome the above-mentioned failure of an initial circuit being stuck in a product state optimization loop.

In this sense, maximal expressivity can be traded for a simpler set of gates, and dimensional expressivity analysis allows us to quantify this gate set deficiency.

5 Translationally invariant circuits

The circuit parameter reduction method described in Sec. 3 is of particular importance if we aim to incorporate (rather than remove) problem-specific symmetries into the quantum circuit as well. These symmetries often partition the search space into physically relevant sectors, which can be searched separately. As such, a quantum circuit that automatically satisfies the symmetries can greatly improve the performance of a quantum simulation, because it only searches the physically relevant subspace of the state space. However, since the measure of deficiency is the codimension of the manifold of reachable states within each physical sector, it is imperative to compute the dimension of these sectors.

In this section, we will explicitly treat the case of translational invariance. This provides an explicit example that can be used to treat other symmetries analogously. Furthermore, translational invariance is a common property in local Hamiltonians. In this case, the physical sectors are the intersections of the unit sphere of the Hilbert space with the eigenspaces of the translation operator. In Sec. 5.1, we will construct the physical sectors and compute their dimension. As such, we will prove the following Theorem 5.1. In particular, the real dimension of the physical sectors can then be compared to the circuit manifold dimension as obtained in Sec. 3, in order to compute the codimension deficiency.

Theorem 5.1. *Let Q be the number of qubits, ω be a root of unity of order d, i.e., $\omega^d = 1$ and $\forall n \in \mathbb{N}_{<d} : \omega^n \neq 1$. If $d \mid Q$ ("d divides Q"), then ω is an eigenvalue of the translation operator.*

Furthermore, let $\#(d) := 2^d - \sum_{\substack{d'\mid d \\ d'<d}} \#(d')$ with $\#(1) = 2$ and $[\#](d) = \frac{\#(d)}{d}$. Then, the complex dimension $\dim_{\mathbb{C}} \mathrm{Eig}(\omega)$ of the eigenspace $\mathrm{Eig}(\omega)$ is given by

$$\dim_{\mathbb{C}} \mathrm{Eig}(\omega) = \sum_{d\mid k\mid Q} [\#](k). \quad (55)$$

The real dimension of the physical sector corresponding to ω is therefore $2\dim_{\mathbb{C}} \mathrm{Eig}(\omega) - 1$.

Furthermore, we will show in App. A that the physical sector corresponding to each root of unity ω is path connected. Thus, given ω, any two states in the physical sector corresponding to ω, and any maximally expressive circuit for the physical sector, there exists a path in parameter space connecting the two states. This is particularly important for variational quantum simulations, as it implies that the optimal state can be reached by a path from arbitrary initial states. However, since it is not directly relevant for the circuit analysis, we will provide the proof and discussion of path connectedness in the appendix.

To illustrate the construction and Theorem 5.1, we will consider the 4-qubit case in Sec. 5.2. We will apply these results to the analysis and design of a 4-qubit translationally invariant circuit for the momentum 1 sector in Sec. 4.

5.1 The general case

Let us consider a Q-qubit system. A basis of the Hilbert space is given by $B := \{|j\rangle ;\ j \in \mathbb{N}_{0,<2^Q}\}$, where $|j\rangle$ denotes the tensor product state $|b_Q(j)_{Q-1}\ldots b_Q(j)_0\rangle$ if and only if $b_Q(j) := b_Q(j)_{Q-1}\ldots b_Q(j)_0$ is the binary representation of j. For instance, $b_3(5) = 101$ with $b_3(5)_2 = b_3(5)_0 = 1$ and $b_3(5)_1 = 0$.

The translation operator τ_Q is defined as the linear operator mapping B to itself and satisfying

$$b_Q(\tau_Q(j)) := b_Q(j)_{Q-2}\ldots b_Q(j)_0 b_Q(j)_{Q-1} \quad (56)$$

for every $j \in \mathbb{N}_{0,<2^Q}$. The set of possible momenta is then given by the set of eigenvalues of τ_Q and the momentum sectors are the eigenspaces of τ_Q. In order to compute the dimension of each momentum sector, we define the following equivalence relation on $\mathbb{N}_{0,<2^Q}$:

$$j \sim_Q k \ :\Longleftrightarrow\ \exists t \in \mathbb{N} :\ j = \tau_Q^t(k). \quad (57)$$

For example, we observe

- $1 \sim_2 2$ since $01 = \tau_2(10)$ in binary,

- $5 \sim_3 3$ since $101 = \tau_3^2(011)$ in binary,

- $\forall j \in \mathbb{N}_{0,<2^Q} :\ j \sim_Q j$ since $j = \tau_Q^Q(j)$.

We denote the equivalence class of j in $\mathbb{N}_{0,<2^Q}/\sim_Q$ by $[j]_Q$. Then we define the order of j to be

$$\operatorname{ord}_Q(j) := \min\{t \in \mathbb{N};\ \tau_Q^t(j) = j\}. \qquad (58)$$

By Lagrange's theorem, we know

$$\forall j \in \mathbb{N}_{0,<2^Q} :\ \operatorname{ord}_Q(j) \text{ divides } Q, \qquad (59)$$

e.g., we observe $\operatorname{ord}_4(8) = 2$ and $[8]_4 = \{8,5\}$ since $b_4(8) = 1010$ and $b_4(5) = 0101$.

Going forward, it will be advantageous to choose a canonical ordering of the elements in $[j]_Q$. For simplicity, we will order them by value, that is

$$[j]_Q = \{j_0, \ldots, j_{\operatorname{ord}_Q(j)-1}\} \qquad (60)$$

where $m < n$ implies $j_m < j_n$, i.e., $[8]_4 = \{8_0, 8_1\}$ with $8_0 = 5$ and $8_1 = 8$. We can use this ordering to define

$$\omega_Q(j_n) := \exp\left(2\pi i \frac{n}{\operatorname{ord}_Q(j_n)}\right) \qquad (61)$$

which is a bijection between the elements of $[j]_Q$ and the roots of unity of order $\operatorname{ord}_Q(j)$. This implies that we can construct an eigenvector $e_Q(j_n)$ of τ_Q for each element of each $[j]_Q$ by setting

$$\tilde{e}_Q(j_n) := \sum_{k=0}^{\operatorname{ord}_Q(j)-1} \omega_Q(j_n)^k \left|\tau_Q^k(j_0)\right\rangle \qquad (62)$$

and

$$e_Q(j_n) := \frac{\tilde{e}_Q(j_n)}{\|\tilde{e}_Q(j_n)\|_{\ell_2(Q)}}. \qquad (63)$$

For example, $[1]_2 = \{1,2\}$ implies $e_2(1) = \frac{|01\rangle+|10\rangle}{\sqrt{2}}$ and $e_2(2) = \frac{|01\rangle-|10\rangle}{\sqrt{2}}$.

By construction, the $e_Q(j_n)$ are independent of all $e_Q(k_m)$ if $[j]_Q \neq [k]_Q$ because they are linear combinations of linearly independent sets. Furthermore, the $e_Q(j_n)$ are independent of the remaining $e_Q(j_m)$ since they are eigenvectors of different eigenvalues of τ_Q. In other words, we have found a basis transformation between B and a basis of eigenvectors of τ_Q. In particular, we have found a complete set of eigenvectors for τ_Q.

Additionally, we can pigeon hole the number of eigenvectors to a given eigenvalue ω. Let ω be a root of unity of order N, i.e., $\omega^N = 1$ and $\forall n \in \mathbb{N}_{<N}:\ \omega^n \neq 1$. Then, the number $\nu(\omega, j)$ of

eigenvectors corresponding to ω generated from elements in $[j]_Q$ is

$$\nu(\omega, j) = \begin{cases} 1 & ,\ N \mid \operatorname{ord}_Q(j), \\ 0 & ,\ N \nmid \operatorname{ord}_Q(j), \end{cases} \qquad (64)$$

where $a \mid b$ denotes "a divides b". Hence, the real dimension of the eigenspace $\operatorname{Eig}(\omega)$ of τ_Q with respect to the eigenvalue ω is given by

$$\dim_{\mathbb{C}} \operatorname{Eig}(\omega) = \#\{[j]_Q;\ N \mid \operatorname{ord}_Q(j)\}. \qquad (65)$$

In particular, since 1 divides every $\operatorname{ord}_Q(j)$, we directly conclude

$$\dim_{\mathbb{C}} \operatorname{Eig}(1) = \#\{[j]_Q;\ j \in \mathbb{N}_{0,<2^Q}\} \qquad (66)$$

and

$$\dim_{\mathbb{C}} \operatorname{Eig}(1) \geq \dim_{\mathbb{C}} \operatorname{Eig}(\omega) \qquad (67)$$

for every eigenvalue ω of τ_Q. In fact, since $|0\ldots0\rangle$ and $|1\ldots1\rangle$ are only in $\operatorname{Eig}(1)$ and every generating j_k in $\operatorname{Eig}(\omega)$ is also in $\operatorname{Eig}(1)$, we obtain

$$\dim_{\mathbb{C}} \operatorname{Eig}(1) \geq \dim_{\mathbb{C}} \operatorname{Eig}(\omega) + 2. \qquad (68)$$

It should be noted that Eq. 68 is optimal. If Q is a power of 2, then all $\operatorname{ord}_Q(j)$ have to be even or 1. Thus, every j with $\operatorname{ord}_Q(j) > 1$ has to give an element of $\operatorname{Eig}(1)$ and $\operatorname{Eig}(-1)$ with only $|0\ldots0\rangle$ and $|1\ldots1\rangle$ having order 1, i.e.,

$$\dim_{\mathbb{C}} \operatorname{Eig}(1) = \dim_{\mathbb{C}} \operatorname{Eig}(-1) + 2. \qquad (69)$$

To compute the dimension of each momentum sector explicitly, let $1 = d_1 < d_2 < \ldots < d_N = Q$ be the divisors of Q. Then, each possible order has to be one of the d_n. Furthermore, for each d_n there exists a $j \in \mathbb{N}_{0,<2^Q}$ with $\operatorname{ord}_Q(j) = d_n$ since, setting $r_n := \frac{Q}{d_n}$, we can define

$$j := \sum_{k=0}^{r_n-1} 2^{kd_n}$$

$$= \underbrace{\overbrace{\underbrace{0\ldots0}_{d_n-1}1 \underbrace{0\ldots0}_{d_n-1}1 \ldots \underbrace{0\ldots0}_{d_n-1}1}^{r_n \text{blocks of size } d_n}} \qquad (70)$$

to obtain an element j of order d_n. Hence, for every divisor d of Q and for every root of unity ω of order d, there exists a state of momentum ω. This also allows us to count the number of elements $j \in \mathbb{N}_{0,<2^Q}$ which have order d since their binary representation is built up of $\frac{Q}{d}$ copies of

binary strings of length d. There are 2^d such binary strings, however this is counting all j with $\mathrm{ord}_Q(j) \mid d$ as well. Hence,

$$\#(d) := \#\{j; \ \mathrm{ord}_Q(j) = d\}$$
$$= 2^d - \sum_{\substack{d' \mid d \\ d' < d}} \#(d') \tag{71}$$

gives a recursion with $\#(1) = 2$.

Since the elements of $\mathbb{N}_{0, < 2^Q}$ which are of order d are grouped into equivalence classes of d elements each, we directly obtain that the number of equivalence classes $[\#](d)$ of order d is given by

$$[\#](d) = \frac{\#(d)}{d}. \tag{72}$$

Hence, we obtain

$$\dim_{\mathbb{C}} \mathrm{Eig}(\omega) = \#\{[j]_Q; \ d \mid \mathrm{ord}_Q(j)\}$$
$$= \sum_{d \mid k \mid Q} [\#](k) \tag{73}$$

which completes the proof of Theorem 5.1.

5.2 Example: the 4-qubit case

In this section, we will consider the $Q = 4$-qubit case as a fully worked example. The construction of the eigenspaces of the 4-qubit translation operator is based on the construction of a suitable mapping between vectors in the 4-qubit Hilbert space $\mathbb{C}^{16}$ and translational equivalence classes of the computational basis of a 4-qubit quantum device. The equivalence classes $[j]_4$ with respect to

$$j \sim_4 k \ :\Longleftrightarrow \ \exists t \in \mathbb{N}: \ j = \tau_4^t(k) \tag{74}$$

are

- $[0]_4 = \{0 = 0000\}$,

- $[1]_4 = \{1 = 0001, 2 = 0010, 4 = 0100, 8 = 1000\}$,

- $[3]_4 = \{3 = 0011, 6 = 0110, 12 = 1100, 9 = 1001\}$,

- $[5]_4 = \{5 = 0101, 10 = 1010\}$,

- $[7]_4 = \{7 = 0111, 14 = 1110, 13 = 1101, 11 = 1011\}$,

- $[15]_4 = \{15 = 1111\}$,

for which we observe the orders $\mathrm{ord}_4(0) = 1$, $\mathrm{ord}_4(1) = 4$, $\mathrm{ord}_4(3) = 4$, $\mathrm{ord}_4(5) = 2$, $\mathrm{ord}_4(7) = 4$, and $\mathrm{ord}_4(15) = 1$.

Each eigenspace corresponds to a root of unity whose order divides 4 since these are precisely the eigenvalues of the translation operator on 4 qubits. Furthermore, the equivalence class $[j]_4$ contributes exactly one basis vector for the eigenspace $\mathrm{Eig}(\omega)$ corresponding to the root of unity ω if and only if the order of j is divisible by the order of ω. This implies that we can obtain the dimension of each eigenspace by counting.

- $\dim \mathrm{Eig}(1) = \#\{[0]_4, [1]_4, [3]_4, [5]_4, [7]_4, [15]_4\} = 6$ (orders divisible by 1),

- $\dim \mathrm{Eig}(-1) = \#\{[1]_4, [3]_4, [5]_4, [7]_4\} = 4$ (orders divisible by 2),

- $\dim \mathrm{Eig}(i) = \#\{[1]_4, [3]_4, [7]_4\} = 3$ (orders divisible by 4),

- $\dim \mathrm{Eig}(-i) = \#\{[1]_4, [3]_4, [7]_4\} = 3$ (orders divisible by 4).

Furthermore, the construction of Sec. 5.1 explicitly allows us to identify a basis for each eigenspace.

- $\mathrm{Eig}(1)$:

    1. $e_4(0_0) = |0000\rangle$,
    2. $e_4(1_0) = \frac{|0001\rangle + |0010\rangle + |0100\rangle + |1000\rangle}{\sqrt{4}}$,
    3. $e_4(3_0) = \frac{|0011\rangle + |0110\rangle + |1100\rangle + |1001\rangle}{\sqrt{4}}$,
    4. $e_4(5_0) = \frac{|0101\rangle + |1010\rangle}{\sqrt{2}}$,
    5. $e_4(7_0) = \frac{|0111\rangle + |1110\rangle + |1101\rangle + |1011\rangle}{\sqrt{4}}$,
    6. $e_4(15_0) = |1111\rangle$.

- $\mathrm{Eig}(-1)$:

    1. $e_4(1_1) = \frac{|0001\rangle - |0010\rangle + |0100\rangle - |1000\rangle}{\sqrt{4}}$,
    2. $e_4(3_1) = \frac{|0011\rangle - |0110\rangle + |1100\rangle - |1001\rangle}{\sqrt{4}}$,
    3. $e_4(5_1) = \frac{|0101\rangle - |1010\rangle}{\sqrt{2}}$,
    4. $e_4(7_1) = \frac{|0111\rangle - |1110\rangle + |1101\rangle - |1011\rangle}{\sqrt{4}}$.

- $\mathrm{Eig}(i)$:

    1. $e_4(1_2) = \frac{|0001\rangle + i|0010\rangle - |0100\rangle - i|1000\rangle}{\sqrt{4}}$,
    2. $e_4(3_2) = \frac{|0011\rangle + i|0110\rangle - |1100\rangle - i|1001\rangle}{\sqrt{4}}$,

3. $e_4(7_2) = \frac{|0111\rangle + i|1110\rangle - |1101\rangle - i|1011\rangle}{\sqrt{4}}$.

- Eig($-i$):

  1. $e_4(1_3) = \frac{|0001\rangle - i|0010\rangle - |0100\rangle + i|1000\rangle}{\sqrt{4}}$,

  2. $e_4(3_3) = \frac{|0011\rangle - i|0110\rangle - |1100\rangle + i|1001\rangle}{\sqrt{4}}$,

  3. $e_4(7_3) = \frac{|0111\rangle - i|1110\rangle - |1101\rangle + i|1011\rangle}{\sqrt{4}}$.

Finally, we can compare these explicit results to the abstract results of Sec. 5.1. Using Eq. 71 and Eq. 72, we obtain

- $\#(1) = 2$ (0 and 15),

- $\#(2) = 2^2 - 2 = 2$ (5 and 10),

- $\#(4) = 2^4 - 2 - 2 = 12$ (everything else).

as well as

- $[\#](1) = 2$ (corresponding to $[0]_4$ and $[15]_4$),

- $[\#](2) = \frac{\#(2)}{2} = 1$ (corresponding to $[5]_4$),

- $[\#](4) = \frac{\#(4)}{4} = 3$ (corresponding to $[1]_4$, $[3]_4$, and $[7]_4$).

and thus with Theorem 5.1

- $\dim_{\mathbb{C}} \mathrm{Eig}(1) = [\#](1) + [\#](2) + [\#](4) = 2 + 1 + 3 = 6$,

- $\dim_{\mathbb{C}} \mathrm{Eig}(-1) = [\#](2) + [\#](4) = 1 + 3 = 4$,

- $\dim_{\mathbb{C}} \mathrm{Eig}(i) = [\#](4) = 3$,

- $\dim_{\mathbb{C}} \mathrm{Eig}(-i) = [\#](4) = 3$.

As such, we can see that the eigenspaces of the momentum operator have the correct dimensions (summing to $\dim_{\mathbb{C}} \mathbb{C}^{2^4} = 16$). Furthermore, the inductive procedure of Theorem 5.1 enables us to quickly identify the required number of parameters for a maximally expressive quantum circuit for each of these momentum sectors. As a corollary, we also obtain structural information about the momentum sectors, which can be used in custom designing quantum circuits similar to the example provided in Sec. 4.

6 Hardware efficient implementation

In this section, we will discuss possible hardware efficient implementation and automation for the proposed dimensional expressivity analysis. Since the analysis requires testing the Jacobian of the quantum circuit for its rank, a classical implementation requires exponential memory in the number of qubits Q (the Jacobian is a matrix of dimension (#parameters)$\times 2^{Q+1}$). We are therefore looking for either a pure quantum algorithm or a hybrid quantum-classical algorithm that can reduce the classical part of the algorithm to scale polynomially in the number of parameters. Since the dimensional expressivity analysis is performed for each parameter at a time, it requires at least polynomial cost in the number of parameters. As such, "efficiency" is also related to the scaling factor between the number of qubits Q and the dimension of the physical state space for a given number of qubits.[7]

In Sec. 6.1, we will derive a hybrid quantum classical algorithm that efficiently performs the dimensional expressivity analysis at a given set of parameters. As such, it can be used for circuit design if tested on sufficiently many parameter sets, as well as for on the fly circuit reduction, e.g., to only update relevant parameters in a variational quantum simulation. In Sec. 6.2, we will briefly discuss the termination condition for the automated dimensional expressivity analysis as well as error estimates. This ensures that even if a generic scheme producing quantum circuits with arbitrarily many parameters is used, the automated procedure will reject any additional parameters. Finally, in Sec. 6.3, we will discuss hardware implementations on IBMQ.

---

[7] For many physically interesting examples, such as local Hamiltonians, polynomial scaling of the state space dimension in Q implies polynomial scaling of the dimensional expressivity analysis in Q as well. However, if the state space grows exponentially in Q, then the dimensional expressivity analysis will also scale exponentially if the entire state space needs to be parametrized.

6.1 A hybrid quantum-classical algorithm for dimensional expressivity analysis

Given a quantum circuit C with N parameters, we recall that the real Jacobian J is given by

$$J = \begin{pmatrix} | & & | \\ \Re\partial_1 C & \cdots & \Re\partial_N C \\ | & & | \\ & & \\ | & & | \\ \Im\partial_1 C & \cdots & \Im\partial_N C \\ | & & | \end{pmatrix}. \qquad (75)$$

We will perform the dimensional expressivity analysis for each parameter at a time starting with the first. In other words, we will start checking whether the second parameter is independent of the first. Then, we test the third parameter against the first and second, and continue checking each new parameter against its predecessors. Thus, in order to check the kth parameter, we need consider the reduced Jacobian

$$\begin{aligned} J_k &= \begin{pmatrix} | & & | \\ \Re\partial_1 C & \cdots & \Re\partial_k C \\ | & & | \\ & & \\ | & & | \\ \Im\partial_1 C & \cdots & \Im\partial_k C \\ | & & | \end{pmatrix} \\ &= \begin{pmatrix} | & & | \\ | & & \Re\partial_k C \\ | & & | \\ J_{k-1} & & \\ | & & | \\ | & & \Im\partial_k C \\ | & & | \end{pmatrix}, \end{aligned} \qquad (76)$$

where we may assume that J_{k-1} has full rank (because we will have removed any prior dependent parameters). J_1 is always full rank, as long as the first parameter does not correspond to a trivial gate. We assume that checking non-triviality is at least possible for the first gate with respect to the ordering of parameters. Any trivial gates corresponding to later parameters will be found by the dimensional expressivity analysis.

Since we are unaware of a quantum algorithm that can check linear independence or the rank of a given dense real matrix, we expect to have to test the rank of each J_k classically. In order to remove exponential memory scaling, we note

that

$$S_k := J_k^* J_k \qquad (77)$$

is a real $k \times k$-matrix with the same rank as J_k. In particular, assuming J_{k-1} has full rank, S_k can have at most a rank deficiency of 1. Hence, S_k is invertible if and only if the new parameter is independent. Invertibility of S_k can be checked in various ways, e.g., by computing its determinant or its eigenvalues. In either case, since S_k is a square matrix of size k, this operation can be performed classically at a cost of $O(k^3)$. Hence, checking invertibility for all S_k with $1 \le k \le N$ has a cost of $O(N^4)$, independent of the number of qubits.

Of course, this requires efficient access to the matrices S_k. Considering the (m, n)-entry $(S_k)_{m,n}$ of S_k, we note

$$\begin{aligned} (S_k)_{m,n} &= \Re\partial_m C^* \Re\partial_n C + \Im\partial_m C^* \Im\partial_n C \\ &= \Re\langle \partial_m C, \partial_n C \rangle. \end{aligned} \qquad (78)$$

In other words, we need to efficiently estimate the $\Re\langle \partial_m C, \partial_n C \rangle$. Following [35, 36], this is generally possible using quantum RAM calls. However, for many applications in variational quantum simulations, a more direct approach is possible.

To efficiently compute the matrices S_k, we will restrict the type of parametric gates to be rotation gates of the form

$$R_G(\vartheta) := \exp\left(-\frac{i}{2}\vartheta G\right), \qquad (79)$$

where G is a gate. This class of rotation gates includes the common R_X, R_Y, and R_Z gates but may also include entangling gates such as R_{CNOT}, $R_{X_q X_{q'}}$, or even more complicated gates R_G with $G = \sum_q X_q X_{q+1} X_{q+2} Y_{q+3}$ as discussed in Sec. 4 (i.e., gates necessary to enforce physical symmetries).

Using a circuit C that comprises only gates R_G as described above and non-parametric gates, we can write

$$\Re\langle \partial_m C, \partial_n C \rangle = \frac{1}{4}\Re\langle \gamma_m, \gamma_n \rangle, \qquad (80)$$

where each γ_j coincides with C except for an additional gate G_j inserted after the application of $R_{G_j}(\vartheta_j)$. For example, if

$$C(\vartheta_1, \vartheta_2) = R_Z(\vartheta_2) R_X(\vartheta_1) |0\rangle, \qquad (81)$$

then

$$\gamma_1 = R_Z(\vartheta_2) X R_X(\vartheta_1) |0\rangle \qquad (82)$$

and

$$\gamma_2 = Z R_Z(\vartheta_2) R_X(\vartheta_1) |0\rangle . \qquad (83)$$

In particular, we note that for all such circuits C, the diagonal elements of S_k are given by

$$\Re \langle \partial_n C, \partial_n C \rangle = \frac{1}{4}. \qquad (84)$$

In order to compute $\Re \langle \gamma_m, \gamma_n \rangle$ with $m \neq n$ on a quantum device, we need to construct the state [35]

$$|\varphi_{m,n}\rangle = \frac{|0\rangle \otimes |\gamma_m\rangle + |1\rangle \otimes |\gamma_n\rangle}{\sqrt{2}}. \qquad (85)$$

Using the setup and circuit assumptions laid out above, this requires an ancillary qubit and insertion of controlled gates CG_j after R_{G_j} where CG_m is controlled by the ancilla being $|0\rangle$ and CG_n is controlled by the ancilla being $|1\rangle$. The former can be achieved using a standard "control if $|1\rangle$" gate CG_j if it is conjugated with X gates on the ancilla, i.e., inserting $X_{\mathrm{anc}} CG_j X_{\mathrm{anc}}$ after R_{G_j}. For example, the state $|\varphi_{2,1}\rangle$ for

$$C(\vartheta_1, \vartheta_2) = R_Z(\vartheta_2) R_X(\vartheta_1) |0\rangle \qquad (86)$$

is obtained using the following circuit:
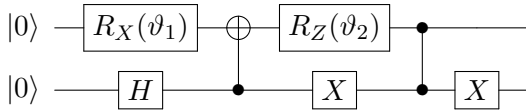
Once the state $|\varphi_{m,n}\rangle$ is prepared, we can apply a final Hadamard gate on the ancilla. This yields the state

$$
\begin{aligned}
&H_{\mathrm{anc}} |\varphi_{m,n}\rangle \\
&= \frac{|0\rangle \otimes (|\gamma_m\rangle + |\gamma_n\rangle) + |1\rangle \otimes (|\gamma_m\rangle - |\gamma_n\rangle)}{2}
\end{aligned}
\qquad (87)
$$

and measuring the ancilla yields

$$
\begin{aligned}
\mathrm{p}(\mathrm{anc} = 0) &= \frac{(\langle \gamma_m | + \langle \gamma_n |)(|\gamma_m\rangle + |\gamma_n\rangle)}{4} \\
&= \frac{1 + \Re \langle \gamma_m, \gamma_n \rangle}{2}
\end{aligned}
\qquad (88)
$$

for the ancilla being measured in $|0\rangle$.

Returning to the $C = R_Z R_X |0\rangle$ example, we thus obtain

$$\Re \langle \gamma_2, \gamma_1 \rangle = 2\mathrm{p}(\mathrm{anc} = 0) - 1, \qquad (89)$$

retrieving $\mathrm{p}(\mathrm{anc} = 0)$ from the circuit shown in Fig. 5.

In general, this means that we can compute all of the matrices S_k ($1 \leq k \leq N$) using $O(N^2 \varepsilon^{-2})$ quantum device calls. Here, ε is the acceptable error for the estimates of $\Re \langle \gamma_m, \gamma_n \rangle$; e.g., a suitable multiple of the maximal standard deviation. The complete cost of the hybrid quantum-classical algorithm is therefore in $O(N^4 \varepsilon^{-2})$. Furthermore, the quantum algorithm part only requires an additional ancilla qubit, two Hadamard gates, two X gates, and two controlled gates CG depending on the specific rotation gates R_G used in the quantum circuit.

6.2 Termination of the automated dimensional expressivity analysis and error estimates

Since the proposed hybrid quantum-classical algorithm is an inductive procedure, we would like to say a few words about the base case, the termination condition, and error estimates.

Given the assumed parametric gates for the quantum circuit, the base case is trivial. S_1 is a 1×1-matrix and all diagonal elements of each S_k are $\frac{1}{4}$. Hence, S_1 is always invertible.

To discuss the termination condition, it is advantageous to recognize that

$$S_k = \begin{pmatrix} S_{k-1} & A_k \\ A_k^* & \frac{1}{4} \end{pmatrix} \qquad (90)$$

holds with

$$A_k := J_{k-1}^* \begin{pmatrix} \Re \partial_k C \\ \Im \partial_k C \end{pmatrix} = \begin{pmatrix} \frac{1}{4} \Re \langle \gamma_1, \gamma_k \rangle \\ \vdots \\ \frac{1}{4} \Re \langle \gamma_{k-1}, \gamma_k \rangle \end{pmatrix}. \qquad (91)$$

As such, S_k is invertible if and only if $\frac{1}{4} - A_k^* S_{k-1}^{-1} A_k$ is non-zero (using the Schur complement and the assumption that S_{k-1} is invertible). Noting that J_{k-1} has independent columns, we conclude that the Moore-Penrose pseudo-inverse[8] J_{k-1}^+ is a left-inverse of J_{k-1} and satisfies

$$J_{k-1}^+ = S_{k-1}^{-1} J_{k-1}^*. \qquad (92)$$

---

[8] The Moore-Penrose pseudo-inverse of a matrix A is defined as a matrix A^+ that satisfies $AA^+A = A$, $A^+AA^+ = A^+$, $(AA^+)^* = AA^+$, and $(A^+A)^* = A^+A$.
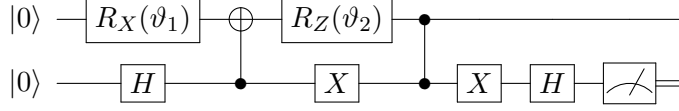
Figure 5: Circuit to compute $\Re\langle\partial_2 C, \partial_1 C\rangle$ using an ancilla qubit (lower quantum wire) for the circuit in Eq. 86.

Furthermore, we obtain

$$A_k^* S_{k-1}^{-1} A_k = \begin{pmatrix} \Re\partial_k C \\ \Im\partial_k C \end{pmatrix}^* J_{k-1} J_{k-1}^+ \begin{pmatrix} \Re\partial_k C \\ \Im\partial_k C \end{pmatrix}. \tag{93}$$

Since J_{k-1}^+ is only a left-inverse (unless in the case which will become the termination condition), the central $J_{k-1} J_{k-1}^+$ will not reduce to the identity. However, if J_{k-1}^+ is a right-inverse as well, then we observe

$$\begin{aligned} \frac{1}{4} - A_k^* S_{k-1}^{-1} A_k &= \frac{1}{4} - \begin{pmatrix} \Re\partial_k C \\ \Im\partial_k C \end{pmatrix}^* \begin{pmatrix} \Re\partial_k C \\ \Im\partial_k C \end{pmatrix} \\ &= \frac{1}{4} - \frac{1}{4}\langle\gamma_k, \gamma_k\rangle \\ &= 0. \end{aligned} \tag{94}$$

This termination condition should hold if the number of independent parameters coincides with the dimension d of the physical state space, i.e., the sub-manifold of the quantum-device state space subject to all physical symmetries that the quantum circuit satisfies. In this case, the tangent space of the physical state space is d-dimensional and we can choose a representation $\mathcal{J}_k$ of J_k such that $\mathcal{J}_k$ is a $d \times k$-matrix. In particular, $\mathcal{J}_d$ is a $d \times d$-matrix with independent columns. In other words, $\mathcal{J}_d$ is invertible and we conclude $\mathcal{J}_d^+ = \mathcal{J}_d^{-1}$. Eq. 94 thus implies by induction that all following parameters will be rejected as dependent. A summary of the procedure in pseudo code is provided in Algorithm 1, where we generate both the independent and redundant parameters for pedagogical reasons.

With respect to error estimates, we assume that all $\frac{1}{4}\Re\langle\gamma_m, \gamma_n\rangle$ are known up to an error of ε (e.g., the corresponding multiple of the standard deviation for a given confidence level). The matrices S_k are therefore estimated by matrices $S_k + \delta_k$, where each matrix element of δ_k is less than ε in absolute value (at least up to the chosen confidence level). If we choose to test S_k for invertibility by computing its eigenvalues, then

---

Algorithm 1: Hybrid quantum-classical expressivity analysis

Input: Parametric circuit, $\vartheta_1, \ldots, \vartheta_N$
Output: Independent and redundant parameters

$I \leftarrow \{1\}$
$R \leftarrow \emptyset$
$S \leftarrow \frac{1}{4}$
for $k \leftarrow 2$ **to** N **do**
  for $l \leftarrow 1$ **to** $\#I$ **do**
    $l' \leftarrow I[l]$
    Measure $\frac{1}{4}\Re\langle\gamma'_l, \gamma_k\rangle$ on the quantum device
    $A_k[l] \leftarrow \frac{1}{4}\Re\langle\gamma'_l, \gamma_k\rangle$
  end
  $\tilde{S} \leftarrow \begin{pmatrix} S & A_k \\ A_k^* & \frac{1}{4} \end{pmatrix}$
  if $\tilde{S}$ *is invertible* **then**
    $I \leftarrow I \cup \{k\}$
    $S \leftarrow \tilde{S}$
  else
    $R \leftarrow R \cup \{k\}$
  end
end
return
  independent parameters: $\{\vartheta_i; \ i \in I\}$,
  redundant parameters: $\{\vartheta_r; \ r \in R\}$

---

we can use the eigenvalue stability theorem

$$|\text{error}(\lambda)| \leq \|\delta_k\|_F \leq k\varepsilon \leq N\varepsilon \tag{95}$$

for all eigenvalues λ of S_k, where

$$\|T\|_F := \sqrt{\sum_{i,j} |T_{i,j}|^2} \tag{96}$$

denotes the Frobenius norm[9] of the matrix T. Hence, if the smallest eigenvalues $\hat{\lambda}_{\min}$ of $S_k + \delta_k$ is compatible with zero, i.e., $\hat{\lambda}_{\min} \leq k\varepsilon$ (note

---

[9]The Frobenius norm is also known as the Hilbert-Schmidt norm and sometimes called Euclidean norm.

that $S_k = J_k^* J_k$ is non-negative), then we would reject the currently tested parameter as being dependent on the previous set of parameters.

With this estimate $\hat{\lambda}_{\min} \leq k\varepsilon$, we could in principle falsely identify a parameter as dependent when it is indeed independent. However, as increasing the number of measurements for the evaluation of $\frac{1}{4}\Re\langle\gamma_m, \gamma_n\rangle$ reduces ε, we can identify independent parameters once sufficiently many quantum device calls have been used to ensure that $\hat{\lambda}_{\min} > k\varepsilon$. Of course, this is still subject to the chosen confidence level but it shows that the probability of falsely identifying a dependent parameter as independent can be made arbitrarily small. Conversely, a dependent parameter should always have a minimal eigenvalue estimate $\hat{\lambda}_{\min} \leq k\varepsilon$.

Choosing other tests for invertibility, such as $\frac{1}{4} - A_k^* S_{k-1}^{-1} A_k > 0,$[10] also allow for error estimates of the form $ck\varepsilon$ where c is an ε-dependent constant larger than $c_0 := 2\left\|S_{k-1}^{-1}\right\| \|A_k\|_2 + \left\|S_{k-1}^{-1}\right\|^2 \|A_k\|_2^2$ and c_0 holds as a bound in the $\varepsilon \to 0$ limit. Choosing $\det(S_k) > 0$ as a test is not recommended since $0 \leq \det(S_k) \leq 4^{-k}$ holds. This follows by induction using $\det(S_1) = \frac{1}{4}$, $\det(S_k) = \det(S_{k-1}) \det\left(\frac{1}{4} - A_k^* S_{k-1}^{-1} A_k\right)$, and $0 \leq \frac{1}{4} - A_k^* S_{k-1}^{-1} A_k \leq \frac{1}{4}$. In any case, we see that there are numerous ways to identify independent parameters.

Since we compute S_k using a quantum device, any test for parameter dependence/independence needs to be resistant to perturbations introduced by the quantum part of the algorithm. The parameter *independence* test satisfies this perturbation resistance because $\hat{\lambda}_{\min} > k\varepsilon$ will be satisfied if we can sufficiently reduce the noise level ε. The parameter *dependence* test, on the other hand, needs to test $\det(S_k) = 0$ or, equivalently, test that the final column vector in J_k is contained in the subspace spanned by the first $k-1$ columns of J_k. Both of these tests ask whether some mathematical object is an element of a zero-measure set. Once the perturbations from the quantum

---

[10]Note that $J_{k-1}J_{k-1}^+$ satisfies $(J_{k-1}J_{k-1}^+)^* = J_{k-1}J_{k-1}^+$ and $(J_{k-1}J_{k-1}^+)^2 = J_{k-1}J_{k-1}^+ J_{k-1}J_{k-1}^+ = J_{k-1}J_{k-1}^+$. Hence, $J_{k-1}J_{k-1}^+$ is an orthoprojector onto the column space of J_{k-1}. In particular, we obtain $0 \leq J_{k-1}J_{k-1}^+ \leq 1$ and, setting $v = \begin{pmatrix} \Re\partial_k C \\ \Im\partial_k C \end{pmatrix}$, we conclude $0 \leq A_k^* S_{k-1}^{-1} A_k = v^* J_{k-1}J_{k-1}^+ v \leq \|v\|_2^2 = 1/4$.

device are introduced, the probability of positively identifying a dependent parameter as being dependent is therefore zero. In other words, without perturbation resistant sufficient conditions for parameter dependence, we need to introduce threshold conditions and either accept that we may erroneously consider a dependent parameter to be independent or vice versa.

Erring on the side of caution, in line with any chosen confidence level (interpreting ε as a suitable multiple of the standard deviation of the quantum algorithm part), we therefore might exclude an independent parameter by erroneously labeling it as dependent. This has the advantage that we do not consider this parameter to increase expressivity, i.e., if it does increase expressivity then the circuit is indeed more expressive than we have concluded. The alternative would be to erroneously allow dependent parameters to be declared independent. This would lead to the problem of constructing a supposedly maximally expressive circuit, but some states cannot be generated because we falsely counted the dimension of the circuit manifold. Of course, while this choice ensures maximal expressivity, it may lead to suboptimal circuit reductions.

6.3 Implementation on IBMQ

In order to experimentally test our approach, we implement the algorithm on IBM's quantum hardware using QISKIT [33]. As described above, we can measure the entries of the vectors A_k on a quantum chip, which, starting from $S_1 = 1/4$, allows us to recursively compute the matrices S_k using Eq. 90. Subsequently, we can numerically determine the spectra of these matrices which allows us to identify superfluous parameters. For comparison, we also construct the exact S_k matrices in order to benchmark our hardware results.

As an example, we consider the single-qubit circuit

$$
\begin{aligned}
C(\vartheta_1,\vartheta_2,\vartheta_3,\vartheta_4) = \\
R_Y(\vartheta_4)R_Z(\vartheta_3)R_X(\vartheta_2)R_Z(\vartheta_1)\left|0\right\rangle.
\end{aligned} \tag{97}
$$

Inspecting Eq. 97, it is apparent that the first R_Z gate can generate arbitrary phases, while the following two rotations R_X and R_Z allow us to map the initial state $|0\rangle$ to any point on the Bloch sphere. As a result, the last R_Y gate is superfluous and we expect the matrix S_4 to have a single
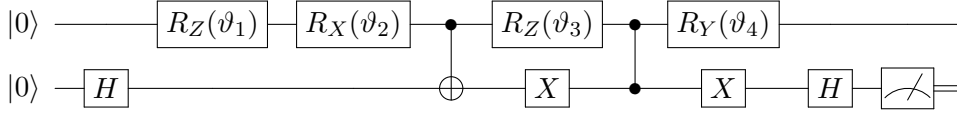
Figure 6: Circuit for obtaining $\Re\langle\gamma_2,\gamma_3\rangle$ on quantum hardware for the ansatz in Eq. 97. The ancillary qubit is initially put into superposition and acts as a control for the additional CNOT and CZ gates. After applying again a Hadamard on the ancilla, a final measurement reveals the probability for the qubit to be in state $|0\rangle$, which is related to $\Re\langle\gamma_2,\gamma_3\rangle$ according to Eq. 88.
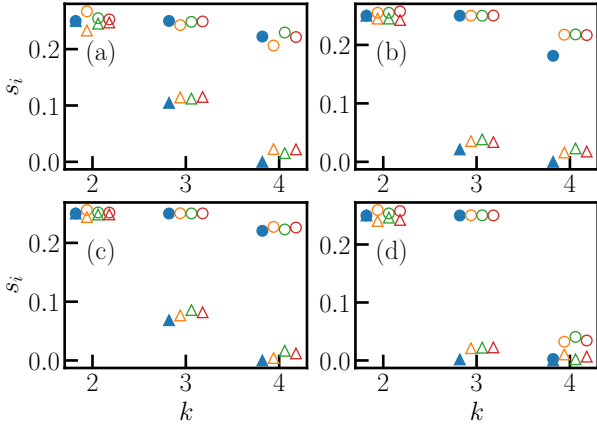

Figure 7: Smallest (triangles) and second smallest (dots) eigenvalues of the matrices S_k, $k \geq 2$, for the circuit in Eq. 97. The different panels each correspond to a different, randomly drawn parameter set $\vartheta_1,\ldots,\vartheta_4$. The filled markers indicate the exact solution, the open markers the results obtained from ibmq_vigo with 1000 measurements (orange markers), 4000 measurements (green markers) and 8000 measurements (red markers). The data points for each k are slightly horizontally shifted for better visibility.

vanishing eigenvalue, while the spectrum for all other S_k, $k \leq 3$, should consist of nonzero values. To check if this is indeed the case we choose various random parameter sets $(\vartheta_1, \vartheta_2, \vartheta_3, \vartheta_4)$, where each angle is drawn uniformly from the interval $[0, 2\pi)$, and evaluate the spectra of the S_k matrices (note that S_1 is always $1/4$ and thus does not need to be measured). To construct the matrices S_k, we measure the vector A_k on the quantum device as outlined in Sec. 6.1. An example for a circuit allowing us to obtain one of the entries of A_k for our ansatz in Eq. 97 is shown in Fig. 6.

Our data for the low-lying spectrum obtained for different parameter sets are shown in Fig. 7 and Tab. 1. Looking at the exact solution in Fig. 7 and Tab. 1, we see indeed that S_4 has a single vanishing eigenvalue for all the parameter sets we study, in agreement with the theoretical

consideration. Focusing on the cases in Fig. 7(a) - Fig. 7(c), we observe that the smallest eigenvalue of S_3 is clearly distinct from zero, thus allowing us to reliably identify the parameter ϑ_4 as superfluous. Fig. 7(d) shows the results for a parameter set which is close to a singular point and at first glance both S_3 and S_4 seem to have a zero eigenvalue. A closer examination of the numerical values in Tab. 1 reveals that for the exact solution the smallest eigenvalue of S_3 is on the order of 10^{-3}, whereas S_4 has a vanishing eigenvalue. Thus, we can still confidently identify the redundant gate in the circuit. By construction, S_k has a single vanishing eigenvalue if and only if ϑ_k is first redundant parameter in the circuit. Our data also show that, apart from the parameter set close to a singular one, there is a clear separation between the smallest and the second smallest eigenvalue for S_k, $k \geq 3$. Hence, even if one can only resolve the spectra of the S_k matrices to a finite precision (e.g. due to a finite number of measurements or noise on a real device), we expect the signal is not immediately lost.

Looking at our hardware results obtained on ibmq_vigo, we see that this is indeed the case and, in general, we observe good agreement between the data from the quantum chip and the exact solution. The data from the quantum device do not show a strong dependence on the number of measurements, and already with a modest number of 1000 measurements, we are able to determine the lowest eigenvalues of the S_k matrices with satisfactory precision. Due to the noise in the quantum device, some of the eigenvalues are slightly shifted compared to the exact result.

In particular, the spectrum of S_4 does in general not contain an eigenvalue that is exactly zero, as Tab. 1 shows. Looking at Fig. 7(b) and Fig. 7(d), we see that these shifts in the spectrum might affect the determination of the redundant parameter. In Fig. 7(d), the shifts due to noise

	(a)	(b)	(c)	(d)
s_2 exact	0.250	0.250	0.250	0.250
s_2 ibmq	0.247	0.243	0.248	0.242
s_3 exact	0.105	0.021	0.068	0.002
s_3 ibmq	0.115	0.034	0.082	0.022
s_4 exact	0	0	0	0
s_4 ibmq	0.022	0.018	0.012	0.007

Table 1: Numerical values for the smallest eigenvalues s_k of the S_k matrices for the exact case and the hardware results obtained with 8000 shots. The different columns correspond the different parameter sets shown in the panels of Fig. 7.

lead to a clear separation between the smallest eigenvalues of S_3 and S_4, although the parameter set chosen is close to a singular one. On the contrary, in Fig. 7(b), the effects of noise cause the smallest eigenvalues of S_3 and S_4 to slightly deviate from zero and become more similar (see Tab. 1).

As a result, noise in current quantum devices might lead to misidentifying the superfluous parameter in certain cases. Nevertheless, considering all our hardware results for the different parameter sets, the smallest eigenvalue of S_4 is consistently the closest one to zero. Thus, despite the noise in ibmq_vigo, we can confidently identity ϑ_4 as superfluous parameter. We also observe a similar performance for different IBM quantum devices (see App. B for details).

Although we focus on a simple single-qubit example here for illustration, our algorithm can readily be applied to more complex circuits, as for example QISKIT's EfficientSU2 circuit, which we discussed previously in Sec. 2 (see App. B for details).

7 Conclusion

In this paper, we proposed a dimensional expressivity analysis for the study and design of parametric quantum circuits. This is of particular interest if the quantum circuit is intended to obey certain physical symmetries and to avoid others. The dimensional expressivity analysis relies on the comparison between the tangent spaces of the physical state space and the manifold of reachable states by the quantum circuit. Fundamentally, a quantum circuit is maximally expressive if and only if the tangent spaces coincide, as then the circuit manifold coincides with the physical state space. Furthermore, identification of a basis in the tangent space allows for the removal of redundant parameters in the quantum circuit. This removal of redundant parameters can also be extended to remove unwanted symmetries from the states generated by the quantum circuit.

We have applied the dimensional expressivity analysis (developed in Sec. 3) to various quantum circuits. As a simple introductory example, we have removed all redundant parameters from QISKIT's EfficientSU2(3, reps=N) circuit in Sec. 2. The analysis therefore provided a maximally expressive 3-qubit quantum circuit that is free from redundant parameters. Additionally, we have optimized QISKIT's EfficientSU2(3, reps=N) circuit by removing its ability to generate an arbitrary phase.

Moreover, we have developed a hybrid-classical algorithm, which allows us to efficiently perform the dimensional expressivity analysis using quantum hardware. We have implemented the algorithm on IBM's quantum hardware and demonstrated its performance for the single (Sec. 6.3) and two-qubit (App. B.2) cases. Although we have not applied any error mitigation [8–12], we can identify the superfluous parameters reasonably well in both cases we study.

Furthermore, we provided an example for dimensional expressivity analysis in the presence of physical symmetries. Using translationally invariant Hamiltonians as an illustration, we analyzed the structure of the physical state space in Sec. 5 and App. A. The specific details of this analysis serve as a guide to similar analyses for general symmetries generated from other spectral invariants.

In addition to the knowledge of maximal expressivity of a quantum circuit, or deficiency thereof, the analysis of physical sectors provides valuable information that can guide the custom design of maximally expressive quantum circuits. We have provided such an example in Sec. 4, where we analyzed a non-performing 4-qubit translationally invariant circuit, understood why the VQS failed to find the ground state, and used the information to design a maximally expressive, translationally invariant 4-qubit circuit. We then showed that this custom circuit overcame the difficulties encountered by the initial ansatz.

In this paper, we mainly used dimensional

expressivity analysis with regard to VQS. The method developed is however completely general and can be applied to any parametric quantum circuit. It allows us to (i) decide whether or not a quantum circuit is maximally expressive, (ii) remove any redundant parameters from a given quantum circuit improving circuit efficiency, (iii) obtain a measure of deficiency in the number of parameters necessary to obtain a maximally expressive quantum circuit, (iv) design custom gate sets satisfying the physical symmetries of the physical model, and (v) remove any unwanted symmetries that a candidate quantum circuit may contain.

Acknowledgments

A Path connectedness of momentum sectors

The incorporation of problem specific symmetries and thus the reduction of the full state space (unit sphere in a complex Hilbert space) to physical sectors generated by the symmetry (here the intersection of the eigenspaces of the translation operator with the unit sphere) may disconnect the physical sectors into multiple connected components. In other words, there may be multiple topological sectors in each physical sector. For example, if there is a "hidden" $\mathbb{Z}_2$ symmetry, then a given starting position may prevent the optimizer from finding the global optimum if the quantum circuit fixes parity. In App. A.1, we will show that this is not the case for the momentum sectors. In particular, we will show that each momentum sector is path connected and that there is a path connecting any two states in any given momentum sector. In fact, the proof makes no reference to the translation operator per se. As such, it holds for all non-trivial spectrally invariant sectors, that is, physical sectors that are intersections of eigenspaces (with dimension ≥ 2) of a given operator and the unit sphere in the Hilbert space. Furthermore, assuming the parameter space is path connected and the quantum circuit is surjective (with respect to the physical sector), then any state in the physical sector can be connected to an initial state generated by an initial parameter set using a path in parameter space. This will be discussed in App. A.2.

These results are not only important as they guarantee the ability of a maximally expressive quantum circuit to reach any state in the physical sector. They are particularly interesting if we suspect the quantum circuit to contain a "hidden" discrete symmetry. While a hidden continuous symmetry reduces the circuit manifold dimension, thus preventing the construction of a maximally expressive circuit, a discrete symmetry is more likely to result in the physical sector splitting into multiple connected components. If we are confronted with a situation in which a maximally expressive quantum circuit cannot generate a path connecting two arbitrary states in the physical sector, then this may be due to an unintentional discrete symmetry all states generated by the circuit obey.

Unfortunately, this does not always have to be caused by a symmetry. The set of singular points, in which there are jumps from lower-semicontinuity, can theoretically form a lower dimensional manifold which splits the physical sector into different regions. A gradient flow would not be able to pass through such a lower dimensional manifold of singular points, i.e., any gradient based solver is unlikely to cross between such regions as well (even if each step is subject to small perturbations unless the solver is trapped on the manifold of singular points).

In either case, failure of path connectedness of a maximally expressive quantum circuit mapping into a spectrally invariant physical sector in-

dicates a fundamental necessity for an in-depth analysis of the quantum circuit in relation to the physical model.

A.1 Proof of path connectedness

To prove the path connectedness claims, let T be an operator, λ an eigenvalue of T, $|\varphi\rangle$ and $|\psi\rangle$ be two normalized eigenvectors of T with respect to λ. First, we need to construct a path γ such that $\gamma(0) = |\varphi\rangle$ and $\gamma(1) = |\psi\rangle$ to show that the physical sector is path connected. Since the dimension of the eigenspace is ≥ 2, we can assume without loss of generality that $|\varphi\rangle$ and $|\psi\rangle$ are linearly independent (otherwise, we can find a linearly independent $|\chi\rangle$ and connect $|\chi\rangle$ to both $|\varphi\rangle$ and $|\psi\rangle$). To construct a path from $|\varphi\rangle$ to $|\psi\rangle$, we need to orthonormalize, that is, we need to find $|\xi\rangle = \alpha |\varphi\rangle + \beta |\psi\rangle$ with

$$
\begin{aligned}
1 &= \langle \xi | \xi \rangle \\
&= |\alpha|^2 + \alpha^* \beta \langle \varphi | \psi \rangle + \alpha \beta^* \langle \psi | \varphi \rangle + |\beta|^*
\end{aligned}
\tag{98}
$$

and

$$
0 = \langle \varphi | \xi \rangle = \alpha + \beta \langle \varphi, \psi \rangle. \tag{99}
$$

Inserting $\alpha = -\beta \langle \varphi, \psi \rangle$ into the first equation yields

$$
1 = |\beta|^2 \left(1 - |\langle \varphi, \psi \rangle|^2 \right). \tag{100}
$$

Setting $\alpha := -\frac{\langle \varphi, \psi \rangle}{\sqrt{1 - |\langle \varphi, \psi \rangle|^2}}$ and $\beta := \frac{1}{\sqrt{1 - |\langle \varphi, \psi \rangle|^2}}$ provides a suitable choice of $|\xi\rangle$. The function

$$
\begin{aligned}
\gamma_0(t) &:= \cos(t) |\varphi\rangle + \sin(t) |\xi\rangle \\
&= \left(\cos(t) - \frac{\langle \varphi, \psi \rangle \sin(t)}{\sqrt{1 - |\langle \varphi, \psi \rangle|^2}} \right) |\varphi\rangle \\
&\quad + \frac{\sin(t)}{\sqrt{1 - |\langle \varphi, \psi \rangle|^2}} |\psi\rangle
\end{aligned}
\tag{101}
$$

thus satisfies $\gamma_0(0) = |\varphi\rangle$ and $\gamma_0(t_\psi) = |\psi\rangle$ for $t_\psi := \arcsin \left(\sqrt{1 - |\langle \varphi, \psi \rangle|^2} \right)$. We therefore observe that $\gamma(t) := \gamma_0(t_\psi t)$ is a path connecting $|\varphi\rangle$ and $|\psi\rangle$, every $\gamma(t)$ has norm 1, and every $\gamma(t)$ is an eigenvector of T with respect to λ. This proves that the physical sectors are path connected.

A.2 Paths in parameter space

Let us now assume that the quantum circuit C maps the parameter space into the physical sector surjectively and that the parameter space is path connected. Given a state $|\psi\rangle$, surjectivity of C ensures the existence of a parameter set ϑ such that $|\psi\rangle = C(\vartheta)$. Fixing an initial parameter set ϑ_0, we can find a path γ_0 in the parameter space connecting ϑ_0 to ϑ. By continuity of C, we therefore obtain a path $\gamma := C \circ \gamma_0$ in the physical sector connecting the initial state $C(\vartheta_0)$ to $|\psi\rangle$.

Although this guarantees the existence of a path in parameter space and therefore the theoretical ability for the circuit C to find the optimum using a local optimizer, if we try to construct this path using information obtained in the physical sector (as we do in variational quantum simulations), we can see the problem of singular points. To illustrate this observation, we will first describe the construction of such a path if there are no singular points on the path.

Starting with a path γ in the physical sector, we would like to find a path γ_0 in parameter space such that $\gamma = C \circ \gamma_0$; i.e., we need to invert C. This is possible if we restrict the circuit to its minimal independent set of parameters at each point in parameter space. If there are no singular points on the path, then the perturbation estimates of Sec. 3.2 provide a radius of validity for each restriction at points on γ. By compactness of $[0, 1]$, the image of γ can be split into finitely many pieces. On each of these pieces, we can use one restriction $\check{C}$ of C to obtain $\gamma_0 = \check{C}^{-1} \circ \gamma$.

Some care has to be taken at the point switching from one restriction to another because the parameters held constant may not coincide. But since the parameters held constant are dependent on the independent parameters, they can be varied to match up without changing the state at which the switch is performed. Once the parameters held constant coincide, the remaining parameters can be made to match up. This inserts a "detour" into the path γ but since we assume that C does not map outside the physical sector, we still have a path connecting the two initial states. Ensuring that the first restriction corresponds to the initial parameter set and the final restriction to the final parameter set, we have found a path in parameter space from ϑ_0 to ϑ that connects $|\varphi\rangle$ to $|\psi\rangle$.

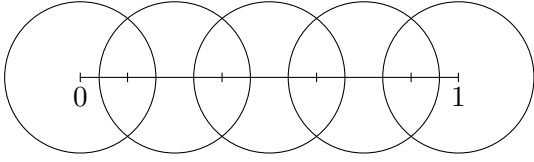The construction is therefore intimately relies

Figure 8: Visualization of a path in physical space with radii of validity for restricted circuits. No singular points are present. Thus there are finitely many overlapping areas of validity (circles). Points to swap to the next restriction are in the interior of the intersection of two areas of validity.
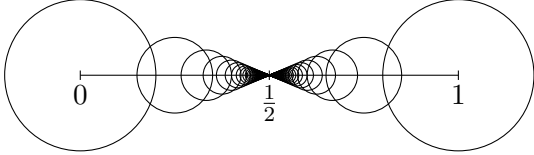

Figure 9: Visualization of a path in physical space with radii of validity for restricted circuits for construction. The point $\frac{1}{2}$ is singular. The radii of validity therefore converge to zero on either side. This separates the interval into two pieces without overlapping areas of validity (circles).

on overlapping radii of validity for the restricted circuits in Fig. 8. At a singular point, these radii converge to zero. The situation for a singular point at $\frac{1}{2}$ is visualized in Fig. 9 instead. In the case without singular points, the points at which restricted circuits are changed lie within the area of validity of both restricted circuits. This allows for local information to be utilized. At the singular point, however, only pointwise information is available if the path is not contained in the lower dimensional manifold generated by the restricted circuit at the singular point.

Luckily, having a singular point on the path is not always a calamity. Although a gradient based optimizer will be particularly vulnerable, it is not guaranteed to fail. Due to discrete step sizes, there is a chance that an update will step from one side of the singularity to the other. If this is the case, it is unlikely to be any more noticeable than a slowing down of the optimizer near the singularity. However, if we consider a gradient descent, in which the update step is a scaled multiple of the gradient, a scaling too small would ensure that this stepping over the singularity will not happen. A possible result of this would be that the gradient descent follows the manifold of singular points, slightly off-set on one side, never crossing to the singularities.

Similarly, if the path is orthogonal to the tangent space of the restricted circuit at the singular point, then too small a gradient scaling would result in the gradient descent approximating the singular point and terminating because the norm of the gradient is below the termination condition. A gradient based optimizer would therefore terminate "successfully" in this case and proclaim the singular point to be the optimum. Having no further insight into the expressivity of the quantum circuit, this situation would be indistinguishable from the optimizer terminating with a false solution due to a local optimum.

Although counter-intuitive, this suggests that increasing the scaling for the parameter update of a local optimizer near an apparent solution can be useful. If it is truly a local optimum and the scaling "overshoots" the update, then we expect to see updates bouncing around the local optimum. If, on the other hand, the slowing down is due to a singular point nearby, then "overshooting" means the artificial optimum is passed and the optimizer stops behaving like it would close to terminating.

The radius of validity estimate $R \geq \min\{R_0, \delta^{-1} |\det(C_0(\vartheta_0))|\}$ derived in Sec. 3.2 can also be used to identify singular points nearby. Approaching a singular point implies convergence of R to 0. Being in a neighborhood of a non-singular local optimum should not show $R \to 0$. In fact, R can be estimated using $\vartheta \mapsto |\det(C_0(\vartheta))|$ or $\vartheta \mapsto \frac{|\det(C_0(\vartheta))|}{\|(\det \circ C_0)'(\vartheta)\|_{\ell_2}}$. Both of these expressions converge to 0 in a neighborhood of a singular point as well ($|\det(C_0(\vartheta))| = 0$ is characteristic for loss of independent parameters due to lower-semicontinuity, and $\frac{|\det(C_0(\vartheta))|}{\|(\det \circ C_0)'(\vartheta)\|_{\ell_2}}$ is a first order approximation of R at ϑ).

Hence, while path connectivity of spectrally invariant sectors is a very useful property ensuring that a VQS with local update routine is theoretically capable of resolving the solution state, provided the quantum circuit is maximally expressive, constructing this path from an initial parameter set to a parameter set corresponding to the solution can be more difficult. In particular, singular points in the maximal expressivity of a quantum circuit can become problematic if this behavior is not monitored.

To monitor proximity to singular points and thereby induced failure of the VQS, we can es-

timate the radius of validity for the current reduced quantum circuit. In general, if we are have a quantum circuit whose reductions have singular points, we are therefore looking for multiple reduced circuits $\check{C}_j$. Each of the $\check{C}_j$ should have an open and path connected domain P_j and each point ϑ in parameter space should be contained in at least one of the P_j. Since the radius of validity at each point ϑ in P_j is bounded from below by the distance $\mathrm{dist}(\vartheta, \partial P_j)$ of ϑ to the boundary of P_j and this distance is a lower bound for the distance to any singular points of $\check{C}_j$, practical methods of avoiding singular points exist if this path connectivity monitoring is included in the VQS. For example, if only finitely many $\check{C}_j$ are used, then one could always choose the circuit $\check{C}_j$ for which the current point ϑ is contained in P_j and $\mathrm{dist}(\vartheta, \partial P_j)$ is maximal.

In this way, monitoring the path of the VQS as well as distances to singular points can circumvent many problems arising from circuit induced local minima.

B Details of the implementation of the hardware efficient algorithm

In the main text, we provided a proof of principle demonstration for a single-qubit case using ibmq_vigo. Here we give further information about the simulation parameters. In addition, we also assess the performance of various IBM quantum hardware in greater detail and also apply our hardware efficient algorithm to QISKIT's `EfficientSU2` circuit for 2 qubits.

In general, the topology of the IBM devices does not offer all-to-all connectivity of the qubits. The topology of the devices we use is given by a linear chain of 3 qubits of which the middle qubit is connected to another linear chain of 2 qubits, except for ibmq_santiago which is a linear chain of 5 qubits [37]. Thus, in a first step we use the transpiler of QISKIT to translate the circuits into a format compatible with the topology of the device. For all the results in the main text and this section we use the highest optimization level to achieve an optimal layout.

B.1 Single-qubit example, various quantum hardware

In Sec. 6.3 in the main text, we focused on our results for ibmq_vigo. To assess the performance of different quantum hardware, we repeat the single-qubit experiment using the circuit in Eq. 97 for randomly drawn parameter sets on different chips. Our results are shown in Fig. 10 and Tab. 2.

Comparing the results obtained from the different hardware backends, we observe only minor
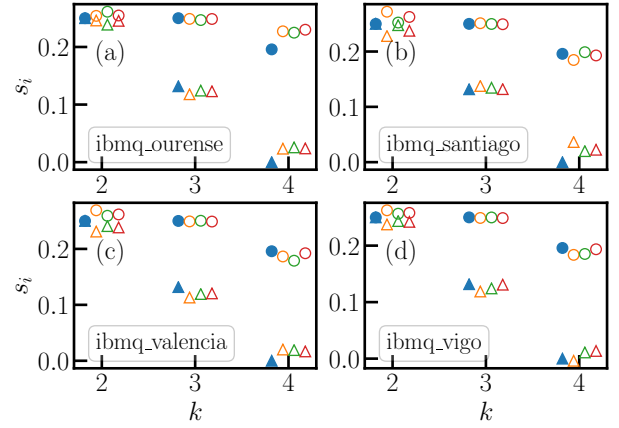

Figure 10: Smallest (triangles) and second smallest (dots) eigenvalues of the matrices S_k, $k \geq 2$, for the circuit in Eq. 97. The different panels correspond to a randomly drawn parameter set evaluated on the different hardware backends ibmq_ourense (a), ibmq_santiago (b), ibmq_valencia (c), and ibmq_vigo (d). The filled markers indicate the exact solution, the open markers the results obtained from the quantum hardware with 1000 measurements (orange markers), 4000 measurements (green markers) and 8000 measurements (red markers). The data points for each k are slightly horizontally shifted for better visibility.

	(a)	(b)	(c)	(d)
s_2 exact	0.250	0.250	0.250	0.250
s_2 ibmq	0.245	0.238	0.238	0.242
s_3 exact	0.132	0.132	0.132	0.132
s_3 ibmq	0.123	0.132	0.120	0.131
s_4 exact	0	0	0	0
s_4 ibmq	0.024	0.022	0.017	0.014

Table 2: Numerical values for the smallest eigenvalues s_k of the S_k matrices for the exact case and the hardware results obtained with 8000 shots. The different columns correspond the different parameter sets shown in the panels of Fig. 10.

differences. In general, the data from ibmq_ourense (Fig. 10(a)), imbq_santiago (Fig. 10(b)), ibmq_valencia (Fig. 10(c)) and ibmq_vigo (Fig. 10(d)) are in good agreement with the exact solution and, similar to our observations in the main text, do not show a strong dependence on the number of measurements. The shifts in the spectra of the S_k matrices due to noise are also comparable for all the hardware backends we examine (see Tab. 2). While certain eigenvalues are slightly closer to the exact solution on some devices, none of them is clearly performing better than the others and we can confidently identify the parameter ϑ_4 as the redundant one (see Tab. 2).

B.2 EfficientSU2 for two qubits

Our hardware efficient algorithm can also be readily applied to more complicated setups such as QISKIT's `EfficientSU2` circuit. In the following we focus on the `EfficientSU2` circuit for two-qubits and one repetition, $N = 1$. This circuit has 8 parameters $\vartheta_1, \ldots, \vartheta_8$ originating from two blocks of R_Y and R_Z rotations, with a CNOT gate in between (see Sec. 2). Applying the analysis from Sec. 3.4, we find that the first 7 parameters are independent and the final $R_Z(\vartheta_8)$ is superfluous. To check for that behavior with our algorithm, we proceed analogously to the single-qubit example. We draw again various random parameter sets, measure the entries of the A_k vectors for those on quantum hardware and construct the S_k matrices from these data. An example of the circuit used to obtain the entries of the vector A_k is given in Fig. 11. Afterward we examine the spectra of the S_k matrices using a classical computer.

Our results obtained on ibmq_vigo for low-lying spectrum for these matrices are shown in Fig. 12 and Tab. 3. Focusing first on the exact results, we see that in this case it is harder to unambiguously identify the superfluous parameter in the circuit. While S_8 has as expected a vanishing eigenvalue for all the parameter combinations we study, Tab. 3 shows that the matrix S_7 has in general a very small, albeit non-vanishing eigenvalue of the order of $10^{-4} - 10^{-5}$. While we can easily resolve such a small difference for the exact results, the effects of noise on a quantum device can possibly obscure this small difference, thus making it hard to unambiguously detect the

	(a)	(b)	(c)	(d)
$\frac{s_5}{10^{-4}}$ exact	1267	1212	1955	1405
$\frac{s_5}{10^{-4}}$ ibmq	1525	1282	1544	1021
$\frac{s_6}{10^{-4}}$ exact	1.632	371.5	1496	120.9
$\frac{s_6}{10^{-4}}$ ibmq	414.0	465.3	1307	366.1
$\frac{s_7}{10^{-4}}$ exact	1.078	4.064	1.522	0.4057
$\frac{s_7}{10^{-4}}$ ibmq	256.1	242.1	263.7	257.2
$\frac{s_8}{10^{-4}}$ exact	0	0	0	0
$\frac{s_8}{10^{-4}}$ ibmq	171.9	240.7	218.5	193.2

Table 3: Numerical values for the smallest eigenvalues s_k of the S_k matrices for the exact case and the hardware results obtained with 8000 shots. The different columns correspond the different parameter sets shown in the panels of Fig. 12.

	(a)	(b)	(c)	(d)
$\frac{s_5}{10^{-4}}$ exact	1955	1955	1955	1955
$\frac{s_5}{10^{-4}}$ ibmq	1949	1877	1553	1544
$\frac{s_6}{10^{-4}}$ exact	1496	1496	1496	1496
$\frac{s_6}{10^{-4}}$ ibmq	1524	1729	1381	1307
$\frac{s_7}{10^{-4}}$ exact	1.522	1.522	1.522	1.522
$\frac{s_7}{10^{-4}}$ ibmq	484.0	379.8	483.5	263.7
$\frac{s_8}{10^{-4}}$ exact	0	0	0	0
$\frac{s_8}{10^{-4}}$ ibmq	449.2	218.2	450.7	218.5

Table 4: Numerical values for the smallest eigenvalues s_k of the S_k matrices for the exact case and the hardware results obtained with 8000 shots. The different columns correspond the different parameter sets shown in the panels of Fig. 13.

redundant gate.

Looking at the results obtained from ibmq_vigo, we observe reasonable agreement with the exact solution. Compared to the single-qubit case, the dependence on the number of measurements is slightly stronger and, in general, the results deviate from the exact solution a little more (see Tab. 3). The latter is to be expected because for the two-qubit case, the circuits necessary for computing the S_k matrices are deeper and contain an additional CNOT gate compared to the single-qubit case. Consequently, our results are more hindered by noise. In particular, Fig. 12(b) and Fig. 12(c) reveal that noise can indeed obscure the identification of the superfluous parameter. In those cases the smallest eigenvalues we obtain from our hardware data for S_7 and S_8 are quite small and very similar (see Tab. 3), thus having the
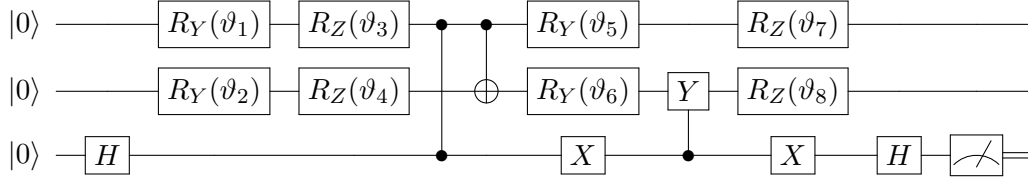
Figure 11: Circuit for measuring $\Re\langle\gamma_3, \gamma_6\rangle$ on quantum hardware. The ancillary qubit is initially put into superposition and acts as a control for the additional CZ and CY gates. A final measurement on the ancilla reveals the probability for obtaining 0 which allows for obtaining $\Re\langle\gamma_3, \gamma_6\rangle$ according to Eq. 88.
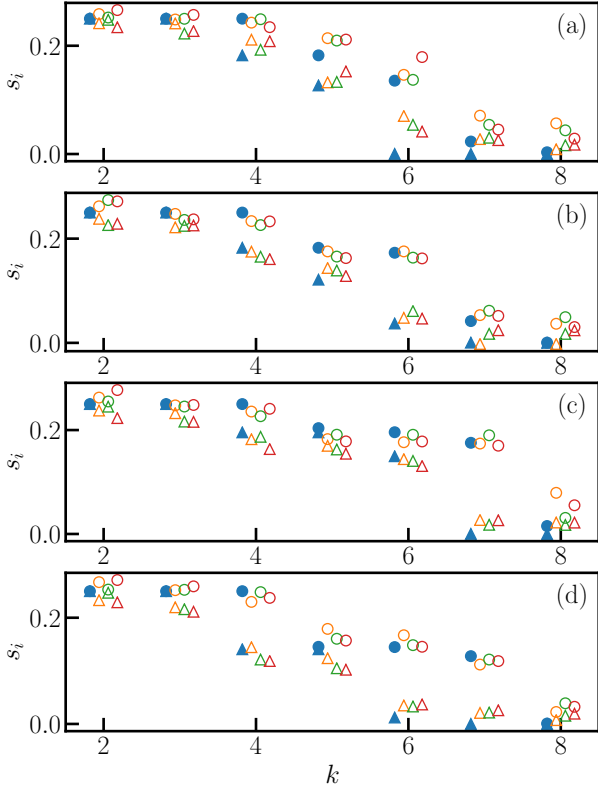

Figure 12: Smallest (triangles) and second smallest (dots) eigenvalues of the matrices S_k, $k \geq 2$, for the `EfficientSU2` circuit for two qubits and a single repetition. The different panels each correspond to a randomly drawn parameter set $\vartheta_1, \ldots, \vartheta_8$. The filled markers indicate the exact solution, the open markers the results obtained from evaluated on ibmq_vigo with 1000 measurements (orange markers), 4000 measurements (gren markers) and 8000 measurements (red markers). The data points for each k are slightly horizontally shifted for better visibility.

Figure 13: Smallest (triangles) and second smallest (dots) eigenvalues of the matrices S_k for $k \geq 2$. The different panels correspond to a randomly drawn parameter set evaluated on the different chips (a) ibmq_ourense, (b) ibmq_santiago, (c) ibmq_valencia, and (d) ibmq_vigo. The filled markers indicate the exact solution, the open markers the results obtained from the quantum hardware with 1000 measurements (orange markers), 4000 measurements (green markers) and 8000 measurements (red markers). The data points for each k are slightly horizontally shifted for better visibility.

potential of misidentifying ϑ_7 the interdependent parameter. The results for the parameter sets shown in Fig. 12(a) and Fig. 12(d) hint that ϑ_8 is the superfluous parameter, however the difference between the smallest eigenvalue of S_7 and S_8 is enhanced by noise in those cases.
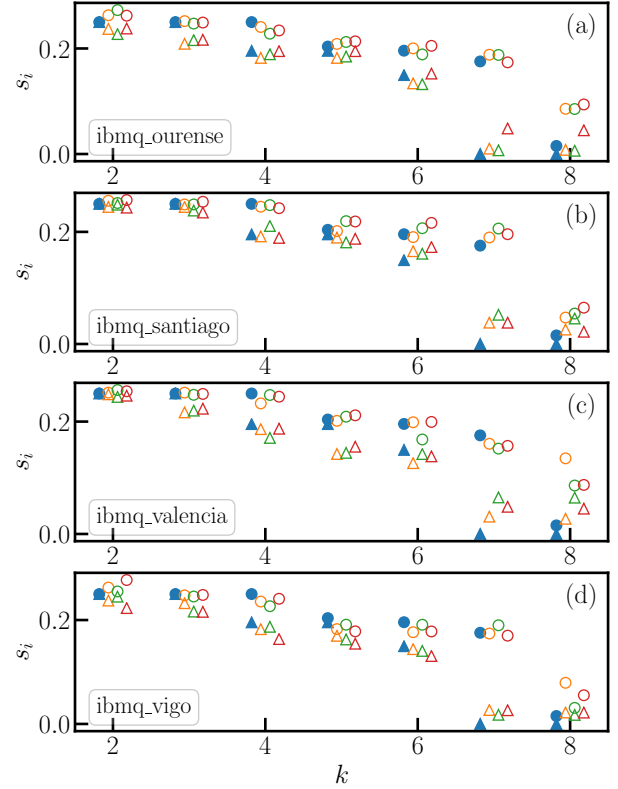
All in all, the data obtained from ibmq_vigo to a certain degree still gives an indication that ϑ_8 is interdependent, but compared to the single-qubit experiments the results are a lot less clear.

This picture also does not change using dif-

ferent hardware backends, as Fig. 13 and Tab. 4 demonstrate. Running the same randomly drawn parameter set on different quantum hardware, we observe rather similar performance. While the data from ibmq_valencia in Fig. 13(c) seems slightly worse for the parameter set we study, none of the other backends can produce the exact solution with notably better accuracy than the others. Again, our results give a hint that the rotation gate associated with ϑ_8 might be superfluous, but this is not unambiguously clear from the data, as Tab. 4 shows.

Our results indicate that with the current noise levels in NISQ devices, it might not be possible to definitely identify the independent parameters in a circuit. Nevertheless, as we approach the interdependent parameters, there is still some signal in the data that might allow to get an idea which parameter is superfluous.

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: 195, 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: 15023, 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*, 574:505, 2019.

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

[9] 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*, 549:242, 2017.

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

[11] 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*, 6:63, 2020.

[12] 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.

[13] 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.

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

[15] Y. Wang et al. Quantum simulation of helium hydride cation in a solid-state spin register. *ACS Nano*, 9:7769, 2015.

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

[17] C. Kokail et al. Self-verifying variational quantum simulation of the lattice schwinger model. *Nature*, 569:355, 2019.

[18] T. Hartung and K. Jansen. Zeta-regularized vacuum expectation values. *J. Math. Phys.*, 60:093504, 2019.

[19] K. Jansen and T. Hartung. Zeta-regularized vacuum expectation values from quantum computing simulations. *PoS (LATTICE 2019) 363*, page 153, 2020.

[20] 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.

[21] J. F. Haase, L. Dellantonio, A. Celi, D. Paul-

son, 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.

[22] M. R. Geller. Sampling and scrambling on a chain of superconducting qubits. *Physical Review Applied*, 10:024052, 2018.

[23] S. Sim, P. D. Johnson, and A. Aspuru-Guzik. Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms. *Adv. Quantum Technol.*, 2:1900070, 2019.

[24] M. Bataille. Quantum circuits of CNOT gates. *arXiv:2009.13247*, 2020.

[25] S. Sim, J. Romero, J. F. Gonthier, and A. A. Kunitsa. Adaptive pruning-based optimization of parameterized quantum circuits. *Quantum Sci. Technol.*, 6:025019, 2020.

[26] S. E. Rasmussen, N. J. S. Loft, T. Bækkegaard, M. Kues, and N. T. Zinner. Reducing the amount of single-qubit rotations in vqe and related algorithms. *Adv. Quantum Technol.*, 3:2000063, 2020.

[27] T. Hubregtsen, J. Pichlmeier, P. Stecher, and K. Bertels. Evaluation of parameterized quantum circuits: on the relation between classification accuracy, expressibility and entangling capability. *arXiv:2003.09887*, 2020.

[28] M. Schuld, R. Sweke, and J. J. Meyer. The effect of data encoding on the expressive power of variational quantum machine learning models. *arXiv:2008.08605*, 2020.

[29] E. Fontana, N. Fitzpatrick, D. Muños Ramo, R. Duncan, and I. Rungger. Evaluating the noise resilience of variational quantum algorithms. *arXiv:2011.01125*, 2020.

[30] B. T. Gard, L. Zhu, G. S. Barron, N. J. Mayhall, S. E. Economou, and E. Barnes. Efficient symmetry-preserving state preparation circuits for the variational quantum eigensolver algorithm. *npj Quantum Information*, 6:10, 2020.

[31] G. S. Barron, B. T. Gard, O. J. Altman, N. J. Mayhall, E. Barnes, and S. E. Economou. Preserving symmetries for variational quantum eigensolvers in the presence of noise. *arXiv:2003.00171*, 2020.

[32] J. Kim, J. Kim, and D. Rosa. Universal effectiveness of high-depth circuits in variational eigenproblems. *arXiv:2010.00157*, 2020.

[33] H. Abraham et al. Qiskit: An Open-source Framework for Quantum Computing. *Zenodo*, 2019.

[34] Qiskit documentation, 2020. `https://qiskit.org/documentation/_modules/qiskit/circuit/library/n_local/efficient_su2.html`. accessed on 2020/07/14.

[35] L. Zhao, Z. Zhao, P. Rebentrost, and J. Fitzsimons. Compiling basic linear algebra subroutines for quantum computers. *arXiv:1902.10394*, 2019.

[36] S. Lloyd, M. Mohseni, and P. Rebentrost. Quantum algorithms for supervised and unsupervised machine learning. *arXiv:1307.0411*, 2013.

[37] IBM Quantum team. *Retrieved from https://quantum-computing.ibm.com,* 2020.