

# Connecting geometry and performance of two-qubit parameterized quantum circuits

Amara Katarwa<sup>1</sup>, Sukin Sim<sup>2,1</sup>, Dax Enshan Koh<sup>3</sup>, and Pierre-Luc Dallaire-Demers<sup>1</sup>

<sup>1</sup>Zapata Computing, Inc., 100 Federal Street, 20th Floor, Boston, Massachusetts 02110, USA

<sup>2</sup>Harvard University

<sup>3</sup>Institute of High Performance Computing, Agency for Science, Technology and Research (A\*STAR), 1 Fusionopolis Way, #16-16 Connexis, Singapore 138632, Singapore

Parameterized quantum circuits (PQCs) are a central component of many variational quantum algorithms, yet there is a lack of understanding of how their parameterization impacts algorithm performance. We initiate this discussion by using principal bundles to geometrically characterize two-qubit PQCs. On the base manifold, we use the Mannoury-Fubini-Study metric to find a simple equation relating the Ricci scalar (geometry) and concurrence (entanglement). By calculating the Ricci scalar during a variational quantum eigensolver (VQE) optimization process, this offers us a new perspective to how and why Quantum Natural Gradient outperforms the standard gradient descent. We argue that the key to the Quantum Natural Gradient's superior performance is its ability to find regions of high negative curvature early in the optimization process. These regions of high negative curvature appear to be important in accelerating the optimization process.

## 1 Introduction

Parameterized quantum circuits (PQCs) are central components of variational quantum algorithms, a class of algorithms well-suited for being implemented on near-term quantum devices. In these algorithms, parameters of PQCs are tuned or optimized, using a classical computer, to prepare quantum states on a quantum computer that encode solutions of a problem, e.g. the ground state of a quantum system or a target probability distribution [1, 2]. While PQCs can be constructed leveraging physical insights, e.g. unitary coupled-cluster [3–5], in applications such as quantum machine learning or implementation of proof-of-principle experimental demonstrations of variational quantum algorithms, PQCs follow a more heuristic design. These PQCs comprise repeated layers of a particular low-depth configuration of single-qubit and two-qubit gate operations [6, 7]. Despite the rapid developments in variational quantum algorithms, PQCs are not yet well understood nor effectively designed. For instance, while

Amara Katarwa: [amara@zapatacomputing.com](mailto:amara@zapatacomputing.com), These authors contributed equally to this work.

Sukin Sim: [sukin.sim@zapatacomputing.com](mailto:sukin.sim@zapatacomputing.com), These authors contributed equally to this work.

Dax Enshan Koh: [dax\\_koh@ihpc.a-star.edu.sg](mailto:dax_koh@ihpc.a-star.edu.sg)

Pierre-Luc Dallaire-Demers: [pierre-luc@zapatacomputing.com](mailto:pierre-luc@zapatacomputing.com)

“hardware-efficient” circuits may correspond to very low depths, it was shown that many of the parameters (and their corresponding gates) are often unnecessary or redundant [8–10]. In our work, to develop a more concrete understanding of the role of parameters in PQCs, we begin with a mathematical description of PQCs.

Formally, parameterized quantum circuits are maps  $\Psi(\theta)$  between a set of continuous parameters  $\theta$  and the output statistics of a set of observables on a given system. In recent

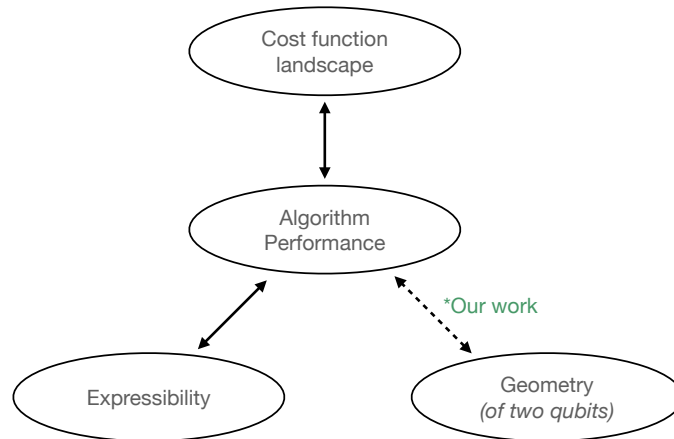


Figure 1: Progress on the study of parameterized quantum circuits (PQCs) and their connections to the algorithm performance. Past works have investigated how use of particular circuits leads to features in the cost function landscape that hinder algorithm performance (e.g. barren plateaus). Other works have introduced circuit metrics such as expressibility and started correlating them to algorithm performance. Our work, while limited to two qubits, initiates the discussion to connect the geometry of PQCs to the algorithm performance and provides valuable insights into why and how Quantum Natural Gradient accelerates the optimization step of near-term quantum algorithms.

years, significant progress has been made to better understand PQCs and their connection to the algorithm performance as illustrated in Fig. 1. For instance, Ref. [11] introduced the phenomenon of “barren plateaus” or regions of non-informative gradients in cost function landscapes resulting from employing PQCs that form approximate 2-designs. Several works extended this work, connecting particular circuit structures to cost function landscape features such as barren plateaus and narrow gorges [12] that hinder algorithm performance. For a better understanding and evaluation of PQCs, Ref. [13] introduced “expressibility” as a quantity to compare among PQCs and rule out circuits with limited capabilities. Since then, works such as Ref. [14] have started correlating expressibility to performance metrics of particular variational quantum algorithms. Additionally, Ref. [15] connected high expressibility to the presence of barren plateaus in cost function landscapes.

Among past works connecting PQCs to algorithm performance, Ref. [16] introduced and investigated the use of Quantum Natural Gradient (QNG) descent to accelerate the optimization of parameters making connections to imaginary time evolution. This modification of the gradient descent which involves an inverse of a metric that is one fourth of the Quantum Fisher Information. The Fubini-Study metric tensor employed in QNG does not contain information about the objective function (unlike gradients or Hessian) and yet has been shown to significantly accelerate optimization by accounting for the geometry of the wave function. Since its discovery, the QNG has been a subject of additional study. For example, while it might

suffer from the same *barren plateau* [11] problem, there is numerical evidence that at least for shallow circuits the variance of the cost function is orders of magnitude bigger than if vanilla gradient descent [17] was used. In [18], methods for using the QNG and adaptively choosing learning rates that can be derived from the QNG were shown to outperform traditional methods like ADAM and LBFGS in learning quantum states. These performance gains come with an overhead cost of calculating what the QNG is, but work has been done showing that for quantum simulation there exist efficient algorithms to calculate it [19] and even in the setting where one is required to calculate it from measurements coming from quantum hardware the overall cost is asymptotically negligible in the number of iterations and qubits [20]. The QNG has also been extended to the non-unitary case where some depolarizing noise is allowed in the circuit [21]. To gain a deeper understanding of the connection between the geometry of the wave function and algorithm performance, we initiate the discussion by providing a geometric characterization of two-qubit PQCs.

The structure of the paper is as follows:

- (i) In Sec. 2, we introduce the geometrical formalism we shall use to analyze PQCs. We consider four specific two-qubit PQCs and take some time to carefully geometrically characterize the circuits.
- (ii) In Sec. 3, we introduce the notions of concurrence and the scalar curvature and uncover a simple and remarkable relationship between the two concepts.
- (iii) In Sec. 4, we capitalize on the relationship between concurrence and curvature in order to provide geometrical insights to the VQE optimization process in the PQCs considered in this work.
- (iv) In Sec. 5, we summarize our work, providing future directions of research and open questions.

## 2 Geometry, quantum mechanics, and parameterized quantum circuits

A geometric reformulation of quantum mechanics has largely been a mathematical and theoretical pursuit while the algebraic view of quantum mechanics appears to be more practical for the everyday quantum mechanician. The question that may arise is whether the geometric approach remains a mere mathematical happenstance or can be leveraged in more practical day-to-day settings. For completeness sake, we provide a quick introduction to the geometric structure of quantum mechanics and hopefully lay bare the complications that arise for the projective nature of quantum states. The discussion will hopefully explain more clearly the geometric connection and what goes wrong if it is not fully incorporated in one's understanding of parameterized quantum circuits. The Kähler structure of quantum mechanics furnishes us with a Riemannian metric and symplectic structure encompassed in what is called the *Quantum Geometric Tensor* (for more in-depth pedagogical introductions to this large and rich area, we refer the readers to [22, 23]).

### 2.1 Quick introduction to fiber bundles

Most topological spaces have a complicated geometry for which there most likely is no easy intuitive picture. One way of addressing this problem is instead to think about the local

geometric structure of this complicated topological space. The local geometry will be thought about as being simply a Cartesian product of lower dimensional geometries. For the rest of the paper, we are going to assume that the complicated topological space  $\mathcal{E}$  can be endowed with a manifold structure. The idea of a *fiber bundle* is to imagine the geometry of  $\mathcal{E}$  as locally being a Cartesian product of two lower dimensional manifolds  $\mathcal{B}$  and  $\mathcal{F}$ . If in fact the *total space*  $\mathcal{E}$  can be thought of as merely  $\mathcal{B} \times \mathcal{F}$  then the fiber bundle is called a *trivial bundle*. Otherwise, in general, the local Cartesian structure will be patched together with some twisting that will make the global geometry look very different from the local geometry. The local picture of  $\mathcal{E}$  can be thought of as there being a *base space*,  $\mathcal{B}$  (which we go to by a projection map  $\pi$  from the total space,  $\mathcal{E}$ ) and then at each point on the base space attaching the second geometry called the *fiber*,  $\mathcal{F}$ . We denote the fiber bundle as

$$\mathcal{F} \rightarrow \mathcal{E} \xrightarrow{\pi} \mathcal{B}. \quad (1)$$

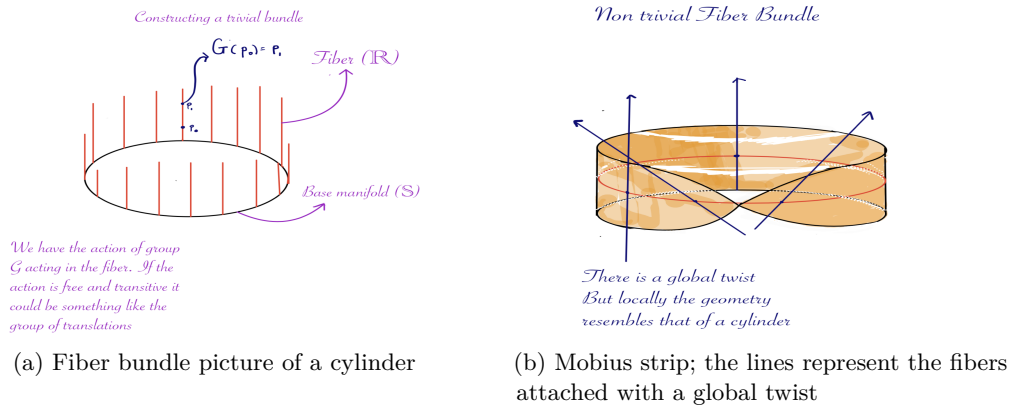


Figure 2: Fiber bundle perspective of a cylinder (a) and a Möbius strip (b). The lines in both pictures represent the fibers, while the circle represents the base manifold. The fiber bundle perspective helps us think locally of complicated manifolds in terms of smaller dimensional manifolds.

For quantum mechanics, we are interested in fiber bundle for which we have an action of a group  $\mathcal{G}$ . How does the group act on our fiber bundle? For all  $g \in \mathcal{G}$ , and for any point  $b \in \mathcal{B}$ ,  $bg$  is merely another point in the fiber attached at the base point  $b$ . We demand we have continuous right action of the group so that in fact  $\mathcal{G}$  is a lie group and that this action is *free* and *transitive* i.e. that respectively only the identity element preserves a point in the fiber,  $\mathcal{F}$  and that for any two points,  $f, f'$  in the fiber there is a unique group element  $g$  such that  $f = f'g$ . What this amounts to is that the group  $\mathcal{G}$  is in fact isomorphic to the fiber  $\mathcal{F}$ . When this happens the fiber bundle is called a *Principal Fiber Bundle*. This principal fiber bundle is then denoted as

$$\mathcal{G} \rightarrow \mathcal{E} \xrightarrow{\pi} \mathcal{B}. \quad (2)$$

### 2.1.1 Quantum mechanics and geometry

We are often taught early that global phases do not matter. While this is usually peppered over since these phases have no physical consequences, it has major consequences from our point of view. Consider a wave function in a Hilbert space,  $|\psi\rangle \in \mathbb{H}$ . The fact that the global

phase does not physically matter means that we are instead considering equivalence classes denoted as  $[|\psi\rangle]$ . This means that the physical space is in fact a projectivized Hilbert space  $\mathbb{P}(\mathbb{H})$ . Note that we can also think of this as a principal fiber bundle where the group action is given by  $U(1)$  so that the principal fiber bundle we have is

$$U(1) \rightarrow \mathbb{H} \xrightarrow{\pi} \mathbb{P}(\mathbb{H}). \quad (3)$$

On the way to introducing a metric on the Hilbert Space  $\mathcal{H}$ , we recall that we have a Hermitian inner product,  $h : \mathbb{H} \times \mathbb{H} \rightarrow \mathbb{C}$  defined by

$$h(\phi, \psi) = \langle \phi | \psi \rangle = G(\phi, \psi) + iF(\phi, \psi), \quad (4)$$

where  $G(\phi, \psi) = \text{Re}(\langle \phi | \psi \rangle)$  and  $F(\phi, \psi) = \text{Im}(\langle \phi | \psi \rangle)$ . The real part is what produces a Riemannian metric and the imaginary part is what gives us a symplectic structure. To arrive at the metric we shall need to consider the tangent space of  $\mathbb{P}(\mathbb{H})$ . The tangent space at point on  $\mathbb{P}(\mathbb{H})$ ,  $T_{[|\psi\rangle]}\mathbb{P}(\mathbb{H})$  is the quotient vector space  $T_{[|\psi\rangle]}\mathbb{H} / \simeq \mathbb{H}$  where the equivalence is  $|\phi_1\rangle \sim |\phi_2\rangle$  whenever  $|\phi_1\rangle - |\phi_2\rangle = a|\psi\rangle$  where  $a \in \mathbb{C}$ . By choosing an orthogonal complement to the space spanned by  $|\psi\rangle$  i.e  $\langle \psi | \phi \rangle = 0$ , we don't have an ambiguity in the tangent space. To ensure that this condition is always met, we use the projection to the orthogonal complement:

$$P_{\psi}^{\perp} = \mathbb{I} - \frac{|\psi\rangle\langle\psi|}{\langle\psi|\psi\rangle}. \quad (5)$$

On the tangent space remembering  $T_{[|\psi\rangle]}\mathbb{H} \simeq \mathbb{H}$ , we see that Eq. (4) may be written in the form

$$h(\phi_1, \phi_2) = \langle \phi_1 | P_{\psi}^{\perp} | \phi_2 \rangle = \langle \phi_1 | \phi_2 \rangle - \langle \phi_1 | \psi \rangle \langle \psi | \phi_2 \rangle. \quad (6)$$

This is in fact what is called the *Quantum Geometric Tensor* in the literature. To see this, consider the wave function  $|\psi\rangle$ , then we go to the tangent space with bases  $\{|\partial_i\psi\rangle\}$  and then plugging this into (6) we get

$$h(\partial_i\psi, \partial_j\psi) = \langle \partial_i\psi | P_{\psi}^{\perp} | \partial_j\psi \rangle = \langle \partial_i\psi | \partial_j\psi \rangle - \langle \partial_i\psi | \psi \rangle \langle \psi | \partial_j\psi \rangle. \quad (7)$$

*Remark:* Note the intrinsic nature of the above derivation for the Quantum Geometric Tensor. Usually this is arrived at by considering the transition between a state  $|\psi_{\vec{u}}\rangle$  and infinitesimally close-by state  $|\psi_{\vec{u}+\delta\vec{u}}\rangle$  then imposing ‘gauge invariance’ of quantum mechanics [24] or noting that the squared differential,  $ds^2 = g_{ij}dx^i dx^j$ , can in quantum mechanics be written as  $ds^2 \propto (1 - F^2)$  where  $F = |\langle \phi | \psi \rangle|$  [16] to lowest order in the parameters that appear in the wave function. The latter derivations, [24] and [16], while correct obscure its intimate origins to the projective nature of Quantum mechanics. The real part of the Quantum Geometric Tensor in (6) is the *Fubini-Study metric*.

## 2.2 Fiber bundles and PQC

A bit of care must be taken in applying the information geometric view to PQC. In order to see why, we carefully define what a PQC is. To do this we follow Ref. [25] and define a variational ansatz or PQC as follows:

**Definition 1.** A *parameterized quantum circuit* is the image of the following map  $\Psi : \mathcal{U} \in \mathbb{R}^m \rightarrow \mathbb{P}(\mathbb{H})$ , i.e.

$$\mathcal{M} = \Psi(\mathcal{U}) := \{[\Psi(\boldsymbol{\theta})]\} \quad (8)$$

where  $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_m) \in \mathcal{U} \subset \mathbb{R}^m$ .

To get at the metric we go to the tangent space and as a consequence define the following push-forward map:

$$d\Psi(\mathbf{v}) : T_p\mathbb{R} \rightarrow T_{[\Psi(\vec{\theta})]}\mathcal{M} : v^i \partial_{\theta_i} \longrightarrow v^i \partial_{\theta_i} [\Psi(\boldsymbol{\theta})] = v^i [\partial_{\theta_i} \Psi(\boldsymbol{\theta})] \quad (9)$$

where  $\mathbf{v} = v^i \partial_{\theta_i}$ . This allows us to define the metric on  $\mathbb{P}(\mathbb{H})$  using our points in  $\mathbb{R}^m$  and this is done by using the pull-back by  $\Psi$ , back to  $T_p\mathbb{R}^m$  i.e.

$$g_{\boldsymbol{\theta}} : T_p\mathbb{R}^m \times T_p\mathbb{R}^m \rightarrow \mathbb{R} : (\mathbf{v}_1, \mathbf{v}_2) = g(\mathbf{v}_1, \mathbf{v}_2) = g(d\Psi(\mathbf{v}_1), d\Psi(\mathbf{v}_2)) \quad (10)$$

where we define the value of the metric as

$$g(\mathbf{v}_1, \mathbf{v}_2) = v_1^i v_2^j (\langle \partial_i \Psi | \partial_j \Psi \rangle - \langle \partial_i \Psi | \Psi \rangle \langle \Psi | \partial_j \Psi \rangle). \quad (11)$$

Here, we have suppressed the dependence on  $\boldsymbol{\theta}$  for ease of notation; we have implicitly chosen a point on a ray in evaluating the value for the metric.

Now the map  $\Psi$  need not be injective so that in fact the metric in Eq. (11) is, in general, degenerate. This is in fact what happens with the circuits we study in this work. As a consequence, no unique inverse metric exists. The problem derives from the fact that the parameters in the wave function are not chosen with the projective nature of quantum mechanics and thus when we ask physical questions like expectation values of hermitian operators we will in fact live on some constrained surface in  $\mathbb{P}(\mathbb{H})$ ; in other words our tangent space at some point will in general be of smaller dimension than  $T_p\mathbb{R}^m$ .

Naively, we have the following Quantum Geometric Tensor for our PQC:

$$G_{\theta_i \theta_j} = \langle \partial_{\theta_i} \Psi(\boldsymbol{\theta}) | \partial_{\theta_j} \Psi(\boldsymbol{\theta}) \rangle - \langle \partial_{\theta_i} \Psi(\boldsymbol{\theta}) | \Psi(\boldsymbol{\theta}) \rangle \langle \Psi(\boldsymbol{\theta}) | \partial_{\theta_j} \Psi(\boldsymbol{\theta}) \rangle. \quad (12)$$

### 2.2.1 Two-qubit parameterized quantum circuits and geometry

In this section we study how the PQCs fit in the geometry of  $\mathbb{P}(\mathbb{H})$  and find the constrained surfaces they live on. This exercise elucidates what part of the geometry the PQCs have access to.

There will be four major circuits we shall consider in studying the information-geometric view of PQCs, namely:

1. the *Hardware-Efficient Ansatz* (HEA) introduced by [7] and studied in the context of the Fubini-Study metric by [26],
2. the *Low Depth Circuit Ansatz* (LDCA) introduced in [27],
3. an ansatz introduced in the context of quantum generative adversarial networks (QGANs) [28], and

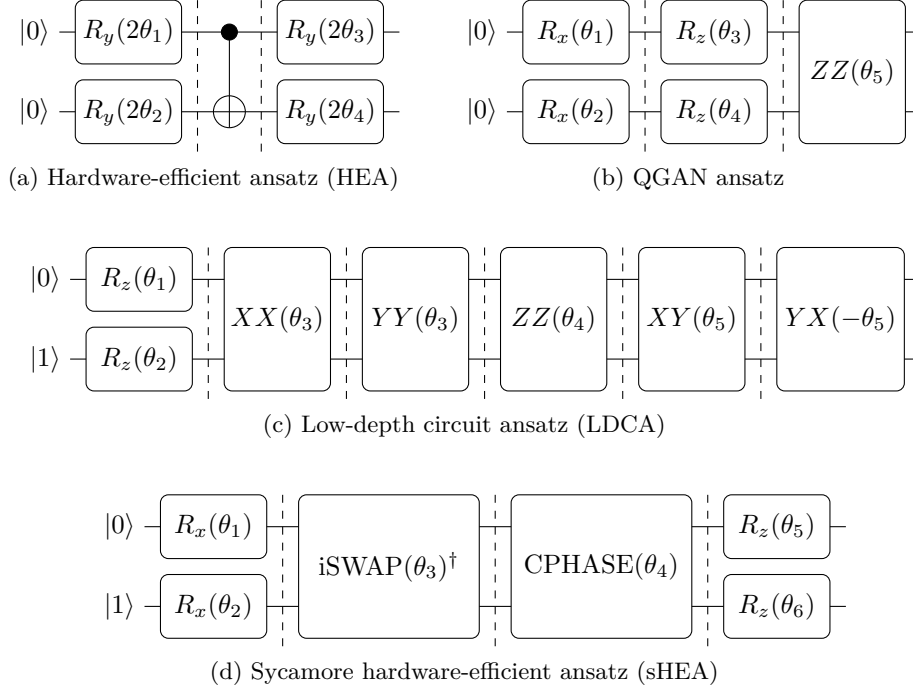


Figure 3: Two-qubit circuit blocks considered in this work. Dashed lines indicate different circuit layers or moments. Gate definitions are provided in Appendix A.

4. an ansatz used in [29] composed of gates native to the Sycamore chip [30], e.g. the fermionic simulation (fSim) operation. We will refer to this ansatz as the Sycamore HEA (sHEA).

In general, we have the following isomorphisms

$$\mathbb{P}(\mathbb{H}^{n+1}) \simeq S^{2n-1}/S^1 \simeq \mathbb{C}P^n \simeq G_{1,n+1} \quad (13)$$

where  $\mathbb{P}(\mathbb{H}^{n+1})$  is the projectivized version of an  $n + 1$  dimensional Hilbert space,  $\mathbb{C}P^n$  is the complex projective space and  $G_{1,n+1}$  is the Grassmanian of dimension  $n$ . For two qubits, the string of isomorphisms in Eq. (13) become specifically the following isomorphisms:

$$\mathbb{P}(\mathbb{H}^4) \simeq S^7/S^1 \simeq \mathbb{C}P^3 \simeq G_{1,4}. \quad (14)$$

Luckily, for two and incidentally for three qubits, we have a fiber bundle picture namely the Hopf Fibration. For two qubits, the Hopf Fibration can be thought of as an  $SU(2)$  principal fiber bundle. We now concentrate on  $S^7$  and think of it having a Hopf Fibration. It is a well-known fact that the fibration has the following character:

$$S^3 \rightarrow S^7 \xrightarrow{\pi} S^4 \quad (15)$$

The fiber  $S^3$  is isomorphic to  $SU(2)$ .

From a quantum mechanical point of view  $S^4$ , which can locally be split into two spheres, i.e  $S^2 \times S^2$ , can be given the following interpretation: the first sphere is a *quasi*-Bloch sphere

representing the degrees of freedom one observer has access to for a given two-qubit entangled state and the second sphere parameterizes the amount of entanglement shared by the two qubits [31].

For each of the major circuits, we calculate the version of the metric that is in general degenerate, a metric that is specifically tailored to the geometry of  $S^7$  and provide explicit parametrizations of how the two qubit circuits sit inside the geometry.

The Hopf base and fiber parameters are calculated as follows [31, 32]. For a 2-qubit state:

$$|\psi\rangle = \alpha |00\rangle + \beta |01\rangle + \gamma |10\rangle + \delta |11\rangle \quad (16)$$

where  $\alpha, \beta, \gamma, \delta \in \mathbb{C}$ , compute the following:

1. First, calculate the  $S^4$  Hopf base parameters  $(\theta_A, \phi_A)$  and  $(\chi, \xi)$  and the equivalent Cartesian coordinates  $(x_0, x_1, x_2, x_3, x_4)$  as follows:

- (a)  $x_0 = |\alpha|^2 + |\beta|^2 - |\gamma|^2 - |\delta|^2$
- (b)  $\theta_A = \arccos(x_0)$ .
- (c)  $x_1 = 2 \operatorname{Re}(\bar{\alpha}\gamma + \bar{\beta}\delta)$
- (d)  $x_4 = 2 \operatorname{Im}(\bar{\alpha}\gamma + \bar{\beta}\delta)$
- (e)  $\phi_A = \arccos(x_1 \csc(\theta_A))$
- (f)  $\chi = \arccos(x_4 \csc(\theta_A) \csc(\phi_A))$
- (g)  $x_3 = 2 \operatorname{Re}(\alpha\delta - \beta\gamma)$
- (h)  $x_2 = -2 \operatorname{Im}(\alpha\delta - \beta\gamma)$
- (i)  $\xi = \arctan(x_3/x_2)$

In calculating how the PQC's fit inside the fibers using intrinsic co-ordinates, one needs to solve highly non-linear trigonometric equations. We simplify our task, by switching co-ordinates to the constrained extrinsic co-ordinates. We take advantage of the following parameterization equivalence:

$$|\psi_q\rangle = \begin{pmatrix} \cos \theta_A \\ \sin \theta_A e^{t\phi_A} \end{pmatrix} q = \frac{1}{\sqrt{2}} \left( \begin{array}{c} \sqrt{1 \pm \sqrt{1 - |z|^2 - |w|^2}} \\ \sqrt{1 \mp \sqrt{1 - |z|^2 - |w|^2}} \frac{z+wj}{\sqrt{|z|^2+|w|^2}} \end{array} \right) q, \quad (17)$$

where  $q$  is a unit quaternion used to parameterize the fiber and  $e^{t\phi_A}$  is a unit quaternion used to parameterize a point on the entanglement sphere.

2. Next, calculate the  $S^3$  Hopf fiber parameters, which are represented by the quaternion  $q$ .

- (a)  $z = \frac{1}{2}(x_1 + ix_4)$
- (b)  $w = \frac{1}{2}(x_3 - ix_2)$
- (c)  $\gamma_{\pm} = \sqrt{1 \pm \sqrt{1 - |z|^2 - |w|^2}}$
- (d)  $|\psi_H\rangle = (\alpha + \beta j) |0\rangle + (\gamma + \delta j) |1\rangle$



$$(e) |c_{\pm}\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} \gamma_{\pm} \\ \gamma_{\mp} \frac{z+wj}{\sqrt{|z|^2+|w|^2}} \end{pmatrix}$$

$$(f) q_{\pm} = \langle c_{\pm} | \psi_H \rangle$$

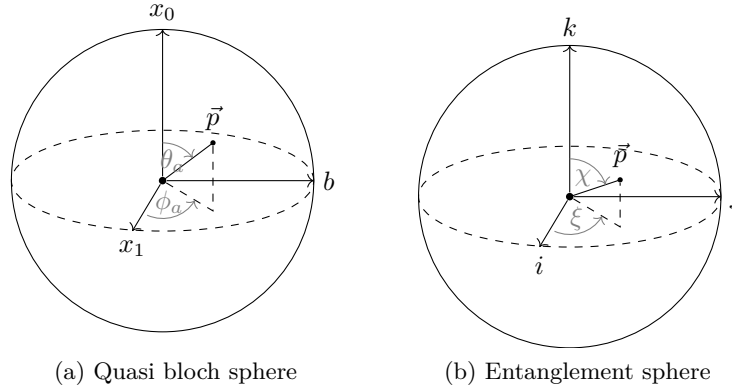


Figure 4: Local geometry of  $S^4$ .

We compute the above Hopf base and fiber parameters for the circuit blocks in Fig. 3 and present our results in Appendix B.

### 2.2.2 Local geometric expressibility

The notion of expressibility for PQCs has been explored in [13]. By thinking of the Hilbert space from the geometric point of view, namely as  $S^7$ , we can arrive at a picture of *local geometric expressibility*. From this perspective we ask ourselves which points in  $S^7$  can our PQCs reach, and which constrained surfaces do they live on in terms of some un-constrained co-ordinates. As discussed earlier, the local degrees of freedom accessible to an observer are parameterized by one of the spheres with intrinsic coordinates  $(\theta_A, \phi_A)$  on the base manifold while the entanglement properties are parameterized by the second sphere parameterized by intrinsic coordinates  $(\chi, \xi)$ .

Table 1: PQC base manifold geometry

| Quantum Circuit | $S^2$ (local sphere)               | $S^2$ (entanglement sphere)              | Constrained Base Space Geometry |
|-----------------|------------------------------------|--|---------------------------------|
| HEA             | Explores with $(\phi_A, \theta_A)$ | Does not explore (Stuck to just a point) | $S^2$                           |
| LDCA            | Explores with $(\theta_A)$         | Explores with $\xi$                      | $S^2$                           |
| QGAN            | Explores with $(\phi_A, \theta_A)$ | Explores with $\chi$                     | $S^3$                           |
| sHEA            | Explores with $(\phi_A, \theta_A)$ | Explores with $(\chi, \xi)$              | $S^4$                           |

The sHEA circuit covers the largest surface area of points, while we see that both HEA and LDCA cover points on a two sphere. Interestingly although both HEA and LDCA cover points constrained to live on  $S^2$ , LDCA uses one of its co-ordinates to explore the entanglement sphere. This, in principle, allows LDCA to explore more sub-manifolds of different entanglement in  $S^4$ .

Next we consider how the different circuits differ in the fiber space. Points in the fiber space appear from one observer's point of view as a difference in the gauge i.e they do not change the entanglement properties on the quantum state (This can be made more explicit by introducing a connection). To help abstract out the minutiae we first re-express (65), (79), (92) that represent the parameterization in the fiber space in a way that makes physical interpretation easier. For a fixed fiber, these functions will in general depend on a subset of the parameters in the quantum circuit as one moves in that fiber. Using the algebraic equivalence between quaternions and the Lie algebra of  $SU(2)$  namely:

$$i \mapsto i\sigma^x, j \mapsto i\sigma^y, k \mapsto i\sigma^z$$

We re-write (65), (79), (92) respectively as follows:

$$q_{\pm HEA} = \frac{1}{2\lambda_2} (e_{\pm}(\theta_4)I + if_{\pm}(\theta_4)\sigma^y) \quad (18)$$

$$q_{\pm LDCA} = \frac{1}{2\mu_3} i(g_{\pm}(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5)\sigma^y + h_{\pm}(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5)\sigma^z) \quad (19)$$

$$q_{\pm QGANs} = \frac{1}{\sqrt{2}} \left( \sin\left(\frac{\theta_1}{2}\right) \gamma_{\mp} + \cos\left(\frac{\theta_1}{2}\right) \gamma_{\pm} \right) i(a_{\pm}(\theta_2, \theta_3, \theta_4, \theta_5)I + b_{\pm}(\theta_2, \theta_3, \theta_4, \theta_5)\sigma^x + c_{\pm}(\theta_2, \theta_3, \theta_4, \theta_5)\sigma^y + d_{\pm}(\theta_2, \theta_3, \theta_4, \theta_5)\sigma^z) \quad (20)$$

$$q_{\pm sHEA} = i(m_{\pm}(\theta_1, \theta_2, \theta_3, \theta_5, \theta_6)I + m_{\pm}(\theta_1, \theta_2, \theta_3, \theta_5, \theta_6)\sigma^x + r_{\pm}(\theta_1, \theta_2, \theta_3, \theta_5, \theta_6)\sigma^y + s_{\pm}(\theta_1, \theta_2, \theta_3, \theta_5, \theta_6)\sigma^z). \quad (21)$$

In the equations above only the  $\theta'_i$ s shown are those that can vary in a specific fiber on a point in the base manifold. Expressions for  $a - g, m, n, s$  can be found in Appendix B.

Table 2: PQC Geometry Summary Table

| Quantum Circuit | Fiber Space   |
|-----------------|---|
| HEA             | $\frac{1}{2\lambda_2} (e_{\pm}(\theta_4)I + if_{\pm}(\theta_4)\sigma^y)$  |
| LDCA            | $\frac{1}{2\mu_3} i(g_{\pm}(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5)\sigma^y + h_{\pm}(\theta_1, \theta_2, \theta_3, \theta_4, \theta_5)\sigma^z)$   |
| QGAN            | $\frac{1}{\sqrt{2}} \left( \sin\left(\frac{\theta_1}{2}\right) \gamma_{\mp} + \cos\left(\frac{\theta_1}{2}\right) \gamma_{\pm} \right) i(a_{\pm}(\theta_2, \theta_3, \theta_4, \theta_5)I + b_{\pm}(\theta_2, \theta_3, \theta_4, \theta_5)\sigma^x + c_{\pm}(\theta_2, \theta_3, \theta_4, \theta_5)\sigma^y + d_{\pm}(\theta_2, \theta_3, \theta_4, \theta_5)\sigma^z)$ |
| sHEA            | $i(m_{\pm}(\theta_1, \theta_2, \theta_3, \theta_5, \theta_6)I + m_{\pm}(\theta_1, \theta_2, \theta_3, \theta_5, \theta_6)\sigma^x + r_{\pm}(\theta_1, \theta_2, \theta_3, \theta_5, \theta_6)\sigma^y + s_{\pm}(\theta_1, \theta_2, \theta_3, \theta_5, \theta_6)\sigma^z)$   |

### 2.3 Quantum Natural Gradient descent

We have discussed the possibility of the metric derived for directly applying Eq. (12) being generally degenerate. In this section, we calculate the ‘‘Fubini-Study’’ metrics (by following Eq. (12)) and see that most are indeed degenerate. Although this is the case, it has been shown in [16] that these degenerate matrices, in practice, lead to improvements in the number of iterations for the optimization process, through approximating to a block-diagonal or diagonal form, adding numerically small values to make matrices invertible, and/or taking the pseudo-inverse. A more recent study, however, notes that (block-)diagonal approximations to the metric tensor may not be necessary as the cost of computing the metric tensor is asymptotically negligible [20]. Using the standard gradient descent method, the parameter update rule for the parameters in the quantum circuit is:

$$\boldsymbol{\theta}^{(t+1)} = \boldsymbol{\theta}^{(t)} - \eta \frac{\partial \mathcal{L}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}. \quad (22)$$

where  $\eta$  is the step size or learning rate and  $\mathcal{L}(\boldsymbol{\theta})$  is the objective function to be minimized.

For the four circuit types considered, we have the following Fubini-Study metrics assuming (12) and naively calculating them. Matrix elements that are included in the block-diagonal approximation to the metric tensor are in blue.

$$g^{\text{HEA}} = \begin{pmatrix} 1 & 0 & \langle \sigma_2^x \rangle_1 & 0 \\ 0 & 1 & 0 & \langle \sigma_1^z \rangle_1 \\ \langle \sigma_2^x \rangle_1 & 0 & 1 & \frac{\langle \{\sigma_1^y, \sigma_2^y\} \rangle_2}{2} \\ 0 & \langle \sigma_1^z \rangle_1 & \frac{\langle \{\sigma_{(1)}^y, \sigma_{(2)}^y\} \rangle_2}{2} & 1 \end{pmatrix} \quad (23)$$

$$g^{\text{LDCA}} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 + \frac{\langle \{\sigma_1^z, \sigma_2^z\} \rangle_4}{2} \end{pmatrix} \quad (24)$$

$$g^{\text{QGAN}} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 - \langle \sigma_1^z \rangle_1^2 & 0 & \langle \sigma_2^z \rangle_2 - \langle \sigma_1^z \rangle_1 \langle \sigma_1^z \sigma_2^z \rangle_1 & 0 \\ 0 & 0 & 0 & 1 - \langle \sigma_2^z \rangle_1^2 & \langle \sigma_1^z \rangle_2 - \langle \sigma_2^z \rangle_1 \langle \sigma_1^z \sigma_2^z \rangle_1 & \langle \sigma_1^z \rangle_2 - \langle \sigma_2^z \rangle_1 \langle \sigma_1^z \sigma_2^z \rangle_1 \\ 0 & 0 & \langle \sigma_2^z \rangle_2 - \langle \sigma_1^z \rangle_1 \langle \sigma_1^z \sigma_2^z \rangle_1 & \langle \sigma_1^z \rangle_2 - \langle \sigma_2^z \rangle_1 \langle \sigma_1^z \sigma_2^z \rangle_1 & 1 - \langle \sigma_1^z \sigma_2^z \rangle_1^2 & 0 \end{pmatrix} \quad (25)$$

$$g^{\text{sHEA}} = \begin{pmatrix} 1 & 0 & 0 & 0 & A & -A \\ 0 & 1 & 0 & 0 & B & -B \\ 0 & 0 & C & D & E & -E \\ 0 & 0 & D & F & -G & G \\ A & B & E & -G & H & I \\ -A & -B & -E & G & I & H \end{pmatrix} \quad (26)$$

$$(27)$$

where for  $g^{sHEA}$  we have

$$\begin{aligned}
A &= {}^1\langle\sigma_1^x \mathcal{S}(\theta_3)^\dagger \sigma_1^z\rangle_2 - \langle\sigma_1^x\rangle_0 \langle\sigma_1^z\rangle_2 \\
B &= {}^1\langle\sigma_2^x \mathcal{S}(\theta_3)^\dagger \sigma_1^z\rangle_2 - \langle\sigma_2^x\rangle_0 \langle\sigma_1^z\rangle_2 \\
C &= \frac{1}{2}(1 - \langle\sigma_1^z \sigma_2^z\rangle) - \frac{1}{4}(\langle\sigma_1^x \sigma_2^x + \sigma_1^y \sigma_2^y\rangle_1)^2 \\
D &= \frac{1}{8}(-\langle(\sigma_1^x \sigma_2^x + \sigma_1^y \sigma_2^y)(I_1 - \sigma_1^z)(I_2 - \sigma_2^z)\rangle_2 + \langle\sigma_1^x \sigma_2^x + \sigma_1^y \sigma_2^y\rangle_1 \langle(I_1 - \sigma_1^z)(I_2 - \sigma_2^z)\rangle_2) \\
E &= \frac{1}{2}(\langle(\sigma_1^x \sigma_2^x + \sigma_1^y \sigma_2^y)\sigma_1^z\rangle_2 + \langle\sigma_1^x \sigma_2^x + \sigma_1^y \sigma_2^y\rangle_1 \langle\sigma_1^z\rangle_2) \\
F &= \frac{1}{4}\langle(I_1 - \sigma_1^z)(I_2 - \sigma_2^z)\rangle_2 - \frac{1}{16}\langle(I_1 - \sigma_1^z)(I_2 - \sigma_2^z)\rangle_2^2 \\
G &= \frac{1}{4}\langle(I_1 - \sigma_1^z)(I_2 - \sigma_2^z)\rangle_2 - \frac{1}{4}\langle(I_1 - \sigma_1^z)(I_2 - \sigma_2^z)\rangle_2 \langle\sigma_1^z\rangle_2 \\
H &= 1 - \langle\sigma_2^z\rangle_2^2 I = \langle\sigma_1^z \sigma_2^z\rangle_2 - \langle\sigma_1^z\rangle_2 \langle\sigma_2^z\rangle_2.
\end{aligned}$$

**Note:** The notation used above  $\langle\mathcal{O}\rangle_i$  is used to mean the expectation value of the operator  $\mathcal{O}$  with respect to the wave function at the  $i^{\text{th}}$  layer of the circuit, while  ${}^j\langle\mathcal{O}\rangle_i = \langle\psi_j|\mathcal{O}|\psi_i\rangle$  and  $|\psi_i\rangle$  is the wave function at the  $i^{\text{th}}$  layer.

Of the four metrics, ones for the QGAN circuit and sHEA are non-degenerate. From the previous section, we can see why this is true; we have 3 independent parameters in the base space, 3 parameters in the fiber space but we also have a global constraint that all the amplitudes must add to 1 so that we have a total of 5 independent global parameters. This matches the number of parameters in the quantum circuit. The same can be seen with the sHEA geometry where we get an extra degree of independence from the base space.

In practice,  $g$  may be singular due to redundant parameterization, or inverting  $g$  becomes computationally challenging with increasing number of parameters. Thus, in the case of natural gradient descent for training classical neural networks with many parameters, the inverse of the Fisher Information Matrix (FIM) has been approximated as an inverse of a block-diagonal matrix (and further approximated as Kronecker products of inverses of smaller matrices) to reduce the computational overhead [33]. The FIM of deep linear networks are also often singular due to parameter redundancies. Thus, generalized inverses are used. Ref. [34] showed that for deep linear networks, any choice of generalized inverse was effective in accelerating natural gradient descent.

### 3 Concurrence and the Ricci Scalar

Unlike its classical analogue, the Quantum Natural Gradient has no obvious connection to the geometry encountered in the optimization process. One might have considered calculating the curvature from the Fubini-Study metric derived from the real part of the Quantum Geometric Tensor, but as has been previously argued, one does not get a legitimate metric since the metric used in Quantum Natural Gradient is in general degenerate. As a consequence, one could consider regularizing the metric either by considering the pseudo-inverse or by adding a small constant multiplied by the identity matrix ( $\epsilon I$ ), a kind of Tikhonov regularization. The downside to this path is that the value of the curvature depends on the regularization procedure and in fact for the Tikhonov regularization the curvature in some cases is ill-defined

since the value at a point depends on how this small constant is brought to zero. Nevertheless, in practice, carefully setting regularization has been shown to reduce the measurement costs of QNG [20].

What we opt for therefore, is a rather indirect way of seeing the Quantum Natural Gradient geometrically at work. We will calculate another metric, a *quaternionic* Fubini-Study metric that can be placed on the base manifold of two qubits,  $S^4$ . Amazingly enough this metric connects the concurrence (entanglement) of the quantum circuit and the curvature of the base manifold.

### 3.1 Concurrence

In the case of two qubits, the entanglement entropy is the unique measure of entanglement. In work by Hill and Wootters [35], it was noted that the entanglement entropy can be thought of as a function of the concurrence which is defined in the following manner:

$$C(\psi) = 2|\alpha\delta - \beta\gamma|, \quad (28)$$

where  $|\psi\rangle = \alpha|00\rangle + \beta|01\rangle + \gamma|10\rangle + \delta|11\rangle$ . The concurrence has been measured in experiments [36, 37]. Extensions to mixed density matrices can be considered by calculating the convex roof extension, but in this work we stick to pure states when calculating the measure.

We can also formulate the notion of concurrence within the geometric picture thus far outlined. First note that the concurrence can be written in terms of the extrinsic co-ordinates of  $S^4$  as:

$$C(\psi) = \sqrt{|x_2|^2 + |x_3|^2}. \quad (29)$$

The concurrences for the four circuits in terms of their corresponding circuit parameters are:

$$C_{\text{HEA}} = |\sin(2\theta_1)\cos(2\theta_2)|, \quad (30)$$

$$C_{\text{LDCA}} = \frac{1}{2}\sqrt{3 - 2\cos(4\theta_3)\cos^2(2\theta_5) - \cos(4\theta_5)}, \quad (31)$$

$$C_{\text{QGAN}} = |\sin(\theta_1)\sin(\theta_2)\sin(\theta_5)|, \quad (32)$$

$$C_{\text{sHEA}} = \frac{1}{2}\left\{\sin^2(\theta_1)\sin^2(\theta_2)\left(\cos(\theta_3) - \cos\left(\frac{\theta_4}{4}\right)\right)^2\right. \quad (33)$$

$$\left. + \left(\sin(\theta_3) - \sin(\theta_1)\sin(\theta_2)\sin\left(\frac{\theta_4}{4}\right) + \sin(\theta_3)\cos(\theta_1)\cos(\theta_2)\right)^2\right\}^{1/2}. \quad (34)$$

### 3.2 Ricci Scalar

In Riemannian geometry, the scalar curvature, also known as the *Ricci scalar*, is an invariant that characterizes the curvature of a Riemannian manifold. It may be defined in terms of the metric tensor  $g$  as follows: Let  $g_{ab}$  denote the components of  $g$ , and let  $g^{ab}$  denote the components of its inverse  $g^{-1}$ . The *scalar curvature* is defined as

$$\mathcal{R} = g^{ab}\left(\Gamma_{ab,c}^c - \Gamma_{ac,b}^c + \Gamma_{ab}^d\Gamma_{cd}^c - \Gamma_{ac}^d\Gamma_{bd}^c\right), \quad (35)$$

where

$$\Gamma^c_{ab} = \frac{1}{2}g^{cd}(g_{da,b} + g_{db,a} - g_{ab,d}) \quad (36)$$

are the *Christoffel symbols of the first kind*. In the equations above, note that we have used Einstein's summation convention and that the commas in the subscripts indicate a partial derivative: for example,  $\Gamma^c_{ab,d} = \partial_d \Gamma^c_{ab}$ . For more details on the scalar curvature, we refer the reader to [38].

Apart from its importance in areas in differential geometry and in cosmology, the Ricci scalar is slowly finding applications in hitherto unexpected places for example the study of phase-transitions in quantum many-body systems [39–41] and in classical machine learning, a formulation of neural nets in the context of Riemannian Geometry has been explored [42] while specific use of Riemannian curvature has been explored in [43, 44].

Using the isomorphism  $HP^1 \simeq S^4$  where  $HP^1$  is the quaternionic projective space we can incorporate the concurrence as part of the calculation of the *Mannoury-Fubini-Study* metric [32] as follows:

$$g_{\mu\nu} = \frac{1}{1-C^2}dC^2 + C^2d\chi^2 + (1-C^2)(d\Phi^2 + \sin^2\Theta d\Theta^2) \quad (37)$$

where  $w = |C|e^{i\chi}$  (defined in Sec. 2.3), and  $0 < \Phi \leq 2\pi, 0 < \Theta \leq \pi$  are the usual polar co-ordinates for  $B^3$  (three dimensional ball). We may make an amusing observation that this is the same metric as the Euclidean Schwarzschild metric for fixed time.

The metrics for the four circuits are therefore

$$g_{\text{HEA}} = \begin{pmatrix} \frac{1}{1-C_{\text{HEA}}^2} & 0 & 0 & 0 \\ 0 & C_{\text{HEA}}^2 & 0 & 0 \\ 0 & 0 & (1-C_{\text{HEA}}^2) & 0 \\ 0 & 0 & 0 & (1-C_{\text{HEA}}^2) \end{pmatrix} \quad (38)$$

$$g_{\text{LDCA}} = \begin{pmatrix} \frac{1}{1-C_{\text{LDCA}}^2} & 0 & 0 & 0 \\ 0 & C_{\text{LDCA}}^2 & 0 & 0 \\ 0 & 0 & (1-C_{\text{LDCA}}^2) & 0 \\ 0 & 0 & 0 & (1-C_{\text{LDCA}}^2) \end{pmatrix} \quad (39)$$

$$g_{\text{QGAN}} = \begin{pmatrix} \frac{1}{1-C_{\text{QGAN}}^2} & 0 & 0 & 0 \\ 0 & C_{\text{QGAN}}^2 & 0 & 0 \\ 0 & 0 & (1-C_{\text{QGAN}}^2) & 0 \\ 0 & 0 & 0 & (1-C_{\text{QGAN}}^2) \sin \Theta_{\text{QGAN}} \end{pmatrix} \quad (40)$$

$$g_{\text{sHEA}} = \begin{pmatrix} \frac{1}{1-C_{\text{sHEA}}^2} & 0 & 0 & 0 \\ 0 & C_{\text{sHEA}}^2 & 0 & 0 \\ 0 & 0 & (1-C_{\text{sHEA}}^2) & 0 \\ 0 & 0 & 0 & (1-C_{\text{sHEA}}^2) \sin \Theta_{\text{sHEA}} \end{pmatrix} \quad (41)$$

where:

$$\sin \Theta_{\text{QGAN}} = \sqrt{1 - (\sin \theta_1 (\sin \theta_3 \cos \theta_5) + \sin \theta_5 \cos \theta_2 \cos \theta_3)^2} \quad (42)$$

for the QGAN ansatz and

$$\sin \Theta_{\text{sHEA}} = \sqrt{1 - y^2}, \quad (43)$$

where

$$\begin{aligned} y = & \frac{1}{2} \left( \sin(\theta_1) \cos\left(\frac{\theta_3}{2}\right) \left( - \left( -\cos(\theta_2) + \left( \cos(\theta_2) + 1 \right) \cos\left(\frac{\theta_4}{4}\right) + 1 \right) \cos(\theta_5) \right. \right. \\ & - \sin\left(\frac{\theta_4}{4}\right) \sin(\theta_5) \left( \cos(\theta_2) + 1 \right) \left. \right) - \sin(\theta_2) \sin\left(\frac{\theta_3}{2}\right) \left( 2 \sin\left(\frac{\theta_4}{4}\right) \sin^2\left(\frac{\theta_1}{2}\right) \cos(\theta_5) \right. \\ & \left. \left. + \sin(\theta_5) \left( \cos(\theta_1) + \left( \cos(\theta_1) - 1 \right) \cos\left(\frac{\theta_4}{4}\right) + 1 \right) \right) \right) \end{aligned} \quad (44)$$

for the sHEA ansatz.

The scalar curvatures  $\mathcal{R}_{\mathcal{A}}$  (see Eq. (35)) of all the PQC metrics (Eqs. (38)–(41))  $g_{\mathcal{A}}$  (for  $\mathcal{A} = \text{HEA}, \text{LDCA}, \text{QGAN}, \text{sHEA}$ ) are calculated to be:

$$\mathcal{R}_{\mathcal{A}} = \frac{2(6C_{\mathcal{A}}^2 - 5)}{C_{\mathcal{A}}^2 - 1}. \quad (45)$$

Note that the Ricci scalar can be calculated by just knowing the metric. Since the metric in our case is just a function of the concurrence of a general 2-qubit circuit, Eq. (45) is completely general for the two-qubit case.

We note that the formula is a simple rational function of the concurrence, and from the formula, we can see that singularities on the base manifold correspond to maximally entangled states with concurrence value of one.

The scalar curvature of each of the circuit written in terms of their circuit parameters are:

$$\mathcal{R}_{\text{HEA}} = 12 - \frac{1}{\sin(2\theta_1) \cos(2\theta_2) + 1} + \frac{1}{\sin(2\theta_1) \cos(2\theta_2) - 1}, \quad (46)$$

$$\mathcal{R}_{\text{LDCA}} = 12 - 2 \sec^2(2\theta_3) \sec^2(2\theta_5), \quad (47)$$

$$\mathcal{R}_{\text{QGAN}} = 12 - \frac{1}{\sin(\theta_1) \sin(\theta_2) \sin(\theta_5) + 1} + \frac{1}{\sin(\theta_1) \sin(\theta_2) \sin(\theta_5) - 1}, \quad (48)$$

$$\begin{aligned} \mathcal{R}_{\text{sHEA}} = & \frac{\left( 12 \sin^2(\theta_1) \sin^2(\theta_2) \left( \cos(\theta_3) - \cos\left(\frac{\theta_4}{4}\right) \right)^2 \right. \\ & \left. + 12 \left( \sin(\theta_3) (\cos(\theta_1) \cos(\theta_2) + 1) - \sin(\theta_1) \sin(\theta_2) \sin\left(\frac{\theta_4}{4}\right) \right)^2 - 40 \right)}{\sin^2(\theta_1) \sin^2(\theta_2) \left( \cos(\theta_3) - \cos\left(\frac{\theta_4}{4}\right) \right)^2} \\ & + \left( \sin(\theta_3) (\cos(\theta_1) \cos(\theta_2) + 1) - \sin(\theta_1) \sin(\theta_2) \sin\left(\frac{\theta_4}{4}\right) \right)^2 - 4. \end{aligned} \quad (49)$$

These scalar curvatures are additionally plotted in Fig. 5. For HEA and LDCA, their Ricci scalars are each a function of two circuit parameters, i.e. only two parameters influence the entanglement in the system. On the other hand, the Ricci scalars of QGAN and sHEA are functions of three and four parameters, respectively. To visualize how parameter values impact the curvature, values of  $\theta_1$  and  $\theta_2$  are scanned over range  $[0, 2\pi]$  while values of other parameters, if they appear in the expression of the Ricci scalar, are fixed at particular values. With QGAN, by tuning the value of  $\theta_5$  from 0 to  $\frac{\pi}{2}$ , we observe a gradual emergence of “wells” of negative curvature. With sHEA, the landscape depends significantly on values of  $\theta_3$  and  $\theta_4$ . For instance, when  $\theta_3 = \pi$ , there are wells of negative curvature similar to those

from QGAN’s landscape. However, when  $\theta_3$  is decreased to approximately  $\frac{\pi}{2}$ , the wells are replaced by “valleys” of negative curvature. Comparing these landscapes provides some insight into the (relative) entangling capabilities of the circuit blocks and the ease of generating high entanglement states by tuning the circuit parameters. Circuits corresponding to curvature landscapes with extensive or large regions of negative curvatures (i.e. high concurrences) such as LDCA, are able to more readily generate states with high entanglement compared to circuits with landscapes with limited regions of negative curvature similar to that of QGAN. Additionally, it appears easier to generate high entanglement states with LDCA than to do so using circuits such as QGAN and sHEA, in which one must tune multiple circuit parameters, several of which need to be set near specific values to generate states with high entanglement.

Why might one be interested in calculating the Ricci scalar? After all, from the face of it, it seems to contain the same amount of information as the concurrence. The answer lies in the goal of this study, namely understanding PQCs from a geometric point of view and their optimization performance with regards to the Quantum Natural Gradient. We have the intuition that the Quantum Natural Gradient somehow incorporates the geometric information of the projective Hilbert space in the optimization procedure in order improve the speed of convergence. The question is that can we map out the geometry of the quantum circuits and see *geometrically* what the quantum natural gradient is doing differently? This is the question we take up in next section and explore.

## 4 Insights into VQE performance through scalar curvature: a two-qubit study

In this section, we numerically demonstrate the connection between the geometry and performance of two-qubit parameterized quantum circuits. To quantify the performance, we consider a toy problem instance of the Variational Quantum Eigensolver (VQE) algorithm [45] which, despite its simplicity, shows how the scalar curvature of a parameterized quantum circuit for two qubits may help inform the effectiveness of an ansatz prior to execution of the algorithm. Namely, the task-at-hand is to estimate the ground state energy of molecular hydrogen at a particular bond length, employing each of the two-qubit circuits. We consider the two-qubit Hamiltonian for molecular hydrogen from an early VQE experimental study [46]:

$$H = \nu_1 \mathbb{I} + \nu_2 Z_1 + \nu_3 Z_2 + \nu_4 Z_1 Z_2 + \nu_5 X_1 X_2 + \nu_6 Y_1 Y_2, \quad (50)$$

where  $\boldsymbol{\nu} = \{\nu_1, \nu_2, \dots, \nu_6\}$  corresponds to Hamiltonian coefficients. For this problem, the ground state wave function is of the form:  $|\psi_g\rangle = \alpha(r_{H-H}) |01\rangle + \beta(r_{H-H}) |10\rangle$ , where the wave function coefficients  $\alpha$  and  $\beta$  are functions of  $r_{H-H}$ , the inter-atomic distance between the two Hydrogen atoms. At  $r_{H-H} = 3.19 \text{ \AA}$ , the ground state wave function is a highly entangled state of the form:

$$|\psi_g\rangle = \alpha |01\rangle + \beta |10\rangle, \quad (51)$$

for  $\alpha, \beta \in \mathbb{C}$  where  $|\alpha|^2 \approx 0.47$  and  $|\beta|^2 \approx 0.53$ . In this case, the ability to generate (highly) entangled states using a parameterized quantum circuit is important for representing the solution state. We executed wave function simulations for the VQE calculations. For the optimization in VQE, we consider the standard gradient descent optimizer as well as the Quantum Natural Gradient optimizer using the block-diagonal and diagonal matrix approximations for the



Fubini-Study metric tensor [16]. We fix the step size of each type of gradient descent to 0.05 and terminate each optimization run based on a convergence threshold of  $10^{-6}$ .

Using this toy problem, we discuss two main observations:<sup>1</sup>

1. Particular circuit structures lead to precise and/or accurate solutions. Using 50 independent optimizations, we show that circuits like LDCA initialize states at high concurrences and consistently and rapidly converge to the solution. Others have wider spread in the final energies and their accuracies. We argue that further insight about these results can be understood from the scalar curvature.
2. We investigate the role of the fiber space by considering the QGAN circuit as a case study. This circuit, as constructed, is unable to reach the ground state but is able to do so by appending local rotation gates. By calculating the curvature, we are also able to see how the Quantum Natural Gradient is able to take advantage of the geometry while the standard gradient descent does not.

#### 4.1 Curvature landscape and optimization

We first investigate the connection between the curvature landscape and VQE optimization. Fig. 6 tracks the energy error, concurrence, and scalar curvature of optimization paths using the four circuit blocks. For each circuit, 50 independent optimization trials were performed using random parameter initialization. For three of the four circuits, namely HEA, LDCA, and sHEA, using Quantum Natural Gradients enabled the optimizations to converge to errors below the chemical accuracy threshold ( $\approx 10^{-3}$  Ha) within 200 descent steps. Referring back to Fig. 5, these three circuits correspond to curvature landscapes with extensive regions of negative curvature or high concurrence. For sHEA,  $\theta_3$  and  $\theta_4$  were tuned over the course of each optimization such that highly entangled states became accessible (e.g.  $\theta_3 \approx \pi/2$ ).

In particular, optimization runs for LDCA started at higher values of concurrence on average and rapidly converged to accurate ground state energies within around 20 descent steps for QNG methods. In addition, the standard deviation over 50 runs is smaller than those corresponding to other circuits.

Additional insights into the optimization procedure can be gathered by looking at the curvature landscapes and also the scalar curvatures reached during the optimization process.

- (i) For the circuits that reach the ground state, we observe more regions of negative curvature; these regions in the Hilbert space are useful in accelerating the optimizations. Having more of these regions seems to correlate with performance of the optimization process. Consider the curvature landscape of LDCA, in which we observe that there are small, repeating hills of positive curvature that are surrounded by regions of negative curvature. This implies that LDCA may have a greater access to highly entangled states that are additionally constrained to a subspace spanned by  $\{|01\rangle, |10\rangle\}$ , both of which led to superior performance in the VQE algorithm instance. On the other hand, if one looks at the QGAN curvature landscapes, we see that for majority of parameter settings the curvature is positive.

---

<sup>1</sup>We repeat the VQE simulations and analyses for a bond length corresponding to an unentangled ground state and obtain similar results. These results are shown in Appendix D.

- (ii) Because the Quantum Natural Gradient is tailored for the wave function geometry, we see that it is able to reach these areas of negative curvature faster than standard gradient descent. Once it reaches these regions, the optimization process is sped up earlier on in the process allowing for a lower number of iterations. In all the cases looked at, the standard gradient descent takes longer to find these regions.
- (iii) Lastly, we briefly comment on the relative costs of Quantum Natural Gradient, which depend on the number of non-zero elements in the metric tensor or its approximations. From Eqs. 23- 26, we see that the diagonal or block-diagonal approximation of the metric tensor for LDCA captures all of the non-zero elements. This is in contrast to the other circuits which have at least two unique elements that are not captured by the approximations to the metric tensor. To make better use of QNG, these other circuits may require computations of elements that are not captured by the block-diagonal or diagonal approximations. For example, in Fig. 6d, we ran the VQE optimizations using the dense or full metric tensor for sHEA, a case in which the block-diagonal approximation does not capture many of the non-zero elements. Shown using purple lines, we observe that the optimization significantly improves in efficiency though at the cost of more function calls needed at each QNG descent step to compute the full metric tensor. On the other hand, LDCA only requires one non-zero element of the metric tensor to be computed at each QNG descent step. This appears to imply that certain circuits, by the way that they are parameterized, are better suited for QNG methods than others.

## 4.2 Non-trivial role of adding single-qubit gates

As observed in Fig. 6b, QGAN consistently fails to reach the ground state. QGAN, as constructed, is unable to reach states corresponding to high entanglement, or equivalently negative curvature. Often, in similar situations in which the circuit seems insufficient, one adds a set of gates or a circuit layer in the hopes of providing greater flexibility to reach the solution state. Thus, we tried augmenting the QGAN circuit with local rotations, adding  $RX$  then  $RZ$  single-qubit gates to each of the two qubits. This adds four new parameters to the circuit. Appending local rotations should not impact the concurrence and thus the curvature. It was surprising to observe, however, that adding local rotations greatly improved the optimization as shown in Fig. 7a. While the average value of concurrence at the start of the optimization using the augmented QGAN was lower than that corresponding to the original QGAN circuit, the updated QGAN circuit was able to estimate the ground state energy with high accuracy using quantum natural gradient methods. The puzzle is then the following: while the concurrence expression in Eq. (32) should not change for the augmented QGAN (since single qubit gates should not increase entanglement), the additional parameterized single-qubit gates somehow appear to have aided in guiding the other circuit parameters to values such that highly entangled states are accessible (shown in Fig. 7c). In other words, without increasing the entanglement in the circuit, we are able to turn an optimization procedure that does not get to an entangled state to one that does. We suspect that this is where another part of the geometry plays a part, namely the fibers. Locally, we have the geometry being  $S^4 \times S^3$  but this is not true globally, which means that how we move through the fibers has a non-trivial impact on the optimization process. This case also illustrates the role of geometry plays in the optimization process because the standard gradient descent still fails to find the

ground state, at least within a reasonable number of optimization steps. From the point of view of just the optimization process, we simply see the concurrence fails to reach the target value and the energy fails to reach low enough values, but as we look at what happens to the scalar curvature, we see the Quantum Natural Gradient descent finds pockets of high negative curvature allowing for the optimization path to more easily move. In other words, as we change parts of the available geometry, the Quantum Natural Gradient is able to better leverage that.

## 5 Concluding remarks

Often, investigations or uses of PQCs are limited to considering their inputs (parameters) and outputs (resulting quantum states or observables). Our work provided an investigation into the inner workings of two-qubit PQCs, which we argue are the simplest instances of PQCs and yet can still provide valuable insights for better understanding these circuits. We first defined PQCs from a geometric perspective. With the help of specifically two-qubit geometry, we explicitly worked out the intrinsic co-ordinates for four examples of two-qubit PQC blocks, at least for the base manifold, which is important for characterizing entanglement in the circuit. As a consequence of our ability to parameterize our circuits in terms of two-qubit geometry, we introduced a notion of local geometric expressibility, which describes how much of the two-qubit geometry can be explored by a two-qubit PQC.

In trying to understand the connection between the geometry of the projective Hilbert space and how it is used by the Quantum Natural Gradient in the optimization process, we used the Mannoury-Fubini-Study metric as a way to calculate the curvature of the base manifold,  $S^4$ . This provided a simple and remarkable connection between the curvature of the base manifold and the amount of entanglement in the ansatz.

With this connection we were able to establish a bridge between the amount of entanglement and quality of the optimization procedure to the geometry of the projective Hilbert Space. This allowed us to notice a correlation between the ability of the ansatz to find earlier on in the optimization process regions of high negative curvature and acceleration of the optimization process. We strongly suspect that this connection is not merely a correlation but represents a chain of causation, for this we give two pieces of evidence:

- (i) The performance of the QGAN circuit which could not find either the entangled ground state or the product ground state was enhanced by creating the augmented QGAN ansatz i.e. by adding single qubit rotations. The optimization process was able to reach regions of high negative curvature that were not accessible before and after which the circuit was able to reach the entangled ground state. Furthermore, in the numerical simulations for which the ground state is a product state and hence lies in a region of positive curvature (Appendix D), the augmented QGAN performed better by initially finding regions that are less positive than the original QGAN.
- (ii) Secondly, by inspecting the performance of standard gradient descent we see that for the circuits that did find the ground state, the regions of high negative curvature were found later in the optimization process than in the case for QNG. Indeed the behavior of the standard gradient descent for the augmented QGAN we believe is highly suggestive. The standard gradient descent in the augmented QGAN ansatz is not able to find these regions of high negative curvature and is not able to reach the ground state for a circuit

we know can find the ground state. On the other extreme end, we can look at the performance of sHEA and consider the performance of the dense QNG which finds region of negative curvature the earliest, one sees that this correlates with finding the ground state much more efficiently than the other approximations of QNG and the standard gradient descent.

In summary, the geometric perspective through the use of the scalar curvature has provided us with a tool to explain why, for example, why LDCA outperforms other circuits other than just explaining via tracking energies or entanglement. Overall, we show that the number of parameters of a given PQC does not necessarily correspond to high circuit flexibility or capability; it matters *how* the circuit is parameterized. In connection with this point, we observed how single qubit gates can significantly impact the circuit performance, i.e. do more than just provide extra parameters. The single qubit gates allow us to explore the fibers of the geometry in such ways as to find regions of negative or less positive curvature. This effect is ultimately possible because the geometry is not just a cartesian product of the base manifold and the fibers but some complicated twist. The geometric perspective also provides insights into how and why the QNG is better than the standard gradient descent; the QNG can find the regions of negative curvature, and if there are none, it opts for regions with less positive curvature early in the optimization process. We stress that the geometric perspective is necessary in order to understand the QNG since the QNG does not depend on the cost function.

While we provided a extensive study on two-qubit PQCs, there remain several puzzles or open questions, which we outline in the following subsection for future work.

## 5.1 Future directions and open questions

### Decomposition Problem

Refs. [16, 26] numerically showed the potential for Quantum Natural Gradients in optimizing parameterized quantum circuits in the context of VQE. A main challenge will be to scale up this method for larger and deeper quantum circuits. In an effort to formulate a hopefully simpler scenario that may be able to be scaled, we describe the following “decomposition problem”:

Suppose we have a four-qubit circuit, the first and second qubits are entangled using one of the PQC building blocks, e.g. LDCA block while the third and fourth qubits are also entangled using a two-qubit block. Overall, we have 2 two-qubit subsystems, each of which we know the metric tensor, call them  $g^{(12)}$  and  $g^{(34)}$ . However, suppose we add a static/non-parametric entangler (e.g. CNOT) to the second and third qubits. How is the metric tensor of the four-qubit system  $g^{(1234)}$  constructed, compared to structures of  $g^{(12)}$  and  $g^{(34)}$ ? The idea being, rather than analyzing the  $n$  qubit geometry from the ground up in order to understand PQCs for  $n$  qubits, can we use the simpler geometry of 2 qubits to bootstrap our way to understanding higher qubit number PQCs?

### Three-qubit case

For this work, we crucially relied on the fact that the geometry of 2 qubits could easily be understood since it was simply the Hopf fibration  $S^7$ . As a consequence of Hopf Invariant

One theorem [47], there is one more Hopf fibration that can be studied in great detail namely the Hopf fibration  $S^7 \rightarrow S^{15} \xrightarrow{\pi} S^8$ . In this case, the use of octonions should be helpful in parameterizing the geometry.

## Acknowledgements

All the authors would like to thank their colleagues at Zapata Computing for their feedback. The metric tensors and scalar curvatures for each PQC were computed using EinsteinPy [48] and Mathematica [49]. The VQE simulations were implemented using OpenFermion [50] and PennyLane [51]. DEK acknowledges funding support from the National Research Foundation, Singapore, through Grant NRF2021-QEP2-02-P03.

## A Gate definitions

We define the rotation operations in the following way:

$$R_P(\theta) = \exp(-i\theta P/2), \quad (52)$$

where  $P \in \{X, Y, Z\}$  for single-qubit gates and  $P \in \{XX, YY, ZZ, XY, YX\}$  for two-qubit gates.

We define the iSWAP<sup>†</sup> gate as:

$$\text{iSWAP}(\theta)^\dagger = e^{-i\theta(X \otimes X + Y \otimes Y)/2}, \quad (53)$$

and the CPHASE gate as:

$$\text{CPHASE}(\phi) = e^{-i\phi(I-Z) \otimes (I-Z)/4}. \quad (54)$$

## B Hopf base and fiber parameters for various circuit blocks

### B.1 Hardware-efficient ansatz (HEA)

The output state of the quantum hardware-efficient ansatz (HEA) depicted by Fig. 3(a) is

$$\begin{aligned} |\psi\rangle_{HEA} = & \cos(\theta_1) \cos(\theta_3) \cos(\theta_2 + \theta_4) - \sin(\theta_1) \sin(\theta_3) \sin(\theta_2 - \theta_4) |00\rangle \\ & + \sin(\theta_2 + \theta_4) \cos(\theta_1) \cos(\theta_3) - \sin(\theta_1) \sin(\theta_3) \cos(\theta_2 - \theta_4) |01\rangle \\ & + \sin(\theta_3) \cos(\theta_1) \cos(\theta_2 + \theta_4) + \sin(\theta_1) \sin(\theta_2 - \theta_4) \cos(\theta_3) |10\rangle \\ & + \sin(\theta_1) \cos(\theta_3) \cos(\theta_2 - \theta_4) + \sin(\theta_3) \sin(\theta_2 + \theta_4) \cos(\theta_1) |11\rangle. \end{aligned} \quad (55)$$

The angle parameters of the  $S^4$  Hopf base  $(\theta_A, \phi_A)$  and  $(\chi, \xi)$  and the equivalent Cartesian coordinates  $(x_0, x_1, x_2, x_3, x_4)$  are calculated to be

$$x_0 = \cos(2\theta_1) \cos(2\theta_3) - 8 \sin(\theta_1) \sin(\theta_2) \sin(\theta_3) \cos(\theta_1) \cos(\theta_2) \cos(\theta_3) \quad (56)$$

$$x_1 = \sin(2\theta_1) \sin(2\theta_2) \cos(2\theta_3) + \sin(2\theta_3) \cos(2\theta_1) \quad (57)$$

$$x_2 = 0 \quad (58)$$

$$x_3 = \sin(2\theta_1) \cos(2\theta_2) \quad (59)$$

$$x_4 = 0 \quad (60)$$

Table 3: Abbreviations and symbols

| Symbol/Abbreviation         | Description                            |
|-----------------------------|--|
| $G, G(\boldsymbol{\theta})$ | Fubini-Study metric tensor             |
| $\mathcal{R}$               | Scalar curvature                       |
| $C, C(\psi)$                | Concurrence                            |
| $\boldsymbol{\theta}$       | Parameter vector                       |
| $\theta_j$                  | $j$ -th parameter                      |
| PQC                         | Parameterized quantum circuit          |
| HEA                         | Hardware-efficient ansatz              |
| LDCA                        | Low-depth circuit ansatz [27]          |
| QGAN                        | Quantum generative adversarial network |

and

$$\theta_A = \arccos(\cos(2\theta_1)\cos(2\theta_3) - 8\sin(\theta_1)\sin(\theta_2)\sin(\theta_3)\cos(\theta_1)\cos(\theta_2)\cos(\theta_3)) \quad (61)$$

$$\phi_A = \arccos\left(\frac{\sin(2\theta_1)\sin(2\theta_2)\cos(2\theta_3) + \sin(2\theta_3)\cos(2\theta_1)}{\sqrt{1 - (\cos(2\theta_1)\cos(2\theta_3) - 8\sin(\theta_1)\sin(\theta_2)\sin(\theta_3)\cos(\theta_1)\cos(\theta_2)\cos(\theta_3))^2}}\right) \quad (62)$$

$$\chi = \frac{\pi}{2} \quad (63)$$

$$\xi = \frac{\pi}{2}. \quad (64)$$

The quaternion  $q_{\pm} = \langle c_{\pm} | \psi_H \rangle$  is calculated to be

$$q_{\pm} = \frac{1}{2\lambda_2}(e_{\pm} + jf_{\pm}) \quad (65)$$

where

$$e_{\pm} = (\cos(\theta_1)\cos(\theta_3)\cos(\theta_2 + \theta_4) - \sin(\theta_1)\sin(\theta_3)\sin(\theta_2 - \theta_4))(\lambda_3\gamma_{\mp} + \lambda_1\gamma_{\pm}) \quad (66)$$

$$f_{\pm} = (\sin(\theta_1)\sin(\theta_3)\cos(\theta_2 - \theta_4) - \sin(\theta_2 + \theta_4)\cos(\theta_1)\cos(\theta_3))(-\lambda_3\gamma_{\mp} - \lambda_1\gamma_{\pm}) \quad (67)$$

where

$$\lambda_1 = \sqrt{2\sin(4\theta_1)\sin(2\theta_2)\sin(4\theta_3) + 4\sin^2(2\theta_1)(\cos^2(2\theta_2) + \sin^2(2\theta_2)\cos^2(2\theta_3)) + 4\sin^2(2\theta_3)\cos^2(2\theta_1)}$$

$$\lambda_2 = \sqrt{\sin(4\theta_1)\sin(2\theta_2)\sin(4\theta_3) + 2\sin^2(2\theta_1)(\cos^2(2\theta_2) + \sin^2(2\theta_2)\cos^2(2\theta_3)) + 2\sin^2(2\theta_3)\cos^2(2\theta_1)}$$

$$\lambda_3 = 2\sin(2\theta_1)\sin(2\theta_2)\sin(2\theta_3) - 2\cos(2\theta_1)\cos(2\theta_3) + 2$$

$$\gamma_{\pm} = \sqrt{1 \pm \sqrt{1 - \frac{1}{4}\sin^2(2\theta_1)\cos^2(2\theta_2) - \frac{1}{4}(\sin(2\theta_1)\sin(2\theta_2)\cos(2\theta_3) + \sin(2\theta_3)\cos(2\theta_1))^2}}. \quad (68)$$

## B.2 Low-depth circuit ansatz (LDCA)

The output state of the quantum hardware-efficient ansatz (HEA) depicted by Fig. 3(a) is

$$\begin{aligned} |\Psi\rangle_{LDCA} = & e^{-\frac{1}{2}i(\theta_1-\theta_2-\theta_4)} (\cos(\theta_3) \cos(\theta_5) - i \sin(\theta_3) \sin(\theta_5)) |01\rangle \\ & - e^{-\frac{1}{2}i(\theta_1-\theta_2-\theta_4)} (\sin(\theta_5) \cos(\theta_3) + i \sin(\theta_3) \cos(\theta_5)) |10\rangle. \end{aligned} \quad (69)$$

The angle parameters of the  $S^4$  Hopf base  $(\theta_A, \phi_A)$  and  $(\chi, \xi)$  and the equivalent Cartesian coordinates  $(x_0, x_1, x_2, x_3, x_4)$  are calculated to be

$$x_0 = \cos(2\theta_3) \cos(2\theta_5) \quad (70)$$

$$x_1 = 0 \quad (71)$$

$$x_2 = \sin(\theta_1 - \theta_2 - \theta_4) \sin(2\theta_5) - \sin(2\theta_3) \cos(\theta_1 - \theta_2 - \theta_4) \cos(2\theta_5) \quad (72)$$

$$x_3 = \sin(2\theta_3) \sin(\theta_1 - \theta_2 - \theta_4) \cos(2\theta_5) + \sin(2\theta_5) \cos(\theta_1 - \theta_2 - \theta_4) \quad (73)$$

$$x_4 = 0 \quad (74)$$

and

$$\theta_A = \arccos(\cos(2\theta_3) \cos(2\theta_5)) \quad (75)$$

$$\phi_A = \frac{\pi}{2} \quad (76)$$

$$\chi = \frac{\pi}{2} \quad (77)$$

$$\xi = -\arctan\left(\frac{\sin(2\theta_3) \cos(2\theta_5) + \sin(2\theta_5) \cot(\theta_1 - \theta_2 - \theta_4)}{\sin(2\theta_3) \cos(2\theta_5) \cot(\theta_1 - \theta_2 - \theta_4) - \sin(2\theta_5)}\right). \quad (78)$$

The quaternion  $q_{\pm} = \langle c_{\pm} | \psi_H \rangle$  is calculated to be

$$q_{\pm} = \frac{1}{\sqrt{\mu_3}} (jg_{\pm} + kh_{\pm}), \quad (79)$$

where

$$\begin{aligned} g_{\pm} = & \left( \cos(\theta_3) \cos\left(\frac{1}{2}(\theta_1 - \theta_2 - \theta_4)\right) \cos(\theta_5) - \sin(\theta_3) \sin\left(\frac{1}{2}(\theta_1 - \theta_2 - \theta_4)\right) \sin(\theta_5) \right) (\mu_1 \gamma_{\pm} + \mu_2 \gamma_{\mp}), \\ h_{\pm} = & \left( \sin\left(\frac{1}{2}(\theta_1 - \theta_2 - \theta_4)\right) \cos(\theta_3) \cos(\theta_5) + \sin(\theta_3) \sin(\theta_5) \cos\left(\frac{1}{2}(\theta_1 - \theta_2 - \theta_4)\right) \right) (-\mu_1 \gamma_{\pm} - \mu_2 \gamma_{\mp}). \end{aligned} \quad (80)$$

where

$$\begin{aligned} \mu_1 = & \sqrt{-2 \cos(4\theta_3) \cos^2(2\theta_5) - \cos(4\theta_5) + 3}, \\ \mu_2 = & 2 - 2 \cos(2\theta_3) \cos(2\theta_5), \\ \mu_3 = & 2 \left( -2 \cos(4\theta_3) \cos^2(2\theta_5) - \cos(4\theta_5) + 3 \right), \\ \gamma_{\pm} = & \frac{1}{2} \sqrt{4 \pm \sqrt{\cos(4\theta_5) + \cos(4\theta_3) (\cos(4\theta_5) + 1) + 13}}. \end{aligned} \quad (81)$$

### B.3 Quantum generative adversarial network (QGAN) ansatz

The output state of the quantum generative adversarial network (QGAN) depicted by Fig. 3(b) is

$$\begin{aligned}
 |\psi\rangle_{QGAN} = & e^{-\frac{1}{2}i(\theta_3+\theta_4+\theta_5)} \cos\left(\frac{\theta_1}{2}\right) \cos\left(\frac{\theta_2}{2}\right) |00\rangle - ie^{-\frac{1}{2}i(\theta_3-\theta_4-\theta_5)} \sin\left(\frac{\theta_2}{2}\right) \cos\left(\frac{\theta_1}{2}\right) |01\rangle \\
 & - ie^{\frac{1}{2}i(\theta_3-\theta_4+\theta_5)} \sin\left(\frac{\theta_1}{2}\right) \cos\left(\frac{\theta_2}{2}\right) |10\rangle - e^{\frac{1}{2}i(\theta_3+\theta_4-\theta_5)} \sin\left(\frac{\theta_1}{2}\right) \sin\left(\frac{\theta_2}{2}\right) |11\rangle
 \end{aligned} \tag{82}$$

The angle parameters of the  $S^4$  Hopf base  $(\theta_A, \phi_A)$  and  $(\chi, \xi)$  and the equivalent Cartesian coordinates  $(x_0, x_1, x_2, x_3, x_4)$  are calculated to be

$$x_0 = \cos(\theta_1) \tag{83}$$

$$x_1 = \sin(\theta_1) [\sin(\theta_3) \cos(\theta_5) + \sin(\theta_5) \cos(\theta_2) \cos(\theta_3)] \tag{84}$$

$$x_2 = -\sin(\theta_1) \sin(\theta_2) \sin(\theta_5) \tag{85}$$

$$x_3 = 0 \tag{86}$$

$$x_4 = \sin(\theta_1) [\sin(\theta_3) \sin(\theta_5) \cos(\theta_2) - \cos(\theta_3) \cos(\theta_5)] \tag{87}$$

and

$$\theta_A = \theta_1 \tag{88}$$

$$\phi_A = \arccos[\sin(\theta_3) \cos(\theta_5) + \sin(\theta_5) \cos(\theta_2) \cos(\theta_3)] \tag{89}$$

$$\chi = \arccos\left(\frac{\sin(\theta_3) \sin(\theta_5) \cos(\theta_2) - \cos(\theta_3) \cos(\theta_5)}{\sqrt{1 - (\sin(\theta_3) \cos(\theta_5) + \sin(\theta_5) \cos(\theta_2) \cos(\theta_3))^2}}\right) \tag{90}$$

$$\xi = 0 \tag{91}$$

The quaternion  $q_{\pm} = \langle c_{\pm} | \psi_H \rangle$  is calculated to be

$$q_{\pm} = \frac{1}{\sqrt{2}} \left( \sin\left(\frac{\theta_1}{2}\right) \gamma_{\mp} + \cos\left(\frac{\theta_1}{2}\right) \gamma_{\pm} \right) (a_{\pm} + ib_{\pm} + jc_{\pm} + kd_{\pm}) \tag{92}$$

where

$$a_{\pm} = \cos\left(\frac{\theta_2}{2}\right) \cos\left(\frac{\theta_3 + \theta_4 + \theta_5}{2}\right) \tag{93}$$

$$b_{\pm} = -\cos\left(\frac{\theta_2}{2}\right) \sin\left(\frac{\theta_3 + \theta_4 + \theta_5}{2}\right) \tag{94}$$

$$c_{\pm} = -\sin\left(\frac{\theta_2}{2}\right) \sin\left(\frac{\theta_3 - \theta_4 - \theta_5}{2}\right) \tag{95}$$

$$d_{\pm} = -\sin\left(\frac{\theta_2}{2}\right) \cos\left(\frac{\theta_3 - \theta_4 - \theta_5}{2}\right) \tag{96}$$

where

$$\gamma_{\pm} = \frac{1}{2} \sqrt{4 \pm \sqrt{14 + 2 \cos(2\theta_1)}}. \tag{97}$$

Where for example  $\langle \sigma_2^z \rangle_2$  is the expectation value of  $\sigma_2^z$  after the second layer of layer of the circuit.



#### B.4 Sycamore hardware-efficient ansatz (sHEA)

The output state of the Sycamore hardware-efficient ansatz (sHEA) depicted by Fig. 3(d) is

$$\begin{aligned}
|\psi\rangle_{sHEA} = & -ie^{-\frac{1}{2}i(\theta_5+\theta_6)} \sin\left(\frac{\theta_2}{2}\right) \cos\left(\frac{\theta_1}{2}\right) |00\rangle \\
& + e^{-\frac{1}{2}i(\theta_5-\theta_6)} \left( \cos\left(\frac{\theta_1}{2}\right) \cos\left(\frac{\theta_2}{2}\right) \cos\left(\frac{\theta_3}{2}\right) - i \sin\left(\frac{\theta_1}{2}\right) \sin\left(\frac{\theta_2}{2}\right) \sin\left(\frac{\theta_3}{2}\right) \right) |01\rangle \\
& + e^{\frac{1}{2}i(\theta_5-\theta_6)} \left( -\sin\left(\frac{\theta_1}{2}\right) \sin\left(\frac{\theta_2}{2}\right) \cos\left(\frac{\theta_3}{2}\right) + i \sin\left(\frac{\theta_3}{2}\right) \cos\left(\frac{\theta_1}{2}\right) \cos\left(\frac{\theta_2}{2}\right) \right) |10\rangle \\
& - ie^{-\frac{1}{4}i(\theta_4-2(\theta_5+\theta_6))} \sin\left(\frac{\theta_1}{2}\right) \cos\left(\frac{\theta_2}{2}\right) |11\rangle
\end{aligned} \tag{98}$$

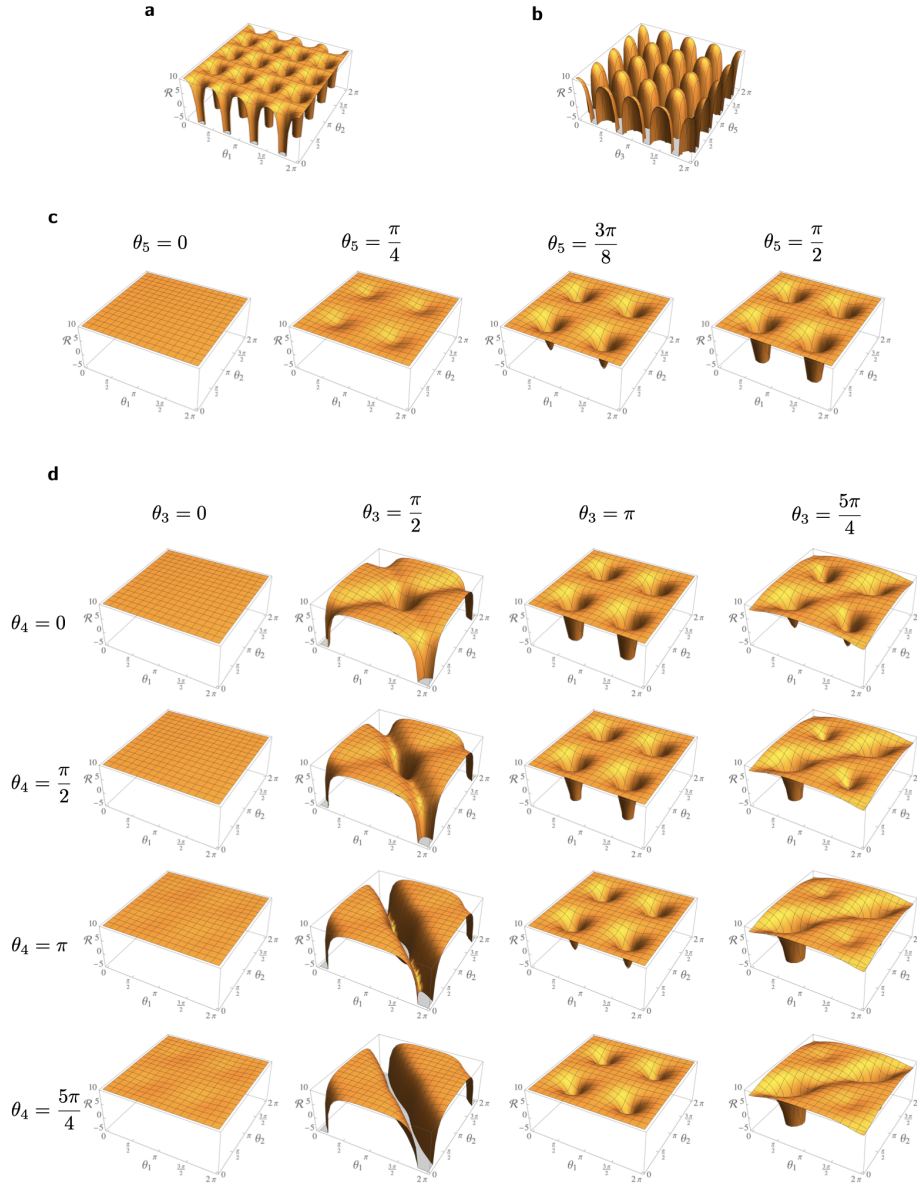


Figure 5: Scalar curvature “landscapes” of the four circuit blocks: (a) HEA, (b) LDCA, (c) QGAN, and (d) sHEA. In the case of (a) and (b), the scalar curvature (denoted using  $\mathcal{R}$ ) is a function of only two parameters:  $\theta_1$  and  $\theta_2$  for HEA and  $\theta_3$  and  $\theta_5$  for LDCA. For the other two circuit blocks, in which  $\mathcal{R}$  is a function of more than two circuit parameters,  $\mathcal{R}$  is plotted as a function of  $\theta_1$  and  $\theta_2$ , while scanning specific values of the remaining parameters to produce snapshots of the landscapes. We limit the range of  $\mathcal{R}$  to  $[-5, 10]$  for ease of visualization.

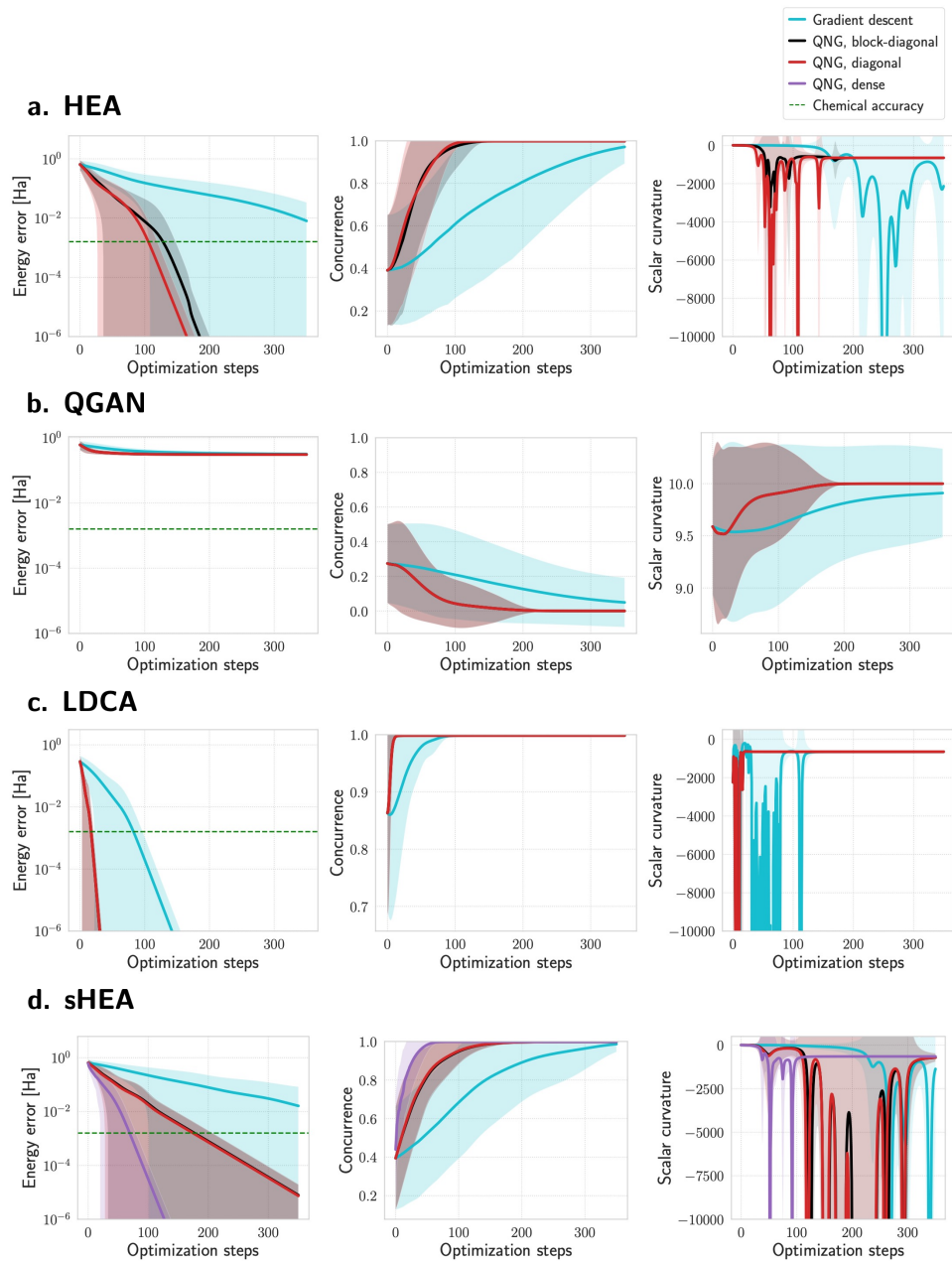


Figure 6: VQE results over 50 random initial points using the four two-qubit ansatz for estimating the ground state energy of molecular hydrogen at  $R_{\text{H-H}} = 3.19 \text{ \AA}$ . Ground state at this bond length is highly entangled. Solid lines show the average values of energy error, concurrence, and scalar curvature.

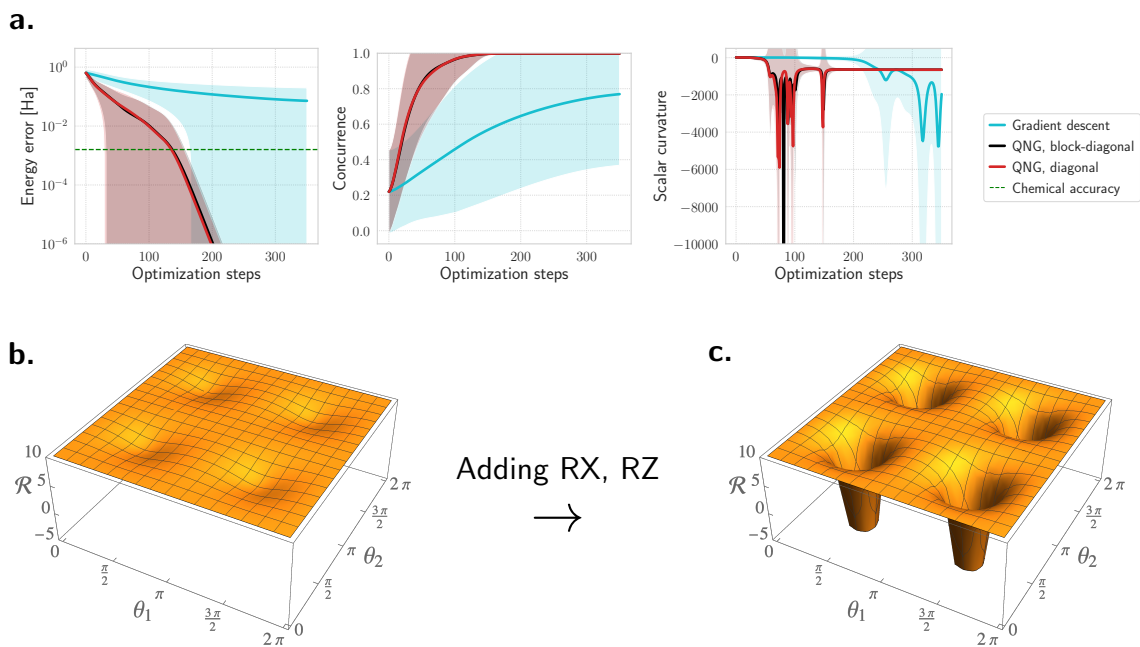


Figure 7: (a) Energy error, concurrence, and curvature tracked over 50 independent optimization trials of QGAN augmented with local rotations (RX and RZ applied to each qubit). (b) Curvature landscape corresponding to final/optimized parameter value of  $\theta_5$  of a particular optimization run using the original QGAN circuit. (c) Curvature landscape corresponding to final parameter value of  $\theta_5$  of a particular optimization run using *augmented* QGAN circuit.

The angle parameters of the  $S^4$  Hopf base  $(\theta_A, \phi_A)$  and  $(\chi, \xi)$  and the equivalent Cartesian coordinates  $(x_0, x_1, x_2, x_3, x_4)$  are calculated to be

$$x_0 = \frac{1}{2} (\cos(\theta_2) (\cos(\theta_3) - 1) + \cos(\theta_1) (\cos(\theta_3) + 1)) \quad (99)$$

$$x_1 = \sin(\theta_2) \sin\left(\frac{\theta_3}{2}\right) \left( \sin^2\left(\frac{\theta_1}{2}\right) \cos\left(\frac{\theta_4}{4} - \theta_5\right) - \cos^2\left(\frac{\theta_1}{2}\right) \cos(\theta_5) \right) \\ + \sin(\theta_1) \cos\left(\frac{\theta_3}{2}\right) \left( \sin^2\left(\frac{\theta_2}{2}\right) \sin(\theta_5) - \sin\left(\frac{\theta_4}{4} - \theta_5\right) \cos^2\left(\frac{\theta_2}{2}\right) \right) \quad (100)$$

$$x_2 = \frac{1}{2} \left( \sin(\theta_3) - \sin(\theta_1) \sin(\theta_2) \sin\left(\frac{\theta_4}{4}\right) + \sin(\theta_3) \cos(\theta_1) \cos(\theta_2) \right) \quad (101)$$

$$x_3 = \frac{1}{2} \sin(\theta_1) \sin(\theta_2) \left( \cos(\theta_3) - \cos\left(\frac{\theta_4}{4}\right) \right) \quad (102)$$

$$x_4 = \frac{1}{2} \left( \sin(\theta_1) \cos\left(\frac{\theta_3}{2}\right) \left( - \left( -\cos(\theta_2) + \left( \cos(\theta_2) + 1 \right) \cos\left(\frac{\theta_4}{4}\right) + 1 \right) \cos(\theta_5) \right. \right. \\ \left. \left. - \sin\left(\frac{\theta_4}{4}\right) \sin(\theta_5) \left( \cos(\theta_2) + 1 \right) \right) - \sin(\theta_2) \sin\left(\frac{\theta_3}{2}\right) \left( 2 \sin\left(\frac{\theta_4}{4}\right) \sin^2\left(\frac{\theta_1}{2}\right) \cos(\theta_5) \right. \right. \\ \left. \left. + \sin(\theta_5) \left( \cos(\theta_1) + \left( \cos(\theta_1) - 1 \right) \cos\left(\frac{\theta_4}{4}\right) + 1 \right) \right) \right) \quad (103)$$

and

$$\theta_A = \arccos \left[ \frac{1}{2} (\cos(\theta_2) (\cos(\theta_3) - 1) + \cos(\theta_1) (\cos(\theta_3) + 1)) \right] \quad (104)$$

$$\phi_A = \arccos \left[ \frac{\sin(\theta_2) \sin\left(\frac{\theta_3}{2}\right) \left( \sin^2\left(\frac{\theta_1}{2}\right) \cos\left(\frac{\theta_4}{4} - \theta_5\right) - \cos^2\left(\frac{\theta_1}{2}\right) \cos(\theta_5) \right) + \sin(\theta_1) \cos\left(\frac{\theta_3}{2}\right) \left( \sin^2\left(\frac{\theta_2}{2}\right) \sin(\theta_5) - \sin\left(\frac{\theta_4}{4} - \theta_5\right) \cos^2\left(\frac{\theta_2}{2}\right) \right)}{\sqrt{1 - \frac{1}{4} (\cos(\theta_2) (\cos(\theta_3) - 1) + \cos(\theta_1) (\cos(\theta_3) + 1))^2}} \right] \quad (105)$$

$$\chi = \arccos \left[ \frac{\sin(\theta_1) \cos\left(\frac{\theta_3}{2}\right) \left( -(-\cos(\theta_2) + (\cos(\theta_2) + 1) \cos\left(\frac{\theta_4}{4}\right) + 1) \cos(\theta_5) - \sin\left(\frac{\theta_4}{4}\right) \sin(\theta_5) (\cos(\theta_2) + 1) \right) - \sin(\theta_2) \sin\left(\frac{\theta_3}{2}\right) \left( 2 \sin\left(\frac{\theta_4}{4}\right) \sin^2\left(\frac{\theta_1}{2}\right) \cos(\theta_5) + \sin(\theta_5) (\cos(\theta_1) + (\cos(\theta_1) - 1) \cos\left(\frac{\theta_4}{4}\right) + 1) \right)}{2 \left( \left( 1 - \frac{1}{4} (\cos(\theta_2) (\cos(\theta_3) - 1) + \cos(\theta_1) (\cos(\theta_3) + 1))^2 \right) \times \left( 1 - \frac{(\sin(\theta_2) \sin\left(\frac{\theta_3}{2}\right) (\sin^2\left(\frac{\theta_1}{2}\right) \cos\left(\frac{\theta_4}{4} - \theta_5\right) - \cos^2\left(\frac{\theta_1}{2}\right) \cos(\theta_5)) + \sin(\theta_1) \cos\left(\frac{\theta_3}{2}\right) (\sin^2\left(\frac{\theta_2}{2}\right) \sin(\theta_5) - \sin\left(\frac{\theta_4}{4} - \theta_5\right) \cos^2\left(\frac{\theta_2}{2}\right)) \right)^2}{1 - \frac{1}{4} (\cos(\theta_2) (\cos(\theta_3) - 1) + \cos(\theta_1) (\cos(\theta_3) + 1))^2} \right)^{-1}} \right] \quad (106)$$

$$\xi = \arctan \left[ \frac{\sin(\theta_1) \sin(\theta_2) \left( \cos(\theta_3) - \cos\left(\frac{\theta_4}{4}\right) \right)}{\sin(\theta_3) - \sin(\theta_1) \sin(\theta_2) \sin\left(\frac{\theta_4}{4}\right) + \sin(\theta_3) \cos(\theta_1) \cos(\theta_2)} \right] \quad (107)$$

The quaternion  $q_{\pm} = \langle c_{\pm} | \psi_H \rangle$  is calculated to be

$$q_{\pm} = m_{\pm} + im_{\pm} + jr_{\pm} + ks_{\pm} \quad (108)$$

where

$$m_{\pm} = \frac{\sin\left(\frac{\theta_2}{2}\right) \sin\left(\frac{1}{2}(\theta_5 + \theta_6)\right) \cos\left(\frac{\theta_1}{2}\right) \left( \frac{2(\cos(\theta_2)(\cos(\theta_3)-1) + \cos(\theta_1)(\cos(\theta_3)+1) - 2)\gamma_{\mp}}{\sqrt{-4\cos(2\theta_1)\cos^4\left(\frac{\theta_3}{2}\right) - \cos(2\theta_3) - 4\sin^4\left(\frac{\theta_3}{2}\right)\cos(2\theta_2) + 4\sin^2(\theta_3)\cos(\theta_1)\cos(\theta_2) + 5}} - \sqrt{2}\gamma_{\pm} \right)}{\sqrt{2}} \quad (109)$$

$$r_{\pm} = \frac{\left(\cos\left(\frac{1}{2}(\theta_1 + \theta_2 - \theta_3 + \theta_5 - \theta_6)\right) + \cos\left(\frac{1}{2}(\theta_1 - \theta_2 + \theta_3 + \theta_5 - \theta_6)\right) + 2\cos\left(\frac{1}{2}(\theta_2 + \theta_3)\right)\cos\left(\frac{1}{2}(\theta_1 - \theta_5 + \theta_6)\right)\right) \left(-B_{\mp} + \sqrt{2}\gamma_{\pm}C\right)}{A} \quad (110)$$

$$s_{\pm} = \frac{\left(\sin\left(\frac{1}{2}(\theta_1 + \theta_2 - \theta_3 + \theta_5 - \theta_6)\right) + \sin\left(\frac{1}{2}(\theta_1 - \theta_2 + \theta_3 + \theta_5 - \theta_6)\right) - 2\sin\left(\frac{1}{2}(\theta_1 - \theta_5 + \theta_6)\right)\cos\left(\frac{1}{2}(\theta_2 + \theta_3)\right)\right) \left(B_{\mp} - \sqrt{2}\gamma_{\pm}C\right)}{A} \quad (111)$$

where

$$A = 4\sqrt{2}\sqrt{-4\cos(2\theta_1)\cos^4\left(\frac{\theta_3}{2}\right) - \cos(2\theta_3) - 4\sin^4\left(\frac{\theta_3}{2}\right)\cos(2\theta_2) + 4\sin^2(\theta_3)\cos(\theta_1)\cos(\theta_2) + 5}, \quad (112)$$

$$B_{\mp} = 2(\cos(\theta_2)(\cos(\theta_3) - 1) + \cos(\theta_1)(\cos(\theta_3) + 1) - 2)\gamma_{\mp}, \quad (113)$$

$$C = \sqrt{-4\cos(2\theta_1)\cos^4\left(\frac{\theta_3}{2}\right) - \cos(2\theta_3) - 4\sin^4\left(\frac{\theta_3}{2}\right)\cos(2\theta_2) + 4\sin^2(\theta_3)\cos(\theta_1)\cos(\theta_2) + 5}, \quad (114)$$

and

$$\gamma_{\pm} = \frac{1}{2}\sqrt{\frac{\pm\sqrt{4\cos(2\theta_1)\cos^4\left(\frac{\theta_3}{2}\right) + \cos(2\theta_3) + 4\sin^4\left(\frac{\theta_3}{2}\right)\cos(2\theta_2) - 4\sin^2(\theta_3)\cos(\theta_1)\cos(\theta_2) + 27}}{\sqrt{2}} + 4}. \quad (115)$$

## C Geometric origins of the quantum natural gradient

Considering the geometric point of view, how might one arrive at the concept of the quantum natural gradient? The idea is to simply find out how the gradient operation would change in different curved geometries. For the quantum mechanics, we would need to consider the geometry of  $\mathbb{P}(H)$ .

Using the correspondence between a vector space  $V$  and its dual  $V^*$  for finite dimensional spaces, we can pair the co-vector  $df$  with the vector  $\nabla f$  i.e

$$df(w) = \langle \nabla f, w \rangle \quad (116)$$

where  $\langle \cdot, \cdot \rangle$  is a chosen bilinear form. Now consider the following calculation:

$$\langle v, w \rangle = \sum_{ij} v^i \langle e_i, e_j \rangle w^j = \nu(w),$$

where  $v, w \in V$  and  $\nu \in V^*$  and  $\{e_i\}$  are basis for  $V$ . We can expand  $\nu$  in a basis  $\{\sigma^i\}$  for  $V^*$  so that  $\nu = \sum_i v_i \sigma^i$  with  $v_i = \nu(e_i) = \langle v, e_j \rangle$ . From the fact  $v^i = \sum_j g^{ij} v_j = \sum_i g^{ij} \langle v, e_j \rangle$ , we have

$$v = \sum_j v^j e_j = \sum_i \left( \sum_j g^{ij} \langle v, e_j \rangle \right) e_i. \quad (117)$$

By the discussion above, we have that

$$\nabla f = \sum_i (\nabla f)^i e_i = \sum_i \left( \sum_j g^{ij} \langle \nabla f, e_j \rangle \right) e_i. \quad (118)$$

Considering (116) we have that  $df(w) = \langle \nabla f, w \rangle = w(f) = \sum_j w^j \frac{\partial f}{\partial x^j}$  so that combining knowledge from (117) and (118) we have  $\langle \nabla, e_j \rangle = e_j(f) = \frac{\partial f}{\partial x^j}$  and thus arrive at

$$(\nabla f)^i = \sum_j g^{ij} \frac{\partial f}{\partial x^j}. \quad (119)$$

Now in Euclidean geometry  $df$  and  $\nabla f$  have the same co-ordinates but in general they do not. For  $\mathbb{P}(H)$  we pick the metric to be Fubini-Study metric and thus arrive at the following update rule for the parameters:

$$\theta^{t+1} = \theta^t - \eta g^{-1} \frac{\partial \mathcal{L}(\theta)}{\partial \theta}. \quad (120)$$

## D Further insights into VQE performance through scalar curvature: unentangled ground state

In this section, we consider another regime of the same VQE problem from Sec. 4. At  $r_{H-H} = 0.2\text{\AA}$ , the ground state wave function is a product state:

$$|\psi_g\rangle \approx \beta |10\rangle, \quad (121)$$



with  $|\beta|^2 \approx 1$ . We repeat the simulation procedure, running 50 independent VQE optimizations for each of the four two-qubit circuits, as shown in Fig. 8. We observe similar results for  $r_{H-H} = 0.2\text{\AA}$  as those for  $r_{H-H} = 3.19\text{\AA}$ ; using LDCA, the optimizations are both precise and accurate. While initial concurrence values start at high values, they rapidly decrease to near 0. This corresponds to starting in a negative curvature region but quickly moving up to a hill of positive curvature. QGAN, in its original structure, is again insufficient for converging to the ground state energy. However, with added local rotations (Fig. 8e), the optimizations reach sufficient accuracy.

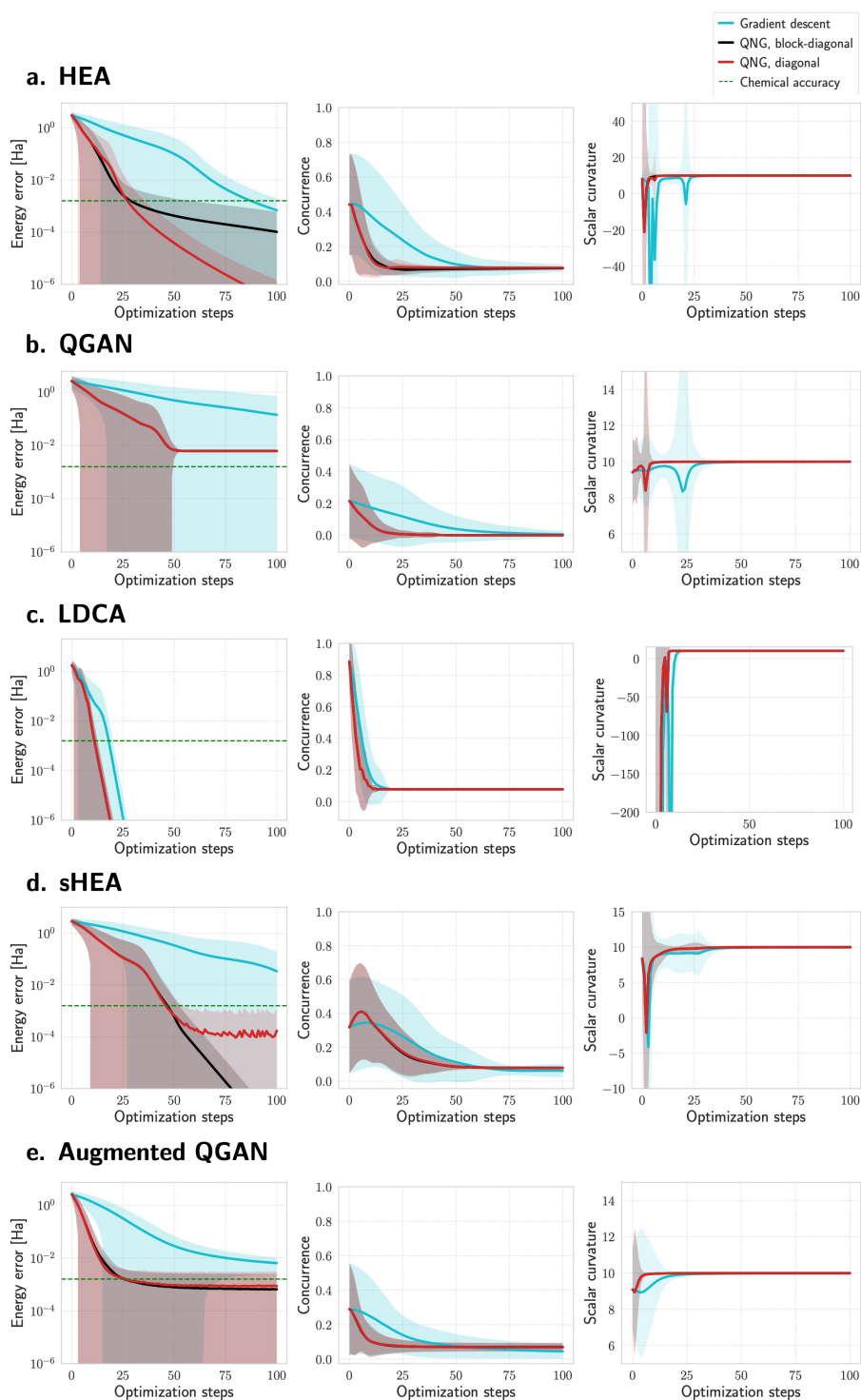


Figure 8: VQE results over 50 trials using the four two-qubit ansatzes for molecular hydrogen where  $R_{H-H} = 0.2 \text{ \AA}$ . Each ansatz is optimized using gradient descent and quantum natural gradient (QNG) using the block-diagonal and diagonal approximations to the metric tensor. Lines in each plot indicate the average quantity, and shaded regions indicate quantities within one standard deviation.

## References

- [1] Marco Cerezo, Andrew Arrasmith, Ryan Babbush, Simon C Benjamin, Suguru Endo, Keisuke Fujii, Jarrod R McClean, Kosuke Mitarai, Xiao Yuan, Lukasz Cincio, et al. Variational quantum algorithms. *Nature Reviews Physics*, 3:625–644, 2021. doi:10.1038/s42254-021-00348-9. [p. 1]
- [2] Kishor Bharti, Alba Cervera-Lierta, Thi Ha Kyaw, Tobias Haug, Sumner Alperin-Lea, Abhinav Anand, Matthias Degroote, Hermanni Heimonen, Jakob S. Kottmann, Tim Menke, Wai-Keong Mok, Sukin Sim, Leong-Chuan Kwek, and Alán Aspuru-Guzik. Noisy intermediate-scale quantum algorithms. *Rev. Mod. Phys.*, 94:015004, Feb 2022. doi:10.1103/RevModPhys.94.015004. [p. 1]
- [3] M.-H. Yung, J. Casanova, A. Mezzacapo, J. McClean, L. Lamata, A. Aspuru-Guzik, and E. Solano. From transistor to trapped-ion computers for quantum chemistry. *Sci. Rep.*, 4:3589, may 2015. doi:10.1038/srep03589. [p. 1]
- [4] Yudong Cao, Jonathan Romero, Jonathan P. Olson, Matthias Degroote, Peter D. Johnson, Mária Kieferová, Ian D. Kivlichan, Tim Menke, Borja Peropadre, Nicolas P.D. Sawaya, Sukin Sim, Libor Veis, and Alán Aspuru-Guzik. Quantum Chemistry in the Age of Quantum Computing. *Chemical Reviews*, 119(19):10856–10915, oct 2019. doi:10.1021/acs.chemrev.8b00803. [p. 1]
- [5] Abhinav Anand, Philipp Schleich, Sumner Alperin-Lea, Phillip W. K. Jensen, Sukin Sim, Manuel Díaz-Tinoco, Jakob S. Kottmann, Matthias Degroote, Artur F. Izmaylov, and Alán Aspuru-Guzik. A quantum computing view on unitary coupled cluster theory. *Chem. Soc. Rev.*, 51:1659–1684, March 2022. doi:10.1039/D1CS00932J. [p. 1]
- [6] Vojtěch Havlíček, Antonio D. Córcoles, Kristan Temme, Aram W. Harrow, Abhinav Kandala, Jerry M. Chow, and Jay M. Gambetta. Supervised learning with quantum-enhanced feature spaces. *Nature*, 567:209–212, mar 2019. doi:10.1038/s41586-019-0980-2. [p. 1]
- [7] Abhinav Kandala, Antonio Mezzacapo, Kristan Temme, Maika Takita, Markus Brink, Jerry M. Chow, and Jay M. Gambetta. Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets. *Nature*, 549:242–246, Sept 2017. doi:10.1038/nature23879. [pp. 1, 6]
- [8] Stig Elkjær Rasmussen, Niels Jakob Søe Loft, Thomas Bækkegaard, Michael Kues, and Nikolaj Thomas Zinner. Reducing the Amount of Single-Qubit Rotations in VQE and Related Algorithms. *Advanced Quantum Technologies*, 3(12):2000063, dec 2020. doi:10.1002/qute.202000063. [p. 2]
- [9] Sukin Sim, Jonathan Romero, Jérôme F. Gonthier, and Alexander A. Kunitsa. Adaptive pruning-based optimization of parameterized quantum circuits. *Quantum Science and Technology*, 6(2):025019, apr 2021. doi:10.1088/2058-9565/abe107. [p. 2]
- [10] Lena Funcke, Tobias Hartung, Karl Jansen, Stefan Kühn, and Paolo Stornati. Dimensional Expressivity Analysis of Parametric Quantum Circuits. *Quantum*, 5:422, March 2021. doi:10.22331/q-2021-03-29-422. [p. 2]

- [11] Jarrod R. McClean, Sergio Boixo, Vadim N. Smelyanskiy, Ryan Babbush, and Hartmut Neven. Barren plateaus in quantum neural network training landscapes. *Nat. Commun.*, 9:4812, 2018. doi:10.1038/s41467-018-07090-4. [pp. 2, 3]
- [12] Andrew Arrasmith, Zoë Holmes, M Cerezo, and Patrick J Coles. Equivalence of quantum barren plateaus to cost concentration and narrow gorges. *Quantum Science and Technology*, 7(4):045015, aug 2022. doi:10.1088/2058-9565/ac7d06. [p. 2]
- [13] Sukin Sim, Peter D. Johnson, and Alán Aspuru-Guzik. Expressibility and entangling capability of parameterized quantum circuits for hybrid quantum-classical algorithms. *Advanced Quantum Technologies*, 2(12):1900070, 2019. doi:10.1002/qute.201900070. [pp. 2, 9]
- [14] Thomas Hubregtsen, Josef Pichlmeier, Patrick Stecher, and Koen Bertels. Evaluation of parameterized quantum circuits: on the relation between classification accuracy, expressibility, and entangling capability. *Quantum Machine Intelligence*, 3:9, 2021. doi:10.1007/s42484-021-00038-w. [p. 2]
- [15] Zoë Holmes, Kunal Sharma, M. Cerezo, and Patrick J. Coles. Connecting ansatz expressibility to gradient magnitudes and barren plateaus. *PRX Quantum*, 3:010313, Jan 2022. doi:10.1103/PRXQuantum.3.010313. [p. 2]
- [16] James Stokes, Josh Izaac, Nathan Killoran, and Giuseppe Carleo. Quantum natural gradient. *Quantum*, 4:269, 2020. doi:10.22331/q-2020-05-25-269. [pp. 2, 5, 11, 17, 20]
- [17] Tobias Haug, Kishor Bharti, and M.S. Kim. Capacity and quantum geometry of parametrized quantum circuits. *PRX Quantum*, 2:040309, Oct 2021. doi:10.1103/PRXQuantum.2.040309. [p. 3]
- [18] Tobias Haug and M. S. Kim. Optimal training of variational quantum algorithms without barren plateaus. *arXiv preprint arXiv:2104.14543*, 2021. doi:10.48550/arXiv.2104.14543. [p. 3]
- [19] Tyson Jones. Efficient classical calculation of the quantum natural gradient. *arXiv preprint arXiv:2011.02991*, 2020. doi:10.48550/arXiv.2011.02991. [p. 3]
- [20] Barnaby van Straaten and Bálint Koczor. Measurement cost of metric-aware variational quantum algorithms. *PRX Quantum*, 2:030324, Aug 2021. doi:10.1103/PRXQuantum.2.030324. [pp. 3, 11, 13]
- [21] Bálint Koczor and Simon C Benjamin. Quantum natural gradient generalised to non-unitary circuits. *arXiv preprint arXiv:1912.08660*, 2019. doi:10.48550/arXiv.1912.08660. [p. 3]
- [22] Hoshang Heydari. Geometric formulation of quantum mechanics. *arXiv preprint arXiv:1503.00238*, 2015. doi:10.48550/arXiv.1503.00238. [p. 3]
- [23] Robert Geroch. *Robert Geroch, Geometrical Quantum Mechanics: 1974 Lecture Notes*. Minkowski Institute Press, Montreal 2013, 2013. [p. 3]
- [24] Ran Cheng. Quantum geometric tensor (Fubini-Study metric) in simple quantum system: A pedagogical introduction. *arXiv preprint arXiv:1012.1337*, 2010. doi:10.48550/arXiv.1012.1337. [p. 5]

- [25] Jutho Haegeman, Michaël Marien, Tobias J. Osborne, and Frank Verstraete. Geometry of matrix product states: Metric, parallel transport, and curvature. *J. Math. Phys.*, 55(2):021902, 2014. doi:10.1063/1.4862851. [p. 5]
- [26] Naoki Yamamoto. On the natural gradient for variational quantum eigensolver. *arXiv preprint arXiv:1909.05074*, 2019. doi:10.48550/arXiv.1909.05074. [pp. 6, 20]
- [27] Pierre-Luc Dallaire-Demers, Jonathan Romero, Libor Veis, Sukin Sim, and Alán Aspuru-Guzik. Low-depth circuit ansatz for preparing correlated fermionic states on a quantum computer. *Quantum Sci. Technol.*, 4(4):045005, Sept 2019. doi:10.1088/2058-9565/ab3951. [pp. 6, 22]
- [28] Pierre-Luc Dallaire-Demers and Nathan Killoran. Quantum generative adversarial networks. *Phys. Rev. A*, 98:012324, jul 2018. doi:10.1103/PhysRevA.98.012324. [p. 6]
- [29] Pierre-Luc Dallaire-Demers, Michał Stęchły, Jerome F Gonthier, Ntwali Tous-saint Bashige, Jonathan Romero, and Yudong Cao. An application benchmark for fermionic quantum simulations. *arXiv preprint arXiv:2003.01862*, 2020. doi:10.48550/arXiv.2003.01862. [p. 7]
- [30] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando GSL Brandao, David A Buell, et al. Quantum supremacy using a programmable superconducting processor. *Nature*, 574:505–510, 2019. doi:10.1038/s41586-019-1666-5. [p. 7]
- [31] Chu-Ryang Wie. Two-qubit Bloch sphere. *Physics*, 2(3):383–396, 2020. doi:10.3390/physics2030021. [p. 8]
- [32] Péter Lévay. The geometry of entanglement: metrics, connections and the geometric phase. *Journal of Physics A: Mathematical and General*, 37(5):1821–1841, jan 2004. doi:10.1088/0305-4470/37/5/024. [pp. 8, 14]
- [33] James Martens and Roger Grosse. Optimizing neural networks with kronecker-factored approximate curvature. In Francis Bach and David Blei, editors, *Proceedings of the 32nd International Conference on Machine Learning*, volume 37 of *Proceedings of Machine Learning Research*, pages 2408–2417, Lille, France, 07–09 Jul 2015. PMLR. [p. 12]
- [34] Alberto Bernacchia, Máté Lengyel, and Guillaume Hennequin. Exact natural gradient in deep linear networks and application to the nonlinear case. In *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, NIPS’18, page 5945–5954, Red Hook, NY, USA, 2018. Curran Associates Inc. [p. 12]
- [35] Sam A. Hill and William K. Wootters. Entanglement of a pair of quantum bits. *Phys. Rev. Lett.*, 78:5022–5025, Jun 1997. doi:10.1103/PhysRevLett.78.5022. [p. 13]
- [36] Li Chen, Ming Yang, Li-Hua Zhang, and Zhuo-Liang Cao. Directly measuring the concurrence of two-atom state via detecting coherent lights. *Laser Phys. Lett.*, 14(11):115205, oct 2017. doi:10.1088/1612-202X/aa8582. [p. 13]
- [37] Lan Zhou and Yu-Bo Sheng. Concurrence measurement for the two-qubit optical and atomic states. *Entropy*, 17(6):4293–4322, 2015. doi:10.3390/e17064293. [p. 13]
- [38] Sean M. Carroll. *Spacetime and Geometry: An Introduction to General Relativity*. Cambridge University Press, 2019. doi:10.1017/9781108770385. [p. 14]

- [39] Anshuman Dey, Subhash Mahapatra, Pratim Roy, and Tapobrata Sarkar. Information geometry and quantum phase transitions in the Dicke model. *Phys. Rev. E*, 86(3):031137, Sept 2012. doi:10.1103/PhysRevE.86.031137. [p. 14]
- [40] Rıza Erdem. Quantum lattice model with local multi-well potentials: Riemannian geometric interpretation for the phase transitions in ferroelectric crystals. *Physica A: Statistical Mechanics and its Applications*, 556:124837, 2020. doi:10.1016/j.physa.2020.124837. [p. 14]
- [41] Michael Kolodrubetz, Vladimir Gritsev, and Anatoli Polkovnikov. Classifying and measuring geometry of a quantum ground state manifold. *Phys. Rev. B*, 88:064304, Aug 2013. doi:10.1103/PhysRevB.88.064304. [p. 14]
- [42] Michael Hauser and Asok Ray. Principles of Riemannian geometry in neural networks. In I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 30. Curran Associates, Inc., 2017. [p. 14]
- [43] T. Yu, H. Long, and J. E. Hopcroft. Curvature-based comparison of two neural networks. In *2018 24th International Conference on Pattern Recognition (ICPR)*, pages 441–447, 2018. doi:10.1109/ICPR.2018.8546273. [p. 14]
- [44] P. Kaul and B. Lall. Riemannian curvature of deep neural networks. *IEEE Trans. Neural Netw. Learn. Syst.*, 31(4):1410–1416, 2020. doi:10.1109/TNNLS.2019.2919705. [p. 14]
- [45] Alberto Peruzzo, Jarrod McClean, Peter Shadbolt, Man-Hong Yung, Xiao-Qi Zhou, Peter J. Love, Alán Aspuru-Guzik, and Jeremy L. O’Brien. A variational eigenvalue solver on a photonic quantum processor. *Nat. Commun.*, 5:4213, Sept 2014. doi:10.1038/ncomms5213. [p. 16]
- [46] Peter JJ O’Malley, Ryan Babbush, Ian D Kivlichan, Jonathan Romero, Jarrod R McClean, Rami Barends, Julian Kelly, Pedram Roushan, Andrew Tranter, Nan Ding, et al. Scalable quantum simulation of molecular energies. *Physical Review X*, 6(3):031007, 2016. doi:10.1103/PhysRevX.6.031007. [p. 16]
- [47] John Frank Adams. On the nonexistence of elements of Hopf invariant one. *Bull. Am. Math. Soc.*, 64(5):279–282, 1958. [p. 21]
- [48] Shreyas Bapat, Ritwik Saha, Bhavya Bhatt, Hrushikesh Sarode, Gaurav Kumar, and Priyanshu Khandelwal. einsteinpy/einsteinpy: EinsteinPy 0.1a1 (Alpha Release - 1), March 2019. doi:10.5281/zenodo.2582388. [p. 21]
- [49] Wolfram Research, Inc. Mathematica, Version 12.0. Champaign, IL, 2019. [p. 21]
- [50] Jarrod R McClean, Nicholas C Rubin, Kevin J Sung, Ian D Kivlichan, Xavier Bonet-Monroig, Yudong Cao, Chengyu Dai, E Schuyler Fried, Craig Gidney, Brendan Gimby, et al. Openfermion: the electronic structure package for quantum computers. *Quantum Science and Technology*, 5(3):034014, 2020. doi:10.1088/2058-9565/ab8ebc. [p. 21]
- [51] Ville Bergholm, Josh Izaac, Maria Schuld, Christian Gogolin, Shahnawaz Ahmed, Vishnu Ajith, M. Sohaib Alam, Guillermo Alonso-Linaje, B. AkashNarayanan, Ali Asadi, et al. PennyLane: Automatic differentiation of hybrid quantum-classical computations. *arXiv preprint arXiv:1811.04968*, 2018. doi:10.48550/arXiv.1811.04968. [p. 21]