# Geometry and Quantum Complexity

Sita Gakkhar

December 2018

## Contents

### Abstract

We survey some geometric ideas that give information and computing on quantum machines a distinct flavor compared to classical setting, and introduce connections between quantum complexity and high energy physics.

## 1 Introduction

A quantum computation is described by application of an unitary operator. We would like to formalize the idea that similar/close unitaries[1] have similar complexity. One approach is via geometry where we try to metricize complexity: closeness arises from a metric on the space of unitaries and distance to a fixed reference serves as a complexity measure. As the application of identity corresponds to simplest possible computation, one which trivially maps input to output, the identity unitary is an intuitive choice for the reference.

The space of unitary operators on $n$ qubits is given by the groups $U(2^n)$ and $SU(2^n)$ (the group of $2^n \times 2^n$ matrices with determinant of norm 1, and its subgroup formed by matrices of determinant 1). Being defined by the preimages of the determinant as a map from $GL(2^n, \mathbb{C})$ to $S^1$ allows one to give the structure of real manifolds to $U(2^n)$ and $SU(2^n)$. So we have a natural geometric setting on which to formalize the notion of "distance". This can be done by considering Riemannian metrics on manifolds $M = U(2^n), SU(2^n)$. Starting with this initial germ of the idea introduced in [7, 17, 18], we explore theme of viewing complexity through a geometric perspective building towards a sketch of how these connect with fundamental physics.

[1] Throughout we refer to an unitary matrix (of dimension clear from context) as unitary/unitaries

In what follows, we will need basic ideas about topology of smooth manifolds, Riemannian geometry and Lie groups. The discussion of high energy physics aspects that is given here is from a bird's eye view and the reader is referred to the cited articles for detailed analysis. Where helpful we give intuition behind ideas underlying the mathematical machinery, however, in the interest of accessibility and succinctness defer the full technical glory to references[2].

Informally, a manifold (in our setting we are primarily concerned with $SU(2^n)$ which additionally is a *compact*) is a *nice*[3] topological space covered by a collection of *charts*, each chart a pair of an open set $U_i$ and homeomorphism $\phi_i$ of this set into the euclidean space, $\mathbb{R}^k$ of the same dimension as the manifold, $k$. The smooth qualifier comes from transition functions, $\phi_i \circ \phi_j^{-1}$ defined whenever $U_{ij} = U_i \cap U_j$ is non-empty, being smooth as maps from $\mathbb{R}^k \to \mathbb{R}^k$ for all $i, j$. By Whitney Embedding Theorem, any smooth real manifold of dimension $k$ smoothly embeds inside an euclidean space of dimension $2k$ (it's just that things are nicer if we can remain agnostic of the higher dimensional ambient space). So we can think of these as surfaces, and get an intuitive approximation of ideas like tangent spaces, the formalism for which is slightly nuanced.

One can give additional structures to smooth manifolds: a smooth Riemannian manifold is a smooth manifold that has an inner product $g_p$ on the tangent space at each point $p$, and over the manifold the inner product varies smoothly. $g_p$ is called the Riemannian metric tensor. Both $SU(2^n)$ and $U(2^n)$ are also Lie groups, i.e. they are smooth manifolds with the topology compatible with the group product and inverses, meaning the group operations as maps from $M \times M \to M$ are smooth[8].

# 2 Quantum complexity and geodesics

We relate complexity of quantum computation to geodesics on Riemannian manifolds as developed in [7, 17, 18, 19]. We will work with $M = SU(2^n)$ unless otherwise indicated, starting in the setting without ancilla qubits. This means that we can work with $SU(2^n)$ and not an embedding in a higher dimensional $SU(2^{n+m})$ with $m$ indicating $m$ ancilla qubits and possibly unbounded.

The significance of the geometric perspective is that it gives a different insight into circuits $-$ circuits are arbitrary while geodesics can be described by an equation. In particular, if a part of an optimal (classical or quantum) circuit is known, it doesn't really give any information about what the rest of the circuit could be; but once a position and velocity is fixed then the geodesic is determined by the geodesic equation.

This doesn't trivialize the problem because while geodesics are locally length minimizing, beyond a certain point they stop being so (though all distance minimizing curves are geodesics: geodesics are the stationary points of the length functional, not always minimizing). For instance, consider two points on a great circle of a 2-sphere (great circles are geodesics for the sphere). If the two points aren't antipodal, then only one direction on the great circle yields the path of minimal length. Finding minimizing geodesic means pinning down the correct tangent, and may be non trivial.

## 2.1 Metricizing complexity

The Lie group $SU(2^n)$ has associated Lie algebra $su(2^n)$ which consists of traceless skew hermitian matrices, a basis for which is given by the Pauli matrices after multiplying by $i$ to make them skew hermitian. A Lie algebra is basically a vector space (so fits in the place of a tangent space) with an additional structure called the Lie bracket[4].

[2] Relevant sections of [13] provide an in-depth account of topology of smooth manifolds, and the same for Riemannian geomentry with [24], [6]

[3] We brush second countability, Hausdorff-ness, compatibility of charts, and other things under the rug to avoid getting too bogged in details

[4] Lie bracket is a non-associative, alternating bilinear map satisfying the Jacobi identity. Any associative algebra can be turned into a Lie algebra by using the commutator as the Lie bracket. Recall that when one *quantizes* the classical theory the Poisson bracket of Hamiltonian dynamics goes to the commutator

There's a map from the Lie agebra to Lie group called the exponential map, which for matrix Lie groups/algebras is the usual matrix exponential. A particularly nice property of this map is that it restricts to a diffeomorphism from some neighborhood of zero in the Lie algebra to a neighborhood of identity in the Lie group.

Now a tangent vector at $U \in M$ can be associated with $H \in su(2^n)$ by associating it with tangent to curve $\gamma(t) = e^{iHt}U$ at $t = 0$. We consider right invariant inner product on $M$: one that's independent of $U \in M$[7] (precisely, the inner product is invariant w.r.t. the right multiplication map). This means we can do everything about the identity $Id \in M$ (by moving things around using the right multiplication) where the tangent space is by definition the Lie algebra and exponential map plays nice. Thus, the inner product is a bilinear form $\langle H, J \rangle$ on $su(2^n)$. Note while the inner product on $M$ is invariant as a function of $U \in M$, the vector itself may change when moving from one tangent space to another (this is captured by the idea of the affine connection we introduce later).

For a curve $U(t)$ on $M$ generated by Hamiltonian $H(t)$ evolving according to Schrödinger's equation, the length is given by integrating the length element $\langle H(t), H(t) \rangle^{1/2}$ along the curve, and the distance between $U, V \in M$, $d(U, V)$, is the infimum over all curves from $U$ to $V$.

Since quantum computation complexity measures the circuit size in terms of 1 and 2 qubit unitaries, so we decompose $su(2^n)$ into subspaces $P$, of $1, 2$ qubit unitaries (that is, one and two body hamiltonians), and $Q$, such that $P + Q = su(2^n)$, penalizing paths for straying into $Q$ by using a right invariant inner product of form

$$\langle H, J \rangle = 2^{-n}(\mathrm{Tr}(H\mathcal{P}(J)) + q\,\mathrm{Tr}(H\mathcal{Q}(J))) \tag{1}$$

where $\mathcal{P}, \mathcal{Q}$ are projections operators on subspaces $P, Q$ and $q$ is a penalty parameter. Equation (1) has a general form

$$\langle H, J \rangle = 2^{-n}\,\mathrm{Tr}(H\mathcal{G}(J)) \tag{2}$$

where $\mathcal{G}$ is strictly positive superoperator (a linear operator on vector space of linear operators satisfying $\mathrm{Tr}(H\mathcal{G}(H)) > 0$ for $H \neq 0$). In equation (1), $\mathcal{G} = \mathcal{P} + q\mathcal{Q}$. We call a metric like this a metric of standard form.

This induces a metric on $SU(2^n)$ which satisfies the following inequalities in terms of approximate and exact gate complexities, $G(U, \epsilon), G(U)$ where $G(U, \epsilon)$ is the number of one and two qubit gates requires to approximate $U$ within $\epsilon$ w.r.t matrix norm (without ancilla qubits), and $G(U)$ the number of the one and two qubit gates required to synthesize it exactly:

$$b_0 G(U, \epsilon)^{b_i} \epsilon^{b_2} n^{-b_3} \leq d(Id, U) \leq G(U)$$

where $b_i$'s are constants. A subtlety here is that $d$ depends on $\epsilon$ (for metrics of standard form via the choice of penalty parameter).

The results carry through if $P$ is replaced by a space generated by different universal gate set of dimension $poly(n)$[7] (say $P'$), and we still have an inequalities of form:

$$poly(n, \epsilon, G(U, \epsilon)) \leq d(Id, H) \leq G(U)$$

where $d, G$ are now with respect to $P'$.

Note that we use "metric" interchangably to mean the Riemannian metric tensor on $SU(2^n)$ (that is, the family of inner products on the manifold) and the metric $d$ which turns $SU(2^n)$ into a metric space when it's clear from context.

### 2.2 The geodesic equation

In analogy with euclidean space where the shortest path between two points is a straight line which is determined by the endpoints and constant directional derivatives, a geodesic on a Riemannian manifold is defined as a curve for which the covariant derivative of the velocity vector field is zero[24]. For a curve passing through origin with tangent $Y$, the covariant derivative of a vector field $Z$ is given by

$$\frac{\partial z^j}{\partial x^k} y^k + \Gamma^j_{kl} y^k z^l \overset{\text{Einstein notation}}{\equiv} \sum_k \frac{\partial z^j}{\partial x^k} y^k + \sum_k \sum_l \Gamma^j_{kl} y^k z^l$$

where $\nabla_Y Z$ is the Levi-Civita connection[5], $\Gamma^j_{kl}$ are Christoffel symbols (defined in terms of partial derivatives of the metric tensor w.r.t. local coordinates), and we are using Einstein summation convention, so repeated indices in matched up/down (contravariant/covariant) positions are implicitly summed.

Now the geodesic equation is in local coordinates $-$ for a point $p$ in the chart $(U, \phi)$, local coordinates are just the identification of the open set $U$ via $\phi$ with the euclidean space so we can work as if we are in euclidean space.

However, the metric we defined in equation (2) was in terms of Hamiltonian $H$ which is an element of the tangent space. So we need to pass to local coordinates. The exponential being a diffeomorphism about idenity helps, and since the metric is right invariant, it is enough to work in the local coordinates about origin then use right invariance to get it everywhere on the manifold.

The way to get a nice set of local coordinates is by using the Lie structure about $Id \in SU(2^n)$. First note that about origin we can associate a coordinate vector $\hat{x}$ with a tangent vector $H = \hat{x} \cdot \hat{\sigma} = \sum x_\sigma \sigma$ (recall the tangent space at identity is the $su(2^n)$: the space of skew Hermitian matrices with trace zero, we can write $H$ in terms of generalized Pauli matrices, which can be taken as tensor products of the single-qubit Pauli matrices). $\hat{x}$ is the Pauli representation of $H$ corresponding to local coordinates.

Let $U = e^{iX}$ be in a small neighborhood about $Id$, and $H$ be a Hamiltonian representing a tangent vector. We are interested in the Pauli representation of $H$ in the tangent space at $U$. This follows from a computation using Baker–Campbell–Hausdorff formula. We only care about a small neighborhood of $U$ as we just need to be able to know how tangent vectors change at origin in terms of local coordinates to evaluate the covariant derivative.

After all the dust settles, we have superoperators $\mathcal{E}_X, \mathcal{D}_X$ connecting the Pauli coordinates $J$ at $X$ with the Hamiltonian $H$, $H = \mathcal{E}_X(J)$, with inverse $\mathcal{E}_X^{-1} = \mathcal{D}_X$ (the inverse only exists in a small neighborhood about origin), $\mathcal{E}_X^\dagger = \mathcal{E}_{-X}$ and $\mathcal{D}_X^\dagger = \mathcal{D}_{-X}$. Extending these results, along similar lines, we can compute the metric, the Christoffel symbols, and the covariant derivative (which were again all in local coordinates). Putting everything together we get:

**Theorem 1.** *Given the Riemannian metric of equation (1) with $\mathcal{G} = \mathcal{P} + q\mathcal{Q}$ on $SU(2^n)$, let $M = (1 - q^{-1})\mathcal{G}(H)$, then the geodesic equation (assuming $q > 1$) is give by:*

$$\dot{M} = i[M, \mathcal{P}(M)]$$

This equation belongs to a much studied class called *Lax equation.* The availability of tools of Riemannian geometry allows for studying questions about analytic form and numerical solutions of geodesic equations, as well as the variational problem where we can consider how geodesics change on varying the metric, for example, by varying the penalty (as noted this is quite relevant here).

[5] Intuitively Levi-Civita connection captures how a vector field changes on the manifold, with constraints of being compatible with metric and having zero torsion

### 2.3 Complexity upper and lower bounds

We sketch how the quantum complexity of applying $U$ can be bounded using the metric approach developed − precisely we outline the argument for showing:

**Theorem 2.**

$$poly(n, \epsilon, G(U, \epsilon)) \leq d([U]) \leq G(U)$$

where $G(U)$ is the complexity of implementing $U$ exactly, and $G(U, \epsilon)$ the complexity of computing an $\epsilon$ approximation w.r.t. to a set of unitaries $\mathcal{H}$, and $d$ is a metric of standard form (dependent on $\epsilon$), $d([U])$ the distance between $U$ and $Id$ with respect to $d$.

The general ideas for proof are derived from the problem of *time optimal Hamiltonian control* [7, 18, 19, 12] where the goal is to find a time-dependent Hamiltonian $H(t)$ synthesizing $U$. Note that a time dependent Hamiltonian can be written as $H(t) = \sum \gamma_\sigma(t)\sigma$ in Pauli representation. An important object that we will work with is the *control curve*.

**Definition 2.1.** For a Hamiltonian $H(t) = \sum \gamma_\sigma(t)\sigma$, the curve $\gamma(t)$ defined as the vector of curves corresponding to each $\sigma$, $(\gamma_\sigma(t))$, is called the control curve[18] which controls the synthesis according to the Schrödinger's equation with boundary conditions:

$$\dot{V} = iH(t)V \text{ and } V(0) = I, V(1) = U \tag{3}$$

The natural setting for this problem is of Finsler geometry which is a generalization of Riemannian geometry where instead of an inner product we have a *cost* function on the tangent space. We follow [18] and start by defining *local metric*.

**Definition 2.2.** A manifold $M$ with local metric[6] is a manifold with a function $F : TM \to [0, \infty)$ (where $TM$ denotes the tangent bundle) such that for each $x \in M$, and all $y \in T_xM$, $F(x, y) \geq 0$ with $F(x, y) = 0$ iff $y = 0$, $F$ is positively homogeneous in second coordinate, and $F(x, \cdot)$ satisfies the triangle inequality for each $x$.

So $F$ gives a notion of length at each point on the manifold: it gives an asymmetric norm (that is, $F(x, v) = F(x, -v)$ may fail).

**Definition 2.3.** A manifold with a local metric, $F$, that is smooth with the hessian of $F(x, \cdot)^2$ at each tangent vector $v \neq 0$ positive definite, is a Finsler manifold. We call such $F$ a Finsler function.

The positive definiteness is related to *strict* triangle inequality [18] − and since for complexity setting that is what we want: we think of $F$ as the cost of applying $H$ and we want two operations to be strictly higher cost than single operation. Also, if we want to do calculus of variations with $F$ then smoothness is reasonable to require.

Given a cost function on the tangent space, $c$, we can associate a cost to applying a Hamiltonian $H(t)$ for time $[0, T]$ by

$$C(H(t)) = \int_0^T dtc(H(t))$$

We define the cost of a unitary $U$ as the cost infimum over all permissible Hamiltonians synthesizing $U$:

$$C(U) = \inf_{T,H} C(H(t))$$

Note how $C(U)$ relates to the lengths of curves when the cost function defines a metric.

[6] As noted in [18], *local metric* is not a standard term

We restrict to the setting of where the cost function is given by a metric of form $\mathcal{G} = \mathcal{P} + p\mathcal{Q}$ (ignoring $2^{-n}$ overall scaling for simplicity), with $\mathcal{P}, \mathcal{Q}$ same projection operators as equation (1) :

$$c(H) = \sqrt{\langle H, H\rangle} = \sqrt{\mathrm{Tr}(\mathcal{G}(H)H)} \tag{4}$$

That is, all Hamiltonians are permissible, although weighed differently[7]. For a Hamiltonian $H$, let $H_P$ be the Hamiltonian obtained by projecting onto the subspace $P$.

To get the $poly(n, \epsilon, G(U, \epsilon)) \leq d([U])$ bound where $d$ is the metric (equivalently, the cost function) of form (4) where we will show how to fix the penalty $p$, we proceed using following lemmas:

**Lemma 3.** *For $H$, an $n$-qubit one or two body Hamiltonian (that is, $H = \mathcal{P}(H) \equiv H_P$), let $V$ be the unitary generated by $H$ over $[s, s+\Delta]$. Then for the $\Delta$-averaged Hamiltonian, $\tilde{H}(t)$,*

$$\tilde{H}(t) = \frac{1}{\Delta}\int_0^\Delta dt H(t)$$

*we have*

$$||V - e^{-i\tilde{H}\Delta}|| = O(N_P^2\Delta^2)$$

*where $N_P$ is the maximum norm over one or two body Hamiltonians:*

$$N_P = \sup_{H\in P} ||H||$$

**Lemma 4.** *Let $g(\Delta, \delta)$ be the complexity of approximating an arbitrary $\Delta$-averaged $n$-qubit unitary (i.e one that comes from a $\Delta$-averaged Hamiltonian) to an accuracy better than $\delta$ in matrix norm using one and two qubit gates. Then $g(\Delta, \delta) = O(p(n)\Delta^2/\delta)$ where $p$ is a polynomial.*

Starting with a unitary $U$ given in terms of a Hamiltonian $H(t)$ as per equation 3, we break an interval $[0, T]$ into sub-intervals of length $\Delta = T/N$, then we have

$$||U_P^j - U_M^j|| \leq O(N_P^2\Delta^2)$$

where $U_P^j$ is the unitary generated by $H_P$ over $j^{th}$ interval, $U_M^j$ the unitary generated by the mean Hamiltonian. We define the final approximating unitary $U_A$ as $U_M^j$ applied in sequence.

Now we also have $T \leq C(U)/c_A$ where $c_A$ is the minimal cost associated to any unitary: $c_A = \inf_H c(H)$. So by construction $U_A$ has complexity at most $Ng(\Delta, \delta) = Tg(\Delta, \delta)/\Delta$

Then using $||U - U_A|| \leq ||U - U_P|| + ||U_P - U_A||$, by repeated application of lemma 3, we get

$$||U - U_A|| \leq RC(U) + O\left(\frac{N_P^2 C(U)\Delta}{c_A}\right) + \frac{C(U)\delta}{c_A\Delta}$$

where the parameter $R = \sup_H ||H - H_P||/c(H)$ measures the quality of approximation w.r.t. the cost function.

For a metric of form $\mathcal{G} = \mathcal{P} + p\mathcal{Q}$,

$$c(H) = \sqrt{\sum_P h_\sigma^2 + p\sum_Q h_\sigma^2}$$

we get that $R \leq 2^n/p$ since $R$ is maximized when the target Hamiltonian only has three or higher body interactions, $c_A = 1$. It can be argued that $N_P = O(n)$, $g(\Delta, O(n^4, \Delta^3)) \leq O(n^2/\Delta)$ (with some technicalities

[7] See Nielsen et al[19] for more general results

and qualifiers involved, see [19]).

And after some fine tuning

$$||U - U_A|| \leq 2^n C(U)/p + O(\epsilon)$$

for $U_A$ synthesized from $O(C(U)^3 n^6 \epsilon^4)$. Nielsen et al [19] note that $C(U) \leq 4^n$ is known for all unitaries, so choosing $p = 8^n/\epsilon$ works to get $O(\epsilon)$ error. This establishes the lower bound since in this setting $C(U)$ is just the length of curve from $I$ to $U$.

The upper bound follows by integrating a smoothed out version of the control curve[18] on $SU(2^n)$ equipped with a Finsler manifold structure via a Finsler function $F$.

We begin by looking at the minimal sequence of gates that synthesizes $U$, say $U = \Pi_1^{G(U)} U_k$ for $U_k = e^{-iH_k}$ where $U_k$ come from a set $\mathcal{A}$. Then considering the control curve (scaled to run on $[0,1]$) $\gamma$ that applies $H_k$'s one after the other for the fixed size interval each: $\gamma(t)\cdot\sigma = G(U)H_i$ for $t \in [(i-1)/G(U), i/G(U)]$

Now we require that $\mathcal{A}$ be identified with a subset $\mathcal{H}$ of $su(2^n)$ (that is, the mapping $\mathcal{H} \rightarrow \mathcal{A}$ given by $H \rightarrow e^{-iH}$ is bijective) and that $F(V, H) \leq 1$ for all $V \in SU(2^n)$ and $H \in \mathcal{H}$. We call such $F$ $\mathcal{A}$-bounding. We further require that $\mathcal{A}$ can exactly generate $U$[8].

For technical reasons we need to regularize $\gamma(t)$ to make it smooth, and we can do this by multiplying with a positive function $r(t)$ which has unit integral ($r(t)$ is a essentially a *mollifier*). Now from positive homogeneity of $F$ along with $\mathcal{A}$-boundedness, it follows easily that

$$d([U]) \leq \int_0^1 F(V(t), r(t)\gamma(t)\sigma)dt = \int_0^1 r(t)F(V(t), \gamma(t)\sigma)dt \leq \int_0^1 r(t)G(U)dt = G(U)$$

The standard metric from equation (1) relates to Finsler functions $F_2, F_q$ introduced in [18] and so the upper bound holds for it as well.

## 2.4 Accounting for ancilla qubits

For a general unitary, bounds on complexity in the setting where an unbounded number of ancilla qubits are allowed are not known. But for "nice" unitaries, something can be said.

Suppose we have $V(|\psi\rangle|0\rangle) = U|\psi\rangle|A\rangle$. Such $V$'s are defined to be $m$-fold extensions of $U$ where $m$ is the number of ancialla qubits. A special extension is where the ancilla qubits end up in the same state they start: $|0\rangle$.The idea is to note that the circuit of figure 1 for any $m$-fold special extension $V$ of $U$ is a $m+1$ fold extension of $U$ that is independent of $V$.

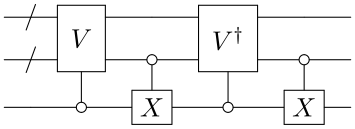

Figure 1: The canonical extension in terms of $V$: The second wire represents $m$ ancillas and $X$ is the Pauli $X$ operator

The action of circuit from figure 1 is equivalent to the figure 2 circuit, denoted as the $m^{th}$ canonical extension $U_m$[7], which is $m+1$-fold special extension of $U$.

[8] We can always add an $\epsilon$ slack and get a bound on $G(U, \epsilon)$ by making the argument with $U'$ which is $\epsilon$ close to $U$ ($\epsilon$ depending on $\mathcal{A}$)

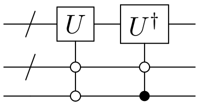

Figure 2: The canonical extension $m+1$ fold extension $U_m$

Let $G(\cdot)$ and $G(\cdot,\epsilon)$ be exact and approximate gate complexities as before without any ancilla qubits. We define $\tilde{G}_m(U)$ be the complexity of exactly synthesizing a m-qubit special extension of $U$ w.r.t. universal set of one and two qubit gates, and $\tilde{G}_\infty(U)$ be the complexity of exactly synthesizing a special extension of $U$ with unbounded ancilla qubits. Also, let $G_\infty(U)$ be the complexity of exactly synthesizing an extension (not necessarily special) of $U$ with unbounded ancilla qubits. Similarly let $G_\infty(\cdot,\epsilon)$ denote approximate complexity with an unbounded number of ancillas where the approximation parameter is evaluated using only the $n$ qubits that correspond to output of $U$.

Nielsen et al in [7] show that there are constants $c_1, c_2$, so that for any $U$ and $m$ the $m+1$ fold canonical extension $U_m$ satisfies

$$\min(m, c_1 G(U_m) - c_2 m) \leq \tilde{G}_\infty(U) \text{ and } G_\infty(U,\epsilon) \leq G(U_m,\epsilon) \tag{5}$$

From this, using bounds on $G(U_m,\epsilon), G(U_m)$ ($U_m$ is just a unitary on $n+m+1$ qubits) in terms of $d$ gives

$$\min(m, c_3 d(I,U_m)) - c_4 m) \leq \tilde{G}_\infty(U) \text{ and } poly(G_\infty(U,\epsilon)) \leq d(I,U_m)$$

for a constant $c_3, c_4$.

The line of argument is fairly standard after noticing that it doesn't help to use more ancillas than gates in the circuit. Using this, it follows that since

$$b_0 G(U,\epsilon)^{b_1} \epsilon^{b_2} n^{-b_3} \leq d(I,U)$$

we also get

$$b_0 G_\infty(U,\epsilon)^{b_1} \epsilon^{b_2} (n+m)^{-b_3} \leq d(I,U_m)$$

This allows for bounding the complexity of approximating a unitary using an unbounded number of ancillas by lengths of geodesics in a finite dimensional space, the latter being a lot more tractable.

Note that the machinery of inequality (5) is especially useful if $G_\infty(U)$ and $\tilde{G}_\infty(U)$ are similar, which indeed holds for unitaries representing functions $\{0,1\}^n \to \{0,1\}$ (so $f(x)$ is the first qubit of output and rest are ignored), permutations, and that which are diagonal in computation basis (e.g. $|x\rangle \to e^{i\theta_x}|x\rangle$).

## 2.5 Connections with classical complexity

In this section we go in the direction opposite to one considered so far: using geometry to illuminate complexity, by using Razborov-Rudich theorem from computational complexity to show that approximating geodesics on $SU(2^n)$ is hard. Razborov-Rudich theorem says that if "good" pseudorandom generators exist then one cannot separate hard-to-compute and easy-to-compute functions. By hard to compute, it means that the minimal non uniform circuit complexity is above a threshold, and "good" is in the Blum-Micali-Yao sense [3].

Random Boolean functions are known to be hard. Now given a good psuedorandom generator it's possible to construct pseudorandom functions that have small circuits, so are easy to compute. The result follows by noting that separating hard and easy would now break pseudorandom generator. As random unitaries have large quantum complexity, this argument translates into quantum setting. Being able to approximate $d(I,U)$ fast means we can separate unitaries with large and small quantum complexity (say about a threshold like $n^{\ln(x)}$). By considering a unitary like $U_f|x\rangle \equiv (-1)^{f(x)}|x\rangle$ where $f$ is either random or pseudorandom (so we have small quantum complexity for $U_f$ when $f$ is pseudorandom and large otherwise), we can similarly break the pseudorandom generator. This implies that it cannot be easy to approximate geodesic lengths on $SU(2^n)$ because of their relationship to bounds on quantum complexity.

# 3 Complexity and random walks on groups

We have been working in the setting where the analysis is aware of the "bad" parts of the $SU(2^n)$ which have high complexity: we have a penalty on the subspace $Q$, but we can still tread there. Now we consider restricting to the subgroup of $SU(2^n)$ generated by a given set of unitaries $S$; in a sense being agnostic of what exists outside this universe. The exposition here follows [14].

Consider the Cayley graph of subgroup $G$ generated by $S$. Assume $S$ is symmetric, i.e, $S = S^{-1}$, but does not contain the identity. The Cayley graph $\Gamma(G, S)$ is constructed as follows:

1. Assign colors $c_s$ to $s \in S$
2. Add edges $g$ to $gs$ of color $c_s$

There's a natural metric on $\Gamma$ which gives the complexity of going from $g_1$ to $g_2$ in $G$: the minimal number of edges needed to connect $g_1$ and $g_2$. This doubles as the measure of complexity. Now the random circuit model corresponds to a random walk: suppose at step $i$ we are are at $g_i$, we pick a random generator $s$ and then apply it, taking the edge to $g_i s$. Starting at time zero from idenity, the string of generators that we applied is a random walk and at the same time a circuit. The distance from the identity to the where we end up gives the complexity in terms of generating set.

Random walks on groups are well studied, and results about random walks can be translated to complexity setting. In particular we are interested in how far does a random walk get by time $t$ as that corresponds to the complexity of a random circuit of length $t$. To start we will work with $G = S_n$ for $n = 2^{K-1}$, the permutation group on $n$ symbols, with all transpositions as $S$.

## 3.1 Complexity geometry of permutation group

From basic theory of permutation groups it follows that the diameter of the Cayley graph is exactly $n-1$: any permutation can be decomposed into disjoint cycles, and a cycle of length $k$ can be written as product of $k-1$ transpositions, so diameter is atleast $n-1$, but $(12..n)$ cannot be decomposed into fewer than $n-1$ transpositions so we get the other inequality. Similarly, if $c(g)$ is the number of disjoint cycles in cycle decomposition of $g$ then the complexity of $g$ is given by, $C(g) = n - c(g)$.

As [14] notes, the average complexity (i.e. the average distance travelled by the random walk) starting at identity as a function of time, in the limit $n$ goes to infinity is given by

$$\frac{C(t)}{n} = 1 - \sum_k \frac{n}{2t} \frac{k^{k-2}}{k!} \left( \frac{2te^{-\frac{2t}{n}}}{n} \right)^k$$

For $t < n/2$, complexity grows linearly at unit speed, then slows down for $t > n/2$ with the second derivative discontinuous at $n/2$. This behaviour at $n/2$ can be considered a phase transition, and is quantified by the geometry of the group using the idea of $\delta$-hyperbolicity.

**Definition 3.1.** For three points in a metric space, a triangle can be defined by pairwise geodesics. A triangle is $\delta$-thin if for any point on the triangle, we can find a point on one of the other two sides within $\delta$. A metric space is $\delta$-hyperbolic if all triangles are $\delta$-thin.

**Definition 3.2.** A group is $\delta$-hyperbolic if the the corresponding Cayley graph is $\delta$-hyperbolic (as defined below).

Informally, $\delta$-hyperbolicity measures how negatively curved the space is, roughly indicated by magnitude of $\delta$.

The value of $\delta$ may depend on the generating set for a Cayley graph, but the existence of a finite $\delta$ does not. With this in hand, we can state the following theorem

**Theorem 5.** *Let $T$ be a triangle formed by the origin and two points sampled independently from the hitting distribution (of the random walk) on the sphere of radius $a2^K$ for $a \in (0, 1)$. If $a < 1/4$, then in limit $K \to \infty$, with probability $1$, $T$ is $\delta$-thin with $\delta \sim O(1)$, whereas $\delta \sim O(2^K)$ for $a > 1/4$*

Given a sequence of spaces $M_n$ of finite diameter $D_n$, one can ask for minimum $\delta_n$ for each $M_n$, so $M_n$ is $\delta_n$-hyperbolic. If diameter of $M_n$ goes to infinity, in the limit this only makes sense[9] if $\delta_n/D_n$ vanishes. So the above theorem is capturing how $\delta$-hyperbolicity starts to break down as $n$ gets large. In some sense, a long enough random walk "starts to feel the compactness" of $SU(2^n)$ and how far it can go from where it's at starts to get limited.

This threshold phenomenon is conjectured (in quantum gravity literature) as indicating "Firewall" [14]. Other ideas like scrambling time and switchback effect are embedded in this toy geometry as well, we refer the reader to references for deeper exploration.

### 3.2 Complexity geometry of generic group

Analyzing random walk on a subgroup generated by a random set of generators can be tricky. However, some geometric properties of analysis on $S_n$ are shared with large classes of Cayley graphs, and these constrain the complexity geometry. A property of hyperbolic geometry that is shared by expander graphs is that area is proportional to volume (for finite groups we can use the counting measure). A random graph is good expander. But that's not the right kind of randomness: we want Cayley graphs coming from a random set of generators. As noted in [14] there's a theorem of Alon and Roichman that gives the existence of a function $s(c)$ such that for any group of order $n$ and for a random generating set $S$ with size $s(c)\log(n)$, the corresponding Cayley graph is a $c$-expander with a probability that goes to 1 as $n$ goes to infinity.

Another geometric property that large number of Cayley graphs are conjectured to satisfy is *cutoff*. This time at which the $L_1$ (equivalently, the total variation distance) distance between the distribution over the vertices of the graph generated by a random walk starting at identity and the equilibrium distribution jumps to zero from close to maximum ([14] cites card shuffling as the canonical example: it can be considered as a random walk on $S_n$ (a $n$ card deck ), then the cutoff is at time $O(n\log(n))$ if each shuffle is a transposition [14, 23]). The cutoff in terms of the random circuit analogy gives the length at which the complexity of the random circuit becomes close to the maximum.

## 4 Quantum complexity and quantum gravity

As noted by [1, 4] the dynamics of a black hole can be represented by a quantum circuit with the growth of geometry behind horizon related to the growth of complexity of the circuit. The relationship is provided by holographic principle which was proposed by Gerard 't Hooft to address the black hole information paradox[10] based on the observation that the entropy of a black hole grows as radius squared and not cubed[11].

[9] otherwise we are looking at "$\infty$-thin" triangles which are actually fat

[10] The paradox where an evaporating black hole via Hawkings' radiation was conjectured to destroy information violating unitary evolution principle of Quantum Mechanics was famously subject of a bet between Stephen Hawkings, Kip Thorne and John Preskill, see https://en.wikipedia.org/wiki/Thorne-Hawking-Preskill_bet

[11] https://en.wikipedia.org/wiki/Holographic_principle

The anti-de Sitter/conformal field theory correspondence (AdS/CFT correspondence) is a realization of the holographic principle. It conjectures the duality between string theories on anti-de Sitter (AdS) spacetime[12] and corresponding quantum field theories (with conformal symmetry) associated to the AdS boundary. Further, it has a discretized formulation as a tensor network ([22] argued that a canonical time slice in an AdS corresponds to a tensor network − a bookkeeping method to describe many body quantum state in terms of network entanglement, given by Multi-scale Entanglement Renormalisation Ansatz (MERA)). The MERA tensor network can be thought of a quantum circuit mapping an input lattice state to an output lattice state[10].

A measure of complexity in this setting is holographic complexity[2] which involves the surface used in computation of entanglement entropy and curvature, hence is purely geometric. Precisely, following Alishahiha[2], for subsystem $A$ in a time slice in the boundary theory, there exists a minimal co-dimension two hyper-surface $\gamma$ in the bulk such that the boundary of $A$ is same as its boundary. The holographic entanglement entropy is proportional to the surface area $\gamma$, and holographic complexity is proportional to the ratio the volume of bulk geometry enclosed by $\gamma$ and the curvature of the space-time.

In the setting of AdS/CFT, holographic complexity relates to evolution of Einstein-Rosen wormhole connecting two CFTs in entangled states described by the eternal AdS black hole[11][13]. An interesting perspective on holographic complexity comes from moving the geometric complexity ideas of Nielsen et al from quantum computation to quantum field theories. We very briefly summarize early ideas[14] in this direction next, before taking a short detour through connections between quantum information/complexity and black hole physics to conclude this survey.

## 4.1 Complexity in Quantum field theories

In [11] Jefferson and Myers consider the complexity of free scalar field in $d$ dimensions described by a lattice of simple harmonic oscillators (SHO). Starting with two coupled SHOs (and then passing to a $\delta$ width lattice), a universal set of operators based on natural operators appearing in SHO analysis (position and momentum operators), [11] considers complexity w.r.t. to the a reference state which is taken to be a factored Gaussian based on arguments from continuous MERA and holographic complexity framework. The analysis is simplified, by noting that the operator set takes Gaussians to Gaussians.

With this, the geometric complexity ideas of Nielsen et al can be directly applied. Comparing the geometric complexity of ground state of a free scalar field derived here, Jefferson and Myers note similarity with analogous computations in holographic complexity (with the caveat that the QFTs being compared are disparate: free scalar theory in this setting vs strongly coupled theory with large number of degrees of freedom in holography, and that similarity is conditional on choices made in geometric complexity calculations). Chapman et al, in [5], also take a similar approach to geometrizing complexity in quantum field theory (using Fubini-Study metric, so we directly work with a metric on the Hilbert space of wave functions instead of Jefferson and Myers' approach of passing to a matrix setting first; also see [21]) and again note similarities with holographic complexity, making it plausible that ideas of geometric complexity can indeed illuminate analysis of holograpy.

## 4.2 Firealls, ER=EPR and computational complexity

Consider what happens when a qubit falls into a black hole. The qubit reappears in scrambled form in Hawking radiation emitted by the black hole. Seemingly the qubit has two copies, one that fell in and one that was emitted. By black hole complimentarity, an external observer can describe everything unitarily without needing to know anything on the other side of the black hole event horizon [20]. This allows for sidestepping the possible violation of the no cloning in this setup known as the *Xeroxing problem*, as long

[12] https://en.wikipedia.org/wiki/Anti-de_Sitter_space

[13] also see [16], where these ideas were introduced

[14] as 2018, this is still a very new subject

as the observer only ever sees the emitted copy and not the one that fell in. But now the natural question becomes what if the external observer decodes the hawking radiation by applying a unitary and then jumps into the black hole? The external observer, Alice, will be able to observe both copies.

A more dramatic version of this scenario happens in a single emission of Hawking radiation when one part of a maximally entangled qubit is emitted as part of Hawking radiation and other part falls into the black hole [25]. For an "old" black hole, after the *Page* time[15], all Hawking radiation emitted is entangled with all Hawking radiation emitted earlier. But emitted particle is also entangled with in-falling particle, violating monogamy of entanglement. AMPS firewall [1, 9] proposal states that at event horizon the entanglement between in-falling and out going particle breaks leading to a large energy release.

Now vacuum in QFT is not empty, it is filled with fluctuating quantum fields. In particular, there is a lot of short range entanglement [1], so a lot of photons in maximally entangled state with another are always falling in; thus, at the event horizon, entanglement is constantly being broken, and this can be thought of firewall made up from energy released with each break.

An the alternative to firewall is the ER=EPR conjecture which says that the in-falling and out going entangled particles are connected by a wormhole (i.e. informally, the EPR state equals the Einstein-Rosen (ER) bridge, while more formally, the cross sectional area of the ER bridge between two entangled black holes relates to the entanglement entropy[15]) and are not independent: Susskind and Maldacena[15] conjecture that "after the scrambling time (but long before the Page time) the interior of the black hole is the Einstein-Rosen bridge system that connects the massively entangled near-horizon system of a black hole."

Consider Scott Aaaronson's description[1] of firewall paradox: Alice starts with $n$ qubits in a known initial state say all zero state, letting them collapse into a blackhole. The time at which $2n/3$ qubits have been radiated away by the black hole, but $n/3$ remain inside is by definition beyond the Page time (assuming nothing else interacts with the black hole). There are three systems: $B$ the next qubit coming out, $H$: qubits still in the black hole excluding $B$ and $R$ that were radiated away. Being beyond Page time $R$ is entangled with $B$. [1] argues that $B$ is also maximally entangled with some qubit in $H$ and that entanglement between $B$ and $H$ is observable. Now the firewall arises if Alice can observe entanglement of $B$ with both $H$ and $R$: Alice can observe $B$ and $R$ entanglement by applying a unitary to the $R$ qubits so that some qubit in $R$ ends up in an EPR state with $B$. This is the Harlow-Hayden decoding task:

**Theorem 6.** *Given as input a description of a quantum circuit $C$, which maps $n$ qubits from initial state $|0\rangle^{\otimes n}$ to a tripartite state $|\psi\rangle_{RBH}$ where $B$ is a single qubit. We're promised that there exists a unitary transformation $U$, acting only on $R$ part of $|\psi\rangle_{RBH}$, and which has the effect of putting $B$ and the rightmost qubit of $R$ into the joint EPR state. The challenge is to apply $U$ to the $R$ part of $|\psi\rangle_{RBH}$.*

Harrow and Hayden in [9] show that the decoding time scales exponentially in $n$. So the black hole would have evaporated before Alice can confirm the entanglement of $B$ and $R$. Aaronson interprets the result of Harrow and Hayden as: If the HH Decoding Task can be done in polynomial time for arbitrary circuits $C$, then $SZK \subset BQP$, tying a possible line of reasoning about black hole physics to relationship between quantum complexity classes.

# References

[1] Scott Aaronson. The Complexity of Quantum States and Transformations: From Quantum Money to Black Holes. *ArXiv e-prints*, page arXiv:1607.05256, July 2016.

[2] Mohsen Alishahiha. Holographic complexity. *Physical Review D*, 92:126009, December 2015.

[15] Page time: the time it takes a black hole to radiate away half of all information that will fall into it. For a black hole with a finite lifetime, i.e. time from when it forms to when it evaporates away, this time is finite; see https://en.wikipedia.org/wiki/Firewall_(physics) at [25].

[3] Boaz Barak and Arora Sanjeev. *Computational Complexity: A Modern Approach*. Cambridge University Press, 2007.

[4] Adam R. Brown, Leonard Susskind, and Ying Zhao. Quantum complexity and negative curvature. *Phys. Rev. D*, 95:045010, Feb 2017.

[5] Shira Chapman, Michal P. Heller, Hugo Marrochio, and Fernando Pastawski. Toward a Definition of Complexity for Quantum Field Theory States. *Physical Review Letters*, 120:121602, March 2018.

[6] Manfredo P. do Carmo. *Riemannian Geometry*. Birkhäuser, 2013.

[7] M. R. Dowling and M. A. Nielsen. The geometry of quantum computation. *eprint arXiv:quant-ph/0701004*, December 2007.

[8] Brian C. Hall. *Quantum Theory for Mathematicians*. Springer-Verlag New York, 2013.

[9] Daniel Harlow and Patrick Hayden. Quantum computation vs. firewalls. *Journal of High Energy Physics*, 2013:85, June 2013.

[10] Markus Hauru. Multiscale Entanglement Renormalisation Ansatz. Master's thesis, University of Helsinki, 2013.

[11] Robert A. Jefferson and Robert C. Myers. Circuit complexity in quantum field theory. *Journal of High Energy Physics*, 2017(10):107, Oct 2017.

[12] Navin Khaneja, Roger Brockett, and Steffen J. Glaser. Time optimal control in spin systems. *Physical Review A*, 63:032308, March 2001.

[13] John M. Lee. *Introduction to Smooth Manifolds*. Springer-Verlag New York, 2003.

[14] Henry W. Lin. Cayley graphs and complexity geometry. *ArXiv e-prints*, page arXiv:1808.06620, August 2018.

[15] J. Maldacena and L. Susskind. Cool horizons for entangled black holes. *Fortschritte der Physik*, 61:781–811, September 2013.

[16] Juan Maldacena. Eternal black holes in anti-de Sitter. *Journal of High Energy Physics*, 2003:021, April 2003.

[17] M. A. Nielsen, M. R. Dowling, M. Gu, and A. C. Doherty. Quantum Computation as Geometry. *Science*, 311:1133–1135, February 2006.

[18] Michael A. Nielsen. A geometric approach to quantum circuit lower bounds. *ArXiv e-prints*, pages quant–ph/0502070, February 2005.

[19] Michael A. Nielsen, Mark R. Dowling, Mile Gu, and Andrew C. Doherty. Optimal control, geometry, and quantum computing. *Physical Review A*, 73:062323, June 2006.

[20] Scott Aaronson. The Cryptographic Hardness of Decoding Hawking Radiation. https://www.scottaaronson.com/talks/hawking.ppt.

[21] Leonard Susskind. Three Lectures on Complexity and Black Holes. *arXiv e-prints*, page arXiv:1810.11563, October 2018.

[22] Brian Swingle. Entanglement renormalization and holography. *Physical Review D*, 86:065007, September 2012.

[23] L. M. Trefethen L. N., Trefethen. How many shuffles to randomize a deck of cards? *Mathematical, Physical and Engineering Sciences*, 456, Oct 2000.

[24] Loring W. Tu. *Differential Geometry*. Springer International Publishing, 2017.

[25] Wikipedia contributors. Wikipedia. https://en.wikipedia.org.