ANALYTICAL AND NUMERICAL SOLUTION OF A SUB-RIEMANNIAN OPTIMAL CONTROL PROBLEM WITH APPLICATIONS TO QUANTUM SPIN SYSTEMS*

AMIT K. SANYAL†, CHRISTOPHER MOSELEY‡, AND ANTHONY BLOCH§

Abstract. Experiments in nuclear magnetic resonance (NMR) spectroscopy and NMR quantum computing require control of ensembles of quantum mechanical systems. The controlled transfer of coherence along a one-dimensional chain of spin systems plays a key role in NMR spectroscopy of proteins, and spin chains have also been proposed for NMR quantum information processing. The problem of time-optimal or energy-optimal control of these systems corresponds to finding optimal paths on Lie groups in which evolution in only certain directions on the group can be directly controlled. In this paper, we consider energy-optimal control of a three-spin system; this turns out to be a sub-Riemannian optimal control problem on SO(4). The goal of this optimal control problem is: given the initial configuration and the desired final configuration as the identity element in SO(4), design three control inputs that steer the system from the initial configuration to the identity $I \in$ SO(4) along an extremal trajectory. We first obtain necessary conditions for the normal extremal trajectories for both the continuous time system, and then for its discrete counterpart obtained from a discrete variational scheme. We also obtain expressions for the control inputs, and provide a numerical algorithm for the system which can be used to carry out accurate numerical simulations.

1. Introduction. Experiments in nuclear magnetic resonance (NMR), such as NMR spectroscopy and NMR quantum computing, require control of ensembles of quantum mechanical systems. In particular, the NMR spectroscopy of proteins involves transfer of coherent states along one-dimensional chains of spin systems [1]. Spin chains have been also proposed for the implementation of solid-state NMR quantum computers [2]. In practice, the transfer time should be as short as possible to minimize the effects of decoherence due to interaction with the laboratory environment, and to optimize the sensitivity of the experiments.

In [3], Yuan, Glaser and Khaneja present a method for efficient transfer of coherence along a linear chain of n-spin systems. The initial step in this method is

represented by the equation

$$(1) \qquad \frac{d}{dt} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & -u & 0 \\ 0 & u & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix},$$

where $x_1, \ldots, x_4$ represent expectation values of unitary operators that play a key role in the transfer, and u is the control. Yuan, Glaser and Khaneja transform this system into an equivalent one-control system in which the optimal control corresponds to finding a geodesic on the 2-sphere endowed with a special Riemannian metric. This is related to previous work by Khaneja et al [4].

Motivated by this prior work, we consider a generalization of the system represented by (1) involving three controls u_1, u_2 and u_3. The optimal control problem is equivalent to finding sub-Riemannian geodesics [5] on the Lie group SO(4). In the special case where u_1 and u_3 are fixed at unity, one recovers the system considered in [3]. We obtain the extremal solutions to this optimal control problem for both continuous and discrete time. The discrete time results are then used to form an algorithm for a numerical solution to this sub-Riemannian optimal control problem. This algorithm provides a way to compute optimal controls for the state transfer problem considered in the first part of [3] and other similar state transfer problems for the three-spin system."

This paper is organized as follows. Section 2 introduces the continuous sub-Riemannian optimal control problem on SO(4). In Section 3, we obtain extremal trajectories for the continuous optimal control problem using Pontryagin's maximum principle [6]. Then in Section 4, we introduce the discrete time sub-Riemannian optimal control problem corresponding to the continuous optimal controls introduced in Section 2. We also obtain discrete extremal trajectories for this problem in this section, using two alternate methods: that of direct discretization using variational principles, and that of applying Pontryagin's maximum principle. Because of the Hamiltonian structure of the underlying system, there are certain scalar quantities that are shown to be conserved in both the continuous time and discrete time extremals. These discrete extremals are used to obtain an iterative numerical algorithm to numerically simulate the (continuous) extremal trajectories in Section 5. Section 6 presents a set of numerical simulation results obtained using this numerical scheme for the extremal trajectory and optimal control inputs. Finally, we summarize the developments made in this paper in the concluding Section 7 and discuss possible future developments.

2. The Continuous Optimal Control Problem. Consider the control system defined by

$$(2) \qquad \frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 & u_1 & 0 & 0 \\ -u_1 & 0 & u_2 & 0 \\ 0 & -u_2 & 0 & u_3 \\ 0 & 0 & -u_3 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix},$$

in which $u_i : \mathbb{R} \to \mathbb{R}$ is differentiable for $i = 1, 2, 3$. The system (2) can be regarded equivalently as defining a family of differentiable curves in the matrix group $SO(4)$. Any such curve $Q : \mathbb{R} \to SO(4)$ satisfies the kinematic equation

$$\dot{Q} = \left[u_1(t)X^1 + u_2(t)X^2 + u_3(t)X^3 \right] Q(t)$$
$$(3) \qquad\qquad = U(t)Q(t)$$

where X^1, X^2, X^3 are elements of the basis $\{X^1, \ldots, X^6\}$ for $\mathfrak{so}(4) = T_I SO(4)$ defined by

$$X^1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad X^2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$X^3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix}, \quad X^4 = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$(4) \qquad X^5 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix}, \quad X^6 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},$$

and

$$(5) \qquad U(t) = u_1(t)X^1 + u_2(t)X^2 + u_3(t)X^3.$$

In Section 3, we obtain necessary conditions for a differentiable path $Q(t)$ satisfying (3) and the boundary conditions

$$Q(0) = Q_0, \quad Q(T) = I,$$

to be a *normal extremal* trajectory (defined below) with respect to the functional

$$(6) \qquad \mathcal{E}(Q) = \int_0^T L(Q(t), U(t)) \, dt,$$

where

$$(7) \qquad L(Q, U) = \frac{1}{2} \left[(u_1)^2 + (u_2)^2 + (u_3)^2 \right].$$

We will refer to this functional as the *action functional,* following Roger Brockett's usage [7]. Extremals of this functional are called *sub-Riemannian geodesics* [5], since they are optimal with respect to a smooth inner product restricted to the 3-plane field spanned by right translations of $\{X^1, X^2, X^3\}$. We also obtain evolution equations for a discrete version of this problem in Section 4, which are then used to obtain a numerical integration algorithm in Section 5 to numerically simulate the normal extremal trajectories in Section 6.

3. Normal extremals for the Continuous Optimal Control Problem. We use the notation $\langle A, B \rangle$ to denote the inner product of two matrices $A, B \in \mathbb{R}^{n \times n}$ defined by

$$(8) \qquad \langle A, B \rangle = \frac{1}{2} \mathrm{tr}(A^T B).$$

The basis $\{X^1, \ldots X^6\}$ is orthonormal with respect to this inner product.

Let U be defined as above, and let V denote a vector of the form $V = v_1 X^4 + v_2 X^5 + v_3 X^6 \in \mathfrak{so}(4)$, so that $\langle U, V \rangle = 0$. Note that for any such V we may rewrite the action Lagrangian as

$$(9) \qquad L(Q, U, V) = \frac{1}{2} \langle U, U \rangle + \langle U, V \rangle.$$

It turns out that this formulation determines a choice of V for each U along extremals of the action functional (Proposition 1).

According to the maximum principle of optimal control theory, there exists a matrix function P, called the *costate* associated with the state function Q, so that the extremals of (6) minimize the Hamiltonian

$$(10) \qquad H(P, Q, U, V) = \langle P, \dot{Q} \rangle - p_0 L(Q, U, V)$$

$$= \langle P, UQ \rangle - p_0 \left(\frac{1}{2} \langle U, U \rangle + \langle U, V \rangle \right)$$

where $p_0 \in \{0, 1\}$. Solutions to the optimal control problem with $p_0 = 1$ are called *normal extremals*, and solutions with $p_0 = 0$ are *abnormal extremals*. In this paper, we obtain necessary conditions for a curve in SO(4) to be a normal extremal.

PROPOSITION 1. *Any extremal Q of* (6) *with costate P satisfies* Hamilton's equations:

$$(11) \qquad \dot{Q} = UQ, \ \dot{P} = -U^\top P.$$

Proof. The first equation is merely (3). The second equation follows from

$$\dot{P} = -\text{grad}_Q H = -\text{grad}_Q \langle P, UQ \rangle$$

(12)
$$= -\text{grad}_Q \langle U^\top P, Q \rangle = -U^\top P.$$

$\square$

By the general theory of the maximum principle, the term $\langle P, UQ \rangle$ in the Hamiltonian H represents the pairing of a right-invariant 1-form with the right-invariant vector field UQ. Indeed,

$$\langle P, UQ \rangle = \frac{1}{2}\text{tr}(P^\top UQ) = \frac{1}{2}\text{tr}(QP^\top U)$$

$$= \frac{1}{4}\text{tr}\big((QP^\top - PQ^\top)U\big)$$

$$= \frac{1}{4}\text{tr}\big((PQ^\top - QP^\top)^\top U\big)$$

(13)
$$= \langle \frac{1}{2}(PQ^\top - QP^\top), U \rangle.$$

Therefore, $\langle P, UQ \rangle$ is equivalent to the pairing of the linear functional $\langle \frac{1}{2}(PQ^\top - QP^\top), \cdot \rangle \in \mathfrak{so}(4)^*$ with $U \in \mathfrak{so}(4)$.

Define $\lambda_i = \langle P, X^i Q \rangle = \langle M, X^i \rangle$, for $i = 1, \dots 6$. Then by (13), we have

(14)
$$\frac{1}{2}\big(PQ^\top - QP^\top\big) = \sum_{i=1}^{6} \lambda_i X^i.$$

PROPOSITION 2. *Along any extremal curve* $Q : [0, T] \to \text{SO}(4)$,

(15)
$$\frac{d}{dt}M = [U, M], \quad M \triangleq \frac{1}{2}\big(PQ^\top - QP^\top\big),$$

where the bracket denotes the matrix commutator.

Proof. If $Q : [0, T] \to \text{SO}(4)$ is an extremal, then

$$\frac{d}{dt}(PQ^\top) = \dot{P}Q^\top + P(\dot{Q^\top})$$

$$= -U^\top PQ^\top + PQ^\top U^\top \quad \text{(by (11))}$$

$$= U(PQ^\top) - (PQ^\top)U$$

$$= [U, PQ^\top],$$

as given by (15). Similarly, one can show that

$$\frac{d}{dt}(QP^\top) = [U, QP^\top].$$

Therefore the result holds. $\square$

For any normal extremal of the action functional, M determines $U \in \text{span}\{X^1, X^2, X^3\} \subset \mathfrak{so}(4)$ and an associated element $V \in \mathfrak{so}(4)$ in the orthogonal complement of $\text{span}\{X^1, X^2, X^3\}$ in the following way.

THEOREM 1. *Any normal extremal* $Q : [0, T] \to \mathrm{SO}(4)$ *necessarily satisfies* $U + V = M$, *so that*

$$
(16) \qquad
\begin{aligned}
u_1 &= \lambda_1, \quad u_2 = \lambda_2, \quad u_3 = \lambda_3, \\
v_1 &= \lambda_4, \quad v_2 = \lambda_5, \quad v_3 = \lambda_6.
\end{aligned}
$$

Moreover, the component functions u_i *and* v_i *satisfy*

$$
(17) \qquad
\begin{aligned}
\dot{u}_1 &= u_2 v_1, & \dot{v}_1 &= u_3 v_2, \\
\dot{u}_2 &= u_3 v_3 - u_1 v_1, & \dot{v}_2 &= -u_3 v_1 + u_1 v_3, \\
\dot{u}_3 &= -u_2 v_3, & \dot{v}_3 &= -u_1 v_2.
\end{aligned}
$$

The system of ordinary differential equations (17) *is the* Euler system *associated with the extremal trajectory.*

Proof. By the maximum principle, a necessary condition for $Q : [0, T] \to \mathrm{SO}(4)$ to be an extremal is that $\mathrm{grad}_U H = M - (U + V) = 0$, so that $U + V = M$. Since $\{X^1, \ldots, X^6\}$ is an orthonormal basis for $\mathfrak{so}(4)$, this establishes equations (16). Equations (17) then follow from Proposition 2 and equations (16). $\qquad\square$

As a corollary, we recover the symmetric equation obtained by Bloch, Crouch and Ratiu in [8] for normal extremals of a generalization of this sub-Riemannian optimal control problem to a semisimple Lie group.

COROLLARY 1. *For any normal extremal,*

$$
(18) \qquad \dot{U} - [U, V] + \dot{V} = 0.
$$

Proof. From Proposition 1 and Proposition 2, $\frac{d}{dt}(U + V) = \frac{d}{dt}(M) = [U, M] = [U, V]$. $\qquad\square$

Note that equations (18) and (17) and equivalent. Systems of equations similar to (18) were obtained in earlier work by Brockett in [7], and also the more recent work in [9].

Although closed-form solutions of the Euler system (17) cannot be obtained, the following conservation laws are easy to check.

PROPOSITION 3. *The quantities*

$$
(19) \qquad \langle U, U \rangle = (u_1)^2 + (u_2)^2 + (u_3)^2 = C_1,
$$

$$
(20) \qquad \langle V, V \rangle = (v_1)^2 + (v_2)^2 + (v_3)^2 = C_2,
$$

$$
(21) \qquad u_1 u_3 - v_1 v_3 + u_2 v_2 = C_3,
$$

are constant along any normal extremal.

Proof. By equations (17), the derivative of each expression on the left is 0. $\qquad\square$

In the following sections, we derive a numerical algorithm to numerically solve for the normal extremal trajectories of this system given boundary conditions.

4. The Discrete Optimal Control Problem. For numerical computations of extremal trajectories, we use direct discretization of the optimal control problem of minimizing the action functional (6). This is an application of the *discrete variational principle* [10, 11], which leads to a symplectic integration algorithm for numerical computation of the extremal trajectories.

Consider a discrete counterpart to the kinematic equation (3):

$$Q_{k+1} = \exp\left(hU_k\right)Q_k$$

(22) $$\text{with }\ Q_0 = Q_I,\ \ Q_N = I$$

where $Q_k \in \mathrm{SO}(4)$, $U_k \in \mathrm{span}\{X^1, X^2, X^3\}$ for each $k = 0, 1, 2, \ldots, N$, for some positive integer N, $Q_I \in \mathrm{SO}(4)$ is the given initial state and h is a time stepsize for the discretization given by $h = T/N$ where $Q(t)$, $t \in [0, T]$ is the corresponding continuous time trajectory. We introduce the action sum

(23) $$\mathcal{E}_d = h \sum_{k=0}^{N} \left(\frac{1}{2}\langle U_k,\, U_k\rangle + \langle U_k,\, V_k\rangle\right)$$

where $V_k \in \mathrm{span}\{X^4, X^5, X^6\}$ for each $k = 0, 1, 2, \ldots, N$. The discrete optimal control problem corresponds to minimizing this cost function subject to (22).

4.1. Discrete Variational Approach. We examine necessary conditions from the variational viewpoint. An admissible variation of Q_k has the form

(24) $$\delta Q_k = Y_k Q_k$$

where $Y_k \in \mathrm{span}\{X^1, X^2, X^3\}$ and $Y_0 = Y_N = 0$.

By equation (22), $\exp\left(U_k\right) = Q_{k+1}\,Q_k^T$, from which we find that

$$h\delta U_k \exp\left(hU_k\right) = \delta(\exp\left(hU_k\right))$$
$$= \delta Q_{k+1} Q_k^T + Q_{k+1}(\delta Q_k)^T$$
$$= Y_{k+1} Q_{k+1} Q_k^T + Q_{k+1} Q_k^T Y_k^T$$
(25) $$= Y_{k+1} \exp\left(hU_k\right) - \exp\left(hU_k\right)Y_k,$$

so that admissible variations of the U_k have the form

(26) $$h\delta U_k = Y_{k+1} - \mathrm{Ad}_{\exp\left(hU_k\right)}Y_k.$$

As in the continuous case, there is a $V_k \in \mathrm{span}\{X^4, X^5, X^6\}$ associated with each U_k, $k = 0, \ldots, N$. Admissible variations of V_k take the form $\delta V_k = Z_k$ where $\langle Y_k,\, Z_k\rangle = 0$.

THEOREM 2. *Normal extremals corresponding to the discrete variational problem of minimizing (23) satisfy*

$$Q_{k+1} = \exp\left(hU_k\right)Q_k$$
(27) $$U_{k+1} - U_k + Ad_{\exp\left(-hU_{k+1}\right)}V_{k+1} - V_k = 0,$$

for $k = 0, 1, 2, \ldots, N - 1$.

Proof. The first of equations (27) is simply the discrete kinematics equation. The first variation of $\mathcal{E}_d$ is given by

$$\delta\mathcal{E}_d = h \sum_{k=0}^{N} \left(\langle U_k, \delta U_k \rangle + \langle V_k, \delta U_k \rangle + \langle U_k, \delta V_k \rangle \right)$$

$$(28) \qquad = \sum_{k=0}^{N} \left(\langle U_k + V_k, Y_{k+1} - \mathrm{Ad}_{\exp(hU_k)} Y_k \rangle + \langle U_k, Z_k \rangle \right),$$

using equation (26) and $\delta V_k = Z_k$. The necessary conditions for extremality are then given by

$$(29) \qquad \delta\mathcal{E}_d = \sum_{k=1}^{N} \langle U_{k-1} + V_{k-1} - \mathrm{Ad}_{\exp(-hU_k)}(U_k + V_k), Y_k \rangle = 0,$$

and $\langle U_k, Z_k \rangle = 0$. Equation (29) is equivalent to the necessary condition in the second of equations (27) for extremal trajectories. $\square$

These discrete extremal trajectories can be used to numerically simulate the continuous extremal trajectories with given initial and final points Q_0 and Q_N on SO(4). In fact, we can show that the first order (in h) approximations of the continuous (18) and discrete (27) extremals are identical. For the discrete extremals (27), we obtain the first order approximation as follows:

$$U_k - U_{k-1} + \left(V_k - h[U_k, V_k] \right) - V_{k-1} \approx 0$$

$$(30) \qquad \Rightarrow \frac{U_k - U_{k-1}}{h} - [U_k, V_k] + \frac{V_k - V_{k-1}}{h} \approx 0.$$

A first order approximation of the continuous extremals (18) will be identical to (30), since the "velocities" to first order are approximated as

$$\dot{U}(t_k) = \frac{U_k - U_{k-1}}{h}, \quad \dot{V}(t_k) = \frac{V_k - V_{k-1}}{h}.$$

Therefore the discrete extremals are equivalent to the continuous extremals at least up to the first order, and hence can be used to approximate the continuous extremals in a numerical simulation. The first-order approximation in (30) can be used to obtain discrete relations that are counterparts of equations (17) in the continuous case.

4.2. Applying Pontryagin's Maximum Principle. We can also obtain the extremals for the discrete optimal control problem of minimizing (23) using the maximum principle. Appending the kinematic and control constraints in (22) to the cost function (23), we can form the discrete Hamiltonian for this problem as below:

$$(31) \qquad H(P_{k+1}, Q_k, U_k, V_k) = H_k = \langle P_{k+1}, \exp(hU_k)Q_k \rangle - \frac{1}{2}\langle U_k, U_k \rangle - \langle V_k, U_k \rangle.$$

This discrete Hamiltonian can also be expressed as:

$$(32) \qquad H(P_{k+1}, Q_k, U_k, V_k) = \langle \exp(-hU_k)P_{k+1}, Q_k \rangle - \frac{1}{2}\langle U_k, U_k \rangle - \langle V_k, U_k \rangle.$$

The discrete extremals can now be obtained by applying the maximum principle to the discrete Hamiltonian, as given in the following result.

THEOREM 3. *Extremals (Q_k, P_k) of the discrete optimal control problem of min-imizing (23) subject to (22) satisfy*

$$Q_{k+1} = \exp(hU_k)Q_k,$$
$$P_{k+1} = \exp(hU_k)P_k, \quad and$$
$$(33) \qquad U_k + Ad_{\exp(-hU_k)}V_k = hM_k, \quad M_k = \frac{1}{2}\left(P_k Q_k^\top - Q_k P_k^\top\right),$$

for $k = 0, 1, \ldots, N-1$.

Proof. The necessary conditions for extremality are obtained from the discrete Hamiltonian (32) as:

$$Q_{k+1} = \operatorname{grad}_{P_{k+1}} H_k = \exp(hU_k)Q_k,$$
$$P_k = \operatorname{grad}_{Q_k} H_k = \exp(-hU_k)P_{k+1}$$
$$\Rightarrow P_{k+1} = \exp(hU_k)P_k,$$

which give the first two of equations (33). Now applying the maximum principle to the discrete Hamiltonian, we get

$$(34) \qquad \operatorname{grad}_{U_k} H_k = \frac{h}{2}\left(P_{k+1}Q_k^\top \exp(-hU_k) - \exp(hU_k)Q_k P_{k+1}^\top\right) - U_k - V_k = 0.$$

Substituting $P_{k+1} = \exp(hU_k)P_k$ into (34), we obtain

$$(35) \qquad \frac{h}{2}Ad_{\exp(hU_k)}\left(P_k Q_k^\top - Q_k P_k^\top\right) = U_k + V_k,$$

which can be rewritten as the third of equations (33). □

From the first two of equations (33), we note that the discrete trajectory expressed as (Q_k, P_k) can be restricted naturally to the product space $SO(4) \times SO(4)$, as was the case for their continuous counterparts in (11). However, such a restriction may include only some of the discrete extremals while other extremals are excluded. Further, Q_k and P_k evolve along the same discrete vector field U_k, $k = 0, 1, \ldots, N$. To relate these equations to equation (27) in the variables U_k and V_k, we rewrite equation (34) as:

$$(36) \qquad hM_{k+1} = U_k + V_k,$$

by substituting the first of equations (33) into equation (34). Therefore, we have

$$(37) \qquad hM_k = U_{k-1} + V_{k-1}.$$

Substituting the above back into the third of equations (33), we get

$$U_k + \mathrm{Ad}_{\exp(-hU_k)}V_k = U_{k-1} + V_{k-1},$$

which is identical to equation (27). The expression (36) also implies that

$$M_k \in \mathfrak{so}(4)^*,$$

or $M_k Q_k \in \mathrm{T}^*_{Q_k}\mathrm{SO}(4)$. Using equations (33), we can show the following result, as the discrete counterpart to Proposition 3.

PROPOSITION 4. *The following quantities*

(38)
$$D_1 = \langle U_k, U_k \rangle, \quad D_2 = \langle V_k, V_k \rangle, \quad and$$
$$D_3 = u_{k_1} u_{k_3} - v_{k_1} v_{k_3} + u_{k_2} v_{k_2}$$

are constant along the extremals (Q_k, U_k, V_k) *of the discrete optimal control problem of minimizing (23) subject to (22).*

Proof. Using equations (35) and (37), we obtain:

$$\langle U_k + V_k, U_k + V_k \rangle = h^2 \langle Ad_{\exp(hU_k)}M_k, Ad_{\exp(hU_k)}M_k \rangle$$
$$= h^2 \langle M_k, M_k \rangle$$
$$= \langle U_{k-1} + V_{k-1}, U_{k-1} + V_{k-1} \rangle.$$

Since the U_k and V_k are orthogonal in the trace inner product, the above equation gives

$$\langle U_k, U_k \rangle + \langle V_k, V_k \rangle = \langle U_{k-1}, U_{k-1} \rangle + \langle V_{k-1}, V_{k-1} \rangle,$$

which leads to the conserved quantities D_1 and D_2 in (38), as U_k and V_k are in trace-orthogonal complements in $\mathfrak{so}(4)$. Similarly, it can be verified using equation (27) that the quantity D_3 is conserved along the discrete extremals. $\qquad\square$

Note that the quantities D_1, D_2 and D_3 are the discrete counterparts of the continuous functions C_1, C_2 and C_3 given by Proposition 3.

These discrete extremal trajectories can be used as part of a numerical algorithm for numerically obtaining the continuous extremal trajectories. This involves solving a discrete two-point boundary value problem. Generally, a numerical algorithm to solve this problem would consist of applying a shooting method so that the initial and final configurations of the computed extremal trajectory are "close to" (up to an error tolerance bound of) the given initial and final configurations ($Q_0 = Q_I$ and $Q_N = I$).

5. Derivation of Numerical Algorithm. In this section, we outline a numerical algorithm for simulating the extremals of the original continuous-time optimal control problem with the cost function (6). This is done by using the corresponding

discrete extremal trajectories given by equations (22) and (27) as a numerical integration scheme. To match the given terminal configurations, we obtain the sensitivity derivative of these configurations to changes in the terminal conditions in the control variables (U and V). These sensitivity derivatives are then used to modify the terminal values of the control variables and numerically integrate them using (27) in an iterative process to match the given terminal configurations.

5.1. Two Point Boundary Value Problem. The two point boundary value problem associated with finding the discrete extremal trajectory is given by the following equations:

$$(39) \qquad Q_k = \exp\left(-hU_k\right)Q_{k+1}, \quad Q_0 = Q_I, \quad Q_N = I,$$

$$(40) \qquad U_{k-1} + V_{k-1} = U_k + \mathrm{Ad}_{\exp\left(-hU_k\right)}V_k.$$

Note that these expressions above give an explicit map

$$(Q_{k+1}, U_k, V_k) \mapsto (Q_k, U_{k-1}, V_{k-1})$$

that can be used for backward time integration. In contrast, the forward time integration map given by equations (22) and (27) is implicit.

Since equations (39) and (40) are explicit, they are fast and easy to implement as a backward time numerical integration scheme. Given Q_{k+1} and U_k, one can obtain Q_k from equation (39). Given U_k and V_k, the component of the right hand side of equation (40) in span$\{X^1, X^2, X^3\}$ is equated to U_{k-1} while the component in span$\{X^4, X^5, X^6\}$ is equated to V_{k-1}. To use this as an integration scheme, we begin from the given final configuration $Q_N = I$ and choose U_{N-1} and V_{N-1}; this means we have to choose 6 scalar parameters as $U_{N-1} + V_{N-1} \in \mathfrak{so}(4)$. Then this set of equations is integrated backwards in time till we reach (Q_1, U_0, V_0). Finally, we check to see that $Q_0 = \exp\left(-hU_0\right)Q_1$ is close to (within the error tolerance bound of) the given initial configuration Q_I.

An alternate method would be to solve this set of equations as a nonlinear root-finding problem with the U_i and V_i ($i = 0, 1, \ldots, N-1$) as variables. However, this method would involve solving for a large number ($= 6(N-1)$) of variables, which would require a large number of iterations by Newton's method and result in a slow and cumbersome algorithm. The shooting method employed here consists of changing the control parameters U_{N-1} and V_{N-1} in an iterative procedure until the initial condition is satisfied up to a preset error tolerance bound. This involves solving the 6 dimensional two-point boundary value problem given by (39) and (40).

The main drawback of a shooting method is that the extremal solutions are very sensitive to changes in the initial values of the parameters to be chosen. The non-linearity of the equations (39) and (40) makes it difficult to construct an accurate estimate of the sensitivity. In addition, the effect of the control variables V_k on the

configuration Q_k is indirect and acts via the control equation (40). This makes it difficult to directly compute the sensitivity of the configuration to these variables. In particular, for the shooting method to work, we require knowledge of the sensitivity derivatives of Q_0 to U_{N-1} and V_{N-1}. A computational scheme to obtain these derivatives numerically is described in what follows next.

5.2. Discrete Linearized System. The idea is to find a linear transformation between the initial and final states of the discrete system of equations (39) and (40). We first obtain the linearized equations of motion for this discrete system. We represent the configuration $Q_k \in \mathrm{SO}(4)$ in the linearized equations by the "exponential coordinates" given by the inverse of the exp map:

$$Q_k = \exp(S_k) \Leftrightarrow S_k = \exp^{-1}(Q_k) = \log(Q_k) \in \mathfrak{so}(4).$$

Therefore, a perturbation of S_k results in a perturbation of Q_k given by:

$$\Delta Q_k = \Delta S_k \exp(S_k) = \Delta S_k Q_k.$$

Using this expression, and taking perturbations of Q_k, U_k and Q_{k+1} in equation (39), we obtain

$$\Delta S_k Q_k = \Big(-h\Delta U_k \exp(-hU_k) + \exp(-hU_k)\Delta S_{k+1} \Big)Q_{k+1},$$

and using equation (39) again on this expression gives us

(41) $$\Delta S_k = -h\Delta U_k + \mathrm{Ad}_{\exp(-hU_k)}\Delta S_{k+1}.$$

A first order (in time step h) expansion of equation (41) gives us the following linear equation

(42) $$\Delta S_k = \Delta S_{k+1} - h[U_k, \Delta S_{k+1}] - h\Delta U_k,$$

which is a linear first-order approximation of the discrete configuration equation (39). The linearized first-order equation corresponding to (40) is obtained directly by linearizing (30) as the following:

(43) $$\Delta U_{k-1} + \Delta V_{k-1} = \Delta U_k - h\Big([\Delta U_k, V_k] + [U_k, \Delta V_k]\Big) + \Delta V_k.$$

Note that the corresponding continuous versions of equations (42) and (43) will be *time-varying*, since the transformation map between $(\Delta S_k, \Delta U_{k-1}, \Delta V_{k-1})$ and $(\Delta S_{k+1}, \Delta U_k, \Delta V_k)$ depends on (U_k, V_k). This transformation is given by the linear

equation below:

$$
\begin{bmatrix} I & 0 & 0 \\ 0 & I & I \end{bmatrix} \begin{bmatrix} \Delta S_k \\ \Delta U_{k-1} \\ \Delta V_{k-1} \end{bmatrix} =
$$

(44)
$$
\begin{bmatrix} I - h\,\mathrm{ad}_{U_k} & -hI & 0 \\ 0 & I + h\,\mathrm{ad}_{V_k} & I - h\,\mathrm{ad}_{U_k} \end{bmatrix} \begin{bmatrix} \Delta S_{k+1} \\ \Delta U_k \\ \Delta V_k \end{bmatrix},
$$

where ad denotes the adjoint representation of $\mathfrak{so}(4)$,

$$
\mathrm{ad}_A B = [A, B].
$$

Since ΔU_{k-1} and ΔV_{k-1} are in orthogonal complements of $\mathfrak{so}(4)$, equation (43) or (44) uniquely determines them. One can use linear projection operators in the linear mapping above to obtain the linear transformation of states

(45)
$$
X_{k-1} = \mathbf{A}_k X_k, \quad X_k = \begin{bmatrix} \Delta S_{k+1} \\ \Delta U_k \\ \Delta V_k \end{bmatrix},
$$

where $\mathbf{A}_k$ is obtained from equation (44). Since $\Delta S_k \in \mathfrak{so}(4)$, $\Delta U_k \in \mathrm{span}\{X^1, X^2, X^3\} \subset \mathfrak{so}(4)$ and $\Delta V_k \in \mathrm{span}\{X^4, X^5, X^5\} \subset \mathfrak{so}(4)$, they can also be expressed as corresponding vectors $\Delta s_k \in \mathbb{R}^6$, $\Delta u_k \in \mathbb{R}^3$, and $\Delta v_k \in \mathbb{R}^3$, as $\mathfrak{so}(4)$ is isomorphic to $\mathbb{R}^6$. In this vector representation, the state transformation equation is

(46)
$$
x_{k-1} = \mathcal{A}_k x_k, \quad x_k = \begin{bmatrix} \Delta s_{k+1} \\ \Delta u_k \\ \Delta v_k \end{bmatrix} \in \mathbb{R}^{12}.
$$

5.3. Numerical Algorithm. From the linearized state equation (46), we obtain the linear map from the final state to the initial state, as follows:

(47)
$$
x_0 = \Phi x_{N-1}, \quad \Phi = \mathcal{A}_1 \mathcal{A}_2 \cdots \mathcal{A}_{N-1}.
$$

The matrix Φ is the sensitivity matrix or sensitivity derivative of the initial conditions with respect to the final conditions. Note that since we begin backwards integration from the final configuration $Q_N = I$ which is known, $\Delta s_N = 0$. Also, the error in the initial configuration Δs_0 can be obtained from Δs_1 and Δu_0 using equation (42), and is given by

$$
\Delta s_0 = L y_0, \quad \text{where } y_0 \triangleq \begin{bmatrix} \Delta s_1 \\ \Delta u_0 \end{bmatrix} \in \mathbb{R}^9,
$$

(48)
$$
L = \begin{bmatrix} \underbrace{I - h\,\mathrm{ad}_{u_0}}_{6\times 6} & \underbrace{\begin{bmatrix} -hI \\ 0 \end{bmatrix}}_{6\times 3} \end{bmatrix}.
$$

The relation between y_0 and $z_{N-1} \triangleq [\Delta u_{N-1}^\top \ \ \Delta v_{N-1}^\top] \in \mathbb{R}^6$ is given by

$$y_0 = \Psi z_{N-1}, \tag{49}$$

where $\Psi \in \mathbb{R}^{9 \times 6}$ is the top-right 9×6 block of Φ. Thus the sensitivity of Δs_0 to z_{N-1} at any iteration step is given by the following map:

$$\Delta s_0 = \mathcal{S} z_{N-1}, \ \ \mathcal{S} \triangleq L\Psi. \tag{50}$$

Equation (50) also defines the sensitivity matrix $\mathcal{S}$. Since the initial configuration $Q_0 = Q_I$ is known, the criterion for convergence to the optimal solution is $\Delta s_0 \to 0$.

Using the sensitivity matrix, an initial guess of the change in the unspecified final control is iterated to satisfy the specified initial configuration. Any iterative scheme of Newton type may be applied. We use a Newton-Armijo iterative scheme [12], which is a line search using the Newton search direction, together with backtracking to ensure sufficient descent of the residual error in initial configuration. A similar scheme has been recently applied to a rigid body attitude control problem in [13]. This numerical algorithm is outlined in steps below.

1. Guess the final controls U_{N-1} and V_{N-1}.
2. Find Q_k, U_{k-1} and V_{k-1} using (39)-(40).
3. Compute the initial condition error; error= $\|\Delta s_0\|$.
4. Set error^t =error, $i = 1$.
5. **while** error$> \epsilon$
6.   Find a line search direction; $D_s = \mathcal{S}^{-1}$.
7.   Set $c = 1$.
8.   **while** error$^t > (1 - 2\alpha c)$error
9.   Choose a trial control $z_{N-1}^t = z_{N-1} + c D_s \Delta s_0$.
10.   Find Q_k, U_{k-1} and V_{k-1} using (39)-(40).
11.   Compute the initial condition error; error$^t = \|\Delta s_0^t\|$.
12.   Set $c = c/10$, $i = i + 1$.
13. **end while**
14. Set $z_{N-1} = z_{N-1}^t$, error=error^t. (accept the trial)
15. **end while**

Here i is the iteration number, and $\epsilon, \alpha \in \mathbb{R}$ are a stopping criterion and a scaling factor, respectively. The trial control vector is expressed as $Z_{N-1} = [u_{N-1}^\top \ \ v_{N-1}^\top]^\top \in \mathbb{R}^6$. While the outer loop finds a search direction by computing the sensitivity derivatives, the inner loop performs a line search to find the largest step size $c \in \mathbb{R}$ along the search direction. The error in satisfying the initial boundary condition is determined in each inner iteration.

6. Numerical Simulation Results. Results from a numerical simulation carried out using the numerical algorithm given in Section 5 are provided here. The
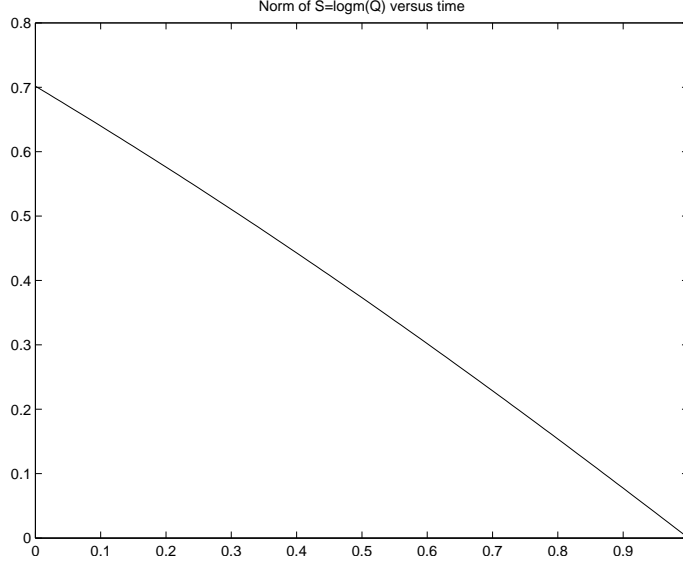
FIG. 1. *Matrix logarithm of optimal state Q with time*

quantities required to carry out this simulation are the given initial configuration

$$
Q_I = \begin{bmatrix}
0.9541 & -0.2965 & -0.0422 & -0.0039 \\
0.2995 & 0.9437 & 0.1401 & -0.0036 \\
0.0017 & -0.1122 & 0.7681 & 0.6304 \\
0.0048 & 0.0939 & -0.6234 & 0.7763
\end{bmatrix},
$$

the known final configuration $Q_N = I$, a simulated time interval of $T = 1$ second and a time stepsize of $h = 0.01$ second. We obtain the following results for this simulation.

Figure 1 plots the matrix logarithm of $Q \in \mathrm{SO}(4)$ against time obtained using our numerical scheme; with $S = \log(Q) \in \mathfrak{so}(4)$ defined so that $\exp(S) = Q$. For the given initial configuration Q_I, we have $= \log(Q_I)$ as below:

$$
S_I = \begin{bmatrix}
0 & -0.3037 & -0.0223 & 0.0004 \\
0.3037 & 0 & 0.1392 & -0.0530 \\
0.0223 & -0.1392 & 0 & 0.6806 \\
-0.0004 & 0.0530 & -0.6806 & 0
\end{bmatrix},
$$

with $\|S_I\| = 0.7$. On backwards time integration, our numerical scheme converges to an extremal trajectory with

$$
Q_0 = \begin{bmatrix}
0.9540 & -0.2969 & -0.0423 & -0.0039 \\
0.2998 & 0.9436 & 0.1402 & -0.0035 \\
0.0017 & -0.1123 & 0.7676 & 0.6310 \\
0.0048 & 0.0941 & -0.6240 & 0.7758
\end{bmatrix},
$$
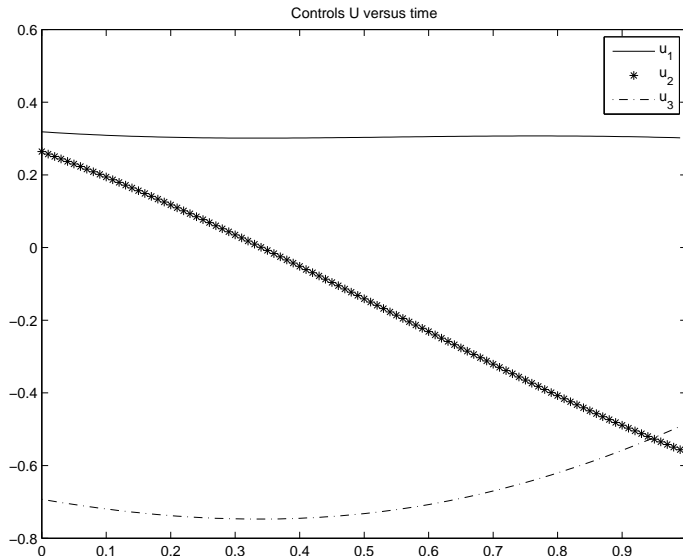
Fig. 2. *Components of optimal control U with time*
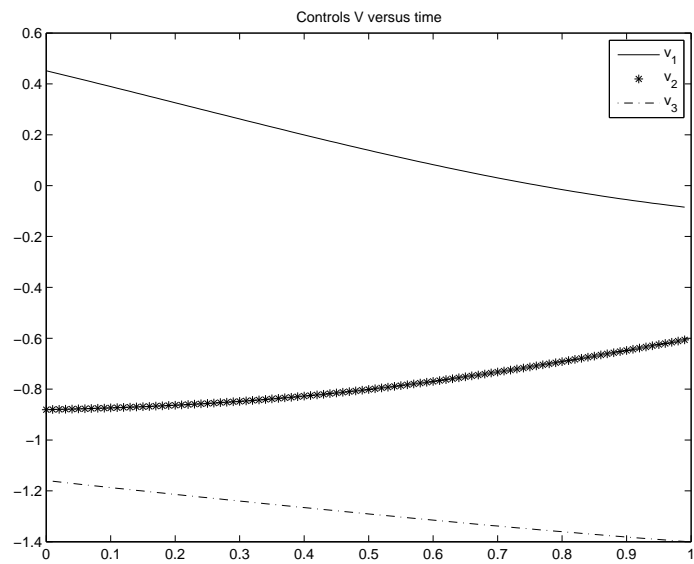
Fig. 3. *Components of V with time*

with $\|S_0\| = 0.7008$. The tolerance level for the initial configuration error was set to $\epsilon = \|\Delta S_0\| = 0.001$. Figures 2 and 3 plot the control components U and the "non-control" directions V respectively with time. As shown in these figures, our numerical integration scheme ensures that the initial configuration error in $Q_0 \in \mathrm{SO}(4)$ is within the required error tolerance level of the given initial configuration Q_I. The scheme converges in 105 iterations with initial guesses of the final control efforts U_N^0 and V_N^0
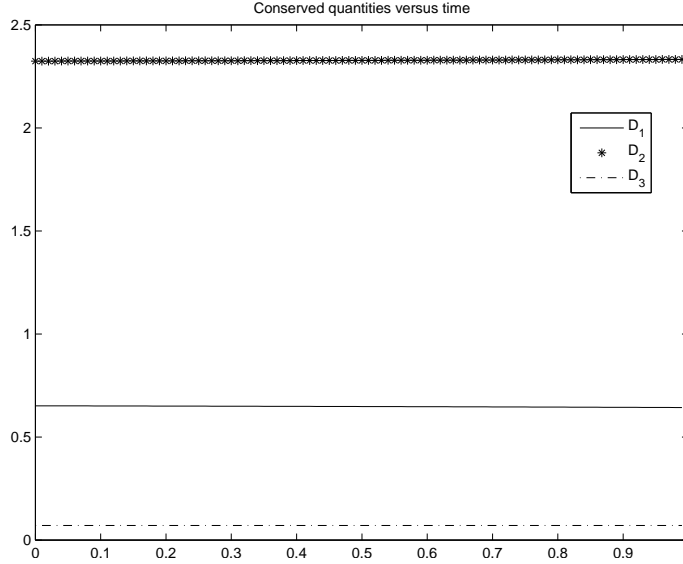
FIG. 4. *Components of conserved quantities* D_1, D_2, D_3 *with time*

$(Z_N^0 = U_N^0 + V_N^0)$ given by :

$$
Z_N^0 = \begin{bmatrix}
0 & 0.3255 & -0.6292 & 0.3973 \\
-0.3255 & 0 & 1.3585 & -0.6699 \\
0.6292 & -1.3585 & 0 & -0.1468 \\
-0.3973 & 0.6699 & 0.1468 & 0
\end{bmatrix}.
$$

The final control efforts U_N and V_N (or $Z_N = U_N + V_N$) obtained for the converged solution are as given below:

$$
Z_N = \begin{bmatrix}
0 & 0.3020 & 0.0846 & -0.6065 \\
-0.3020 & 0 & -0.5564 & 1.3991 \\
-0.0846 & 0.5564 & 0 & -0.4922 \\
0.6065 & -1.3991 & 0.4922 & 0
\end{bmatrix}.
$$

Figure 4 shows that the conserved scalar quantities $D_1 = \langle U_k, U_k \rangle$, $D_2 = \langle V_k, V_k \rangle$ and $D_3 = u_{k_1} u_{k_3} - v_{k_1} v_{k_3} + u_{k_2} v_{k_2}$ given by Proposition 4, which correspond to the conserved quantities C_1, C_2 and C_3 given by Proposition 3 along the normal extremals, are indeed conserved numerically as well.

7. Conclusions. This paper presents a novel numerical approach to solving a sub-Riemannian optimal control problem that arises in certain quantum spin systems. The continuous optimal control is first formulated, and an extremal solution is obtained. Then this continuous optimal control problem is discretized to obtain a numerical scheme that is symplectic in nature, and conserves certain first integrals of the system due to its origin in discrete variational principles. Numerical solutions of

the extremal trajectories using this numerical algorithm have been obtained, and show the efficiency of this numerical scheme and the numerically robust properties arising from its underlying discrete variational structure. Such numerical schemes may be applied to numerically simulate the implementation of specific logic gates using linear chains of spin systems.

REFERENCES

[1] J. Cavanagh, W. J. Fairbrother, A. G. Palmer, and N. J. Skelton, *Protein NMR Spectroscopy: Principles and Practice*, Academic Press, San Diego, 1996.

[2] T. D. Ladd, J. R. Goldman, A. Dãna, F. Yamaguchi, and Y. Yamamoto, *Quantum computation in a one-dimensional crystal lattice with NMR force microscopy*, quant-ph/0009122v4, 12 Jun 2001.

[3] H. Yuan, S.J. Glaser, and N. Khaneja, *Geodesics for efficient creation and propagation of order along Ising Spin Chains*, Phys. Rev. A , 76(2007), 012316.

[4] N. Khaneja, S. Glaser, and R. Brockett, *Sub-Riemannian geodesics and time optimal control of spin systems*, Proceedings of the American Control Conference, pp. 2806-2811, 2002.

[5] R. Montgomery, *A Tour of Subriemannian Geometries, Their Geodesics and Applications*, American Mathematical Society, Providence, RI, 2002.

[6] A. M. Bloch, *Nonholonomic Mechanics and Control*, Springer-Verlag, New York, 2003.

[7] R. W. Brockett, *Lie Theory and Control Systems Defined on Spheres*, SIAM J. Appl. Math., 25:2(1973), pp. 213-225.

[8] A. Bloch, P. Crouch, and T. Ratiu, *Sub-Riemannian Optimal Control Problems*, Fields Institute Communications, 3(1994), pp. 35-48.

[9] A. M. Bloch, P. E. Crouch, J. E. Marsden, and A. K. Sanyal, *Optimal Control and Geodesics on Quadratic Matrix Lie Groups*, Foundations of Computational Mathematics, 8:4(2008), pp. 469-500.

[10] J. E. Marsden and M. West, *Discrete Mechanics and Variational Integrators*, Acta Numerica, 10(2001), pp. 357-514.

[11] E. Hairer, C. Lubich, and G. Wanner, *Geometric Numerical Integration*, Springer-Verlag, Berlin, 2002.

[12] C. T. Kelley, *Iterative Methods for Linear and Nonlinear Equations*, SIAM, 2004.

[13] T. Lee, N. H. McClamroch, and M. Leok, *Optimal Attitude Control for a Rigid Body with Symmetry*, Proceedings of the American Control Conference, pp. 1073-1078, 2007.