
Article:
Russell, Benjamin James and Stepney, Susan orcid.org/0000-0003-3146-5401 (2015) Zermelo Navigation in the Quantum Brachistochrone. Journal of Physics A: Mathematical and Theoretical. 115303. ISSN 1751-8113

https://doi.org/10.1088/1751-8113/48/11/115303

Zermelo Navigation in the Quantum Brachistochrone

Benjamin Russell, Susan Stepney

Department of Computer Science, University of York, UK, Y010 5DD

February 23, 2015

Abstract

We analyse the optimal times for implementing unitary quantum gates in a constrained finite dimensional controlled quantum system. The family of constraints studied is that the permitted set of (time dependent) Hamiltonians is the unit ball of a norm induced by an inner product on $\mathfrak{su}(n)$. We also consider a generalisation of this to arbitrary norms. We construct a Randers metric, by applying a theorem of Shen on Zermelo navigation, the geodesics of which are the time optimal trajectories compatible with the prescribed constraint. We determine all geodesics and the corresponding time optimal Hamiltonian for a specific constraint on the control i.e. $\kappa \operatorname{Tr}(\hat{H}_c(t)^2) = 1$ for any given value of $\kappa > 0$. Some of the results of Carlini et. al. are re-derived using alternative methods. A first order system of differential equations for the optimal Hamiltonian is obtained and shown to be of the form of the Euler Poincaré equations. We illustrate that this method can form a methodology for determining which physical substrates are effective at supporting the implementation of fast quantum computation.

1 Problem and Motivation

1.1 Implementation of Quantum Gates in Constrained Quantum Systems

We study the speed that a quantum system can implement a desired quantum gate. Here, all systems have pure states and finite dimensional Hilbert spaces associated to them.

This question has been discussed from many perspectives before, for example [1, 11, 20, 25, 30, 31, 36, 41, 46]. Notable recent works based on geometric methods are [11, 12, 45]. Our previous work [39] begins an investigation into the application of "Zermelo Navigation" to determining speed limits for implementing quantum gates in systems of the form eqn.(1), by a applying a solution in time optimal control based on Randers geometry [14]. We attempt to solve essentially the same problem as Carlini *et al* [12], but in a way based on intrinsic geometric structures in order to derive a first order equation for the optimal Hamiltonian driving the time evolution operator.

Caneva *et al* [10] address the behavior of a commonly applied numerical algorithm, the Krotov method, near the quantum speed limit [10]. Nielsen [35] has highlighted a connection between Finlser geometry and quantum optimal control. Furthermore, this work indicates an interesting connection between quantum circuit complexity and Finsler geometry. A fuller bibliography on the quantum speed limit more generally can be found in the introduction to [39].

In order to implement a certain quantum information processing (QIP) task in a controlled quantum system, we consider the dynamics of the system (more precisely, the time evolution operator $\hat{U}_t$) as given by the Schrödinger equation:

$$\frac{d\hat{U}_t}{dt} = -i\hat{H}_t\hat{U}_t = -i\left(\hat{H}_0 + \hat{H}_c(t)\right)\hat{U}_t \tag{1}$$

This is a standard form taken by the Schrödinger equation in the case of a quantum optimal control problem [46]. $\hat{H}_0$ is the 'drift' Hamiltonian, which represents the dynamics of the system in the absence of control fields; here we take it to be time independent throughout. $\hat{H}_c(t)$ is the control Hamiltonian, which represents the effect of control fields on the dynamics of $\hat{U}_t$. We say that a quantum system implements a gate $\hat{O}$ at time T if $\hat{U}_T = \hat{O}$ [12,39,45].

1.2 Central Results of the Paper

We show that, under the assumption that $\hat{H}_c(t)$ is constrained such that $h(i\hat{H}_c(t), i\hat{H}_c(t)) = 1$ for some inner product on $\mathfrak{su}(n)$, the time optimal trajectories of the time evolution operator $\hat{U}_t$ are exactly the geodesics of the following right invariant Randers metric on $SU(n)$:

$$F_{\hat{U}}(\hat{A}\hat{U}) = \sqrt{\frac{h(\hat{A},\hat{A})}{\lambda} + \frac{h(\hat{A}, i\hat{H}_0)^2}{\lambda^2}} + \frac{h(\hat{A}, i\hat{H}_0)}{\lambda} \tag{2}$$

where $\lambda = 1 - h(i\hat{H}_0, i\hat{H}_0)$.

We show that, under the assumption that $\hat{H}_c(t)$ is constrained such that $\kappa \operatorname{Tr}(\hat{H}_c(t)^2) = 1$ (that is, choosing h to be some multiple of the Killing form of $\mathfrak{su}(n)$), the time optimal trajectories are given by:

$$\hat{V}_t = \exp(-it\hat{H}_0)\exp(it\hat{D})$$

for some $i\hat{D}$ (depending on the desired gate) in $\mathfrak{su}(n)$ such that $h(i\hat{D}, i\hat{D}) = \kappa \operatorname{Tr}(\hat{D}^2) = 1$.

We show that, under the same constraint, the optimal control Hamiltonian is:

$$\hat{H}_c(t) = -\exp(-it\hat{H}_0)\hat{D}\exp(it\hat{H}_0) \tag{3}$$

We give a series expansion based on the Baker–Campbell–Hausdorff formula for $\hat{D}$ in terms an arbitrary desired gate $\hat{O} \in SU(n)$.

We present a set of equations (88) which determine the time optimal control Hamiltonian for an arbitrary gate in any situation where $\check{F}(i\hat{H}_c(t)) = 1$ for an arbitrary Minkowski norm $\check{F}$ on $\mathfrak{su}(n)$ in scenarios where $\check{F}(i\hat{H}_0) \leq 1$ for that same norm.

2 Speed Limits in Constrained Systems

In this section we outline the problem of Zermelo navigation on a manifold. We illustrate how a Randers metric, which depends on the 'drift' dynamics of a controlled dynamical system on a manifold and the driftless speed of the navigator, has the property that its geodesics are time optimal trajectories for navigating on the same manifold. We discuss Shen's theorem [42] which provides an explicit formula for this Randers metric. We are motivated by the desire to determine time optimal trajectories in quantum systems with constraints in novel ways. This method builds on previous work as it allows one to obtain the optimal time for *arbitrary* trajectories, as well as determining theoretically optimal ones. This is practically relevant as it is not always the case that all trajectories can be practically realised.

In the following we use the notation T_pM to refer to the tangent space at the point p on a manifold M, and TM to refer to the entire tangent bundle of a manifold, that is, all tangent spaces considered together. For example $T_{\hat{U}}SU(n)$ is the tangent space to $SU(n)$ at the point $\hat{U} \in SU(n)$. We use the notation $\Gamma(TM)$ to refer to the set of all smooth sections of the tangent bundle TM, that is, effectively to say, all smooth vector fields on M. For good references using this notation for both general manifolds and Lie groups, see [9,30].

2.1 Randers Metrics and Shen's Theorem on Zermelo Navigation

2.1.1 Definition of a Finsler Metric

A Finsler metric on a manifold M, of which a Randers metric is a special case, is a type of metric structure generalising a Riemannian metric. This generalisation is essentially that the norm induced on each tangent space does not need to be induced by an inner product. Technically, a Finsler metric on a smooth manifold M is a smoothly varying 'Minkowski norm' on each tangent space on M.

Definition A *Minkowski Norm* on a vector space V is a function $|\cdot| : V \to \mathbb{R}$ which satisfies the following axioms:

- $\forall v \in V$, $|v| \geq 0$. $|v| = 0 \Leftrightarrow v = 0$. Positive definite.
- $\forall v \in V, \forall \lambda \in \mathbb{R}^+$, $|\lambda v| = \lambda |v|$. Positive homogeneous.
- $\forall u, v \in V, |u + v| \leq |v| + |u|$. Subadditive.
- $\forall v, x, y \in V$ the Hessian of $|\cdot|^2$ is positive definite. That is: $\frac{1}{2} \frac{\partial^2}{\partial s \partial t} |v + tx + sy|^2 \Big|_{t=0, s=0} \geq 0$

The only Minkowski norms that are true norms are exactly those that are reversible: $|-v| = |v|$ $\forall v \in V$. For a detailed discussion of this definition see [2, 14]. See [23] for an interesting and thorough discussion of variations of Finsler metrics in use in applications.

2.1.2 Zermelo Navigation

We note a type of metric for which the lengths of any curve is the optimal traversal time rather than just a bound, namely Randers metrics. A Randers metric is a specific case of a Finsler metric [14]. A Randers metric on a manifold M is a Finsler metric on M which is a Randers norm on each tangent space.

Definition A *Randers* norm on a real, finite dimensional vector space $\mathbb{R}^N$ is a map $|\cdot| : \mathbb{R}^N \to \mathbb{R}$ which can be written as follows:

$$|v| = \sqrt{\alpha(v, v)} + \beta(v) \tag{4}$$

where α is an inner product on $\mathbb{R}^N$ and β is a one form.

Definition Subsequently, a *Randers metric* on a manifold M can always be written as a Finlser metric with fundamental function $F_p : T_pM \to \mathbb{R}$ of the following form:

$$F_p(v) = \sqrt{\alpha_p(v, v)} + \beta_p(v) \tag{5}$$

for any $v \in T_pM$, where α is a Riemannian metric in the usual sense and β is a differential one form.

A theorem due to Shen relates the problem of Zermelo navigation on a Riemannian manifold (M, h) in the presence of vector field W (such that $h_p(W_p, W_p) \leq 1, \forall p \in M$) to the geodesics of Randers metrics [14, Ch.2]. The results of Shen [42] completely classify Randers metrics and demonstrate that they are in one to one correspondence with Zermelo navigation problems on Riemannian manifolds.

The problem of Zermelo navigation on a Riemannian manifold can be stated as follows [3]. Given:

- A Riemannian manifold (M, h)
- A smooth vector field $W \in \Gamma(TM)$ such that $h_p(W_p, W_p) \leq 1$, $\forall p \in M$

determine the time optimal trajectories to follow in order to navigate from an arbitrary given initial point p_I to an arbitrary given terminal point p_T on M, where the navigator is restricted such that its speed (according to h) in the absence of wind (i.e. $W = 0$) is exactly 1.

Shen's theorem can be stated as follows: the time optimal trajectories are exactly the geodesics of the Randers metric F defined by the following equations:

$$\begin{aligned} F_p(v) &= \sqrt{\alpha_p(v,v)} + \beta_p(v) \\ \alpha_p(u,v) &= \frac{h_p(u,v) + \lambda_p \beta_p(u)\beta_p(v)}{\lambda_p} \\ \beta_p(v) &= -\frac{h_p(v, W_p)}{\lambda_p} \\ \lambda_p &= 1 - h_p(W_p, W_p) \end{aligned} \tag{6}$$

where these definitions hold $\forall u, v \in T_pM, \forall p \in M$. The optimal times are the geodesic lengths of the same metric. (M, h) and W together are know as the navigation data for the specific instance of Zermelo navigation. For a full derivation of these facts and much more detailed discussion providing stronger intuitive aids, see [14, Ch.2]. For a recent application to quantum computing see [39].

2.2 Right Invariance and Right Extension

It is well known that $SU(n)$ is a compact, connected Lie group. Here we state some standard definitions from Lie theory (see for example [9, 18, 26]), relevant for our later derivations.

Definition A function $F : TSU(n) \to \mathbb{R}$ is said to be *Right Invariant* if the following holds:

$$F_{R_{\hat{V}}(\hat{U})}(dR_{\hat{V}}\big|_{\hat{U}}(\hat{A})) = F_{\hat{U}}(\hat{A}) \tag{7}$$

$\forall \hat{U}, \hat{V} \in SU(n), \forall \hat{A} \in T_{\hat{U}}SU(n)$. As $SU(n)$ is a linear algebraic group (or LAG) [22] this definition simplifies to:

$$F_{\hat{U}\hat{V}}(\hat{A}\hat{V}) = F_{\hat{U}}(\hat{A}) \tag{8}$$

Definition A function $F : TSU(n) \to \mathbb{R}$ is said to be the *Right Extension* of $G : T_{\hat{I}}SU(n) \to \mathbb{R}$ if the following holds:

$$F_{\hat{U}}(\hat{A}) = F_{\hat{I}}(\hat{A}\hat{U}^{-1}) \tag{9}$$

One readily checks that right extension always leads to a right invariant function on TM.

This construction allows one to construct all right invariant Randers metrics on $SU(n)$ as they are in one to one correspondence with pairs of the form (α, β) wherein $\alpha : \mathfrak{su}(n) \times \mathfrak{su}(n) \to \mathbb{R}$ is an inner product on $\mathfrak{su}(n)$ (when identified with the tangent space at the identity) and $\beta : \mathfrak{su}(n) \to \mathbb{R}$ is a one form on $\mathfrak{su}(n)$. α and β can be combined into a Randers norm $F : \mathfrak{su}(n) \to \mathbb{R}$ by setting $F(\hat{A}) = \sqrt{\alpha(\hat{A}, \hat{A})} + \beta(\hat{A})$. Now the metric can be set to be the right extension of this function to the entire tangent bundle of $SU(n)$ as follows:

$$F_{\hat{U}}(\hat{A}\hat{U}) = F_{\hat{I}}(\hat{A}) = \sqrt{\alpha(\hat{A}, \hat{A})} + \beta(\hat{A}) \tag{10}$$

One can also easily check that F is a smooth function on $TM/\{0\}$ (the slit tangent bundle) in the required sense to be a Finsler metric. This follows from the fact that right translation in a Lie group is smooth as the group multiplication operation is smooth.

One could equally right translate α and β before adding them, but this would yield the same result. It it readily checked that a Randers metric is right invariant exactly when its Riemannian and linear parts are right invariant individually. This fact precludes the possibility of a bi-invariant Randers metric on $SU(n)$ as there are no bi-invariant one forms on $SU(n)$.

2.3 Shen's Theorem Applied to Right Invariant Metrics on $SU(n)$ and Quantum Mechanics

We now set up the problem of Zermelo navigation on $SU(n)$ and show how it can be applied to quantum mechanics. Suppose that $h : \mathfrak{su}(n) \times \mathfrak{su}(n) \to \mathbb{R}$ is an inner product on $\mathfrak{su}(n)$. Suppose that a controlled quantum system of the form eqn.(1) is constrained such that its control Hamiltonian $\hat{H}_c(t)$ satisfies $h(i\hat{H}_c(t), i\hat{H}_c(t)) = 1, \forall t$. That is, the constraint is time independent and satisfied for all time. It is clear that the right invariance of the Riemannian metric $h_{\hat{U}}$ (formed by right extending h) corresponds to a constraint only on $\hat{H}_t$, not any constraint depending on $\hat{U}_t$ explicitly. This argument is very similar to the argument for the requirement of right invariance presented in [35].

Suppose further that a drift Hamiltonian $\hat{H}_0$ is given. The time evolution operator for our quantum system now satisfies a Schrödinger equation of the form eqn.(1). The tangent vector $\frac{d\hat{U}_t}{dt}$ to the curve $\hat{U}_t$ has two terms: $-i\hat{H}_c(t)\hat{U}_t$ and $-i\hat{H}_0\hat{U}_t$. In order to fix terminology closer to the original formulation of the Zermelo navigation problem, we define the "wind" vector field on $SU(n)$ by $\hat{W}_{\hat{U}} = -i\hat{H}_0\hat{U}$.

In such a setup, there is enough information to construct the "navigation data" for a Zermelo navigation problem. From these ingredients one can construct a Finsler metric (which is in fact a Randers metric) that has the property that its geodesics are the time optimal trajectories for $\hat{U}_t$ to be driven between given endpoints, by applying Shen's theorem, eqn.(6). This Randers metric is $F_{\hat{U}}(\hat{A}) = \sqrt{\alpha_{\hat{U}}(\hat{A}, \hat{A})} + \beta_{\hat{U}}(\hat{A})$. In terms of the navigation data $(h, \hat{W})$ on $SU(n)$, the α and β are found to be:

$$\begin{aligned}\alpha_{\hat{U}}(\hat{A}\hat{U}, \hat{A}\hat{U}) &= \frac{\lambda h_{\hat{U}}(\hat{A}\hat{U}, \hat{A}\hat{U}) + h_{\hat{U}}(\hat{A}\hat{U}, \hat{W}_{\hat{U}})^2}{\lambda^2} \\ &= \frac{h_{\hat{U}}(\hat{A}\hat{U}, \hat{A}\hat{U})}{\lambda} + \frac{h_{\hat{U}}(\hat{A}\hat{U}, -i\hat{H}_0\hat{U})^2}{\lambda^2} \\ &= \frac{h(\hat{A}, \hat{A})}{\lambda} + \frac{h(\hat{A}, i\hat{H}_0)^2}{\lambda^2}\end{aligned} \tag{11}$$

$$\beta_{\hat{U}}(\hat{A}\hat{U}) = \frac{h_{\hat{U}}(\hat{A}\hat{U}, -i\hat{H}_0\hat{U})}{1 - h_{\hat{U}}(-i\hat{H}_0\hat{U}, -i\hat{H}_0\hat{U})} = \frac{h(\hat{A}, -i\hat{H}_0)}{1 - h(i\hat{H}_0, i\hat{H}_0)} = \frac{h(\hat{A}, -i\hat{H}_0)}{\lambda} \tag{12}$$

Thus F is, in full, given by:

$$F_{\hat{U}}(\hat{A}\hat{U}) = \sqrt{\frac{h(\hat{A}, \hat{A})}{\lambda} + \frac{h(\hat{A}, i\hat{H}_0)^2}{\lambda^2}} + \frac{h(\hat{A}, -i\hat{H}_0)}{\lambda} = F(\hat{A}) \tag{13}$$

As stated above, the right invariance of this quantity is clear. Here $\lambda_{\hat{U}} = 1 - h_{\hat{U}}(\hat{A}\hat{U}, \hat{A}\hat{U}) = 1 - h(\hat{A}, \hat{A})$, thus we note that λ is a scalar quantity because it is right invariant, and that all right invariant scalar quantities are constant. This is a simplifying factor of the case when h and W are right invariant compared to the general case.

2.4 The Length Functional when F is Right Invariant

In Riemannian geometry, each Riemannian metric defines a length functional on the space of all curves on a Riemannian manifold. The situation in Finsler geometry is essentially identical.

The length functional $L[\hat{U}_t]$ for a curve $\hat{U}_t : [0, T] \to SU(n)$, for a Finlser metric can be written as follows:

$$L[\hat{U}_t] = \int_{t=0}^{T} F_{\hat{U}_t}\left(\frac{d\hat{U}_t}{dt}\right) dt \tag{14}$$

In the case that F is is right invariant one finds:

$$L[\hat{U}_t] = \int_{t=0}^{T} F\left(\frac{d\hat{U}_t}{dt}\hat{U}_t^{-1}\right) dt$$

In the case that $\hat{U}_t$ solves the Schrödinger equation (1) one finds:

$$\begin{aligned} L[\hat{U}_t] &= \int_{t=0}^{T} F_{\hat{U}_t}\left(\frac{d\hat{U}_t}{dt}\right) dt = \int_{t=0}^{T} F_{\hat{U}_t}\left(-i\hat{H}(t)\hat{U}_t\right) dt \\ &= \int_{t=0}^{T} F\left(-i\hat{H}(t)\right) dt = \int_{t=0}^{T} F\left(-i\hat{H}_0 - i\hat{H}_c(t)\right) dt \end{aligned} \tag{15}$$

The length $L[\hat{U}_t]$ depends only on quantities in $\mathfrak{su}(n)$ rather than on the group in general, as all dependence on $\hat{U}_t$ itself has disappeared. It is possible to formulate the geodesic equation for such a Finsler metric as an ODE in $\mathfrak{su}(n)$, and leads to an equation for the time optimal control Hamiltonian for a controlled quantum system with the type of constraint discussed above.

2.5 Geodesics of Right Invariant Randers Metrics on $SU(n)$

This shows that the desired time optimal trajectories required are the geodesics of a right invariant Randers metric on $SU(n)$. In order to find these geodesics we must determine the extremal curves of the length functional for a Randers metric:

$$L[\hat{U}_t] = \int_0^T \left[\sqrt{\alpha_{\hat{U}_t}\left(\frac{d\hat{U}_t}{dt}, \frac{d\hat{U}_t}{dt}\right)} + \beta_{\hat{U}_t}\left(\frac{d\hat{U}_t}{dt}\right)\right] dt \tag{16}$$

under the assumption of right invariance.

This could be achieved via the usual Euler–Lagrange (EL) equations. However another method exists that exploits the right invariance of the metric. By considering the quantity $\frac{d\hat{U}_t}{dt}\hat{U}_t^{-1} = \frac{d\hat{U}_t}{dt}\hat{U}_t^{\dagger}$, we can determine a first order differential equation for its value along a geodesic, the Euler–Poincaré (EP) equations ([13,19]; see [27] for an application to optimal control) by the procedure of Lagrangian reduction. Furthermore, we have $\frac{d\hat{U}_t}{dt}\hat{U}_t^{\dagger} = -i\hat{H}_t$, by applying the Schrödinger equation for any potentially time dependent quantum system. These yield a first order equation for the Hamiltonian which drives the system along a geodesic of any given Randers metric.

The geodesics of Randers metrics have already been seen to be the time optimal trajectories for the relevant constraint, so the EP equations corresponding to a Randers metric on $SU(n)$ are a first order system of equations for the Hamiltonian driving $\hat{U}_t$ along a time optimal trajectory. This illustrates an alternative understanding of the origin of the first order "Brachistochrone" equation [12], at least in a special case. Applying boundary conditions to such an equation allows us to obtain the Hamiltonian that drives $\hat{U}_t$ from $\hat{I}$ to a desired operator (i.e. a specific quantum gate) in the least time.

In a coordinate-free language (where $\xi \in \mathfrak{su}(n)$) the EP equation reads [13, 19]:

$$\frac{d}{dt}\frac{\partial \ell}{\partial \xi} = -\mathrm{ad}_{\xi}^{*}\left(\frac{\partial \ell}{\partial \xi}\right) \tag{17}$$

where $\ell : \mathfrak{su}(n) \to \mathbb{R}$ is the restriction of an arbitrary right invariant Lagrangian $\mathcal{L} : TSU(n) \to \mathbb{R}$ to $\mathfrak{su}(n)$, and ad^* is the co-adjoint representation of $\mathfrak{su}(n)$ [5]. Note the minus sign here: this is due to the metric here being right invariant rather than left invariant as is more commonly studied in pure mathematics contexts such as [19].

There are some additional conditions on $\mathcal{L}$ for the EP equations to apply; these can be readily found in any mathematical description of the theory of Lagrangian reduction [13, 27]. It is clear that all Finsler metrics meet the required conditions. For example, it is clear that the regularity condition is met, as it is present in the definition of a Finsler metric.

This equation may also been seen with a δ (signifying a functional derivative) in place of the d above; this is the form of the equation which applies to infinite dimensional problems rather than the finite dimensional ones studied here.

On fixing a basis $\{\hat{B}_k\}$ for $\mathfrak{su}(n)$ and expressing an arbitrary element $-i\hat{H}_t$ as $\xi^k\hat{B}_k$, the EP equation takes the form [13, 19]:

$$\frac{d}{dt}\frac{\partial \ell}{\partial \xi^d} = -C^b_{ad}\frac{\partial \ell}{\partial \xi^b}\xi^a \tag{18}$$

where C^d_{ab} are the structure constants of $\mathfrak{su}(n)$. See [18] for details of structure constants in general, and [37] for $\mathfrak{su}(n)$ specifically, where the structure constants of $\mathfrak{su}(2)$ and $\mathfrak{su}(4)$ are given explicitly. The tensor C possesses many symmetries, including $C^a_{bd} = -C^a_{db}$ for example; this follows directly from the antisymmetry of the Lie bracket. The use of these symmetries as a tool for simplifying the EP equations in the case of $\mathfrak{su}(n)$ will be included in further work.

Henceforth the subscripts indicating a point on $SU(n)$ are dropped from α and β, and they are understood to be restricted to the tangent space of $SU(n)$ at the identity, i.e. $\mathfrak{su}(n)$. However, coordinate indicies still appear.

Setting ℓ to be the square (to obtain unit speed geodesics) Randers norm $\ell(\hat{A}) = \frac{1}{2}(F\big|_{\hat{I}}(\hat{A}))^2 = \frac{1}{2}\left(\sqrt{\alpha(\hat{A},\hat{A})} + \beta(\hat{A})\right)^2$, i.e. the restriction of a Randers metric F on $SU(n)$ to $\mathfrak{su}(n)$, we can derive the EP equation associated to the geodesics of F.

Substituting:

$$\frac{\partial \ell}{\partial \xi^d} = \left((\alpha_{ij}\xi^i\xi^j)^{1/2} + \beta_k\xi^k\right)\left(||\xi||_\alpha^{-1}\alpha_{nd}\xi^n + \beta_d\right) \tag{19}$$

Differentiating:

$$\begin{aligned}\frac{d}{dt}\left(\frac{\partial \ell}{\partial \xi^d}\right) &= \frac{d}{dt}\left(||\xi||_\alpha + \beta_k\xi^k\right)\left(||\xi||_\alpha^{-1}\alpha_{nd}\xi^n + \beta_d\right) \\ &= \left(||\xi||_\alpha^{-1}\langle\dot{\xi},\xi\rangle_\alpha + \beta_j\dot{\xi}^j\right)\left(||\xi||_\alpha^{-1}\alpha_{md}\xi^m + \beta_d\right) - \\ &\quad \left(||\xi||_\alpha + \beta_k\xi^k\right)\left(||\xi||_\alpha^{-3}\langle\xi,\dot{\xi}\rangle_\alpha\alpha_{hd}\xi^h - ||\xi||_\alpha^{-1}\alpha_{kd}\dot{\xi}^k\right)\end{aligned} \tag{20}$$

These yield the EP equation of a geodesic:

$$\begin{aligned}&\left(||\xi||_\alpha^{-1}\langle\dot{\xi},\xi\rangle_\alpha + \beta_j\dot{\xi}^j\right)\left(||\xi||_\alpha^{-1}\alpha_{md}\xi^m + \beta_d\right) - \\ &\left(||\xi||_\alpha + \beta_k\xi^k\right)\left(||\xi||_\alpha^{-3}\langle\xi,\dot{\xi}\rangle_\alpha\alpha_{hd}\xi^h - ||\xi||_\alpha^{-1}\alpha_{kd}\dot{\xi}^k\right) \\ &= -C^a_{bd}\left(\left(||\xi||_\alpha + \beta_k\xi^k\right)\left(||\xi||_\alpha^{-1}\alpha_{na}\xi^n + \beta_a\right)\right)\xi^b\end{aligned} \tag{21}$$

where we take the following meanings: $||\xi||_\alpha := \sqrt{\alpha_{ij}\xi^i\xi^j}$ and $\langle\mu,\nu\rangle_\alpha := \alpha_{ij}\mu^i\nu^j$.

We are interested in the geodesics associated to a navigation problem specified in terms of its navigation data. Such an equation can be obtained by substituting in the definitions of α and β in terms of h and W from Shen's solution to the navigation problem. This is an elementary but somewhat tedious computation that provides little insight, so is omitted here. Many numerical/approximate methods exist for solving this type of first order ODE. Thus, many practical approaches could be taken to solving for the optimal Hamiltonian after obtaining the equation in terms of h and $\hat{O}$. We leave obtaining this final equation to the reader.

2.6 Applications of Killing Fields to Determining Geodesics

Choose $h(\hat{A}, \hat{A}) = \kappa \operatorname{Tr}(\hat{A}^\dagger \hat{A})$ (i.e. a constant positive multiple κ of the Killing form); then the right extension of h is the unique bi-invariant metric. This is essentially the case studied in [12,39]. It exhibits a simplifying factor pertaining to the task of determining geodesics. The geodesics can be determined by an application of a special case of Robles [38, thm.2]. We use $\sigma = 0$ in that theorem, as the special case of a Killing field (see [21] for definitions) in place of the infinitesimal homothety. Also, we specialise to $SU(n)$, rather than a general manifold, as this is the case relevant to quantum mechanics. In fact, there are no infinitesimal homotheties that are not Killing fields for the bi-invariant metric on $SU(n)$, so no real restriction has been incurred on which metrics can have their geodesics determined using the following theorem. The theorem states:

Theorem 2.1 (adapted from [38, thm.2]) *Given:*

- *A Riemannian manifold* (M, h)
- *A smooth vector field* $\hat{W} \in \Gamma(TSU(n))$ *on* $SU(n)$ *such that* $\pounds_{\hat{W}}(h) = 0$ *(that is, the Lie derivative of the metric is* 0*, or equivalently* $\hat{W}$ *is a Killing vector field).*

Given that F *is the Randers metric solving the Zermelo navigation problem on* M *for* h *and* $\hat{W}$*, then the unit* F *speed geodesics of* F *are given by* $\hat{V}_t = \phi_t(\hat{S}_t)$*, where:*

- ϕ_t *is the flow associated to* $\hat{W}$
- $\hat{S}$ *is a unit speed geodesic of* h

Any geodesic of F obtained this way is a length minimiser if and only if the associated Riemannian geodesic of h is a length minimiser of h [38].

In the case that h is the bi-invariant metric, the unit speed geodesics $\hat{S}_t$ are the one parameter subgroups of $SU(n)$, parameterised to have unit h speed. These can all be expressed as $\hat{S}_t = \exp(it\hat{D})$ for some $\hat{D} \in \mathfrak{su}(n)$ that is a unit vector for the same h. The flow associated to the vector field $\hat{W}_{\hat{U}} = -i\hat{H}_0\hat{U}$ is $\phi_t(\hat{U}) = \exp(-it\hat{H}_0)\hat{U}$. This follows from the observation that the equation defining the flow is exactly the Schrödinger equation with Hamiltonian $\hat{H}_0$. We thus conclude that the time optimal trajectories are given by:

$$\begin{aligned} \hat{U}_t = \phi_t(\hat{S}_t) &= \phi_t\left(\exp(it\hat{D})\right) \\ &= \exp(-it\hat{H}_0)\exp(it\hat{D}) \end{aligned} \tag{22}$$

This is to be compared with [12, eqn.51] which exhibits a similar product of exponential structure.

We determine the optimal Hamiltonian by assuming $\hat{U}_t$ solves the Schrödinger equation for an as yet unknown Hamiltonian $\hat{H}_t$:

$$\frac{d\hat{U}_t}{dt} = -i\hat{H}_t\hat{V}_t \tag{23}$$

which implies that:

$$\begin{aligned} \hat{H}_t &= i\frac{d\hat{U}_t}{dt}\hat{U}_t^\dagger \\ &= i\left((-i\hat{H}_0)\hat{V}_t + \hat{V}_t(i\hat{D})\right)\hat{V}^\dagger \\ &= i\left((-i\hat{H}_0) + \hat{V}_t(i\hat{D})\hat{V}^\dagger\right) \\ &= \hat{H}_0 - \hat{V}_t(\hat{D})\hat{V}^\dagger \\ &= \hat{H}_0 - \exp(-it\hat{H}_0)(\hat{D})\exp(it\hat{H}_0) \\ &= \hat{H}_0 + i\mathrm{Ad}_{\exp(-it\hat{H}_0)}(i\hat{D}) \end{aligned} \tag{24}$$

In order to conclude that these are the geodesics and their associated Hamiltonians, we check that the given $\hat{W}_{\hat{U}}$ is a Killing field for the metric h. This is achieved by checking $\pounds_{\hat{W}}(h) = 0$ thus:

$$\begin{aligned}
&\frac{d}{dt} h_{\phi_t(\hat{V})} \left(d\phi_t\big|_{\hat{V}}(\hat{A}\hat{V}), d\phi_t\big|_{\hat{V}}(\hat{A}\hat{V}) \right) \Big|_{t=0} \\
&= \frac{d}{dt} h_{\exp(-it\hat{H}_0)\hat{V}} \left(\exp(-it\hat{H}_0)\hat{A}\hat{V}, \exp(-it\hat{H}_0)\hat{A}\hat{V} \right) \Big|_{t=0} \\
&= \frac{d}{dt} h_{\exp(-it\hat{H}_0)} \left(\exp(-it\hat{H}_0)\hat{A}, \exp(-it\hat{H}_0)\hat{A} \right) \Big|_{t=0} \\
&= \frac{d}{dt} h_{\hat{I}}(\hat{A}, \hat{A}) \Big|_{t=0} = 0
\end{aligned} \tag{25}$$

where $\hat{V}$ is an arbitrary group element and $\hat{A}\hat{V}$ is an arbitrary element of $T_{\hat{V}}SU(n)$. We have $d\phi_t\big|_{\hat{V}}(\hat{A}\hat{V}) = \exp(-it\hat{H}_0)\hat{A}\hat{V}$ trivially, as it is the differential of a linear map. Here, both the left and the right invariance of the metric h have been appealed to; this proof would need to be modified, or may not hold, in the case that h is not the unique bi-invariant metric.

We conclude (by subtracting $\hat{H}_0$) that the control Hamiltonian $\hat{H}_c(t)$ driving a system (which meets the required premises) along a time optimal trajectory is given by:

$$\hat{H}_c(t) = -\exp(-it\hat{H}_0)\hat{D}\exp(it\hat{H}_0) \tag{26}$$

This is constant exactly when $\hat{D}$ commutes with $\hat{H}_0$, which, from eqn.(28), is equivalent to saying that $\hat{H}_0$ commutes with $\hat{O}$. Thus we have obtained a necessary and sufficient condition for the time optimality of constant control fields for any system meeting the premises of the above derivation. This condition neatly sidesteps the need for the analysis of "homogenous geodesics" (see [28,29] for much interesting mathematical discussion and [33,34] for more direct physical applications of the concept) of right invariant Randers metrics, which could pose significant mathematical challenges. See section (2.10) for a fuller discussion of this topic of homogeneous geodesics.

What remains to determine is the formula for a geodesic with desired endpoints (connecting the identity $\hat{I}$ to a desired operator $\hat{O} \in SU(n)$) and the corresponding Hamiltonian. This boils down to determining the $\hat{D}$ corresponding to a given $\hat{O} \in SU(n)$. To determine which $\hat{D}$ yields the geodesic with endpoints $\hat{I}$ and $\hat{O}$ such that the system traverses the geodesic in time T, we need to solve:

$$\hat{U}_T = \exp(-iT\hat{H}_0)\exp(iT\hat{D}) = \hat{O} \tag{27}$$

Rearranging and taking logs:

$$\exp(iT\hat{D}) = \exp(iT\hat{H}_0)\hat{O} \tag{28}$$

$$i\hat{D} = \frac{1}{T}\log\left(\exp(iT\hat{H}_0)\hat{O}\right) \tag{29}$$

which yields the desired geodesic and corresponding control Hamiltonian:

$$\begin{aligned}
\hat{U}_t &= \exp(-it\hat{H}_0)\exp\left(\frac{t}{T}\log\left(\exp(iT\hat{H}_0)\hat{O}\right)\right) \\
&= \exp(-it\hat{H}_0)\left(\exp(iT\hat{H}_0)\hat{O}\right)^{t/T}
\end{aligned} \tag{30}$$

$$\begin{aligned}
\hat{H}_c(t) &= \frac{i}{T}\exp(-it\hat{H}_0)\log\left(\exp(iT\hat{H}_0)\hat{O}\right)\exp(it\hat{H}_0) \\
&= \frac{i}{T}\log\left(\exp(-it\hat{H}_0)\exp(iT\hat{H}_0)\hat{O}\exp(it\hat{H}_0)\right) \\
&= \frac{i}{T}\log\left(\exp\left(i(T-t)\hat{H}_0\right)\hat{O}\exp(it\hat{H}_0)\right)
\end{aligned} \tag{31}$$

We can take the $\exp(\pm it\hat{H}_0)$ factors inside the logarithm, because the matrix logarithm is analytic [44, Ch.7], which follows from the fact that any matrix function f which is defined by a power series obeys $f(\hat{V}^{-1}\hat{A}\hat{V}) = \hat{V}^{-1}f(\hat{A})\hat{V}$ for all matrices $\hat{A}$ and all non singular $\hat{V}$.

2.7 Optimal Times

Now we attempt to find T_{opt}, the optimal time for implementing some given $\hat{O}$. Insisting that the left hand side of eqn.(28) has norm 1 according to h, that $h(i\hat{D}, i\hat{D}) = 1$, in accordance with the premise that it is the *unit speed* geodesics of h that are needed, we determine that:

$$1 = h(i\hat{D}, i\hat{D}) = h\left(\frac{1}{T}\log\left(\exp(iT\hat{H}_0)\hat{O}\right), \frac{1}{T}\log\left(\exp(iT\hat{H}_0)\hat{O}\right)\right) \tag{32}$$

which yields the following equation to be solved for T:

$$-\frac{\kappa}{T^2}\operatorname{Tr}\left(\left[\log\left(\exp(iT\hat{H}_0)\hat{O}\right)\right]^2\right) = 1 \tag{33}$$

The smallest positive solution that is truly the optimal time; we refer to this as T_{opt}. At the time of writing, we have found no method for solving this analytically in general; it appears prohibitively difficult by standard means known to the authors. However, once $\hat{H}_0$ and $\hat{O}$ are given, it can easily be solved numerically; some simple cases are illustrated in §3.

One special case that can be solved analytically is where $\hat{O}$ and $\hat{H}_0$ commute. Expanding the matrix logarithm using $\log(\hat{A}\hat{B}) = \log(\hat{A}) + \log(\hat{B})$, rearranging and applying the standard quadratic formula gives:

$$T_{\text{opt}} = \frac{i\kappa\operatorname{Tr}(\hat{H}_0\log(\hat{O}))}{\kappa\operatorname{Tr}(\hat{H}_0^2) - 1} \pm \sqrt{\frac{\kappa\operatorname{Tr}((\log(\hat{O}))^2)}{\kappa\operatorname{Tr}(\hat{H}_0^2) - 1} - \frac{\kappa^2(\operatorname{Tr}(\hat{H}_0\log(\hat{O}))^2)}{(\kappa\operatorname{Tr}(\hat{H}_0^2) - 1)^2}} \tag{34}$$

where, as in [39], the $\pm$ is chosen to ensure a positive time.

Once T_{opt} is known, either analytically or numerically, then the true geodesics and corresponding optimal control Hamiltonian are:

$$\begin{aligned}\hat{U}_t &= \exp(-it\hat{H}_0)\left(\exp(iT_{\text{opt}}\hat{H}_0)\hat{O}\right)^{t/T_{\text{opt}}} \\ \hat{H}_c(t) &= \frac{i}{T_{\text{opt}}}\log\left(\exp\left(i(T_{\text{opt}} - t)\hat{H}_0\right)\hat{O}\exp(it\hat{H}_0)\right)\end{aligned} \tag{35}$$

We can use the well-known BCH formula [18, §3] to evaluate approximations to $i\hat{D}$ as it provides a series type representation for the solution to $\exp(z) = \exp(x)\exp(y)$. Given a certain x, y in a Lie algebra the solution for z is given by:

$$z = x + y + \frac{1}{2}[x, y] + \frac{1}{12}[x, [x, y]] - \frac{1}{12}[y, [x, y]] - \frac{1}{24}[y, [x, [x, y]]] + \cdots \tag{36}$$

We apply this to eqn.(28) to solve $\exp(iT\hat{D}) = \exp(iT\hat{H}_0)\exp(\log(\hat{O}))$, to obtain:

$$\begin{aligned}i\hat{D} = i\hat{H}_0 + \frac{1}{T}\log\hat{O} + \frac{1}{2}[i\hat{H}_0, \log(\hat{O})] + \frac{T}{12}[i\hat{H}_0, [i\hat{H}_0, \log(\hat{O})]] \\ - \frac{1}{12}[\log\hat{O}, [i\hat{H}_0, \log\hat{O}]] - \frac{T}{24}[\log(\hat{O}), [i\hat{H}_0, [i\hat{H}_0, \log(\hat{O})]]] + \cdots\end{aligned} \tag{37}$$

2.8 Scope of Our Approach

The choice that h is the bi-invariant metric in the above derivation of the geodesics given in eqn.(3) allows us to exploit the fact that the geodesics of this metric are the one parameter subgroups. If a different right

invariant metric $h_{\hat{U}}$ were chosen (that is, a different physical constraint on $\hat{H}_c(t)$), then the geodesics of $h_{\hat{U}}$ would need to be found by a different method. In the case that, at the identity, h (in some basis of $\mathfrak{su}(n)$) has the form h_{ij}, then its unit speed geodesics can be found using the standard methods of the Riemannian geodesic equation (or the EP equations for a right invariant Riemannian metric), although they will not necessarily be the one parameter subgroup.

This section depends on being able to apply theorem (2.1) to determine the geodesics of F in terms of the geodesics of h and the flow ϕ_t associated to $\hat{W}_{\hat{U}} = -i\hat{H}_0\hat{U}$. In order to do this we need that $\hat{W}_{\hat{U}}$ is a Killing vector field of h. In such cases the following results apply. At the time of writing, we are unaware of tractable, necessary and sufficient conditions for checking when a right invariant Riemannian metric on $SU(n)$ has any right invariant Killing fields; it is well known that all *left* invariant vector fields are Killing for such a metric. However, there are certainly at least some in addition to the bi-invariant metric, for example:

$$g_{\hat{U}}(\hat{A}\hat{U}, \hat{A}\hat{U}) = \mathrm{Tr}(\hat{A}^\dagger \hat{A}) + \zeta \,\mathrm{Tr}(\hat{A}, \hat{C})^2 \tag{38}$$

for $\zeta \in \mathbb{R}^+$ small enough that the formula for g defines a positive definite Riemannian metric. It is readily checked that this metric has the right invariant vector field $\hat{X}_{\hat{U}} = \hat{C}\hat{U}$ for a Killing field.

The unit speed geodesics of an arbitrary h can be found by using the Euler Poincaré equation:

$$\frac{d}{dt}\frac{\left(\partial h_{ij}\xi^i\xi^j\right)}{\partial \xi^d} = -C^a_{bd}\frac{\left(\partial h_{ij}\xi^i\xi^j\right)}{\partial \xi^a}\xi^b \tag{39}$$

This yields:

$$\dot{\xi}^k = -C^a_{bd}h^{kd}h_{ai}\xi^i\xi^b \tag{40}$$

When this can be solved, one can obtain the unit speed geodesics of h and apply theorem (2.1) again. Once the geodesics of h are known they can be used to find the geodesics of the Randers metric F solving a Zermelo navigation problem on $SU(n)$ where the 'wind' is the vector field $\hat{W}_{\hat{U}} = -i\hat{H}_0\hat{U}$. If the unit speed geodesics of h are given by $\hat{S}_t$, then the geodesics of F are given by $\hat{V}_t = \phi_t(\hat{S}_t) = \exp(-it\hat{H}_0)\hat{S}_t$.

We can now solve for the optimal Hamiltonian in a manner similar to the specific case where h is the bi-invariant metric, by setting:

$$\frac{d\hat{V}_t}{dt} = -i\hat{H}_t\hat{V}_t \tag{41}$$

and deducing that the following is required:

$$\frac{d}{dt}\left(\exp(-it\hat{H}_0)\hat{S}_t\right) = -i\hat{H}_t\exp(-it\hat{H}_0)\hat{S}_t \tag{42}$$

which yields:

$$\hat{H}_t = \hat{H}_0 - \exp(-it\hat{H}_0)\left(-i\frac{d\hat{S}_t}{dt}\hat{S}_t^\dagger\right)\exp(it\hat{H}_0) \tag{43}$$

If $\hat{S}_t$ are the unit speed geodesics of h, then by solving the EP equation associated to h, we can find the Hamiltonian driving along $\hat{S}_t$ at unit speed. Suppose then that this Hamiltonian is known and solves:

$$\frac{d\hat{S}_t}{dt} = -i\hat{Q}_t\hat{S}_t \tag{44}$$

It follows that:

$$\hat{Q}_t = i\frac{d\hat{S}_t}{dt}\hat{S}_t^\dagger \tag{45}$$

Substituting into eqn.(43) gives:

$$\hat{H}_t = \hat{H}_0 - \exp(-it\hat{H}_0)(i\hat{Q}_t)\exp(it\hat{H}_0) \tag{46}$$

The control Hamiltonian is:

$$\hat{H}_c(t) = \exp(-it\hat{H}_0)(i\hat{Q}_t)\exp(it\hat{H}_0) \tag{47}$$

All that is required to obtain the optimal trajectories is to solve the EP equation for a right invariant Riemannian metric on $SU(n)$ and then apply the above procedure. This can be achieved by solving a first order system of ODEs. As this system is the EP equation for a *quadratic* ℓ, it is possible to solve for the time derivative of ξ analytically. This makes numerical integration even simpler.

Solving for a geodesic with specified end points appears difficult without first actually obtaining the h geodesics in closed form. However, in some simplifying cases of the value of h the unit speed geodesics can be obtained in closed form using Jacobi Elliptic functions. For an example of how to do this on the group $SO(3)$, see any mathematical mechanics textbook (eg [34]) that includes a derivation of the solutions of the Euler equations for a falling rigid body. The $SU(2)$ case proceeds similarly and is tractable, essentially because $\mathfrak{so}(3) \cong \mathfrak{su}(2)$, as is familiar from the theory of spin half particles under spatial rotations in quantum mechanics.

2.9 Physical Constraints Encompassed and not Encompassed by Shen's Theorem

It is worth noting what is *not* achieved by this approach. The only constraints that can be studied this way are those which restrict the control Hamiltonian $i\hat{H}_c$ to the unit sphere of a norm induced by a given inner product h on $\mathfrak{su}(n)$. This is because such inner-products are in one-to-one correspondence with right invariant metrics on $SU(n)$ (by right translation). This allows only quadratic constraints to be studied. An interesting constraint not included in this class is the restriction that the energy expectation (in some specific state $|\psi\rangle$) associated to the control Hamiltonian alone is equal to a fixed constant for all time:

$$\kappa\langle\psi|\hat{H}_c(t) - \frac{1}{n}E_0(t)\hat{I}|\psi\rangle = 1 \tag{48}$$

Future work will include potentially applying the results in [21, 42] to a generalisation of the setup described here. The desired generalisation would be to relax the condition that the function representing the constraint on the control Hamiltonian is an inner product, and to allow more general Minkowski norms to take this role instead. This leads to a desire to solve the problem of Zermelo navigation on Finsler manifolds rather than Riemannian ones, which is currently not solved. See [3] for an exhaustive account of the status of Zermelo navigation of Riemannian manifolds.

The main result of [21] is particularly relevant as it generalises [38] and allows one to replace the role of h with a general Finsler metric that has $i\hat{H}_0\hat{U}$ as a Killing field. One class of Finsler metrics with this property is the bi-invariant ones, of which there are many. The proof is similar to the bi-invariant Riemannian cases already presented, and so is omitted. Examples of such constraints are found in the Finsler metrics formed from the right translation of the Shatten p-Norms on $\mathfrak{su}(n)$. These correspond the constraint that $F^{(p)}(\hat{H}_c(t)) := \kappa\left(\sum_n |E_n|^p\right)^{1/p} = 1\ \forall t$, thus generalising the case of the bi-invariant Riemannian metric studied above, wherein $p = 2$ (the only value of p yielding a Riemannian metric on $SU(n)$). Solving the navigation problem in general has not been achieved by the mathematics community, as far as the authors are aware. However, there are other cases besides the Riemannian case that have been solved; the Kropina metric case [47] is notable. In the absence of a solution to the navigation problem analogous to the role of Randers metrics in the Riemannian case, alternative methods must been sought. The central result of [21]

allows one to determin the geodesics of the Finsler metric solving the navigation problem on $SU(n)$ for which $\hat{H}_c(t)$ is constrained such that $F^{(p)}(\hat{H}_c(t)) = 1$. The geodesics of the (yet unknown) metric solving the navigation problem in such a generalised case are:

$$\hat{V}_t = \exp(-it\hat{H}_0)\exp(it\hat{D}) \tag{49}$$

where $F^{(p)}(i\hat{D}) = 1$.

One can also find the time optimal Hamiltonian for implementing gate $\hat{O}$ to be the same as the Riemannian ($p = 2$) case:

$$\hat{H}_c(t) = \frac{i}{T}\exp(-it\hat{H}_0)\log\left(\exp(iT\hat{H}_0)\hat{O}\right)\exp(it\hat{H}_0) \tag{50}$$

except with T_{opt} taking a different value. The requirement is now that T_{opt} is the value of T that solves:

$$F^{(p)}\left(\log(\exp(iT\hat{H}_0)\hat{O})\right) = T \tag{51}$$

Further work will also include an investigation into application of [24]. This work extends the results of [21, 38] to so called 'conic'-Finsler metrics and obtains a similar form for the geodesics under broadly analogous conditions.

2.10 Homogeneous Geodesics and the Optimality of Constant Controls

A *homogeneous geodesic* (through the identity) on a Lie group with a Finsler metric F is a one-parameter subgroup which is also a geodesic [17,28]. Here, it is a curve of the form $\hat{U}_t = \exp(-it\hat{A})$ for some $i\hat{A} \in \mathfrak{su}(n)$, which is also a geodesic of F. The $\hat{A} \in \mathfrak{su}(n)$ is a *geodesic vector*. These are exactly the curves that can be trajectories of a controlled quantum systems (of the type discussed throughout) for which only constant controls are permitted, as discussed in [39].

Theorem 3.1 in [28] presents a condition, which need to be mildly adapted to our situation. Adapted to the present example of $SU(n)$ (rather than a more general homogeneous space), the condition for $\hat{X}$ to be a geodesic vector is:

$$g_{\hat{X}}(\hat{X}, [\hat{X}, \hat{Z}]) = 0, \forall \hat{Z} \in \mathfrak{su}(n) \tag{52}$$

where g is the fundamental tensor of F restricted to $\mathfrak{su}(n)$. Here the fundamental tensor of a Finsler metric is the Hessian of the same metric point-wise. For the definition of the fundamental tensor, much more detailed discussion of its role in Finsler geometry, and many specifics of Randers metrics, see [14].

Many of the results about homogeneous geodesics of Finsler metrics on Lie groups are applicable to *left* invariant metrics. However, these results can be easily adapted to the right invariant case, which arises naturally in quantum mechanics. The construction of an "opposite group" allows one to adapt results without difficulty; typically only sign changes are incurred.

Theorem 3.7 in [28] can be applied to establish that any right invariant Randers metric on $SU(n)$ ($n \geq 2$) will have infinitely many homogeneous geodesics. The theorem establishes that a right invariant Finsler metric on a compact Lie group will have infinitely many geodesic vectors. This can be easily seen by direct application of the theorem and by the observation that the rank of $SU(n)$ is $n - 1$, and thus the theorem applies to any qubit system of more than one bit. The theorem requires that the rank is ≥ 2, thus in all cases except $SU(2)$ (the single qubit), there exist infinitely many homogeneous geodesics. This establishes the importance of determining all homogeneous geodesics of right invariant Randers metrics on $SU(n)$. Furthermore, in the case of a Randers metrics (as is our case) theorem (4.2) in [28] establishes a practically simplifying condition on homogeneous geodesics.

3 Example Calculations

In this section we use the results derived above to calculate speed limits; these follow the examples in [39], except here allowing time dependent controls. Throughout this section the metric h representing the constraint on $\hat{H}_c(t)$ is $h(i\hat{A}, i\hat{B}) = \kappa \operatorname{Tr}((i\hat{A})^\dagger i\hat{B}) = \kappa \operatorname{Tr}(\hat{A}^\dagger \hat{B}) = \kappa \operatorname{Tr}(\hat{A}\hat{B})$. The dagger vanishes in the final step since $\hat{A}$ is Hermitian. This yields the constraint $\operatorname{Tr}(\hat{H}_c(t)^2) = \frac{1}{\kappa}$ for all times.

3.1 Single Spin in a Magnetic Field

3.1.1 Pauli σ_y Gate

Following the example calculations performed in [39] we study the speed limit to performing the gate:

$$\hat{O} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = -i\sigma_y \tag{53}$$

in a system constrained such that $h\left((i\hat{H}_c(t))^\dagger i\hat{H}_c(t)\right) = \kappa \operatorname{Tr}(\hat{H}_c(t)^2) = 1$ with drift Hamiltonian:

$$\hat{H}_0 = B^x\sigma_x + B^y\sigma_y \tag{54}$$

such that $h(i\hat{H}_0, i\hat{H}_0) \leq 1$.

Substituting these conditions into eqn.(33) for the optimal time gives

$$-\frac{\kappa}{T^2} \operatorname{Tr}\left(\log(\mu(T))^2\right) = 1 \tag{55}$$

where the matrix μ is defined as

$$\mu(T) := \begin{pmatrix} \frac{\sin(TB)(B^y + iB^x)}{B} & -\cos(TB) \\ \cos(TB) & \frac{\sin(TB)(B^y - iB^x)}{B} \end{pmatrix} \tag{56}$$

and where $B = ||\vec{B}||$ (standard Euclidean norm).

A numerical solution can be achieved using any zero finding algorithm; we found Mathematica's "FindRoot" method to be effective when given suitable starting points. The result (due to the cyclic property of the matrix trace) depends on the eigenvalues of the quantity within the trace. Thus one should diagonalize before attempting to numerically solve eqn.(33). It greatly simplifies the process and can be easily achieved with any good algebra package.

By eqn.(30), the geodesic (of the Randers metric on $SU(2)$ which solves Zermelo's navigation problem) which connects $\hat{I}$ to $\hat{O}$ is given by:

$$\begin{aligned} \hat{V}_t &= \exp(-it\hat{H}_0) \exp\left(\frac{t}{T_{\text{opt}}} \log\left(\exp(iT_{\text{opt}}\hat{H}_0)\hat{O}\right)\right) \\ &= \exp(-it\hat{H}_0) \left(\exp(iT_{\text{opt}}\hat{H}_0)\hat{O}\right)^{t/T_{\text{opt}}} \\ &= \exp\left(-it(B^x\sigma_x + B^y\sigma_y)\right) \left(\exp(iT_{\text{opt}}(B^x\sigma_x + B^y\sigma_y))(-i\sigma_y)\right)^{t/T_{\text{opt}}} \\ &= \exp\left(-it\left(B^x\sigma_x + B^y\sigma_y\right)\right) \left(\mu(T_{\text{opt}})\right)^{t/T_{\text{opt}}} \end{aligned} \tag{57}$$

These geodesics are not generally of the form of any time independent trajectory, i.e. the optimal controls are not constant.

The optimal control Hamiltonian in closed form is:

$$\hat{H}_c(t) = \frac{i}{T_{\text{opt}}} \exp\left(-it\left(B^x\sigma_x + B^y\sigma_y\right)\right) \log(\mu(T_{\text{opt}})) \exp\left(it\left(B^x\sigma_x + B^y\sigma_y\right)\right)$$

where

$$\exp(-it\,(B^x\sigma_x + B^y\sigma_y)) = \begin{pmatrix} \cos(tB) & \frac{-\sin(tB)(B^y + iB^x)}{B} \\ \frac{\sin(tB)(B^y - iB^x)}{B} & \cos(tB) \end{pmatrix} \tag{58}$$

Expanding these quantities in an attempt to find $\hat{H}_c(t)$ does not lead to tractable expressions.

The optimal control fields can also be determined. Consider that the control Hamiltonian can be expressed as $i\hat{H}_c(t) = D_t^k i\sigma_k$. One can now extract the individual control fields D^k from the following, as $i\sigma_k$ are an orthogonal basis (according to $h(i\hat{A}, i\hat{B}) = \kappa\,\mathrm{Tr}(\hat{A}\hat{B})$) for $\mathfrak{su}(n)$:

$$D_t^k = h\left(i\hat{H}_c(t), \frac{i\sigma_k}{2}\right) = \frac{\kappa}{2}\,\mathrm{Tr}(\hat{H}_c(t)\sigma_k) \tag{59}$$

This calculation rapidly becomes intractable analytically, however by substituting the Taylor series for *log* and *exp*, one can find a simple series representation for the optimal control fields. Numerically extracting the control fields to arbitrary precision is straightforward.

3.1.2 Specific Physical Instance

We now specialise to specific values for κ, B^x and B^y in order to numerically obtain the optimal time. We set $B^x = B^y = 1/4$ and $\kappa = 1$; one readily checks that this system meets the "small wind" premise of Shen's theorem.

We solved the specific instance of eqn.(33) using Mathematica. The smallest, real, positive root is $T_{\mathrm{opt}} \approx 3.2043\ldots$. (The actual physical time in seconds can be obtained by reintroducing the physical constants that have been lost after non-dimensionalisation throughout. Specifically, $\hbar$ has been set to 1 throughout.)

In this instance the optimal control Hamiltonian is given by

$$\hat{H}_c(t) = \frac{i}{T_{\mathrm{opt}}} \exp\left(\frac{-it}{4}(\sigma_x + \sigma_y)\right) \log \begin{pmatrix} \frac{1+i}{\sqrt{2}}\sin\left(\frac{T_{\mathrm{opt}}}{\sqrt{8}}\right) & -\cos\left(\frac{T_{\mathrm{opt}}}{\sqrt{8}}\right) \\ \cos\left(\frac{T_{\mathrm{opt}}}{\sqrt{8}}\right) & \frac{1-i}{\sqrt{2}}\sin\left(\frac{T_{\mathrm{opt}}}{\sqrt{8}}\right) \end{pmatrix} \exp\left(\frac{it}{4}(\sigma_x + \sigma_y)\right) \tag{60}$$

by eqn.(30). We have evaluated the logarithm exactly in closed form, but the result is cumbersome and does not provide any physical insight, so it is omitted.

We obtain the optimal control fields from this result. The kth field is obtained by evaluating eqn.(59) $\frac{1}{2}\,\mathrm{Tr}(\hat{H}_c(t)\sigma_k)$ numerically, as shown in fig.(1). As a check, we have numerically confirmed that the control fields have the property that the sum of their squares is $1/2$ $(\forall t)$, which the constraint on $\mathrm{Tr}(\hat{H}_c(t)^2)$ mandates.

3.1.3 Example of Assessing Quantum Systems

Given a choice of quantum systems for the potential implementation of fast quantum computation, one requires methods to assess systems. Here we show how the methods illustrated in this paper can be used to perform such assessment through an example.

As shown in [39], the optimal time-independent control implementation time for the same gate as above $-i\sigma_y$ is:

$$T_{\mathrm{opt}} = \frac{\pi}{2}\frac{B^y}{(D^2 - B^2)}\left(1 \pm \sqrt{1 + \frac{D^2 - B^2}{(B^y)^2}}\right) \tag{61}$$

where $D = ||\vec{D}||$ (standard Euclidean norm).

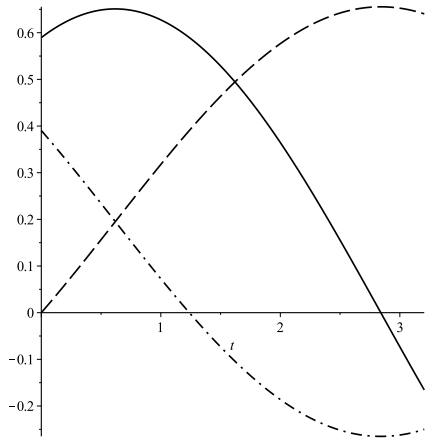

Figure 1: Optimal control fields D^x (dashed), D^y (dot-dashed), D^z (solid) for Pauli y gate, as a function of time, in the $B = 1/4$ case.

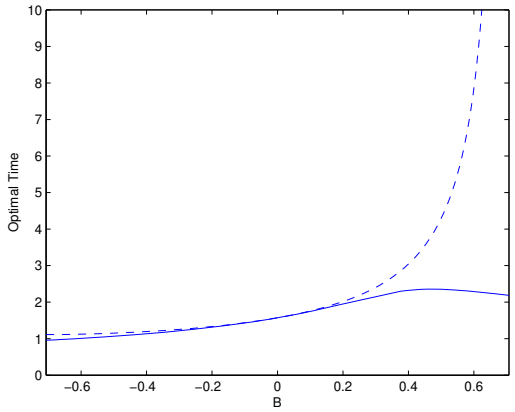

Figure 2: Optimal times for time dependent (solid) and independent (dashed) controls.

To illustrate, we set $\kappa = 1/2$ (achieved through setting the value of $D = 1$, a choice made for ease) and $B^x = B^y = b$ (some real number $b^2 < 1/2$). Then the optimal time in the time independent case is:

$$T_{\text{opt}} = \frac{\pi}{2} \frac{b}{(1-2b^2)} \left(1 \pm \frac{\sqrt{1-b^2}}{b} \right) \tag{62}$$

where the $\pm$ is chosen to make the time always positive [39]. We numerically solve eqn.(55) in Matlab, with the given parameters substituted to find the optimal time (as a function of b also) when time dependent controls are allowed. This yields the results of fig (2). This illustrates that some systems and some types of controls are significantly better for the implementation of specific QIP tasks. The method presented is a powerful tool for assessing such situations. One can clearly see from fig.(2) that scenarios with values of $b < 0.2$ are favorable over the $b > 0.2$ region if only constant controls are permitted.

The computation time to obtain the fig.(2) and the optimal control schemes shown in fig.(1) were both negligible in Matlab, and that this also appears to be the case for two qubit gates.

3.2 Swap Gate in a Length Two Heisenberg Spin Chain

3.2.1 Two Spin Chain

The drift Hamiltonian for a two spin chain with (arbitrary spin coupling) is [40]:

$$\hat{H}_0 = J^x \sigma_x \otimes \sigma_x + J^y \sigma_y \otimes \sigma_y + J^z \sigma_z \otimes \sigma_z \tag{63}$$

Again, we take h to be the Killing form so we can apply theorem (2.1) to obtain the geodesics in closed form. By eqn (30), the geodesic which connects $\hat{I}$ to $\hat{O}$ of the relevant Randers metric on $SU(4)$ are given by eqn. (35). One can exploit the block diagonal form of $\hat{H}_0$ in order to simplify the equation that needs to be solved.

The optimal Hamiltonian is:

$$\begin{aligned}\hat{H}_c(t) =& \frac{i}{T_{\text{opt}}} \exp\left(-it\left(J^x \sigma_x \otimes \sigma_x + J^y \sigma_y \otimes \sigma_y + J^z \sigma_z \otimes \sigma_z\right)\right) \\ & \times \log\left(\exp\left(iT_{\text{opt}}\left(J^x \sigma_x \otimes \sigma_x + J^y \sigma_y \otimes \sigma_y + J^z \sigma_z \otimes \sigma_z\right)\right)\hat{O}\right) \\ & \times \exp(it\left(\left(J^x \sigma_x \otimes \sigma_x + J^y \sigma_y \otimes \sigma_y + J^z \sigma_z \otimes \sigma_z\right)\right))\end{aligned} \tag{64}$$

The actual control fields can be extracted, similarly to before, via:

$$f_{mn}(t) = \frac{\kappa}{4} \operatorname{Tr}\left(\hat{H}_c(t)\sigma_m \otimes \sigma_n\right) \tag{65}$$

This is the expansion of $\hat{H}_c(t)$ in a basis for $\mathfrak{su}(4)$. This basis is:

$$\{i\sigma_n \otimes \sigma_m \| n, m = 0, x, y, z \text{ but not both } n = 0 \text{ and } m = 0\} \tag{66}$$

Here σ^0 is taken to be the 2×2 identity matrix whereas the other σs are all the standard Pauli matrices. One readily checks that this basis is orthogonal w.r.t. the Killing form, which is the key property applied when extracting the control fields in eqn.(65). The origin of the 4 in this formula is the trace of the 4×4 identity matrix. Explicitly, we are representing $\hat{H}_c(t)$ as:

$$\hat{H}_c(t) = \sum f_{mn}(t)\sigma_m \otimes \sigma_n \tag{67}$$

where ths sum is over the basis vectors appearing in eqn (66).

3.2.2 Specific Physical Instance: The XXX-Spin Chain

In order to again illustrate the way in which out method allows us to determine which systems are best suited to quickly implementing a QIP task, we study the case of the isotropic Heisenberg spin chain, the $J^x = J^y = J^z = J$ case of eqn.(63). This leaves only one parameter J to consider, yielding a simple pedagogic example for the method. We set $\kappa = 1$ and consider the optimal time for implementing the (special unitary) swap gate:

$$\hat{O} = e^{i\pi/4} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{68}$$

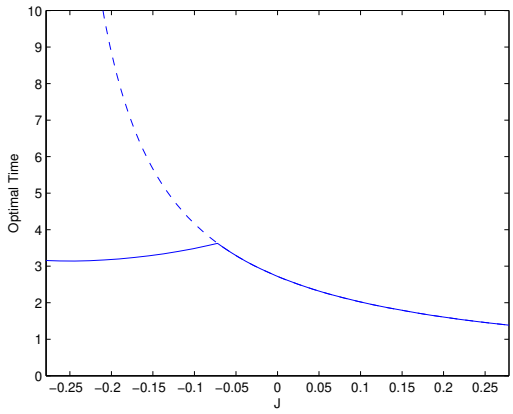

Figure 3: Optimal times for time dependent (solid) and time independent (dashed) controls in an XXX Spin Chain.

Using eqn (30), the optimal Hamiltonian is:

$$\hat{H}_c(t) = \frac{i}{4T_{\rm opt}} \begin{pmatrix} 2e^{-itJ} & 0 & 0 & 0 \\ 0 & e^{-itJ}+e^{3itJ} & e^{-itJ}-e^{3itJ} & 0 \\ 0 & e^{-itJ}-e^{3itJ} & e^{-itJ}+e^{3itJ} & 0 \\ 0 & 0 & 0 & 2e^{-itJ} \end{pmatrix} \tag{69}$$
$$\times \log \left(\frac{1}{2} e^{\frac{i\pi}{4}} \begin{pmatrix} 2e^{iT_{\rm opt}J} & 0 & 0 & 0 \\ 0 & e^{iJT_{\rm opt}}-e^{-3iJT_{\rm opt}} & e^{iJT_{\rm opt}}+e^{-3iJT_{\rm opt}} & 0 \\ 0 & e^{iJT_{\rm opt}}+e^{-3iJT_{\rm opt}} & e^{iJT_{\rm opt}}-e^{-3iJT_{\rm opt}} & 0 \\ 0 & 0 & 0 & 2e^{iJT_{\rm opt}} \end{pmatrix} \right)$$
$$\times \begin{pmatrix} 2e^{itJ} & 0 & 0 & 0 \\ 0 & e^{itJ}+e^{-3iJt} & e^{itJ}-e^{-3itJ} & 0 \\ 0 & e^{itJ}-e^{-3iJt} & e^{itJ}+e^{-3itJ} & 0 \\ 0 & 0 & 0 & 2e^{itJ} \end{pmatrix}$$

We can determine the optimal time $T_{\rm opt}$ as before, by numerically solving eqn (33). Substituting the specifics of the current problem into this equation yields:

$$-\frac{1}{(T_{\rm opt})^2} \,{\rm Tr} \left(\log \left(\frac{1}{2} e^{\frac{i\pi}{4}} \begin{pmatrix} 2e^{iT_{\rm opt}J} & 0 & 0 & 0 \\ 0 & e^{iJT_{\rm opt}}-e^{-3iJT_{\rm opt}} & e^{iJT_{\rm opt}}+e^{-3iJT_{\rm opt}} & 0 \\ 0 & e^{iJT_{\rm opt}}+e^{-3iJT_{\rm opt}} & e^{iJT_{\rm opt}}-e^{-3iJT_{\rm opt}} & 0 \\ 0 & 0 & 0 & 2e^{iJT_{\rm opt}} \end{pmatrix} \right)^2 \right) = 1$$

Using the same procedure as in the previous example, we obtain the optimal execution times. These are shown in fig.(3). The dashed line in fig.(3) plots the function, adapted to this specific scenario, from [39, eqn.21], giving the time independent optimal time:

$$T_{opt} = \frac{\pi}{2} \left(\frac{\sqrt{3}}{2\sqrt{3}J \pm 1} \right) \tag{70}$$

where the $\pm$ is again chosen to make the time positive. The cusp in the time dependent line is due to its behaviour intersecting the time independent line, and then following the time independent behaviour.

3.2.3 Example of Assessing Quantum Systems

It can be seen from fig.(3) that, in the region between $J \approx -0.075$ and $J = \frac{1}{\sqrt{12}}$, constant controls perform exactly as well as time dependent ones. However, outside this region, time dependent controls significantly out-perform constant ones for which the optimal time has an asymptote at $J = -\frac{1}{\sqrt{12}}$. This indicates that our method can be used to assess the choice of the best type of control scheme as well as to determine optimal control fields when a specific physical system is considered.

The fact that the optimal times agree for both types of control scheme on a contiguous region of values for the parameter J suggests that the theorem leading to eqn.(52) could potentially be strengthened (perhaps only in the Randers case) to show that the infinitely many geodesics vectors of a Finsler metric are a connected set. We conjecture that this is the case for Randers metrics on $SU(n)$ ($n \geq 2$). These results justify the physical relevance of that theorem.

4 Further Constraints on $\hat{H}_c$ Using Lagrange Multipliers

4.1 The Need for Further Constraints

The constraints studied thus far are, alone, insufficient for many physical applications. The assumption thus far as been only that the control Hamiltonian is constrained to be such that $h(i\hat{H}_c(t), i\hat{H}_c(t)) = 1$. In comparison to [12], only the roles of the L_T and L_S parts of the Lagrangian have been treated here. The L_T part is analogous to our application of the result of [42] on Zermelo navigation in the case of a right invariant Riemannian metric. The L_S part has no analogue as our work expresses the problem in a geometrically intrinsic way. We consider this to be an advantage of our method as it allows a more mathematical view of the problem; intrinsic geometry has been proven many times to be superior for mathematical analysis of geometric problems over methods based on many constraints or specific coordinate systems. This allows us to formulate the problem as a *first order* system of ODE from the outset by using the EP equation, rather than needing to compute any first variations or use the EL equations. So we can obtain a first order equation for the optimal Hamiltonian, *without* the need to actually determine any geodesics a priori.

The disadvantage of our method compared with [12] is that we can handle fewer constraints, as described above (§2.9). Our method thus far can handle only the cases where the "size" type constraint [12] is representable by an inner product.

To motivate the need for further constraints, we again consider the drift Hamiltonian for a two spin "chain" (with anisotropic couplings J) [40]:

$$\hat{H}_0 = J_x \sigma^x \otimes \sigma^x + J_y \sigma^y \otimes \sigma^y + J_z \sigma^z \otimes \sigma^z \tag{71}$$

Simply constraining the control Hamiltonian to be such that $h(i\hat{H}_c(t), i\hat{H}_c(t)) = 1$ for some inner product h is insufficient for practical applications where the producible set of control Hamiltonians does not include every direction within $\mathfrak{su}(n)$. For example, a common model of a controlled spin chain is one in which the control Hamiltonian takes the form:

$$\hat{H}_c(t) = f_1(t)\sigma_z \otimes \hat{I} + f_2(t)\hat{I} \otimes \sigma_z \tag{72}$$

That is, there is one local control field in the z direction only for each site in the chain. In such a situation no terms like $\sigma^x \otimes \sigma^x$ (or multiples there of) could appear in the control Hamiltonian, as these represent the couplings between sites in the chain, and are not the effect of *any possible* external field. In this case (choosing h to be κ times the Killing form), the constraint $h(i\hat{H}_c(t), i\hat{H}_c(t)) = 1$ evaluates to $\kappa \operatorname{Tr}(\hat{H}_c(t)^2) = 1$, which only constrains the sum of the squares of the control fields. An extra constraint must be added to exclude those terms from the control Hamiltonian that cannot be physically implemented.

This can be done by including Lagrange multipliers to create a new functional:

$$\Lambda\left(\hat{U}_t, \frac{d\hat{U}_t}{dt}, \omega_k(t)\right) = \frac{1}{2}\left[F_{\hat{U}_t}\left(\frac{d\hat{U}_t}{dt}\right)\right]^2 + \sum_k \omega_k(t)\left(f_{k,\hat{U}_t}\left(\frac{d\hat{U}_t}{dt}\right) - c_k\right) \tag{73}$$

where F is the Randers metric solving the relevant navigation problem, f_k represent the additional constraints, and ω_k are the Lagrange multipliers. The values c_k represent the value of f to which the trajectory is constrained.

We consider only the case where f is right invariant; this results in Λ also being right invariant. This corresponds to situations where the additional constraints depend only on the Hamiltonian, rather than the on location of $\hat{U}_t$ on $SU(n)$.

In this situation Λ can be expressed as:

$$\Lambda\left(\frac{d\hat{U}_t}{dt}\hat{U}_t^\dagger, \omega_k(t)\right) = \frac{1}{2}\left[F\left(\frac{d\hat{U}_t}{dt}\hat{U}_t^\dagger\right)\right]^2 + \sum_k \omega_k(t)\left(f_k\left(\frac{d\hat{U}_t}{dt}\hat{U}_t^\dagger\right) - c_k\right) \tag{74}$$

$$= \frac{1}{2}\left[F\left(-i\hat{H}(t)\right)\right]^2 + \sum_k \omega_k(t)\left(f_k\left(-i\hat{H}(t)\right) - c_k\right) \tag{75}$$

where $\hat{H}_t$ is the Hamiltonian such that $\hat{U}_t$ solves the corresponding Schrödinger equation.

4.2 Forbidden Directions

One specific set of f_k and c_k with practical relevance is: $f_k\left(\frac{d\hat{U}_t}{dt}\hat{U}_t^\dagger\right) = \mathrm{Tr}\left(\left(\left(\frac{d\hat{U}_t}{dt}\hat{U}_t^\dagger + i\hat{H}_0\right)i\hat{F}_k\right)\right)$ with $c_k = 0$. This corresponds to the $\hat{F}_k$ spanning a set of "forbidden" terms for the control Hamiltonian. One can check this interpretation of the constraint by noticing that if $\hat{U}_t$ solves the Schrödinger equation with a Hamiltonian of the form of eqn.(1), then variation of eqn.(74) by ω_k yields:

$$\mathrm{Tr}\left(\left(\frac{d\hat{U}_t}{dt}\hat{U}_t^\dagger + i\hat{H}_0\right)i\hat{F}_k\right) \tag{76}$$

which implies:

$$\mathrm{Tr}\left(\hat{H}_c(t)\hat{F}_k\right) = 0$$

and thus the desired "forbidden" directions are trace-orthogonal to the control Hamiltonian, and the control Hamiltonian has no component in any forbidden direction. These are essentially identical to the "linear homogeneous" constraints in [12]. There is a subtle difference however: here the forbidden direction applies only to the control Hamiltonian and not the overall Hamiltonian. This constraint is equivalent to an *affine* constraint on the overall Hamiltonian. Adding too many additional constraints may render the system in question uncontrollable. Existence/uniqueness of optimal trajectories is an issue not addressed in [12].

In order to find the equation satisfied by the optimal Hamiltonian that takes into account some additional constraints, we must modify eqn.(21). In the remainder of this paper we consider only the "forbidden direction" type of additional constraint. The equations satisfied by the optimal Hamiltonian (if any exist) can be found (in a basis for $\mathfrak{su}(n)$) by variation of each dependent variable on which Λ depends. Variation by $\frac{d\hat{U}_t}{dt}\hat{U}_t^\dagger$ yields the EP equation for Λ:

$$\frac{d}{dt}\frac{\partial\Lambda}{\partial\xi^d} = -C^a_{bd}\frac{\partial\Lambda}{\partial\xi^a}\xi^b \tag{77}$$

Variation by ω_k yields:

$$\mathrm{Tr}\left(\left(\frac{d\hat{U}_t}{dt}\hat{U}_t^\dagger + i\hat{H}_0\right) i\hat{F}_k\right) = 0 \tag{78}$$

Equations (77) and (78) need to be solved simultaneously in order to obtain the optimal Hamiltonian.

Closed form solutions for these equations and further numerical solution techniques in specific cases of physical interest form the basis of further work. We intend to perform a complete analysis of common two qubit gates implemented in spin chain systems and other laser driven models. We hope to obtain an exact formula for the initial conditions required (for the system (77) and (78)) in order for numerical solution of the system of ODEs to yield a geodesic connecting $\hat{I}$ to an arbitrary desired gate $\hat{O}$. A method for achieving a very similar goal has be found in a very different context [4]; only the Riemannian case is addressed, but it seems that the technique is easily adaptable.

4.3 Role of Sub-Riemannian Geometry and Sub-Finsler Geometry

Mathematically, the optimal trajectories can be understood as sub-Finsler geodesics, or more specifically, what could be appropriately called sub-Randers geodesics. For all the relevant definitions needed here see [16]. One can find information about sub-Riemannian geometry in optimal control in [43]. For an application of sub-Finsler geometry in optimal control see [32]. For a specific application in quantum control see [25]. The now well known "Hormander's condition" for the controllability of affine linear control systems of the type studied in this work, that is systems of the form of eqn.(1), indicates when too many forbidden direction results in the system no longer being controllable. This would mean that there were unitary gates that could not be implemented using a system constrained in such a way.

The method for handling the constraint $h(i\hat{H}_c(t), i\hat{H}_c(t)) = 1$, alongside additional constraints, shows that the optimal trajectories for $\hat{U}_t$ are geodesics of Randers metric restricted to an affine distribution, say $\mathcal{D}$, on $SU(n)$. $\mathcal{D}$ is the distribution consisting of vectors in $TSU(n)$ of the form $-i\hat{H}_0\hat{U}+\mathrm{Span}\{i\hat{H}_k\hat{U}\big|k=0,\ldots,N\}$. Here $\{i\hat{H}_k\big|k=0,\ldots,N\}\subset\mathfrak{su}(n)$ span the subset of $\mathfrak{su}(n)$ that is h-orthogonal to the span of the subset of $\mathfrak{su}(n)$ spanning the "forbidden directions". This distribution is right invariant in the sense that: $\mathcal{D}_{\hat{U}} = \mathcal{D}\hat{U}$. That is, the optimal trajectories are the length minimising curves that connect given endpoints ($\hat{I}$ to $\hat{O}$) according a Randers metric F (solving the navigation problem in our case), and which are parallel to the distribution $\mathcal{D}$. A curve $\hat{V}_t$ being parallel to $\mathcal{D}$ means that $\frac{d\hat{V}_t}{dt}\in\mathcal{D}_{\hat{V}_t}\ \forall t$. The system is controllable, that is, every unitary gate could be implemented, as long as this distribution is "Bracket Generating". This provides an exact condition, albeit a very technical one, for controlability in the presence of additional constraints. The systems formed by equations (77) and (78) are solved by such curves. Eqn.(77) imposes that a curve is an, at least local, extremal curve of the length functional. Eqn.(78) can then be understood as imposing the curve is parallel according to $\mathcal{D}$.

4.4 Example Equations For Optimal Trajectories

We illustrate how a forbidden direction can be treated in the example of a single spin. For simplicity, we consider the case that there is no drift. This makes the navigation metric F Riemannian, which makes ℓ a quadratic function. This allows us to solve for the time derivative of ξ explicitly in the EP equations, and then integrate the equation in closed form by hand. The case with drift is conceptually identical, except it may not always be possible to solve for $\dot{\xi}$, so the resulting system could be more difficult to solve analytically.

We consider the system with control Hamiltonian constrained such that $\frac{1}{2}\mathrm{Tr}(\hat{H}_c(t)^2) = 1$. Writing $i\hat{H}_c(t) = \xi^k(t)i\sigma_k$ we see that this condition is ${\xi^x}^2+{\xi^y}^2+{\xi^z}^2 = 1$. Suppose further that we are restricted to $\xi^z = c$ for $\forall t$. This is different from the examples worked out in [12], as this is an affine constraint rather than

a linear homogeneous constraint. In a situation with a drift term, it is simple to see that a linear constraint on the control Hamiltonian corresponds to an equivalent affine constraint on the overall Hamiltonian.

In the present case the overall Lagrangian is:

$$\Lambda(\hat{H}_t, \omega(t)) = \frac{1}{2}\left(\xi^{x\,2} + \xi^{y\,2} + \xi^{z\,2}\right) + \omega(t)\left(\xi^z - c\right) \tag{79}$$

The EP equations are:

$$\begin{pmatrix} \dot{\xi^x} \\ \dot{\xi^y} \\ \dot{\xi^z} + \dot{\omega} \end{pmatrix} = \begin{pmatrix} -\omega\xi^y \\ \omega\xi^x \\ 0 \end{pmatrix} \tag{80}$$

and $\xi^z(t) = c$, $\dot{\xi^z} = 0$, which implies:

$$\begin{pmatrix} \dot{\xi^x} \\ \dot{\xi^y} \end{pmatrix} = \begin{pmatrix} -\omega\xi^y \\ \omega\xi^x \end{pmatrix} \tag{81}$$

and that ω is a constant. The general solution (after imposing the unit speed condition) is:

$$\begin{aligned} \xi^x(t) &= A\cos(\omega t) - \sqrt{1 - c^2 - A^2}\,\sin(\omega t) \\ \xi^y(t) &= A\sin(\omega t) + \sqrt{1 - c^2 - A^2}\,\cos(\omega t) \\ \xi^z(t) &= c \end{aligned} \tag{82}$$

wherein A is an arbitrary constant parameter and ω is the Lagrange multiplier. The trajectories in $\mathfrak{su}(n)$ are circles in $\mathfrak{su}(2)$ centered at:

$$\begin{pmatrix} 0 \\ 0 \\ c \end{pmatrix} \tag{83}$$

The unit speed condition imposed ensures that the parameter in $\hat{H}_t$ is physical time as it appears in the Schrödinger equation.

Both A and ω paremeterise possible endpoints of $\hat{U}_t$ after the trajectory on the group is reconstructed via the Schrödinger equation. The optimal Hamiltonians are:

$$\hat{H}_t = (A\cos(\omega t) - B\sin(\omega t))\sigma_x + (A\sin(\omega t) + B\cos(\omega t))\sigma_y + c\sigma_z \tag{84}$$

where $B = \pm\sqrt{1 - c^2 - A^2}$.

4.5 Generalising the Forbidden Directions Equations

The method for handling forbidden directions can be generalised to include a much larger class of constraints replacing the constraint that $h(i\hat{H}_c(t).i\hat{H}_c(t)) = 1$. One can replace the role of h representing the constraint of the size of $i\hat{H}_c(t)$ with an arbitrary right invariant Finsler metric, which we denote by $\check{F}_{\hat{U}}$, i.e. we now, more generally than before, impose $\check{F}(i\hat{H}_c(t)) = 1$ holds at the identity on $SU(n)$. As right invariant Finsler metrics on $SU(n)$ are in one-to-one correspondence with Minkowski norms on $\mathfrak{su}(n)$, this new class of constraints is much larger that the class of right invariant Riemannian metrics employed before.

We adapt Shen's [42] Lemma 3.1 to the case of $SU(n)$ with a right invariant $\check{F}$. In addition to the exact solution for the Riemannian case given by Shen's theorem, this gives an equation for a Finlser metric F the geodesics of which are time optimal the the presence of the constraint $\check{F}(i\hat{H}_c(t)) = 1$.

$$\check{F}\left(\frac{\hat{A}}{F(\hat{A})} + i\hat{H}_0\right) = 1 \qquad \forall \hat{A} \in \mathfrak{su}(n)/\{0\} \tag{85}$$

Note that the roles of F and $\check{F}$ are reversed here compared with the original presentation. One can easily check that the solution for F will be a Randers metric exactly when $\check{F}$ is Riemannian; this is exactly the case solved by Shen's theorem.

The premise that the 'wind'/drift Hamiltonian can be overcome by the control is now $\check{F}(i\hat{H}_0) \leq 1$. This guarantees that the desired time optimal trajectories are the geodesics of the Finsler metric F solving eqn.(85). The solution F is right invariant if both $\check{F}$ and the drift vector field are, as is the case for quantum control problems. This follows directly from substituting right invariant $\check{F}$ and drift vector field into the equation defining F and then right translating to the identity.

Now we give the set of equation that define the optimal trajectories in such a scenario. Time optimality yields:

$$\frac{d}{dt}\frac{\partial \Lambda}{\partial \xi^d} = -C^a_{bd}\frac{\partial \Lambda}{\partial \xi^a}\xi^b \tag{86}$$

Variation by ω_k to impose the forbidden direction constraints, as before, yields:

$$\mathrm{Tr}\left(\hat{H}_c(t)\hat{F}_k\right) = 0 \tag{87}$$

Here Λ is as before, except the F is no longer necessarily a Randers metric, but is now the solution to eqn.(85). This solution is guaranteed to also be a Finsler metric [42].

Together this all yields the system for the time optimal Hamiltonian $\hat{H}_t = \hat{H}_0 + \hat{H}_c(t) = \xi^k \hat{G}_k$:

$$\begin{cases} \check{F}\left(\dfrac{\hat{A}}{F(\hat{A})} + i\hat{H}_0\right) = 1 \quad \forall \hat{A} \in \mathfrak{su}(n)/\{0\} \\ \\ \dfrac{d}{dt}\dfrac{\partial \Lambda}{\partial \xi^d} = -C^a_{bd}\dfrac{\partial \Lambda}{\partial \xi^a}\xi^b \\ \\ \mathrm{Tr}\left(\hat{H}_c(t)\hat{F}_k\right) = 0 \quad \forall k, \forall t \geq 0 \\ \\ \hat{U}_T = \mathcal{T}\exp\left(\displaystyle\int_0^T -i\hat{H}_t dt\right) = \hat{O} \end{cases} \tag{88}$$

These we refer to as the time optimality equations for the gate $\hat{O}$, the drift Hamiltonian $\hat{H}_0$ and the constraint that $\check{F}(\hat{H}_c(t)) = 1\ \forall t \geq 0$. Here, $\hat{G}_k$ are a basis for $\mathfrak{su}(n)$. These equations determine the optimal Hamiltonian.

As in [12], we have not yet found a way to impose the boundary condition $\hat{U}_T = \hat{O}$ (for some T) without solving the other time optimality equations explicitly. It is, however, known which variations at the algebra level correspond to variations of $\hat{U}_t$ that leave the end points of a curve on $SU(n)$ fixed [19]. In quantum mechanical terms these are exactly variations of $-i\hat{H}_t$ of the form: $\delta i\hat{H}_t = i\frac{d\hat{K}_t}{dt} + [i\hat{H}_t, i\hat{K}_t]$. Here $\hat{K}_t$ is any smooth curve in $\mathfrak{su}(n)$ which is 0 at both end points. A method for imposing similar boundary conditions is presented in [4] in a different context. We hope to analyse that method and adapt it to quantum control scenarios, so that end point conditions on $\hat{U}_t$ can be imposed at the algebra level and thus the EP equations can still be applied.

5 Conclusions And Further Work

We have shown that the time optimal control of any quantum system of the form eqn.(1) with the constraint that $F(i\hat{H}_c(t)) = 1$, for some Minkowski norm F on $\mathfrak{su}(n)$, is exactly the problem of finding geodesics of

a right invariant Finsler metric on $SU(n)$. Furthermore, we have shown that this can be achieved, for the Hamiltonian driving $\hat{U}_t$ along a geodesic, using the EP equations, which are first order. We have also shown that, in the presence of forbidden directions for $\hat{H}_c(t)$, the optimal trajectories for $\hat{U}_t$ are the geodesics of a right invariant sub-Finsler metric on $SU(N)$. Here, sub-Finsler geodesics is taken to mean the shortest curves connecting desired endpoints which are parallel to a specific affine distribution. We have also shown that this problem can also be expressed as a system of equations in $\mathfrak{su}(n)$ and that these equations are first order also.

We have shown that the method of quantum optimal control based on Randers geometry is highly effective in the case of constraints on the control Hamiltonian of the form $h(i\hat{H}_c, i\hat{H}_c)$ for some inner product h on $\mathfrak{su}(n)$. We have produced a broadly applicable method that does not depend on the dimension of the Hilbert space that the goal gate $\hat{O}$ acts on. We believe that our approach will be taken into a more practical setting with further analysis of the "forbidden directions" type constraints. A numerical method for solving eqn.(77) will be presented in further work, along with examples of many practically encountered gates and constraints.

We have also shown that the method can be at least partially generalised to an even broader class of problems where the constraint is represented by a Finsler metric rather than a Riemannian one, at least in the case that the Finsler metric representing the constraint has the required Killing field, which includes all bi-invariant ones of which there are uncountably many. One example class of uncountably many bi-invariant Finsler metrics on $SU(n)$ are the formed by the right translation of the Shatten-p norms from the identity. In this case the desired geodesics can be found in close form.

We intend to produce a general purpose Matlab script into which one can insert:

- A drift Hamiltonian
- A norm constraining the control Hamiltonian
- A desired gate $\hat{O}$

and out of which will be produced the optimal control Hamiltonian and control fields, by numerically solving the system of equations (88) and the optimal time. We predict that the main obstacle to this will be solving for the metric F in terms of $\check{F}$.

Recent work [45] contains a large appendix "Euler-Lagrange equation on $SU(n)$" discussing methods for finding geodesics on $SU(n)$. Other recent work [35] also discusses finding Finsler geodesics on $SU(n)$ in the context of quantum optimal control. We feel that the relative simplicity of the EP equations, which hold on a vector space $\mathfrak{su}(n)$, compared to the methods described in [35, 45] that hold on $SU(n)$, justify the usefulness of our approach. Furthermore, they avoid the need to ever determine a geodesic on $SU(n)$ when all that is practically needed is the Hamiltonian that drives $\hat{U}_t$ along that geodesic. Another advantage is the lack of need for the use of any coordinate system on $SU(n)$, as the EP equations directly exploit the right trivisation of $SU(n)$, $TSU(n) \cong SU(n) \times \mathfrak{su}(n)$ (available as Lie groups are all parallelizable manifolds, for definitions see [15]).

The desired optimal trajectories are geodesics of eqn.(14). However, the metric can also be used to obtain optimal times for a system (meeting the appropriate premises) to traverse an *arbitrary curve*. This is illustrated for the time independent trajectories and a specific drift Hamiltonian and value h in [39]. This is in contrast to other methods which exactly determine optimal trajectories, but do not offer any way to obtain optimal times for arbitrary trajectories. In practical physical systems, it is unlikely that arbitrary trajectories can be implemented, so one needs a technique for assessing the trajectories that can be implemented over a method for determining theoretically optimal ones. For example a laser pulse is often described using an envelope function [46] and this imposes a form for the control functions a priori. Our method allows one to assess such systems; a full analysis of a laser driven 2 spin-$\frac{1}{2}$ particle system with a known envelope function will be given in forthcoming work.

Recent work by Brody, Meier and Gibbons [6–8] tackles similar navigation problems. The work in [8] specifically obtains identical results to our equations (26, 27), using an interesting method complementing those in this paper, for the time optimal navigation trajectories and associated control Hamiltonians.

Acknowledgments

We thank Ian Macintosh for his helpful comments on the Killing fields of invariant metrics on $SU(n)$ and other matters of differential geometry. We thank Eli Hawkins for his helpful comments on invariant Lagrangians on Lie groups. We thank Robert Bryant for many helpful comments, answers and references on the MathOverflow Stack Exchange. We further thank our anonymous reviewers who offered many detailed and helpful comments and corrections. Russell is supported by an EPSRC studentship.

References

[1] C. Altafini and F. Ticozzi. Modeling and Control of Quantum Systems: An Introduction. *IEEE Transactions on Automatic Control*, 57:1898–1917, 2012.

[2] P. L. Antonelli, editor. *Handbook of Finsler geometry. vol 1.* Kluwer Academic Publishers, 2003.

[3] David Bao, Colleen Robles, and Zhongmin Shen. Zermelo navigation on Riemannian manifolds. *J. Differential Geometry*, 66:377–435, 2004.

[4] Martin Bauer, Martins Bruveris, and Peter W. Michor. Overview of the geometries of shape spaces and diffeomorphism groups. *Journal of Mathematical Imaging and Vision*, 50(1-2):60–97, 2014.

[5] R. Berndt. *Representations of Linear Groups: An Introduction Based on Examples from Physics and Number Theory.* Vieweg, 2007.

[6] Dorje C. Brody, Gary W. Gibbons, and David M. Meier. Time-optimal navigation through quantum wind. *New Journal of Physics*, 17, 2015. (to appear) arXiv:1410.6724 [quant-ph].

[7] Dorje C. Brody and David M. Meier. Elementary solution to the time-independent quantum navigation problem. *Journal of Physics A*, 48(5):055302, 2015.

[8] Dorje C. Brody and David M. Meier. Solution to the quantum Zermelo navigation problem. *Physical Review Letters*, 114, 2015. (to appear) arXiv:1409.3204 [quant-ph].

[9] Daniel Bump. *Lie Groups.* Springer, 2004.

[10] T. Caneva, M. Murphy, T. Calarco, R. Fazio, S. Montangero, V. Giovannetti, and G. E. Santoro. Optimal control at the quantum speed limit. *Phys. Rev. Lett.*, 103:240501, Dec 2009.

[11] Alberto Carlini, Akio Hosoya, Tatsuhiko Koike, and Yosuke Okudaira. Time-optimal quantum evolution. *Phys. Rev. Lett.*, 96:060503, 2006.

[12] Alberto Carlini, Akio Hosoya, Tatsuhiko Koike, and Yosuke Okudaira. Time-optimal unitary operations. *Phys. Rev. A*, 75:042308, 2007.

[13] Hernán Cendra, Jerrold E. Marsden, and Tudor S. Ratiu. Lagrangian reduction by stages. *Memoirs of the American Mathematical Society*, 152(722), 2001. (updated 2009).

[14] X. Cheng and Z. Shen. *Finsler Geometry: An Approach via Randers Spaces.* Springer, 2013.

[15] L. Conlon. *Differentiable Manifolds*. Birkhäuser Boston, 2008.

[16] Enrico Le Donne. Lecture notes on sub-Riemannian geometry. https://sites.google.com/site/enricoledonne/, 2010.

[17] Parastoo Habibi, Dariush Latifi, and Megerdich Toomanian. Homogeneous geodesics and the critical points of the restricted Finsler function. *Journal of Contemporary Mathematical Analysis*, 46(1):12–16, 2011.

[18] Brian C. Hall. *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*. Springer, 2003.

[19] Darryl D. Holm, Jerrold E. Marsden, and Tudor S. Ratiu. The Euler–Poincaré equations and semidirect products with applications to continuum theories. *Advances in Mathematics*, 137(1):1–81, 1998.

[20] Michael Hsieh and Herschel Rabitz. Optimal control landscape for the generation of unitary transformations. *Phys. Rev. A*, 77:042306, Apr 2008.

[21] Libing Huang and Xiaohuan Mo. On geodesics of Finsler metrics via navigation problem. *Proc. Amer. Math. Soc.*, 139:3015–3024, 2011.

[22] J. E. Humphreys. *Linear Algebraic Groups*. Graduate Texts in Mathematics. Springer, 1975.

[23] M. A. Javaloyes and M. Sánchez. On the definition and examples of Finsler metrics. arXiv:1111.5066 [math.DG], November 2011.

[24] M. A. Javaloyes and H. Vitório. Zermelo navigation in pseudo-Finsler metrics. arXiv:1412.0465 [math.DG], December 2014.

[25] N. Khaneja, S. J. Glaser, and R. Brockett. Sub-Riemannian geometry and time optimal control of three spin systems: Quantum gates and coherence transfer. *Phys. Rev. A*, 65(3):032301, March 2002.

[26] A. W. Knapp. *Lie Groups Beyond an Introduction*. Birkhäuser Boston, 2002.

[27] Wang-Sang Koon and Jerrold E. Marsden. Optimal control for holonomic and nonholonomic mechanical systems with symmetry and Lagrangian reduction, 1995. California Institute of Technology, CaltechCDSTR:1995.022.

[28] Dariush Latifi. Homogeneous geodesics of left invariant Finsler metrics. arXiv:0711.4480 [math.DG], 2007.

[29] Dariush Latifi and Megerdich Toomanian. Invariant naturally reductive randers metrics on homogeneous spaces. *Mathematical Sciences*, 6(1), 2012.

[30] K.-Y. Lee and H. F. Chau. Relation between quantum speed limits and metrics on $U(n)$. *Journal of Physics A Mathematical General*, 46(1):015305, January 2013.

[31] B. Li, Z. Yu, S. Fei, and X. Li-Jost. Time optimal quantum control of two-qubit systems. *Science China Physics, Mechanics, and Astronomy*, 56:2116–2121, 2013.

[32] C. López and E. Martínez. Sub-finslerian metric associated to an optimal control system. *SIAM J. Control Optim.*, 39(3):798–811, September 2000.

[33] Andrzej J. Maciejewski. Reduction, relative equilibria and potential in the two rigid bodies problem. *Celestial Mechanics and Dynamical Astronomy*, 63(1):1–28, 1995.

[34] J. E. Marsden and T. S. Ratiu. *Introduction to Mechanics and Symmetry: A Basic Exposition of Classical Mechanical Systems*. Springer, 1999.

[35] M. A. Nielsen. A geometric approach to quantum circuit lower bounds. *Quant. Info. Comp*, 6:213–262, 2006.

[36] José P. Palao and Ronnie Kosloff. Quantum computing by an optimal control algorithm for unitary transformations. *Phys. Rev. Lett.*, 89:188301, 2002.

[37] Walter Pfeifer. *The Lie Algebras su(N): An Introduction*. Birkhäuser, 2003.

[38] Colleen Robles. Geodesics in Randers spaces of constant curvature. arXiv:math/0501358 [math.DG], 2005.

[39] Benjamin Russell and Susan Stepney. Zermelo navigation and a speed limit to quantum information processing. *Physical Review A*, 90:012303, 2014.

[40] L. Šamaj and Z. Bajnok. *Introduction to the Statistical Physics of Integrable Many-body Systems*. Cambridge University Press, 2013.

[41] T. Schulte-Herbrüggen, A. Spörl, N. Khaneja, and S. J. Glaser. Optimal control for generating quantum gates in open dissipative systems. *Journal of Physics B Atomic Molecular Physics*, 44(15):154013, 2011.

[42] Z. Shen. Finsler Metrics with K = 0 and S = 0. arXiv:math/0109060 [math.DG], September 2001.

[43] G. Stefani. *Geometric Control Theory and Sub-Riemannian Geometry*. Springer, 2014.

[44] J. Stillwell. *Naive Lie Theory*. Springer, 2008.

[45] X. Wang, M. Allegra, K. Jacobs, S. Lloyd, C. Lupo, and M. Mohseni. Quantum brachistochrone curves as geodesics: obtaining accurate control protocols for time-optimal quantum gates. arXiv:1408.2465 [quant-ph], 2014.

[46] J. Werschnik and E. K. U. Gross. Quantum Optimal Control Theory. arXiv:0707.1883 [quant-ph], 2007.

[47] Ryozo Yoshikawa and SorinV. Sabau. Kropina metrics and Zermelo navigation on Riemannian manifolds. *Geometriae Dedicata*, 171(1):119–148, 2014.