

Spin Dynamics: A Paradigm for Time Optimal Control on Compact Lie Groups

G. DIRR¹, U. HELMKE¹, K. HÜPER², M. KLEINSTEUBER¹ and Y. LIU³

¹*Mathematical Institute, Würzburg University, Am Hubland, 97074 Würzburg, Germany (e-mail: {dirr, helmke, kleinsteuber}@mathematik.uni-wuerzburg.de)*

²*National ICT Australia, Canberra Research Laboratory, SEACS Program, Locked Bag 8001, Canberra ACT 2601, Australia, and Dept. of Information Engineering, Research School of Information Sciences and Engineering, The Australian National University, Canberra ACT 0200, Australia (e-mail: knut.hueper@nicta.com.au)*

³*Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong (e-mail: mayliu@inet.polyu.edu.hk)*

(Received 1 October 2004; accepted in revised form 16 December 2005)

Abstract. The development of efficient time optimal control strategies for coupled spin systems plays a fundamental role in nuclear magnetic resonance (NMR) spectroscopy. In particular, one of the major challenges lies in steering a given spin system to a maximum of its so-called transfer function. In this paper we study in detail these questions for a system of two weakly coupled spin- $\frac{1}{2}$ particles. First, we determine the set of maxima of the transfer function on the special unitary group $SU(4)$. It is shown that the set of maxima decomposes into two connected components and an explicit description of both components is derived. Related characterizations for the restricted optimization task on the special orthogonal group $SO(4)$ are obtained as well. In the second part, some general results on time optimal control on compact Lie groups are re-inspected. As an application of these results it is shown that each maximum of the transfer function can be reached in the same optimal time. Moreover, a global optimization algorithm is presented to explicitly construct time optimal controls for bilinear systems evolving on compact Lie groups. The algorithm is based on Lie-theoretic time optimal control results, established in [15], as well as on a recently proposed hybrid optimization method. Numerical simulations show that the algorithm performs well in the case a two spin- $\frac{1}{2}$ system.

Key words: Bilinear systems, Global optimization, Lie groups, NMR spectroscopy, Quantum systems, Spin dynamics, Time optimal control

1. Introduction

In quantum mechanics, a state vector of a closed quantum system is represented by an element in a Hilbert space \mathcal{H} . For a quantum mechanical N -body problem, the total Hilbert space is formed by taking the tensor product over all the single particle spaces, which gives

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \cdots \otimes \mathcal{H}_N.$$

The Hilbert space associated with a single spin- $\frac{1}{2}$ system in NMR spectroscopy is the complex vector space \mathbb{C}^2 . Correspondingly, the Hilbert space of N coupled spin- $\frac{1}{2}$ particles is the tensor product

$$\mathbb{C}^2 \otimes \dots \otimes \mathbb{C}^2 \cong \mathbb{C}^{2^N}.$$

In coherent ensemble spectroscopy such as NMR, the state of a quantum mechanical ensemble is described by its density operator, i.e. by a positive semidefinite operator C on \mathcal{H} normalized to $\text{tr}(C) = 1$, where $\text{tr}(\cdot)$ denotes the trace. The projection of the density operator C onto the observable A , the so-called transfer between C and A , is mathematically expressed by the generalized expectation value $\text{tr}(C^\dagger A)$, where $(\cdot)^\dagger$ denotes the conjugate transposed. However, note that in this setting so-called non-Hermitian detection operators occur by restricting C and A to signal-relevant components or measuring simultaneously non-commuting observables as in quadrature detection, cf. [8].

From the experimenter's point of view, the transfer between C and A is an essential quality for the performance of a spectroscopic measurement. Now, by controlling the time evolution of a quantum mechanical system/ensemble, i.e. by controlling its Schrödinger equation the experimenter can influence the transfers between C and A .

Hence the following two basic problems play a crucial role in this context:

- (I) Characterize and compute all unitary operators that maximize the transfer between a given pair of (detection) operators A and C , representing signal-relevant components of an observable and a density operator, respectively.
- (II) Develop explicit time optimal control strategies for the Schrödinger equation of a spin system which steers the identity operator to a final one (e.g. one maximizing the system's transfer).

While the first problem can be expressed as a global optimization task on the special unitary group $SU(2^N)$, the second one defines a time optimal control problem on $SU(2^N)$. Its solutions, so-called time optimal pulse sequences, are not only fundamental to NMR experiments. They are also an essential ingredient of NMR based quantum computing, cf. [10]. Moreover, somehow related topics arise in engineering areas such as inverse kinematics and robotics, cf. [6]. These are all hard mathematical problems for which a simple or straightforward solution cannot be expected. Thus there is value in analyzing special low dimensional cases to gain further insight into the general situation.

Here, we focus on a prototype of a system, consisting of two weakly coupled spin- $\frac{1}{2}$ particles, the so-called I_1S -spin system, and its transfer from

S^- to I^- . For precise definitions and more details on quantum mechanical terms in NMR spectroscopy see [7, 8]. In this case, the operators C and A have the explicit Kronecker product descriptions

$$C = \sigma_- \otimes I_2, \quad A = I_2 \otimes \sigma_-, \quad \text{with} \quad \sigma_- := \frac{1}{2}(\sigma_x - i\sigma_y) = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad (1)$$

where

$$\sigma_x := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y := \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_z := \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad (2)$$

denote the Pauli matrices and I_n the $(n \times n)$ -identity matrix. Moreover, the time evolution of the system subject to external controls is governed by the time-dependent Schrödinger equation

$$\dot{X}(t) = -2\pi i \left(H_d + \sum_{j=1}^4 u_j(t) H_j \right) X(t), \quad X(0) = I_4, \quad u_j(t) \in \mathbb{R}, \quad (3)$$

which leads to a control system on the Lie group $SU(4)$. Hence the time evolution of the density/detection operator C is given by

$$C(t) = X(t)C(0)X(t)^\dagger,$$

where X is the solution of Eq. (3). The terms H_d and H_j are called according to their physical interpretation drift Hamiltonian and control Hamiltonians, respectively. These are

$$H_d := J \cdot I_z S_z, \quad H_1 := I_x, \quad H_2 := I_y, \quad H_3 := S_x, \quad H_4 := S_y, \quad (4)$$

respectively, where

$$\begin{aligned} I_x &:= \frac{1}{2}\sigma_x \otimes I_2, & I_y &:= \frac{1}{2}\sigma_y \otimes I_2, & I_z &:= \frac{1}{2}\sigma_z \otimes I_2, \\ S_x &:= \frac{1}{2}I_2 \otimes \sigma_x, & S_y &:= \frac{1}{2}I_2 \otimes \sigma_y, & S_z &:= \frac{1}{2}I_2 \otimes \sigma_z. \end{aligned} \quad (5)$$

The scalar J is called the weak coupling constant. Without loss of generality, we will assume that J is normalized to $J = 1$. The time dependent real-valued functions u_j can be chosen arbitrarily by the experimenter, and hence are considered as controls. In [19], controllability of N -spin systems on $SU(2^N)$ for generic weak couplings has been shown. In our case, however, we will provide a more general Lie-theoretic argument to guarantee the controllability of the Schrödinger equation (3) on $SU(4)$.

Now, after we have fixed our basic notation, we can state the main open research tasks for the spin system at hand as follows:

- (I) Characterize all unitary operators X_F that maximize the transfer function

$$f: SU(4) \rightarrow \mathbb{R}, \quad f(X) = \operatorname{Re} \operatorname{tr}(C^\dagger X A X^\dagger).$$

- (II) (a) Given initial state $X(0) = I_4$ and final state $X_F \in SU(4)$, find controls u_1, \dots, u_4 , a so-called pulse sequence, such that the corresponding solution X of the Schrödinger equation (3) satisfies

$$X(0) = I_4 \quad \text{and} \quad X(T) = X_F \quad \text{for optimal time } T \geq 0.$$

- (b) If the final state X_F is a maximizing operator in the sense of (I), compute the optimal time T and investigate whether or not T depends on the choice of X_F .

A detailed discussion of Problem (I) for arbitrary A and C can be found e.g. in [4] where the transfer function is called the C -numerical radius function of A . While the Hermitian case is well understood [5, 13], the non-Hermitian one is much harder to analyse. For matrices A and C of interest in NMR spectroscopy this was first approached in [7], where a gradient algorithm analogously to [13] was developed. Convergence results and specific step size selection rules were obtained in [11].

Problem (II) is a non-standard time optimal control problem in the sense that the control set is unbounded, cf. Section 3. For basics on geometric control theory and time optimal control we refer to [1, 14]. Particular results on time optimal control of spin systems are developed in [15–17]. Therein the optimal time to achieve a particular maximum of the transfer function has been computed to be equal to $\frac{3}{2}J^{-1}$ for $N=2$, where J denotes the weak coupling constant appearing in the drift Hamiltonian H_d in Eq. (4). However, this does not completely answer Problem (IIb). For $N=3$ only partial results are available under additional assumptions on the drift term and for $N \geq 4$ no general formula for the optimal time seems to be known. Moreover, efficient algorithms for the explicit computation of time optimal pulse sequences are missing in any of these cases.

In Section 2, we determine the maxima of the transfer function. In particular, we show that the set of maxima consists of two connected components and obtain an explicit description of both. Similar results are derived for the maxima of the transfer function restricted to the subgroup $SO(4)$ of real orthogonal transformations. In Section 3, supplementing [15], it is shown that any maximum of the transfer function can be reached in the same optimal time $\frac{3}{2}J^{-1}$, although one cannot pass from one connected component of the set of maxima to the other one in zero time. In Section 4, we propose an algorithm from global optimization to determine time optimal controls. While the numerical method can be implemented in

principle for an arbitrary N -spin system, it works effectively only if prior information on the structure of the time optimal controls is known. Thus we apply our results in Section 5 to the simplest case $N = 2$, where such structural information is available from [15]. The computational approach of Section 5 is however not restricted to NMR spectroscopy and may also prove useful for other constructive controllability tasks.

2. Characterization of the Maxima

In this section, we are concerned with a solution of Problem (I), i.e. with the maximization of the transfer function

$$f : SU(4) \rightarrow \mathbb{R}, \quad f(X) = \operatorname{Re} \operatorname{tr}(C^\dagger X A X^\dagger). \quad (6)$$

Our main result is an explicit characterization of its global maxima. First, we fix some notation. Let $SU(4) := \{X \in \mathbb{C}^{4 \times 4} \mid X X^\dagger = I_4, \det X = 1\}$ be the special unitary group of (4×4) -matrices. Moreover, let $\operatorname{Stab}(C) := \{X \in SU(4) \mid X C X^\dagger = C\}$ and $\operatorname{Stab}(A) := \{X \in SU(4) \mid X A X^\dagger = A\}$ be the stabilizer subgroups of the operators

$$C = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad (7)$$

respectively, cf. Eq. (1). Finally, let \mathcal{X}_{\max} denote the set of global maxima of the transfer function (6) and consider an arbitrary maximizing transformation $\widehat{X} \in SU(4)$. Then it is easy to see that the inclusion

$$\operatorname{Stab}(C^\dagger) \widehat{X} \operatorname{Stab}(A) \subset \mathcal{X}_{\max} \quad (8)$$

is always satisfied. Now, Theorem 1 shows that in our case even equality holds.

THEOREM 1. (a) *The maximal value of the transfer function (6) is*

$$\max_{X \in SU(4)} \operatorname{Re} \operatorname{tr}(C^\dagger X A X^\dagger) = 2.$$

(b) *Let*

$$P := \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (9)$$

and

$$X_* := e^{-\frac{i\pi}{4}} P. \tag{10}$$

Then the stabilizer subgroups of C and A in $SU(4)$ are given by

$$\text{Stab}(C) = \{ \text{diag}(U, U) \mid U \in U(2), \det U = \pm 1 \}$$

and

$$\begin{aligned} \text{Stab}(A) &= X_*^\dagger \text{Stab}(C) X_* \\ &= X_*^\dagger \{ \text{diag}(U, U) \mid U \in U(2), \det U = \pm 1 \} X_*, \end{aligned}$$

respectively.

(c) The set of maxima of the transfer function (6) is given by

$$\begin{aligned} \mathcal{X}_{\max} &= \{ \text{diag}(U, U) \mid U \in U(2), \det U = \pm 1 \} X_* \\ &= \text{Stab}(C^\dagger) X_* \text{Stab}(A). \end{aligned} \tag{11}$$

In particular, $\mathcal{X}_{\max} = \mathcal{X}_{\max}^+ \cup \mathcal{X}_{\max}^-$ has exactly two connected components, namely

$$\mathcal{X}_{\max}^+ := \{ \text{diag}(U, U) \mid U \in U(2), \det U = 1 \} X_* \tag{12}$$

and

$$\mathcal{X}_{\max}^- := \{ \text{diag}(U, U) \mid U \in U(2), \det U = -1 \} X_*. \tag{13}$$

Proof. For a proof of (a) we refer to [11]. For (b), a straightforward calculation using the fact that we are dealing with unitary transformations, yields

$$\text{Stab}(C) = \text{Stab}(C^\dagger) = \{ \text{diag}(U, U) \mid U \in U(2), \det U = \pm 1 \}. \tag{14}$$

Moreover, $C = X_* A X_*^\dagger$. Thus we conclude

$$\begin{aligned} \text{Stab}(A) &= X_*^\dagger \text{Stab}(C) X_* \\ &= X_*^\dagger \{ \text{diag}(U, U) \mid U \in U(2), \det U = \pm 1 \} X_*. \end{aligned} \tag{15}$$

To prove (c), let $\widehat{X} \in SU(4)$ be any maximizing transformation and let

$$Z := \widehat{X} X_*^\dagger = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix}.$$

Then

$$2 = \operatorname{Re tr}(C^\dagger Z X_* A X_*^\dagger Z^\dagger) = \operatorname{Re tr}(C^\dagger Z C Z^\dagger) = \operatorname{Re tr}(Z_{22} Z_{11}^\dagger). \tag{16}$$

The unitarity of Z implies $\operatorname{Re tr}(Z_{ii} Z_{ii}^\dagger) = \|Z_{ii}\|^2 \leq 2$ for $i = 1, 2$. By the Cauchy-Schwarz inequality we obtain

$$2 = \operatorname{Re tr}(Z_{22} Z_{11}^\dagger) \leq \|Z_{22}\| \cdot \|Z_{11}\| \leq 2.$$

Therefore, there exists a real number $\lambda \neq 0$ such that $Z_{22} = \lambda Z_{11}$. Substituting $Z_{22} = \lambda Z_{11}$ in Eq. (16) yields $\lambda = 1$. In particular, this shows $\|Z_{22}\|^2 = \|Z_{11}\|^2 = 2$ and $Z_{12} = Z_{21} = 0$. This proves the first equality in Eq.(11). The second equality follows immediately from Eqs. (14) and (15). Moreover, Eqs. (12) and (13) obviously yield a disjoint decomposition of \mathcal{X}_{\max} into two connected components. \square

The above result shows that \mathcal{X}_{\max} does not contain any real orthogonal transformation, i.e. any transformation of the subgroup $SO(4) := \{Y \in \mathbb{R}^{4 \times 4} \mid Y Y^\dagger = I_4, \det Y = 1\}$. Hence a natural question to ask is, what happens if one restricts the transfer function (6) to $SO(4)$ transformations. The following proposition gives an answer.

PROPOSITION 2. (a) *Let A and C be defined as in Eq. (7). Then the maximal value of the transfer function (6) restricted to $SO(4)$ is*

$$\max_{Y \in SO(4)} \operatorname{Re tr}(C^\dagger Y A Y^\dagger) = 1.$$

(b) *Let P be defined as in Eq. (9). Then the stabilizer subgroups of C and A in $SO(4)$ are given by*

$$\operatorname{Stab}_{SO(4)}(C) = \{ \operatorname{diag}(V, V) \mid V \in O(2) \}, \tag{17}$$

and

$$\operatorname{Stab}_{SO(4)}(A) = P^\dagger \operatorname{Stab}_{SO(4)}(C) P, \tag{18}$$

respectively.

(c) *Let*

$$Y_* := \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

Then the set of maxima of the transfer function (6) restricted to $SO(4)$ is given by

$$\mathcal{O}_{\max} = \text{Stab}(C)_{SO(4)} Y_* \text{Stab}(A)_{SO(4)}. \tag{19}$$

In particular, $\mathcal{O}_{\max} = \mathcal{O}_{\max}^+ \cup \mathcal{O}_{\max}^-$ has exactly two connected components, namely

$$\mathcal{O}_{\max}^+ := \{ \text{diag}(V, V) Y_* P^\dagger \text{diag}(\widehat{V}, \widehat{V}) P \mid V, \widehat{V} \in O(2), \det(V\widehat{V}) = 1 \}$$

and

$$\mathcal{O}_{\max}^- := \{ \text{diag}(V, V) Y_* P^\dagger \text{diag}(\widehat{V}, \widehat{V}) P \mid V, \widehat{V} \in O(2), \det(V\widehat{V}) = -1 \}.$$

Proof. (a) Let Σ be defined as in Eq. (A.2) and let $Y \in SO(4)$. Then $Y \mapsto \Sigma^\dagger Y \Sigma$ is a diffeomorphism from $SO(4)$ onto $SU(2) \otimes SU(2)$, cf. Appendix A.1. Now substituting $Y = \Sigma(U_1 \otimes U_2) \Sigma^\dagger$ yields

$$\begin{aligned} \text{tr}(C^\dagger Y A Y^\dagger) &= \text{tr}(C^\dagger \Sigma(U_1 \otimes U_2) \Sigma^\dagger A \Sigma(U_1 \otimes U_2) \Sigma^\dagger) \\ &= \text{tr}(\Sigma^\dagger C^\dagger \Sigma(U_1 \otimes U_2) \Sigma^\dagger A \Sigma(U_1 \otimes U_2)^\dagger). \end{aligned}$$

Using the identities

$$\begin{aligned} \Sigma^\dagger C^\dagger \Sigma &= \frac{1}{2} \left(\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \otimes \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \\ \Sigma^\dagger A \Sigma &= -\frac{i}{2} \left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \otimes \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \otimes \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \right), \end{aligned}$$

one can rewrite Eq. (2) as follows.

$$\begin{aligned} \text{tr}(C^\dagger Y A Y^\dagger) &= \frac{-i}{4} \text{tr} \left\{ \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \otimes \left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} U_2 \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} U_2^\dagger \right) + \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \otimes \left(U_2 \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} U_2^\dagger \right) \right. \\ &\quad + \left(\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} U_1 \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} U_1^\dagger \right) \otimes \left(U_2 \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} U_2^\dagger \right) \\ &\quad \left. + \left(\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} U_1 \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} U_1^\dagger \right) \otimes \left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} U_2 \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} U_2^\dagger \right) \right\} \\ &= \frac{-i}{4} \text{tr} \left(\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} U_1 \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} U_1^\dagger \right) \cdot \text{tr} \left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} U_2 \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} U_2^\dagger \right), \end{aligned}$$

where the last equality holds by the identity $\text{tr}(B \otimes D) = \text{tr} B \cdot \text{tr} D$. Now Eq. (20) implies the estimate $|\text{tr}(C^\dagger Y A Y^\dagger)| \leq 1$, since all matrices in Eq. (20) are unitary. Equality is achieved by

$$Y_* := \Sigma(U_* \otimes U_*)\Sigma^\dagger = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad \text{where } U_* := \frac{1}{\sqrt{2}} \begin{bmatrix} e^{i\pi/4} & e^{-i\pi/4} \\ -e^{-i\pi/4} & e^{i\pi/4} \end{bmatrix}. \tag{20}$$

(b) Eqs. (17) and (18) are an immediate consequence of Eqs. (14) and (15), respectively, and the identity

$$\text{Stab}_{SO(4)}(C) = \text{Stab}(C) \cap SO(4).$$

(c) Let $\widehat{Y} = \Sigma(U_1 \otimes U_2)\Sigma^\dagger$ be any maximizing $SO(4)$ transformation. Then Eq. (20) in part (a) shows that U_1 and U_2^\dagger have to be diagonalizing transformations of

$$\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

satisfying either

$$U_1 \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} U_1^\dagger = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix} \quad \text{and} \quad U_2^\dagger \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} U_2 = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

or

$$U_1 \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} U_1^\dagger = \begin{bmatrix} -i & 0 \\ 0 & i \end{bmatrix} \quad \text{and} \quad U_2^\dagger \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} U_2 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

This yields

$$U_1 = \begin{bmatrix} e^{i\phi} & 0 \\ 0 & e^{-i\phi} \end{bmatrix} U_* \quad \text{and} \quad U_2 = U_* \begin{bmatrix} e^{i\psi} & 0 \\ 0 & e^{-i\psi} \end{bmatrix}$$

or

$$U_1 = \begin{bmatrix} 0 & -e^{i\phi} \\ e^{-i\phi} & 0 \end{bmatrix} U_* \quad \text{and} \quad U_2 = U_* \begin{bmatrix} 0 & -e^{-i\psi} \\ e^{i\psi} & 0 \end{bmatrix},$$

and thus

$$\widehat{Y} = \Sigma \left(\begin{bmatrix} e^{i\phi} & 0 \\ 0 & e^{-i\phi} \end{bmatrix} \otimes \mathbf{I}_2 \right) (U_* \otimes U_*) \left(\mathbf{I}_2 \otimes \begin{bmatrix} e^{i\psi} & 0 \\ 0 & e^{-i\psi} \end{bmatrix} \right) \Sigma^\dagger$$

or

$$\widehat{Y} = \Sigma \left(\begin{bmatrix} 0 & -e^{i\phi} \\ e^{-i\phi} & 0 \end{bmatrix} \otimes I_2 \right) (U_* \otimes U_*) \left(I_2 \otimes \begin{bmatrix} 0 & -e^{-i\psi} \\ e^{i\psi} & 0 \end{bmatrix} \right) \Sigma.$$

Moreover, a straightforward calculation shows

$$U_* \otimes U_* = \left(\begin{bmatrix} e^{i\phi} & 0 \\ 0 & e^{-i\phi} \end{bmatrix} \otimes \begin{bmatrix} \cos \psi & i \sin \psi \\ i \sin \psi & \cos \psi \end{bmatrix} \right) (U_* \otimes U_*) \left(\begin{bmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{bmatrix} \otimes \begin{bmatrix} e^{i\psi} & 0 \\ 0 & e^{-i\psi} \end{bmatrix} \right)$$

and

$$U_* \otimes U_* = \left(\begin{bmatrix} 0 & -e^{i\phi} \\ e^{-i\phi} & 0 \end{bmatrix} \otimes \begin{bmatrix} -i \cos \psi & -\sin \psi \\ \sin \psi & i \cos \psi \end{bmatrix} \right) (U_* \otimes U_*) \left(\begin{bmatrix} i \sin \phi & i \cos \phi \\ i \cos \phi & -i \sin \phi \end{bmatrix} \otimes \begin{bmatrix} 0 & -e^{-i\psi} \\ e^{i\psi} & 0 \end{bmatrix} \right).$$

This implies

$$\begin{aligned} & \left(\begin{bmatrix} e^{i\phi} & 0 \\ 0 & e^{-i\phi} \end{bmatrix} \otimes I_2 \right) (U_* \otimes U_*) \left(I_2 \otimes \begin{bmatrix} e^{i\psi} & 0 \\ 0 & e^{-i\psi} \end{bmatrix} \right) = \\ & \left(I_2 \otimes \begin{bmatrix} \cos \psi & -i \sin \psi \\ -i \sin \psi & \cos \psi \end{bmatrix} \right) (U_* \otimes U_*) \left(\begin{bmatrix} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{bmatrix} \otimes I_2 \right) \end{aligned}$$

and

$$\begin{aligned} & \left(\begin{bmatrix} 0 & -e^{i\phi} \\ e^{-i\phi} & 0 \end{bmatrix} \otimes I_2 \right) (U_* \otimes U_*) \left(I_2 \otimes \begin{bmatrix} 0 & -e^{-i\psi} \\ e^{i\psi} & 0 \end{bmatrix} \right) = \\ & \left(I_2 \otimes \begin{bmatrix} i \cos \psi & \sin \psi \\ -\sin \psi & -i \cos \psi \end{bmatrix} \right) (U_* \otimes U_*) \left(\begin{bmatrix} -i \sin \phi & -i \cos \phi \\ -i \cos \phi & i \sin \phi \end{bmatrix} \otimes I_2 \right). \end{aligned}$$

Hence we obtain

$$\begin{aligned} \mathcal{O}_{\max} = & \left\{ \Sigma \left(I_2 \otimes \begin{bmatrix} \cos \psi & -i \sin \psi \\ -i \sin \psi & \cos \psi \end{bmatrix} \right) (U_* \otimes U_*) \left(\begin{bmatrix} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{bmatrix} \otimes I_2 \right) \Sigma^\dagger \mid \phi, \psi \in [-\pi, \pi] \right\} \\ \cup & \left\{ \Sigma \left(I_2 \otimes \begin{bmatrix} i \cos \psi & \sin \psi \\ -\sin \psi & -i \cos \psi \end{bmatrix} \right) (U_* \otimes U_*) \left(\begin{bmatrix} -i \sin \phi & -i \cos \phi \\ -i \cos \phi & i \sin \phi \end{bmatrix} \otimes I_2 \right) \Sigma^\dagger \mid \phi, \psi \in [-\pi, \pi] \right\} \end{aligned}$$

Using the equalities

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} Y_* P^\dagger = \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \end{bmatrix} P = Y_*$$

and

$$\text{diag} \left(\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \right) Y_* P^\dagger \text{diag} \left(\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \right) P = Y_*$$

we conclude

$$\begin{aligned} \mathcal{O}_{\max} &= \left\{ \text{diag} (R(\psi), R(\psi)) Y_* P^\dagger \text{diag} (R(\phi), R(\phi)) P \mid \phi, \psi \in [-\pi, \pi] \right\} \\ &\cup \left\{ \text{diag} (\widehat{R}(\psi), \widehat{R}(\psi)) Y_* P^\dagger \text{diag} (R(\phi), R(\phi)) P \mid \phi, \psi \in [-\pi, \pi] \right\} \\ &= \mathcal{O}_{\max}^+ \cup \mathcal{O}_{\max}^- = \text{Stab}_{SO(4)}(C^\dagger) X_* \text{Stab}_{SO(4)}(A). \end{aligned}$$

Here

$$R(\tau) := \begin{bmatrix} \cos \tau & -\sin \tau \\ \sin \tau & \cos \tau \end{bmatrix} \quad \text{and} \quad \widehat{R}(\tau) := \begin{bmatrix} \cos \tau & \sin \tau \\ \sin \tau & -\cos \tau \end{bmatrix}.$$

Moreover,

$$\text{diag} \left(\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \right) Y_* \notin \mathcal{O}_{\max}^+$$

implies $\mathcal{O}_{\max}^+ \cap \mathcal{O}_{\max}^- = \emptyset$ and thus \mathcal{O}_{\max} decomposes in two connected components, namely \mathcal{O}_{\max}^+ and \mathcal{O}_{\max}^- . □

Finally, we consider the restriction of the transfer function to $SU(2) \otimes SU(2)$, the so-called fast subgroup of Eq. (3), cf. Section 3. This restriction is of particular interest from an experimenter’s point of view, because its potential maxima correspond to the maximum transfer which can be achieved in time zero.

PROPOSITION 3. *Let A and C be defined as in Eq. (7). Then*

$$\text{Re tr}(C^\dagger X A X^\dagger) = 0 \quad \text{for all } X \in SU(2) \otimes SU(2). \tag{21}$$

Proof. The proof is a straightforward calculation similar to Proposition 2. □

3. Time Optimal Control

The subject of this section is the solution of the control problem (IIa) and (IIb). First, we state two general results on time optimal control which will be our main tools to tackle (IIa) and (IIb). Both results take essentially advantage of the fact that the underlying control system (3) evolves on a

compact Lie group. Therefore, it is convenient and necessary to summarize some facts on Lie theory and geometric control for the reader who is not familiar with these topics. For a detailed study we refer to [12, 18] and [1, 14, 21], respectively. We start with the relevant Lie theoretic terms, followed by some basic definitions on geometric control. For an example illustrating the concepts below we refer the reader to the second part of this section.

Let G be a connected and compact (matrix) Lie group with real semi-simple Lie algebra \mathfrak{g} . Thus \mathfrak{g} is endowed with the *Lie bracket* operation $[X, Y] := XY - YX$. A typical example is the Lie group $SU(n)$ of all unitary $(n \times n)$ -matrices with determinant one and its Lie algebra $\mathfrak{su}(n)$, i.e. the set of all skew-Hermitian $(n \times n)$ -matrices with trace zero. To every Lie subalgebra $\mathfrak{h} \subset \mathfrak{g}$ there corresponds a unique connected Lie subgroup of G denoted by $\exp \mathfrak{h}$ that is generated by the one parameter subgroups

$$\{ \exp(t\Omega) \mid t \in \mathbb{R}, \Omega \in \mathfrak{h} \}$$

and has \mathfrak{h} as its Lie algebra. Here $\exp: \mathfrak{g} \rightarrow G$ denotes the (matrix) exponential map, i.e.

$$\exp(t\Omega) = e^{t\Omega} := \sum_{k=0}^{\infty} \frac{t^k}{k!} \Omega^k.$$

A decomposition of \mathfrak{g} into two vector spaces $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ is called *Cartan-like*, if the Lie bracket relations

$$[\mathfrak{k}, \mathfrak{k}] \subset \mathfrak{k}, \quad [\mathfrak{k}, \mathfrak{p}] \subset \mathfrak{p}, \quad \text{and} \quad [\mathfrak{p}, \mathfrak{p}] \subset \mathfrak{k} \tag{22}$$

hold. Obviously, \mathfrak{k} is a Lie subalgebra by the first relation of Eq. (22). Moreover, the decomposition (22) is orthogonal with respect to the Killing form $\kappa: \mathfrak{g} \times \mathfrak{g} \rightarrow \mathbb{R}, (X, Y) \mapsto \text{tr}(\text{ad}_X \circ \text{ad}_Y)$, where $\text{ad}_X: \mathfrak{g} \rightarrow \mathfrak{g}, Y \mapsto [X, Y]$. Cartan-like decompositions of \mathfrak{g} always exist and have been classified in the literature, cf. [12]. Now, let $\mathfrak{a} \subset \mathfrak{p}$ be a maximal Abelian subalgebra. Then every $X \in G$ decomposes into

$$X = K_1 A K_2 \tag{23}$$

with $K_1, K_2 \in \exp \mathfrak{k}$ and $A \in \exp \mathfrak{a}$, cf. [12], Chapter V, Thm 6.7. The decompositions (23) are called *KAK-decompositions* of G . Moreover, let

$$N_{\exp \mathfrak{k}}(\mathfrak{a}) := \{ K \in \exp \mathfrak{k} \mid KaK^{-1} \in \mathfrak{a} \text{ for all } a \in \mathfrak{a} \}$$

denote the normalizer of \mathfrak{a} in $\exp \mathfrak{k}$ and

$$Z_{\exp \mathfrak{k}}(\mathfrak{a}) := \{ K \in \exp \mathfrak{k} \mid KaK^{-1} = a \text{ for all } a \in \mathfrak{a} \}$$

the centralizer. Then the *Weyl group* of the pair $(G, \exp \mathfrak{a})$ is defined by the quotient group

$$\mathcal{W}(G, \exp \mathfrak{a}) := N_{\exp \mathfrak{k}}(\mathfrak{a}) / Z_{\exp \mathfrak{k}}(\mathfrak{a}). \tag{24}$$

It is a finite subgroup of $\exp \mathfrak{k} \subset G$. Note that the preceding concepts can be developed also for noncompact Lie groups. Actually, the classical theory in the noncompact case is easier since more restrictively defined Cartan decompositions, cf. [18], VI. 6.26 are used. This leads to a *KAK* factorization that is unique up to conjugation by a member of the Weyl group. In the compact Lie group case, however, the uniqueness of general *KAK* decompositions (23) is a delicate issue. We only state a uniqueness result for the relevant case where G is simply connected. In this case, A in Eq. (23) is unique up to a group action of the semidirect product of the Weyl group and a finite group which is isomorphic to the subgroup $\exp \mathfrak{a} \cap \exp \mathfrak{k}$. More precisely, we have the following proposition.

PROPOSITION 4. *Let G be a simply connected compact Lie group with semi-simple Lie algebra \mathfrak{g} and let \mathfrak{k} , \mathfrak{p} , and \mathfrak{a} be defined as above. If $X = K_1AK_2 = K'_1A'K'_2$ are two factorizations of the above type (23), then*

$$A' = WADW^\dagger, \tag{25}$$

with $D \in \exp \mathfrak{a} \cap \exp \mathfrak{k}$ and $W \in \mathcal{W}(G, \exp \mathfrak{a})$.

Proof. For a proof we refer to [12], Chapter VII, Theorems 8.3, 8.5, and 8.6. □

Finally, we need two more Lie algebraic terms for the notation of the subsequent results, namely the *Weyl orbit* of an element $\Delta \in \mathfrak{g}$

$$\mathcal{W} \cdot \Delta = \{W\Delta W^{-1} \mid W \in \mathcal{W}(G, \exp \mathfrak{a})\}$$

and the convex hull of $\mathcal{W} \cdot \Delta$, denoted by $\text{cov}(\Delta)$.

Next, we recall some control-theoretic ideas on Lie groups. Let G and \mathfrak{g} be as above and let

$$\dot{X}(t) = \left(\Delta + \sum_{j=1}^N u_j(t) \Omega_j \right) X(t), \quad X(0) = X_0, \quad u_j(t) \in \mathbb{R} \tag{26}$$

be a *right invariant control system* on G with $\Delta, \Omega_1, \dots, \Omega_N \in \mathfrak{g}$. The set $R_T(X_0)$ of all points which are reachable from X_0 in time $T \geq 0$ is defined as the set of all terminal points $X(T)$ of solutions of (26) originating at $X(0) = X_0$. The *reachable set* $R(X_0)$ itself is defined as the union of these sets $R_T(X_0)$ for all $T \geq 0$, i.e.

$$R(X_0) := \bigcup_{T \geq 0} R_T(X_0).$$

Note, that the right invariance of (26) implies

$$R(X_0) = R(I)X_0 \quad \text{for all } X_0 \in G.$$

Finally, a system on G is said to be *controllable*, if $R(X_0) = G$ for all $X_0 \in G$. Note, that for right invariant systems this is equivalent to $R(I) = G$. Next, we want to introduce the optimal time $T(X_F)$ to reach a final point X_F from an initial point X_0 . For simplicity we restrict ourselves to the case $X_0 = I$. Since the controls of Eq. (26) can be chosen from an unbounded set, this is a bit more subtle than in the classical setting where the controls usually are assumed to be bounded. Let X_F be any final point in G and set

$$T(X_F) := \inf \{ T \geq 0 \mid X_F \in \overline{R_T(I)} \}. \tag{27}$$

Note, that the right hand side of Eq. (27) differs from the standard definition of most textbooks such as [14] in two ways: The infimum and the closure of the reachable sets are used instead of the minimum and the reachable sets itself. Both modifications take into account that the controls of Eq. (26) are unbounded. Nevertheless, in abuse of language, we keep the term optimal time and say that $X \in \exp \mathfrak{k}$ can be reached (approximately) in time $T = 0$. Moreover, we refer to \mathfrak{k} and $\exp \mathfrak{k}$ as the *fast subalgebra* and *fast subgroup* of system (33), respectively.

Now, we are prepared to state two results on right invariant control systems on compact Lie groups which provide an approach to solve Problem (IIa) and (IIb). The first one generalizes Lemma 4 in [3] and guarantees the controllability of Eq. (3). The second one, the Time Optimal Torus Theorem, characterizes the solutions of a time optimal control problem on compact Lie groups as the solutions of a convex optimization task, cf. [15]. This offers an alternative approach to time-optimal control on Lie groups that yields more specific results than an immediate application of the Pontryagin Maximum Principle. Moreover, a numerical algorithm for the explicit computation of time optimal solutions/pulse sequences can be derived from these results, cf. Section 4 and 5.

THEOREM 5. *Let G be a compact, connected, and simple Lie group with Lie algebra \mathfrak{g} and let $\mathfrak{k} := \langle \Omega_1, \dots, \Omega_N \rangle_L \subsetneq \mathfrak{g}$ be a proper subalgebra such that $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$, $\mathfrak{p} := \mathfrak{k}^\perp$, is a Cartan-like decomposition. Then system (26) is controllable if and only if $\Delta \notin \mathfrak{k}$.*

Proof. “ \implies ”: If $\Delta \in \mathfrak{k}$, then the reachable set of the identity is equal to $\exp \mathfrak{k}$, cf. [14]. Note, that $\exp \mathfrak{k}$ is closed, and hence compact, cf. Remark 1

(a). However, as $\mathfrak{k} \neq \mathfrak{g}$ by assumption, $\exp \mathfrak{k}$ is a proper subgroup of G , and thus Eq. (26) is not controllable. Therefore $\Delta \notin \mathfrak{k}$ is necessary for controllability.

“ \Leftarrow ”: Let \mathfrak{g} be a simple Lie algebra and $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ be a Cartan-like decomposition. We show that any subalgebra \mathfrak{k}' such that \mathfrak{k} is a proper subset of \mathfrak{k}' is equal to \mathfrak{g} . Then the controllability of Eq. (26) is implied by [14], Chapter 6, Theorem 3.

Let \mathfrak{k}' be any subalgebra with $\mathfrak{k} \subsetneq \mathfrak{k}'$ and $\mathfrak{r} := \mathfrak{k}' \cap \mathfrak{p}$. First, we prove that

$$i := [\mathfrak{r}, \mathfrak{r}] \oplus \mathfrak{r} \neq 0 \tag{28}$$

is an ideal of \mathfrak{g} . Since $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$ is a Cartan-like decomposition, cf. (22), we have $[\mathfrak{r}, \mathfrak{k}] \subset [\mathfrak{p}, \mathfrak{k}] \subset \mathfrak{p}$. Moreover, since \mathfrak{k}' is a subalgebra, we have $[\mathfrak{r}, \mathfrak{k}] \subset [\mathfrak{k}', \mathfrak{k}] \subset \mathfrak{k}'$ and thus

$$[\mathfrak{r}, \mathfrak{k}] \subset \mathfrak{r}. \tag{29}$$

As $[\mathfrak{r}, \mathfrak{r}] \subset [\mathfrak{p}, \mathfrak{p}] \subset \mathfrak{k}$, then

$$[[\mathfrak{r}, \mathfrak{r}], \mathfrak{r}] \subset [\mathfrak{k}, \mathfrak{r}] \subset \mathfrak{r}. \tag{30}$$

Therefore the Jacobi identity yields

$$[[\mathfrak{r}, \mathfrak{r}], \mathfrak{k}] = [\mathfrak{r}, [\mathfrak{r}, \mathfrak{k}]] \subset [\mathfrak{r}, \mathfrak{r}]. \tag{31}$$

Let κ denote the Killing form on \mathfrak{g} and let $\mathfrak{p}' := \mathfrak{k}'^\perp$. By the invariance property of κ and the relation $[\mathfrak{r}, \mathfrak{k}'] \subset \mathfrak{k}'$ we obtain

$$\kappa([\Omega_{\mathfrak{r}}, \Omega_{\mathfrak{p}'}], \Omega') = -\kappa(\Omega_{\mathfrak{p}'}, [\Omega_{\mathfrak{r}}, \Omega']) = 0$$

for all $\Omega_{\mathfrak{r}} \in \mathfrak{r}$, $\Omega_{\mathfrak{p}'} \in \mathfrak{p}'$, and $\Omega' \in \mathfrak{k}'$. Therefore $[\mathfrak{r}, \mathfrak{p}'] \subset \mathfrak{p}'$. On the other hand, as $\mathfrak{p}' \subset \mathfrak{p}$ then $[\mathfrak{r}, \mathfrak{p}'] \subset [\mathfrak{p}, \mathfrak{p}] \subset \mathfrak{k}$ and hence

$$[\mathfrak{r}, \mathfrak{p}'] = 0. \tag{32}$$

Using Eqs. (29–32), we conclude

$$\begin{aligned} [i, \mathfrak{g}] &= [[\mathfrak{r}, \mathfrak{r}] \oplus \mathfrak{r}, \mathfrak{k} \oplus \mathfrak{r} \oplus \mathfrak{p}'] \\ &= [[\mathfrak{r}, \mathfrak{r}], \mathfrak{k}] + [[\mathfrak{r}, \mathfrak{r}], \mathfrak{r}] + [[\mathfrak{r}, \mathfrak{r}], \mathfrak{p}'] + [\mathfrak{r}, \mathfrak{k}] + [\mathfrak{r}, \mathfrak{r}] + [\mathfrak{r}, \mathfrak{p}'] \subset i. \end{aligned}$$

By assumption, however, \mathfrak{g} has no non-trivial ideals. Hence $i = \mathfrak{g}$, which implies $\mathfrak{r} = \mathfrak{g}$ and thus $\mathfrak{k}' = \mathfrak{g}$. □

THEOREM 6 (Time Optimal Torus Theorem, [15]). *Let*

$$\dot{X}(t) = \left(\Delta + \sum_{j=1}^N u_j(t)\Omega_j\right)X(t), \quad X(0) = I, \quad u_j(t) \in \mathbb{R} \tag{33}$$

be a right invariant control system on a compact, connected, and semisimple Lie group G with Lie algebra \mathfrak{g} . Let $\mathfrak{k} := \langle \Omega_1, \dots, \Omega_N \rangle_L \subset \mathfrak{g}$ be a subalgebra such that $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$, $\mathfrak{p} := \mathfrak{k}^\perp$, is a Cartan-like decomposition and let $\mathfrak{a} \subset \mathfrak{p}$ be a maximal Abelian subalgebra such that $\Delta \in \mathfrak{a}$. Then the following holds:

- (a) *The set $\{K_1 \exp(ta)K_2 \mid K_1, K_2 \in \exp \mathfrak{k}, t \geq 0, a \in \text{cov}(\Delta)\}$ is contained in the closure of the reachable set $R(I)$.*
- (b) *The optimal time $T(X_F)$ is given by*

$$T(X_F) = \min \{t \geq 0 \mid t \cdot \text{cov}(\Delta) \cap \Gamma_{X_F} \neq \emptyset\}, \tag{34}$$

where $\Gamma_{X_F} := \{a \in \mathfrak{a} \mid X_F = K_1 \exp(a)K_2, K_1, K_2 \in \exp \mathfrak{k}\}$.

Remark 1. (a) In [15], the Time Optimal Torus Theorem is stated under the additional assumption that $\exp \mathfrak{k}$ is closed. This assumption, however, is not necessary and follows from the Cartan-like decomposition and the semisimplicity of \mathfrak{g} , cf. [12], Chapter V, p. 230, Exercise (a) together with Chapter IV, Proposition 3.6.

- (b) Note, that we do not assume Eq. (33) to be controllable. In particular, the optimal time $T(X_F)$ given by Eq. (35) can be $+\infty$ which means that X_F does not belong to the closure of the reachable set $R(I)$. However, if $\Delta \neq 0$ and G is simple instead of semisimple then controllability is implied by Theorem 5.

For simply connected G , a simplified version of Theorem 6(b) can be obtained from Eq. (25), where the set Γ_{X_F} in Eq. (34) is replaced by the affine lattice $a_0 + \Gamma_0$ as defined below.

COROLLARY 7. *If G is compact, simple and simply connected and if $X_F = K_1 \exp(a_0)K_2$ with $K_1, K_2 \in \exp \mathfrak{k}$ and $a_0 \in \mathfrak{a}$, then the optimal time $T(X_F)$ is given by*

$$T(X_F) = \min \{t \geq 0 \mid t \cdot \text{cov}(\Delta) \cap (a_0 + \Gamma_0) \neq \emptyset\}, \tag{35}$$

where $\Gamma_0 := \{a \in \mathfrak{a} \mid \exp(a) \in \exp \mathfrak{k}\}$.

Proof. This is an immediate consequence of the invariance of the convex hull $\text{cov}(\Delta)$ under the Weyl group action, Theorem 6(b), and Eq. (25). \square

Now, we return to our particular control problem on $SU(4)$. First, we identify and compute the basic ingredients which are necessary to apply the preceding results. Let

$$\Delta := -2\pi i I_z S_z = \frac{\pi}{2} \begin{bmatrix} -i & 0 & 0 & 0 \\ 0 & i & 0 & 0 \\ 0 & 0 & i & 0 \\ 0 & 0 & 0 & -i \end{bmatrix},$$

$$\Omega_1 := -2\pi i I_x = \pi \begin{bmatrix} 0 & 0 & -i & 0 \\ 0 & 0 & 0 & -i \\ -i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \end{bmatrix}, \quad \Omega_2 := -2\pi i I_y = \pi \begin{bmatrix} 0 & -i & 0 & 0 \\ -i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i \\ 0 & 0 & -i & 0 \end{bmatrix},$$

$$\Omega_3 := -2\pi i S_x = \pi \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad \Omega_4 := -2\pi i S_y = \pi \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

cf. Eqs. (4) and (5). Here, we have normalized the coupling constant to $J = 1$. Furthermore, define

$$\mathfrak{k} := \mathfrak{su}(2) \widehat{\otimes} \mathfrak{su}(2) := \{I_2 \otimes \Omega' + \Omega'' \otimes I_2 \mid \Omega', \Omega'' \in \mathfrak{su}(2)\}, \tag{36}$$

$$\exp \mathfrak{k} := SU(2) \otimes SU(2) := \{U_1 \otimes U_2 \mid U_1, U_2 \in SU(2)\}, \tag{37}$$

$$\mathfrak{p} := \langle iI_x S_x, iI_x S_y, iI_x S_z, \dots, iI_z S_z \rangle, \tag{38}$$

$$\mathfrak{a} := \langle iI_x S_x, iI_y S_y, iI_z S_z \rangle. \tag{39}$$

The intersection $\exp \mathfrak{a} \cap \exp \mathfrak{k}$ is given by the eight elements

$$\exp \mathfrak{a} \cap \exp \mathfrak{k} = \left\{ \pm I_4, \pm \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \pm \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \pm \begin{bmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix} \right\} \tag{40}$$

and the Weyl orbit of Δ is equal to

$$\{\pm 2\pi i I_x S_x, \pm 2\pi i I_y S_y, \pm 2\pi i I_z S_z\}. \tag{41}$$

Note, that its simple form is due to the rather special nature of the drift term in Eq. (3). In general, a Weyl orbit does not possess the above sign symmetry, i.e., for arbitrary $\Delta \in \mathfrak{a}$, there is no Weyl group element such that

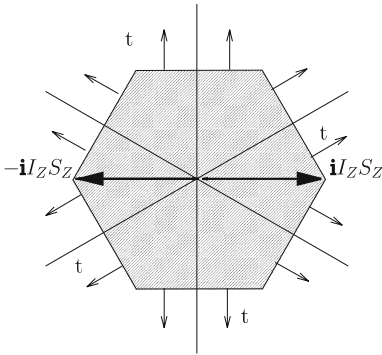


Figure 1. Convex hull of the Weyl Orbit of the symmetric drift term $iI_Z S_Z$.

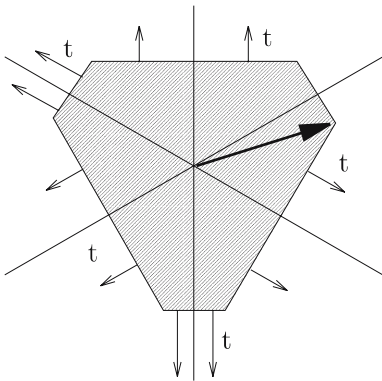


Figure 2. Convex hull of the Weyl Orbit of a nonsymmetric drift term $\Delta \in \mathfrak{a}$.

$W\Delta W^\dagger = -\Delta$, cf. Figures 1 and 2 for illustrations of two different situations.

Now, it is easy to check that \mathfrak{k} and \mathfrak{p} as defined above provide a Cartan-like decomposition of $SU(4)$ with \mathfrak{a} maximal Abelian in \mathfrak{p} . Hence we obtain the following corollaries from Theorems 5 and 6, respectively.

COROLLARY 8. *The Schrödinger equation (3) is controllable.*

Proof. Since $SU(4)$ is simple, this is an immediate consequence from Theorem 5. □

COROLLARY 9. *Let Δ , \mathfrak{k} , \mathfrak{p} , and \mathfrak{a} be defined as above and let*

$$X_F = K_1 e^{-2\pi i(\alpha_1 I_x S_x + \alpha_2 I_y S_y + \alpha_3 I_z S_z)} K_2 \in SU(4)$$

with $K_1, K_2 \in \exp \mathfrak{k}$ and $\alpha := (\alpha_1, \alpha_2, \alpha_3)^\top \in \mathbb{R}^3$ be any final state. Then the optimal time $T(X_F)$ is given by

$$T(X_F) = \min \left\{ \sum_{k=1}^3 |t_k| \mid (t_1, t_1, t_3) \in \alpha + \widehat{\Gamma}_0 \right\}, \tag{42}$$

where $\widehat{\Gamma}_0 := \left\{ (2k_1+2k_2, 2k_1+2k_3, 2k_2+2k_3)^\top \mid k_1, k_2, k_3 \in \mathbb{Z} \right\}$.

Proof. Since $SU(4)$ is simple, Corollary 7 implies

$$T(X_F) = \min \{ t \geq 0 \mid t \cdot \text{cov}(\Delta) \cap (a_0 + \Gamma_0) \neq \emptyset \}. \tag{43}$$

Choosing $\{-2\pi i I_x S_x, -2\pi i I_y S_y, -2\pi i I_z S_z\}$ as a basis of \mathfrak{a} , we can write

$$a_0 = -2\pi i (\alpha_1 I_x S_x + \alpha_2 I_y S_y + \alpha_3 I_z S_z).$$

Moreover, a straightforward calculation using Eq. (40) yields

$$\Gamma_0 = \left\{ 4\pi i \left((k_1+k_2) I_x S_x - (k_1+k_3) I_y S_y - (k_2+k_3) I_z S_z \right) \mid k_1, k_2, k_3 \in \mathbb{Z} \right\}$$

and hence $a_0 + \Gamma_0$ is given by

$$-2\pi i \left((\alpha_1 - 2k_1 - 2k_2) I_x S_x + (\alpha_2 + 2k_1 + 2k_3) I_y S_y + (\alpha_3 + 2k_2 + 2k_3) I_z S_z \right),$$

with $k_1, k_2, k_3 \in \mathbb{Z}$. Now, by Eq. (43) and the particular form of the Weyl orbit of $-2\pi i I_z S_z$ we obtain Eq. (42). \square

This leads at least in principle to a solution of Problem (IIa). Therefore, we next address Problem (IIb). While Proposition 11 shows that it is not possible to steer from one component of \mathcal{X}_{\max} to the other one in time zero, Theorem 12 guarantees that each maximum of the transfer function (6) can be reached in the same optimal time.

LEMMA 10. *Let $X = \text{diag}(U, U) \in SU(4)$. Then*

$$X \in SU(2) \otimes SU(2) \iff U \in SU(2).$$

The proof of Lemma 10 is omitted.

PROPOSITION 11. *Let \mathcal{X}_{\max}^+ and \mathcal{X}_{\max}^- be defined as in Theorem 1. Then the following holds.*

(a) *If $X \in \mathcal{X}_{\max}^+$ then*

$$\mathcal{X}_{\max}^+ \subset \left(SU(2) \otimes SU(2) \right) X \quad \text{and} \quad \mathcal{X}_{\max}^- \cap \left(SU(2) \otimes SU(2) \right) X = \emptyset.$$

(b) If $X \in \mathcal{X}_{\max}^-$ then

$$\mathcal{X}_{\max}^- \subset \left(SU(2) \otimes SU(2) \right) X \quad \text{and} \quad \mathcal{X}_{\max}^+ \cap \left(SU(2) \otimes SU(2) X \right) = \emptyset.$$

Proof. Proposition 11 is an immediate consequence of Theorem 1 and Lemma 10. □

THEOREM 12. *Each element in \mathcal{X}_{\max}^+ and \mathcal{X}_{\max}^- can be reached in the same optimal time $T = \frac{3}{2}$.*

Proof. By Proposition 11 and the Equivalence Theorem, cf. [15], it is sufficient to show that $X_* \in \mathcal{X}_{\max}^+$, defined by Eq. (10), and $iX_* \in \mathcal{X}_{\max}^-$, can be reached in the same optimal time $T = \frac{3}{2}$. Using Corollary 7, we sketch the calculations that have to be done to determine the optimal times for X_* and iX_* . First, we have to find a KAK -decomposition for X_* and iX_* , respectively. According to Eq. (23) this means a factorization of the type

$$K_1 e^{-2\pi i(\alpha_1 I_x S_x + \alpha_2 I_y S_y + \alpha_3 I_z S_z)} K_2, \tag{44}$$

where $K_1, K_2 \in SU(2) \otimes SU(2)$. Possible ones are

$$K_1 = K_2 = I_4, \quad \alpha = \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2} \right) \quad \text{for } X^*$$

and

$$K_1 = K_2 = I_4, \quad \alpha = \left(-\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2} \right) \quad \text{for } iX^*.$$

This factorization can be computed via the results of Appendix A.1. Now, by Eq. (42) we have to choose $k_1, k_2, k_3 \in \mathbb{Z}$ such that the sum

$$\left| \frac{1}{2} - 2(k_1 + k_2) \right| + \left| \frac{1}{2} + 2(k_1 + k_3) \right| + \left| \frac{1}{2} + 2(k_2 + k_3) \right|$$

and

$$\left| -\frac{1}{2} - 2(k_1 + k_2) \right| + \left| -\frac{1}{2} + 2(k_1 + k_3) \right| + \left| -\frac{1}{2} + 2(k_2 + k_3) \right|,$$

respectively, is minimized. In both cases, the minimum is obtained for $k_1 = k_2 = k_3 = 0$ and thus the optimal times are equal to $T = \frac{3}{2}$. □

As mentioned before, the above theorem shows that each maximum of the transfer function can be reached in the same optimal time. However, this is no longer true, if the system's initial state is different from the identity matrix. Therefore, it is of interest to study the set \mathcal{I}_T of all unitary operators $X \in SU(4)$ which can be steered to the set of maxima \mathcal{X}_{\max} in a given but fixed time $T \geq 0$ and in particular, its interplay with the fibres of f . A partial result in this direction is given in Proposition 13.

PROPOSITION 13. *The intersection of the fibre $f^{-1}(0)$ and the set \mathcal{I}_0 of all unitary operators $X \in SU(4)$ which can be steered in time zero to the set of maxima \mathcal{X}_{\max} is equal to*

$$\mathcal{I}_0 \cap f^{-1}(0) = \left\{ \begin{bmatrix} 0 & e^{i\phi} \\ -e^{-i\phi} & 0 \end{bmatrix} \otimes U \mid U \in U(2), \det U = \pm 1 \right\} X_*. \tag{45}$$

Proof. Let $\widehat{X} \in \mathcal{X}_{\max}$ be any maximum of the transfer function f . By the Equivalence Principle, cf. [15], and Theorem 1, any element which can be reached from \widehat{X} in time zero is given by $X = (U_1 \otimes U_2)X_*$, with $U_1 \in SU(2)$, $U_2 \in U(2)$ and $\det U_2 = \pm 1$. Moreover,

$$\begin{aligned} f(X) &= \text{Retr}(C^\dagger X A X^\dagger) \\ &= \text{Retr}(C^\dagger (U_1 \otimes U_2) C (U_1 \otimes U_2)^\dagger) \\ &= \text{Retr} \left\{ \left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) (U_1 \otimes U_2) \left(\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \otimes \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) (U_1^\dagger \otimes U_2^\dagger) \right\} \\ &= 2 \text{Retr} \left\{ \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} U_1 \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} U_1^\dagger \right\}. \end{aligned}$$

Hence

$$f(X) = 0 \iff U_1 = \begin{bmatrix} 0 & e^{i\phi} \\ -e^{-i\phi} & 0 \end{bmatrix},$$

and thus

$$\mathcal{I}_0 \cap f^{-1}(0) = \left\{ \begin{bmatrix} 0 & e^{i\phi} \\ -e^{-i\phi} & 0 \end{bmatrix} \otimes U_2 \mid U_2 \in U(2), \det U_2 = \pm 1 \right\} X_*. \quad \square$$

4. Numerical Approach

In this section we develop a numerical algorithm to solve the time optimal control problem associated with Eq. (33). Again, let G be a connected and compact (matrix) Lie group with semisimple Lie algebra \mathfrak{g} and let $\mathfrak{k} := \langle \Omega_1, \dots, \Omega_N \rangle_L \subset \mathfrak{g}$ be a subalgebra. Furthermore, let $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{p}$, $\mathfrak{p} := \mathfrak{k}^\perp$ be a Cartan-like decomposition and let $\mathfrak{a} \subset \mathfrak{p}$ be an arbitrary but fixed maximal Abelian subalgebra of \mathfrak{p} such that $\Delta \in \mathfrak{a}$. Finally, let

$$\{\Delta_1, \dots, \Delta_p\} := \{W \Delta W^\dagger \mid W \in \mathcal{W}(G, \exp \mathfrak{a})\} \subset \mathfrak{a} \tag{46}$$

be the Weyl orbit of Δ . Now, consider an arbitrary element $X_F \in G$. Our goal is to compute a factorization

$$X_F = K_1 \underbrace{\exp\left(\sum_{k=1}^p t_k \Delta_k\right)}_{=: A(t_1, \dots, t_p)} K_2 = K_1 \prod_{k=1}^p e^{t_k \Delta_k} K_2 \tag{47}$$

with $K_1, K_2 \in \exp \mathfrak{k}$, $t_k \geq 0$ for $k = 1, \dots, p$ such that $T := \sum_{k=1}^p t_k$ is minimal among all possible factorizations. By Theorem 6(b), this factorization problem can be interpreted equivalently as an optimal control problem for (33). Now, we proceed in detail as follows. Let $\Phi : G \rightarrow [0, \infty)$ be a smooth function satisfying

$$\Phi(X) = 0 \iff X = X_F.$$

With it, Eq. (47) can be cast into the following optimization setting.

PROBLEM 1.

$$\begin{aligned} \min \quad & f(t, K_1, K_2) := \sum_{k=1}^p t_k, \\ \text{subject to} \quad & g(t, K_1, K_2) := \Phi(X(t, K_1, K_2)) = 0, \\ & t \geq 0, \end{aligned} \tag{48}$$

where $t := (t_1, \dots, t_p) \in \mathbb{R}^p$ and $X(t, K_1, K_2)$ is defined by the right hand side of Eq. (47).

In general, Problem 1 is a nonlinear programming task. Any direct approach faces numerical difficulties due to the complexity of the feasible region. Therefore, following a well-known procedure from optimal control [21], we propose an indirect method by computing the associated *value functions* W and V which are defined as follows.

PROBLEM 2. (a) Given $T \geq 0$, solve

$$\begin{aligned} \min \quad & g(t, K_1, K_2), \\ \text{subject to} \quad & f(t, K_1, K_2) = T, \\ & t \geq 0. \end{aligned}$$

Set $W(T) := g(t^*, K_1^*, K_2^*)$, where (t^*, K_1^*, K_2^*) is a solution of (a).

(b) Given $T \geq 0$, solve

$$\begin{aligned} \min \quad & g(t, K_1, K_2), \\ \text{subject to} \quad & f(t, K_1, K_2) \leq T, \\ & t \geq 0. \end{aligned}$$

Set $V(T) := g(t^*, K_1^*, K_2^*)$, where (t^*, K_1^*, K_2^*) is a solution of (b).

Note,

$$V(T) = \min_{0 \leq \tau \leq T} W(\tau). \tag{49}$$

Moreover, recall that the value functions $W(T)$ and $V(T)$ for an optimal control problem are the minimal achievable values of the cost function at control time T and up to control time T , respectively. Thus, by Theorem 6(a), $W(T)$ and $V(T)$ can be interpreted as the value functions for the optimal control problem of minimizing $\Phi(X(t))$ subject to solutions $X(t)$ of (33).

Finally, we arrive at the problem of finding the smallest nonnegative T such that $V(T) = 0$.

PROBLEM 3.

$$\begin{aligned} \min \quad & T \\ \text{subject to} \quad & V(T) = 0, \\ & T \geq 0. \end{aligned}$$

where $V(T)$ is defined as in Problem 2(b).

Remark 2. The following properties related to the above optimization problems hold:

- (a) If T^* is the solution of Problem 3 and (t^*, K_1^*, K_2^*) is a solution of Problem 2(b) associated with $T = T^*$, then (t^*, K_1^*, K_2^*) is a solution of Problem 1.
- (b) If (t^*, K_1^*, K_2^*) is a solution of Problem 1, then for $T^* := f(t^*, K_1^*, K_2^*)$, the triple (t^*, K_1^*, K_2^*) is also a solution of Problem 2(b) associated with $T = T^*$, and T^* is the solution of Problem 3.
- (c) The function $V(T)$ is non-increasing over $T \geq 0$. Let (t^*, K_1^*, K_2^*) be a solution of Problem 2(b) associated with some $T = T'$. If $T^* := f(t^*, K_1^*, K_2^*) < T'$, then $V(T)$ remains constant on $[T^*, T']$.

Property (a) in Remark 2 shows that if a solution T^* is found for Problem 3, so (t^*, K_1^*, K_2^*) is a solution for Problem 1. Therefore, our effort will be made for a numerical solution of Problem 3. Our task is a twofold one. Firstly, we need a powerful algorithm for solving Problem 3, assuming $V(T)$ can be evaluated efficiently. Secondly, we need a fast and reliable method to compute $V(T)$, i.e., to solve Problem 2(b) associated with $T \geq 0$.

Note that Problem 3 is a one dimensional problem which involves a non-linear and probably non-smooth constraint function $V(T)$. What we know

about $V(T)$ is that it is a nonnegative, non-increasing function. Moreover, $V(T)$ is also expected to be zero for large T values. Thus we are looking for the smallest nonnegative T^* such that $V(T)$ remains 0 on $[T^*, \infty)$. It is clear that in our case gradient based methods will not work reliably. Hence, using the properties listed in Remark 2 and the idea of golden section, we devise the following algorithm for solving Problem 3.

ALGORITHM 1

Step 0

Choose $M > m \geq 0$ s.t. estimated optimal T is between m and M .

Choose small $\delta > 0$.

Set $T_l = m$ and $T_u = M$.

Step 1

Let $T_0 = T_l + 0.618(T_u - T_l)$.

Step 2

Solve Problem 2(b) for $T = T_0$ to obtain a solution (t', K'_1, K'_2) , using the hybrid descent method detailed in [20].

Step 3

If $g(t', K'_1, K'_2) > 0$

Set $T_l = T_0$.

Go to Step 1.

Else

Set $T_u = f(t', K'_1, K'_2)$ and $T_0 = \max(T_l, T_u - \delta)$.

Go to Step 4.

End

Step 4

If $T_l \geq T_u - \delta$

Output $T^* = T_u$ and $(t^*, K_1^*, K_2^*) = (t', K'_1, K'_2)$.

Take T^* as optimal solution of Problem 3.

Take (t^*, K_1^*, K_2^*) as optimal solution of Problem 1.

Stop.

Else

Set $T_0 = T_u - \delta$.

Go to Step 2.

End

The difficulty of the above algorithm lies in the following two points: (a) Find a reasonable and explicit parametrization $K = K(\mu_1, \dots, \mu_r)$ of the subgroup K . (b) Evaluate $V(T)$, which is the global optimal value of Problem 2(b).

Obviously, the first one depends on the subgroup K . To counter the second one, we choose the hybrid descent algorithm developed in [20], which combines simulated annealing with a gradient based method. The simulated annealing method is employed to escape from local minima, while a gradient descent is used to find a better local minima. The hybrid method has been successfully tested by solving various problems of different dimensions. Simulations show that our method is efficient for solving Problem 2(b).

5. Application to Coupled Spin Systems

We illustrate the above general Algorithm 1 to solve Problem (IIa) for a system of two weakly coupled spin- $\frac{1}{2}$ particles. Calculations have been implemented in MATLAB. To obtain an explicit description of factors K_1 and K_2 , we exploit the Euler parametrization of $SU(2)$, i.e., we express both factors in terms of the one-parameter subgroups generated by iH_1, \dots, iH_4 or, equivalently, by $-2\pi iI_x, -2\pi iI_y, -2\pi iS_x, -2\pi iS_y$, namely

$$K_1(\mu_1, \dots, \mu_6) = e^{-2\pi i\mu_1 I_x} e^{-2\pi i\mu_2 I_y} e^{-2\pi i\mu_3 I_x} e^{-2\pi i\mu_4 S_x} e^{-\pi i\mu_5 S_y} e^{-2\pi i\mu_6 S_x} \quad (50)$$

$$K_2(\mu_7, \dots, \mu_{12}) = e^{-2\pi i\mu_7 I_x} e^{-2\pi i\mu_8 I_y} e^{-2\pi i\mu_9 I_x} e^{-2\pi i\mu_{10} S_x} e^{-2\pi i\mu_{11} S_y} e^{-2\pi i\mu_{12} S_x}. \quad (51)$$

This yields the following parametrization for the right hand side of Eq. (47)

$$X(t, \mu) := K_1(\mu_1, \dots, \mu_6) A(t_1, \dots, t_6) K_2(\mu_7, \dots, \mu_{12}), \quad (52)$$

$$A(t_1, \dots, t_6) := e^{-2\pi i(t_1-t_2)I_x S_x} e^{-2\pi i(t_3-t_4)I_y S_y} e^{-2\pi i(t_5-t_6)I_z S_z}. \quad (53)$$

In order to gain an explicit pulse sequence, it is necessary to express the factors in Eq. (53) purely in terms of the drift Hamiltonian and the control Hamiltonians. Taking advantage of the structure of the Weyl orbit of $-2\pi iI_z S_z$, cf. Eq. (41), we have

$$\pm iI_x S_x = e^{-\frac{\pi i}{2} I_y} e^{\mp \frac{\pi i}{2} S_y} iI_z S_z e^{\pm \frac{\pi i}{2} S_y} e^{\frac{\pi i}{2} I_y}, \quad (54)$$

$$\pm iI_y S_y = e^{\frac{\pi i}{2} S_x} e^{\pm \frac{\pi i}{2} I_x} iI_z S_z e^{\mp \frac{\pi i}{2} I_x} e^{-\frac{\pi i}{2} S_x}, \quad (55)$$

$$-iI_z S_z = e^{-\pi i S_y} iI_z S_z e^{\pi i S_y}. \quad (56)$$

A detailed description of the matrices appearing in Eqs. (50)–(56) is given in the Appendix A.2. Now let $X_F \in SU(4)$ be any final state and

$\Phi : SU(4) \rightarrow [0, \infty)$ be defined by $\Phi(X) := \|X_F - X\|^2$, where $\|\cdot\|$ denotes the Frobenius norm. Due to the identity

$$\Phi(X(t, \mu)) = \|X_F - X(t, \mu)\|^2 = 8 - 2 \operatorname{Re} \operatorname{tr} (X_F^\dagger X(t, \mu))$$

we arrive at the following optimization problem.

PROBLEM 4.

$$\begin{aligned} \min \quad & f(t, \mu) := t_1 + t_2 + t_3 + t_4 + t_5 + t_6, \\ \text{subject to} \quad & g(t, \mu) := 4 - \operatorname{Re} \operatorname{tr} (X_F^\dagger X(t, \mu)) = 0, \\ & t \geq 0, \end{aligned} \tag{57}$$

where $t = [t_1, \dots, t_6]^\top \in \mathbb{R}^6$, $u = [\mu_1, \dots, \mu_{12}]^\top \in \mathbb{R}^{12}$, and $X(t, \mu)$ is defined as in Eq. (52).

Finally, we present two numerical experiments, both of them produce time optimal pulse sequences. Note that the set of unitary (2×2) -matrices having determinant equal to ± 1 can be parameterized as

$$U = \begin{bmatrix} \varepsilon \cos(\alpha) e^{i\beta} & -\varepsilon \sin(\alpha) e^{-i\gamma} \\ \sin(\alpha) e^{i\gamma} & \cos(\alpha) e^{-i\beta} \end{bmatrix}$$

with $\alpha, \beta, \gamma \in [-\pi, \pi]$ and $\varepsilon \in \{+1, -1\}$. Consequently, X_F is a maximizing transformation of the transfer function (6) if and only if

$$X_F(\alpha, \beta, \gamma, \varepsilon) = e^{-\frac{i\pi}{4}} \begin{bmatrix} \varepsilon \cos(\alpha) e^{i\beta} & 0 & -\varepsilon \sin(\alpha) e^{-i\gamma} & 0 \\ \sin(\alpha) e^{i\gamma} & 0 & \cos(\alpha) e^{-i\beta} & 0 \\ 0 & \varepsilon \cos(\alpha) e^{i\beta} & 0 & -\varepsilon \sin(\alpha) e^{-i\gamma} \\ 0 & \sin(\alpha) e^{i\gamma} & 0 & \cos(\alpha) e^{-i\beta} \end{bmatrix}. \tag{58}$$

EXAMPLE 1. The first experiment reproduces a result given in [15] which therefore seems to be well-known to NMR experimenters. Let

$$X_F(0, 0, 0, +1) := X_* = A\left(\frac{1}{2}, 0, \frac{1}{2}, 0, \frac{1}{2}, 0\right).$$

In this case the control parameters μ_1, \dots, μ_{12} can be chosen all equal to zero. The smallest value found for T and the corresponding optimal time are

$$\begin{aligned} T^* &= 1.499996 \\ t^* &= [0.499993; 0; 0.500017; 0; 0.499986; 0] \end{aligned}$$

EXAMPLE 2. For the second experiment, we have chosen

$$X_F \left(\frac{\pi}{2}, 0, -\frac{\pi}{2}, -1 \right) = \begin{bmatrix} 0 & i & 0 & 0 \\ -i & 0 & 0 & 0 \\ 0 & 0 & 0 & i \\ 0 & 0 & -i & 0 \end{bmatrix} X_*$$

In the actual implementation of the algorithm, we set $m = 1$, $M = 3.5$ and $\delta = 0.001$. Problem 4 is solved with only two evaluations of the function $V(T)$. The smallest value found for T and the corresponding (t^*, μ^*) are listed below.

$$T^* = 1.500019$$

$$t^* = [0.500015; 0; 0; 0.499997; 0; 500007, 0]$$

$$\mu^* = [0.999999; 0.750487; 0.999999; 0.825958; 1.000000; 0.867122; 0.829952; 1.000000; 0.863124; -1.000000; 0.749514; 1.000000]$$

The computed optimal solution satisfies $g(t^*, \mu^*) \approx 3.0e - 10$. The matrix $X_F^\dagger X(t, \mu)$ in Eq. (57), evaluated at the computed optimal solution (t^*, μ^*) , is approximately (error $< 10^{-6}$) the identity. The plot of the value function $V(T)$ over the interval $[0, 2]$ is given in Figure 3. Further experiments

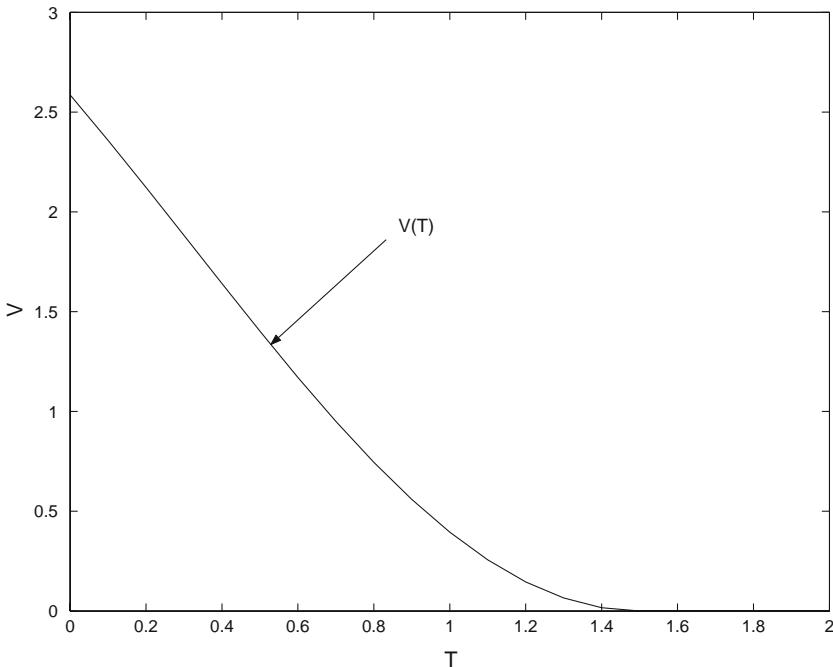


Figure 3. Plot of $V(T)$ over $[0, 2]$.

with randomly generated X_F , i.e. with randomly chosen $\alpha, \beta, \gamma \in [-\pi, \pi]$, $\varepsilon \in \{+1, -1\}$ confirmed our theoretical results.

Appendix A

A.1. SOME FACTS ON $SU(2) \otimes SU(2)$

We present two equivalent representations of $\mathfrak{su}(2) \times \mathfrak{su}(2)$ in $\mathfrak{su}(4)$ to conclude that the subgroup $SU(2) \otimes SU(2)$ coincides up to conjugation with the subgroup

$$SO(4) := \{X \in \mathbb{R}^{4 \times 4} \mid XX^T = I_4, \det(X) = 1\}.$$

Consider the following basis of the Lie algebra $\mathfrak{so}(4)$ of all 4×4 skew-symmetric matrices

$$\begin{aligned} \Omega_x &:= \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, & \Omega_y &:= \begin{bmatrix} 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, & \Omega_z &:= \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix}, \\ \Psi_x &:= \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1 & 0 \end{bmatrix}, & \Psi_y &:= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, & \Psi_z &:= \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{bmatrix}. \end{aligned}$$

It is easily checked that

$$\rho_1 : \mathfrak{su}(2) \times \mathfrak{su}(2) \longrightarrow \mathfrak{so}(4) \subset \mathfrak{su}(4)$$

defines a faithful representation via

$$\begin{aligned} \rho_1(i\sigma_k, 0) &:= \Omega_k, & k \in \{x, y, z\} \\ \rho_1(0, i\sigma_k) &:= \Psi_k, & k \in \{x, y, z\}. \end{aligned}$$

Another faithful representation of $\mathfrak{su}(2) \times \mathfrak{su}(2)$ is given by

$$\begin{aligned} \rho_2 : \mathfrak{su}(2) \times \mathfrak{su}(2) &\longrightarrow \mathfrak{su}(4) \\ \rho_2(\Omega_1, \Omega_2) &= \Omega_1 \otimes I_2 + I_2 \otimes \Omega_2. \end{aligned} \tag{A.1}$$

Note that both representations are irreducible, i.e. there are no $\Lambda_i \in GL(4, \mathbb{C})$, $i = 1, 2$ such that $\Lambda_i \rho_i(\Omega_1, \Omega_2) \Lambda_i^{-1}$ is block diagonal for all $\Omega_1, \Omega_2 \in \mathfrak{su}(2)$, cf. [9]. Using the notation of Section 3 we have

$$\rho_2(\mathfrak{su}(2) \times \mathfrak{su}(2)) = \mathfrak{su}(2) \widehat{\otimes} \mathfrak{su}(2)$$

and

$$\exp\left(\rho_2(\mathfrak{su}(2) \times \mathfrak{su}(2))\right) = SU(2) \otimes SU(2).$$

The relation between ρ_1 and ρ_2 as well as between $SU(2) \otimes SU(2)$ and $SO(4)$ are as follows, cf. [9].

THEOREM A.1. *The representations ρ_1 and ρ_2 are equivalent, i.e. there exists a $\Sigma \in SU(4)$ such that $\Sigma\rho_2\Sigma^\dagger = \rho_1$. In particular, we have the following equalities*

$$(a) \quad \Sigma (\mathfrak{su}(2) \widehat{\otimes} \mathfrak{su}(2)) \Sigma^\dagger = \mathfrak{so}(4)$$

$$(b) \quad \Sigma (SU(2) \otimes SU(2)) \Sigma^\dagger = SO(4).$$

Proof. To prove (a) we explicitly present a special unitary matrix which simultaneously maps any element of $\mathfrak{su}(2) \widehat{\otimes} \mathfrak{su}(2)$ onto $\mathfrak{so}(4)$ via conjugation. Part (b) then follows by the well-known properties of the exponential map.

Let

$$\Sigma := \frac{1}{2} \begin{bmatrix} 1+i & 0 & 0 & 1+i \\ 0 & -1+i & -1+i & 0 \\ 0 & 1+i & -1-i & 0 \\ -1+i & 0 & 0 & 1-i \end{bmatrix}. \tag{A.2}$$

Obviously, iI_x, \dots, iS_z defined as in Eq. (5) form a basis of $\mathfrak{su}(2) \widehat{\otimes} \mathfrak{su}(2)$. Moreover, a straightforward computation yields $i\Sigma I_j \Sigma^\dagger = \Omega_j$ and $i\Sigma S_j \Sigma^\dagger = \Psi_j$ for all $j \in \{x, y, z\}$ and this immediately implies the equivalence of ρ_1 and ρ_2 . \square

For computations, it is also of interest that the above transformation $\Omega \mapsto \Sigma \Omega \Sigma^\dagger$ maps the maximal Abelian subalgebra \mathfrak{a} defined as in Eq. (39) onto the standard maximal Abelian subalgebra of $SU(4)$, i.e., onto the traceless, diagonal, skew-Hermitian matrices.

LEMMA A.2. *Let Σ be defined as in Eq. (60) and \mathfrak{a} as in Eq. (39). Then $\Sigma \mathfrak{a} \Sigma^\dagger$ is diagonal. In particular one has*

$$i\Sigma I_X S_X \Sigma^\dagger = \frac{1}{4} \text{diag}(i, i, -i, -i),$$

$$i\Sigma I_Y S_Y \Sigma^\dagger = \frac{1}{4} \text{diag}(-i, i, -i, i),$$

$$i\Sigma I_Z S_Z \Sigma^\dagger = \frac{1}{4} \text{diag}(i, -i, -i, i).$$

The proof of Lemma A.2 is omitted.

A.2. EXPLICIT DESCRIPTION OF THE FACTORS

Here we present the matrices defined by Eqs. (50)–(56).

$$e^{-2\pi i\mu I_x} = \begin{bmatrix} \cos \pi \mu & 0 & -i \sin \pi \mu & 0 \\ 0 & \cos \pi \mu & 0 & -i \sin \pi \mu \\ -i \sin \pi \mu & 0 & \cos \pi \mu & 0 \\ 0 & -i \sin \pi \mu & 0 & \cos \pi \mu \end{bmatrix}, \quad (\text{A.3})$$

$$e^{-2\pi i\mu I_y} = \begin{bmatrix} \cos \pi \mu & 0 & -\sin \pi \mu & 0 \\ 0 & \cos \pi \mu & 0 & -\sin \pi \mu \\ \sin \pi \mu & 0 & \cos \pi \mu & 0 \\ 0 & \sin \pi \mu & 0 & \cos \pi \mu \end{bmatrix}, \quad (\text{A.4})$$

$$e^{-2\pi i\mu S_x} = \begin{bmatrix} \cos \pi \mu & -i \sin \pi \mu & 0 & 0 \\ -i \sin \pi \mu & \cos \pi \mu & 0 & 0 \\ 0 & 0 & \cos \pi \mu & -i \sin \pi \mu \\ 0 & 0 & -i \sin \pi \mu & \cos \pi \mu \end{bmatrix}, \quad (\text{A.5})$$

$$e^{-2\pi i\mu S_y} = \begin{bmatrix} \cos \pi \mu & -\sin \pi \mu & 0 & 0 \\ \sin \pi \mu & \cos \pi \mu & 0 & 0 \\ 0 & 0 & \cos \pi \mu & -\sin \pi \mu \\ 0 & 0 & \sin \pi \mu & \cos \pi \mu \end{bmatrix}, \quad (\text{A.6})$$

$$e^{2\pi i t I_x S_x} = \begin{bmatrix} \cos \frac{\pi}{2} t & 0 & 0 & -i \sin \frac{\pi}{2} t \\ 0 & \cos \frac{\pi}{2} t & -i \sin \frac{\pi}{2} t & 0 \\ 0 & -i \sin \frac{\pi}{2} t & \cos \frac{\pi}{2} t & 0 \\ -i \sin \frac{\pi}{2} t & 0 & 0 & \cos \frac{\pi}{2} t \end{bmatrix}, \quad (\text{A.7})$$

$$e^{2\pi i t I_y S_y} = \begin{bmatrix} \cos \frac{\pi}{2} t & 0 & 0 & i \sin \frac{\pi}{2} t \\ 0 & \cos \frac{\pi}{2} t & -i \sin \frac{\pi}{2} t & 0 \\ 0 & -i \sin \frac{\pi}{2} t & \cos \frac{\pi}{2} t & 0 \\ i \sin \frac{\pi}{2} t & 0 & 0 & \cos \frac{\pi}{2} t \end{bmatrix}, \quad (\text{A.8})$$

$$e^{2\pi i t I_z S_z} = \begin{bmatrix} e^{-\frac{\pi}{2} i t} & 0 & 0 & 0 \\ 0 & e^{\frac{\pi}{2} i t} & 0 & 0 \\ 0 & 0 & e^{\frac{\pi}{2} i t} & 0 \\ 0 & 0 & 0 & e^{-\frac{\pi}{2} i t} \end{bmatrix}. \quad (\text{A.9})$$

6. Conclusion

For a system of two weakly coupled spin- $\frac{1}{2}$ particles we have shown that the set of maxima \mathcal{X}_{\max} of the system's transfer function decomposes into two connected components \mathcal{X}_{\max}^+ and \mathcal{X}_{\max}^- . They are contained in two different right cosets of the fast subgroup $SU(2) \otimes SU(2)$ of Eq. (3). Nevertheless, each component and hence each element in \mathcal{X}_{\max} can be reached from the identity I_4 in the same optimal time $T = \frac{3}{2}$ ($J = 1$). Moreover, an efficient optimization algorithm for time optimal strategies for bilinear systems of type (26) has been developed. The method is based on Lie theoretic results in [15] and has been established on an arbitrary connected, compact, and semisimple Lie Group. A numerical implementation on $SU(4)$ has been worked out to explicitly compute time optimal pulse sequences for Eq. (3).

Acknowledgements

G. Dirr, U. Helmke and K. Hüper were supported by the DAAD project PPP Hong Kong/Germany D/01/22045. M. Kleinstauber was partially supported by the Marie-Curie PhD programme Control Training Site. Y. Liu was supported by RGC Germany/Hong Kong Joint Research Scheme G-Hk 014/01. The third author's employer, National ICT Australia Limited, is funded by the Australian Government's Department of Communications, Information Technology and the Arts and the Australian Research Council through Backing Australia's Ability and the ICT Centre of Excellence Program.

References

1. Agrachev, A. and Sachkov, Y. (2004), *Control Theory from the Geometric Viewpoint*, Springer Verlag, Berlin.
2. Albertini, F. and D'Alessandro, D. (2003), Notions of Controllability for Bilinear Multilevel Quantum Systems, *IEEE Transactions on Automatic Control* 48(8), 1399–1403.
3. D'Alessandro, D. (2003). Controllability of one spin and two interacting spins, *Mathematics Control Signals Systems* (1994) 16, 1–25.
4. Ando, T. and Li, C.-K. (eds.), *Special Issue: The Numerical Range and Numerical Radius of Linear and Multilinear Algebra* 37(1–3), 1–238, Gordon and Breach.
5. Bloch, A.M., Brockett, R.W. and Ratiu, T.S. (1992). Completely integrable gradient flows, *Communications in Mathematical Physics* 147, 57–74.
6. Brockett, R.W. (1990), Some Mathematical Aspects of Robotics. In: *Robotics, Proceedings of Symposia in Applied Mathematics*, Vol. 41, AMS, Providence.
7. Glaser, S.J., Schulte-Herbrüggen, T., Sieveking, M., Schedletsky, O., Nielsen, N.C., Sørensen, O.W. and Griesinger, C. (1998), Unitary control in quantum ensembles: Maximizing signal intensity in coherent spectroscopy, *Science* 280, 421–424.
8. Ernst, R.R., Bodenhausen, G. and Wokaun, A. (1987), *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, Oxford University Press, New York.

9. Goodman, R. and Wallach, N.R. (1998), *Representations and Invariants of the Classical Groups*, Cambridge University Press.
10. Hammerer, K., Vidal, G. and Cirac, J.I. (2002), Characterization of nonlocal gates, *Physical Review A* 66, 062321.
11. Helmke, U., Hüper, K., Moore, J.B. and Schulte-Herbrüggen, Th. (2002), Gradient flows computing the C-numerical range with applications in NMR spectroscopy, *Journal of Global Optimization* 23(3–4), 283–308.
12. Helgason, S. (1978), *Differential Geometry, Lie Groups, and Symmetric Spaces*, Academic Press, San Diego.
13. Helmke, U. and Moore, J.B. (1994), *Optimization and Dynamical Systems*, Springer Verlag, London.
14. Jurdjevic, V. (1997), *Geometric Control Theory*, Cambridge University Press, New York.
15. Khaneja, N., Brockett, R. and Glaser, S.J. (2001), Time optimal control in spin systems, *Physical Review A*, 63(3)/032308.
16. Khaneja, N., Glaser, S. and Brockett, R. (2002), Sub-Riemannian geometry and time optimal control of three spin systems: Quantum gates and coherence transfer, *Physical Review A*, 65(3)/032301.
17. Khaneja, N. and Glaser, S. (2001), Cartan decomposition of $SU(2^n)$ and control of spin systems, *Chemical Physics* 267, 11–23.
18. Knapp, A.W. (2002). *Lie Groups Beyond an Introduction*, Second Edition. Birkhäuser, Boston.
19. Schulte-Herbrüggen, T. (1998), *Aspects and Prospects of High-Resolution NMR*. PhD thesis, ETH Zürich. Dissertation ETH No. 12752.
20. Yiu, C., Liu, Y. and Teo, K.L. (2004), A hybrid decent method for global optimization. *Journal of Global Optimization* 28(2), 229–238.
21. Zabczyk, J. (1992), *Mathematical Control Theory: An Introduction*. Birkhäuser, Boston.