A Universal Approach to Interpolation on Reductive Homogeneous Spaces

Markus Schlarb

Department of Mathematics, Julius-Maximilians-Universität Würzburg, Germany
Center for Signal Analysis of Complex Systems, Ansbach University of Applied Sciences, Germany

markus.schlarb@uni-wuerzburg.de

Knut Hüper

Department of Mathematics, Julius-Maximilians-Universität Würzburg, Germany

knut.hueper@uni-wuerzburg.de

Fátima Silva Leite

Institute of Systems and Robotics, University of Coimbra, Portugal
Department of Mathematics, University of Coimbra, Portugal

fleite@mat.uc.pt

Abstract

A method for computing a $\mathscr{C}^2$-curve with given initial and final velocity interpolating a finite number of points on a reductive homogeneous space is presented. Here the reductive homogeneous space is assumed to be embedded into some manifold in a suitable way making the proposed approach very general. Building on the notion of intrinsic rolling, the method presented here offers a solution of the interpolation problem in closed form. This is illustrated on the example of matrix Lie groups. Moreover, this method is applied to the (compact) Stiefel manifold, where an efficient algorithm for solving the interpolation problem is also obtained.

Keywords: Interpolation, Lie Groups, Reductive Homogeneous Space, Rolling without slip and without twist, Stiefel Manifolds

2020 Mathematics Subject Classification: 53C30, 53Z30, 53Z50, 65D05, 65D07

Contents

1 Introduction

There exist several approaches to solve interpolation problems on manifolds. We do not give an overview here but mention interpolation by "Riemannian Splines", see for example the survey article [9] and references therein. These curves are defined as solutions of a variational problem. Hence they satisfy an optimality condition which might be also desirable from the point of view of applications. By means of an artificial potential, the variational approach can be extended to incorporate obstacle avoidance [8, 19]. Moreover, in principle, the variational approach can be applied to arbitrary Riemannian manifolds. Despite of this generality, approaching interpolation problems by "Riemannian splines" has the following drawback: In general, closed form expressions for these curves are not known, see for example [9, Sec. 2].

Besides generalizations of the de Casteljau algorithm, see e.g. [14], another approach for solving interpolation problems on (pseudo-)Riemannian manifolds is the so-called "Rolling and Unwrapping technique", see [29] and the recent work [26]. This method relies on knowing explicit expressions for the so-called "extrinsic rollings" of the manifold over an affine tangent space in the sense of [43, Ap. B, Def. 1.1]. For several manifolds, explicit closed-form solutions of an interpolation problem are obtained by this method. Indeed, in [26, 29], interpolation problems on the n-dimensional sphere $\mathrm{S}^n \subseteq \mathbb{R}^{n+1}$, the Grassmann manifold $\mathrm{Gr}_{n,k} \subseteq \mathbb{R}^{n\times n}_{\mathrm{sym}}$ and the special orthogonal group $\mathrm{SO}(n) \subseteq \mathbb{R}^{n\times n}$ are addressed. Curves generated by this method applied to $\mathrm{S}^2 \subseteq \mathbb{R}^3$ are investigated in [28]. Here we also mention [15], where this rolling and unwrapping technique applied to pseudo-orthogonal groups is briefly discussed.

This interpolation method was also used for solving several engineering problems. For instance, it was applied to $\mathrm{Gr}_{n,k}$ for domain adaptation [11]. In this context, we also refer to [4], where an interpolation problem on $\mathrm{Gr}_{n,k}$ was addressed by the de Casteljau algorithm. Very recently, in the important context of big data analysis, the concepts of de Casteljau's algorithm and Bezier curves are applied to polynomial regression problems on

Riemannian manifolds [21,22]. Moreover, supervised multi-class classification problems are addressed in [12] by using extrinsic rollings of $\mathrm{Gr}_{n,k}$. In addition, the interpolation method via extrinsic rolling and unwrapping from [26, 29] was adapted/applied to a product of special orthogonal groups $\mathrm{SO}(3) \times \cdots \times \mathrm{SO}(3)$ in [49] for the recognition of human actions. Also having the application of the rolling and unwrapping technique in mind, in [35], rolling maps for the essential manifold, viewed as the product $\mathrm{Gr}_{3,2} \times \mathrm{SO}(3)$, are studied. Finally, we also mention [44], where the authors propose to use the rolling and unwrapping technique applied to SO(3) for robot motion planning.

In the applications mentioned above, the manifolds to which the rolling and unwrapping technique from [26, 29] were applied, are either Grassmann manifolds, special orthogonal groups or products thereof. In spite of the success of the rolling and unwrapping technique, at least to our best knowledge, it has never been applied to manifolds which cannot be equipped with the structure of a (pseudo-)Riemannian symmetric space.

Nevertheless, interpolation problems on non-symmetric spaces are of interest from an applied point of view, as well. For instance, one faces interpolation problems on the real (compact) Stiefel in the context of domain adaptation [5]. Moreover, there is an approach to parametric model order reduction that leads to interpolation problems on the real (compact) Stiefel manifold [53], see also [52, Sec. 5.5]. Although extrinsic rollings of the Stiefel manifold $\mathrm{St}_{n,k} = \{X \in \mathbb{R}^{n\times k} \mid X^\top X = I_k\} \subseteq \mathbb{R}^{n\times k}$ have been already derived in [25], as far as we know, the rolling and unwrapping technique from [26, 29] has never been applied to $\mathrm{St}_{n,k}$, when $1 < k < n - 1$. It is well-known that $\mathrm{St}_{n,k}$ is not a symmetric space unless $k = 1$, $k = n - 1$ or $k = n$. Nevertheless, it still carries the structure of a reductive homogeneous space. Similarly, many other manifolds that play a role in engineering applications, like flag manifolds, see e.g. [46, 50] and references therein, are reductive homogeneous spaces.

This motivates us to propose a method for solving a specific interpolation problem, see Problem 2.1 in Section 2 below for its precise formulation, which is, at least in principle, applicable to an arbitrary reductive homogeneous space. In view of the applications mentioned above, potentially, the presented interpolation method can be applied to a wide range of engineering problems. Moreover, in many applications, the manifold of interest is not only a reductive homogeneous space but, in addition, embedded into some other manifold (typically a vector space). In this case, it is often convenient to implement algorithms on the embedded submanifold using the coordinates of the embedding space. For instance, usually, one prefers to perform computations on the sphere $\mathrm{S}^n \subseteq \mathbb{R}^{n+1}$ using the coordinates of $\mathbb{R}^{n+1}$ instead of relying on local charts or on a coset representation like $R \cdot \mathrm{SO}(n) \in \mathrm{SO}(n+1)/\mathrm{SO}(n) \cong \mathrm{S}^n$. Therefore the proposed approach for solving the interpolation problem is designed such that it is directly applicable to a reductive homogeneous space identified with an embedded submanifold of some other manifold via a suitable equivariant embedding.

The presented interpolation method is motivated by and closely related to the rolling and unwrapping technique from [26,29] already mentioned above. Analogously to the original rolling and unwrapping technique from [26,29], the method proposed in this text yields an explicit solution of the interpolation problem. Clearly, this is an advantage concerning implementations. Moreover, it requires only the knowledge of an intrinsic rolling in contrast to the rolling and unwrapping technique from [26,29] which relies on the notion of extrinsic rolling. In addition, we do not insist on exploiting a (pseudo-)Riemannian structure but, instead, consider rollings with respect to an arbitrary invariant covariant derivative. This freedom allows for additional design choices which can be used, for example, to reduce the computational cost.

The proposed method can be straightforwardly adapted to particular reductive homogeneous spaces. This is illustrated on the example of matrix Lie groups. Moreover, we apply the interpolation method to the real (compact) Stiefel manifold $\mathrm{St}_{n,k}$. By choosing an appropriate reductive decomposition and a suitable covariant derivative on $\mathrm{St}_{n,k}$, a closed-form solution of the interpolation problem is obtained, which is also efficient from a computational point of view for $k \ll n$.

Although the method proposed in this text provides some solution of the interpolation problem in closed form, we do not claim that this solution is optimal in some sense and investigations going into this direction are out of the scope of this text. Nevertheless, geometric properties of the curves like the length, energy and covariant acceleration can be computed at least numerically as we indicate in Section 7. These properties may serve as a measure for quality of the interpolating curves.

We now give an overview of this text. After the introduction, we state the interpolation problems precisely and explain our setting in more detail. The third section recalls the differential geometric concepts behind reductive homogeneous spaces and intrinsic rolling. To make the paper as self-contained as possible and, in particular, to ensure accessibility to a wider readership, modern but now standard notations are exclusively used. Also, some concepts such as invariant covariant derivatives, group actions, reductive decompositions, etc., are recalled. Because intrinsic rolling has not found yet its way into text books, a subsection recalls the precise mathematical definition as well as some results for rolling reductive homogeneous spaces from [41] which are relevant for this text. In the fourth section solving specific interpolation problems on reductive homogeneous spaces via intrinsic rolling and unwrapping techniques are presented in detail. In particular, an algorithm is proposed which allows from our perspective a direct application to specific reductive homogeneous spaces. For this, certainly some necessary background in differential geometry is needed. The fifth section is devoted to apply our ideas to specific interpolation poblems on matrix Lie groups, followed by another section where Stiefel manifolds are the central object. For both cases, i.e. matrix Lie groups and Stiefel manifolds, algorithms for solving interpolation problems are obtained which can be implemented straightforwardly only making use of standard (numerical) linear algebra and requiring only a modest background in differential geometry. The seventh section discusses some geometric properties of interpolation curves. A conclusion follows.

2 Problem Statement and Setting

We first formulate the interpolation problem which is of our main interest in this text.

Problem 2.1 *Let M be a connected manifold and let $x_0, \dots, x_k \in M$ be $k+1$ points on M. Moreover, let $0 = t_0 < \cdots < t_k = T$ and let $\xi_0 \in T_{x_0}M$ and $\xi_k \in T_{x_k}M$ be given. Compute a $\mathscr{C}^2$-curve $\beta\colon [0,T] \to M$ such that*

$$\beta(t_i) = x_i, \quad i \in \{0, \dots, k\} \tag{2.1}$$

holds and

$$\dot{\beta}(0) = \xi_0 \quad \textit{as well as} \quad \dot{\beta}(T) = \xi_k \tag{2.2}$$

is fulfilled.

To be more precise, we address Problem 2.1, where M is a connected reductive homogeneous space embedded into some manifold in a suitable equivariant way. As already pointed out in the introduction, taking such an embedding in the formulation of the interpolation method into account can be an advantage concerning implementations.

In addition, having a method to solve Problem 2.1 can be helpful to obtain a solution for the following interpolation problem which is called "Hermite manifold interpolation problem" in [52], see also [47].

Problem 2.2 *Let M be a connected manifold. Let $x_0, \dots, x_k \in M$ be $k+1$ points on M and let $0 = t_0 < \dots < t_k = T$. Moreover, let $\xi_0 \in T_{x_0}M, \dots, \xi_k \in T_{x_k}M$ be $k+1$ vectors in the tangent spaces at the given points $x_0, \dots, x_k \in M$. Compute a $\mathscr{C}^1$-curve $\beta\colon [0,T] \to M$ such that*

$$\beta(t_i) = x_i \quad \text{and} \quad \dot{\beta}(t_i) = \xi_i \tag{2.3}$$

holds for all $i \in \{0, \dots, k\}$.

Indeed, knowing how to solve Problem 2.1 for $k = 1$, a solution of Problem 2.2 can be obtained by piecing together the solutions of Problem 2.1. By this approach, an algorithm for solving Problem 2.2 on $\mathrm{St}_{n,k}$ is derived and shown to be computationally efficient for $k \ll n$ in Section 6. Furthermore an algorithm for solving Problem 2.2 on a matrix Lie group is obtained in Section 5, as well.

3 Reductive Homogeneous Spaces and Intrinsic Rolling

In this section, we recall some facts on reductive homogeneous spaces and introduce the notation that is used throughout the whole text if not indicated otherwise.

3.1 Notation and Terminology

If M is a smooth manifold, we denote its tangent bundle by TM. The algebra of smooth functions $M \to \mathbb{R}$ is denoted by $\mathscr{C}^\infty(M)$ and we write $\Gamma^\infty(TM)$ for the $\mathscr{C}^\infty(M)$-module of smooth vector fields on M. Next, let $f\colon M \to N$ be a smooth map between smooth manifolds. Then its tangent map is denoted by $Tf\colon TM \to TN$. If M is equipped with a pseudo-Riemannian metric, i.e. M is a pseudo-Riemannian manifold, the corresponding Levi-Civita covariant derivative on M is denoted by ∇^{LC}. Moreover, a curve $c\colon I \to M$ defined on some interval $I \subseteq \mathbb{R}$ is assumed to be smooth, i.e. $c \in \mathscr{C}^\infty(I, M)$, if not indicated otherwise.

We point out that we follow the convention in [38]. A scalar product on a vector space V is a non-degenerated symmetric bilinear form on V while an inner product on V is a positive definite symmetric bilinear form.

Next, we introduce some notations concerning Lie groups.

Notation 3.1 Let V be a finite dimensional $\mathbb{R}$-vector space. Then $\mathrm{GL}(V)$ denotes the general linear group of V, as usual. Its Lie algebra is denoted by $\mathfrak{gl}(V)$. If V is equipped with a scalar product $\langle \cdot, \cdot \rangle$, the (pseudo-)orthogonal group of V with respect to $\langle \cdot, \cdot \rangle$ is denote by $\mathrm{O}(V, \langle \cdot, \cdot \rangle)$, or simply by $\mathrm{O}(V)$ for short. The Lie algebra of $\mathrm{O}(V)$ is denoted by $\mathfrak{so}(V)$. For $V = \mathbb{R}^n$, equipped with the standard (Euclidean) inner product, we identify the general linear group and the orthogonal group with the matrix Lie groups denoted by

$$\mathrm{GL}(n) = \{A \in \mathbb{R}^{n \times n} \mid \det(A) \neq 0\} \tag{3.1}$$

and

$$\mathrm{O}(n) = \{R \in \mathbb{R}^{n \times n} \mid R^\top R = RR^\top = I_n\}, \tag{3.2}$$

respectively. Moreover, we write

$$\mathrm{SO}(n) = \{R \in \mathrm{O}(n) \mid \det(R) = 1\} \tag{3.3}$$

for the special orthogonal group. The Lie algebras of $\mathrm{SO}(n)$ and $\mathrm{O}(n)$ coincide and are both given by $\mathfrak{so}(n) = \{\Omega \in \mathbb{R}^{n\times n} \mid \Omega^\top = -\Omega\}$. Furthermore, we write $\mathfrak{gl}(n) = \mathbb{R}^{n\times n}$ for the Lie algebra of $\mathrm{GL}(n)$.

Next, let G be a Lie group and let $\mathfrak{g}$ be its Lie algebra. For fixed $g \in G$, we write

$$\ell_g \colon G \to G, \quad h \mapsto \ell_g(h) = gh \tag{3.4}$$

for the left translation by g and the right-translation by g is denoted by

$$r_g \colon G \to G, \quad h \mapsto r_g(h) = hg. \tag{3.5}$$

The neutral element in G is denoted by e, if not indicated otherwise. Moreover, we write

$$\mathrm{Ad} \colon G \to \mathrm{GL}(\mathfrak{g}), \quad g \mapsto (\xi \mapsto \mathrm{Ad}_g(\xi)) = g\xi g^{-1} \tag{3.6}$$

for the adjoint representation of G, where the second equality holds if G is a matrix Lie group. Furthermore, the exponential map of G is denoted by

$$\exp \colon \mathfrak{g} \to G. \tag{3.7}$$

3.2 Invariant Covariant Derivatives

We now consider invariant covariant derivatives on reductive homogeneous spaces. Since reductive homogeneous spaces play an essential role in this text, we recall their definition from [38, Chap. 11, Def. 21], see also [18, Def. 23.8]. We also refer to [18, Sec. 23.4] and [38, Chap. 11] for more details.

Definition 3.2 *Let G be a Lie group and let $H \subseteq G$ be a closed subgroup. Denote by $\mathfrak{g}$ the Lie algebra of G and write $\mathfrak{h}$ for the Lie algebra of H viewed as a Lie subalgebra of $\mathfrak{g}$. Then the homogeneous space G/H is called reductive if there exists a subspace $\mathfrak{m} \subseteq \mathfrak{g}$ such that $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$ holds and*

$$\mathrm{Ad}_h(\mathfrak{m}) \subseteq \mathfrak{m} \tag{3.8}$$

is satisfied for all $h \in H$.

From now on, if not indicated otherwise, we denote by G/H a reductive homogeneous space with a fixed reductive decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$. Let $\xi \in \mathfrak{g}$. Then we write

$$\xi_\mathfrak{m} = \mathrm{pr}_\mathfrak{m}(\xi) \in \mathfrak{m} \tag{3.9}$$

for the projection of ξ onto $\mathfrak{m}$ whose kernel is $\mathfrak{h}$. Analogously, the projection of ξ onto $\mathfrak{h}$, whose kernel is $\mathfrak{m}$, is denoted by $\xi_\mathfrak{h} = \mathrm{pr}_\mathfrak{h}(\xi) \in \mathfrak{h}$. We write

$$\mathrm{pr} \colon G \to G/H, \quad g \mapsto \mathrm{pr}(g) = g \cdot H \tag{3.10}$$

for the canonical projection, where $g \cdot H \in G/H$ denotes the coset defined by $g \in G$. Moreover, the map

$$\tau \colon G \times G/H \ni (g, g' \cdot H) \mapsto (gg') \cdot H \in G/H \tag{3.11}$$

is a transitive smooth G-action on G/H from the left. Following [18, p. 676], we denote by

$$\tau_g\colon G/H \to G/H, \quad g' \cdot H \mapsto (gg') \cdot H \tag{3.12}$$

the diffeomorphism associated with (3.11), where $g \in G$ is fixed. Note that $\tau_g \circ \mathrm{pr} = \mathrm{pr} \circ \ell_g$.

Next we briefly recall some facts on invariant covariant derivatives whose definition is recalled following [39, Def. 4.2].

Definition 3.3 *A covariant derivative* $\nabla\colon \Gamma^\infty\big(T(G/H)\big) \times \Gamma^\infty\big(T(G/H)\big) \to \Gamma^\infty\big(T(G/H)\big)$ *is called invariant if*

$$\nabla_X Y = (\tau_{g^{-1}})_*\big(\nabla_{(\tau_g)_* X}(\tau_g)_* Y\big) \tag{3.13}$$

holds for all $g \in G$ *and* $X, Y \in \Gamma^\infty\big(T(G/H)\big)$*, where* $(\tau_g)_* X = T\tau_g \circ X \circ \tau_{g^{-1}}$ *denotes the well-known push-forward of* X*.*

Definition 3.4 *Let* $\alpha\colon \mathfrak{m} \times \mathfrak{m} \to \mathfrak{m}$ *be a bilinear map. Then* α *is called an* $\mathrm{Ad}(H)$*-invariant bilinear map if*

$$\alpha\big(\mathrm{Ad}_h(X), \mathrm{Ad}_h(Y)\big) = \mathrm{Ad}_h\big(\alpha(X, Y)\big) \tag{3.14}$$

is fulfilled for all $X, Y \in \mathfrak{m}$ *and* $h \in H$*.*

We recall from [37, Thm. 8.1] that there is a one-to-one correspondence between $\mathrm{Ad}(H)$-invariant bilinear maps and invariant affine connections on G/H. These invariant affine connections are in one-to-one correspondence to the invariant covariant derivatives on G/H which are expressed in terms of horizontally lifted vector fields on G in [39, Sec. 4.2].

Here we only state the one-to-one correspondence of invariant covariant derivatives and $\mathrm{Ad}(H)$-invariant bilinear maps $\mathfrak{m} \times \mathfrak{m} \to \mathfrak{m}$ from [39, Def. 4.16] and refer to [37, 39] for more details.

Let $X_{G/H}, Y_{G/H} \in \Gamma^\infty\big(T(G/H)\big)$ denote the fundamental vector fields of $X, Y \in \mathfrak{m}$ associated to the action from (3.11), i.e.

$$X_{G/H}(x) = \tfrac{\mathrm{d}}{\mathrm{d}t} \tau_{\exp(tX)}(x)\big|_{t=0}, \quad x \in G/H \tag{3.15}$$

holds and $Y_{G/H}$ is defined analogously. Then $\nabla^\alpha\colon \Gamma^\infty\big(T(G/H)\big) \times \Gamma^\infty\big(T(G/H)\big) \to \Gamma^\infty\big(T(G/H)\big)$ denotes the unique invariant covariant derivative on G/H corresponding to the $\mathrm{Ad}(H)$-invariant bilinear map $\alpha\colon \mathfrak{m} \times \mathfrak{m} \to \mathfrak{m}$ by requiring

$$\nabla^\alpha_{X_{G/H}} Y_{G/H}\big|_{\mathrm{pr}(e)} = T_e \,\mathrm{pr}\big(-[X, Y]_{\mathfrak{m}} + \alpha(X, Y)\big) \tag{3.16}$$

for all $X, Y \in \mathfrak{m}$. There are two special invariant covariant derivatives, the so-called canonical invariant covariant derivatives of first and second kind which correspond to the canonical invariant affine connections of first and second kind from [37, Sec. 10]. We now recall their definition from [39, Def. 4.34].

Definition 3.5 *Let* G/H *be a reductive homogeneous spaces with reductive decomposition* $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$*.*

1. *The invariant covariant derivative associated to the* $\mathrm{Ad}(H)$*-invariant bilinear map defined by*

$$\alpha\colon \mathfrak{m} \times \mathfrak{m} \to \mathfrak{m}, \quad (X, Y) \mapsto \tfrac{1}{2}[X, Y]_{\mathfrak{m}} \tag{3.17}$$

is called canonical invariant covariant derivative of first kind. It is denoted by ∇^{can1}*.*

2. *The invariant covariant derivative associated to the* $\mathrm{Ad}(H)$*-invariant bilinear map defined by*

$$\alpha\colon \mathfrak{m}\times\mathfrak{m}\to\mathfrak{m},\quad (X,Y)\mapsto 0 \tag{3.18}$$

is called canonical invariant covariant derivative of second kind. It is denoted by $\nabla^{\mathrm{can}2}$.

For $X\in\mathfrak{m}$, the curve $\gamma\colon\mathbb{R}\ni t\mapsto \mathrm{pr}(\exp(tX))\in G/H$ is the geodesic through $\gamma(0)=\mathrm{pr}(e)$ with initial velocity $\dot\gamma(0)=T_e\,\mathrm{pr}\,X$ with respect to both canonical invariant covariant derivatives $\nabla^{\mathrm{can}1}$ and $\nabla^{\mathrm{can}2}$, see e.g. [39, Lem. 4.32].

3.3 Intrinsic Rolling

Next we recall a notion of rolling a manifold M over another manifold $\widehat{M}$ of equal dimension which can be considered as a reformulation of the definition from [20, Sec. 7], [31, p. 35], see also [41, Re. 3.3].

Definition 3.6 *Let* M *and* $\widehat{M}$ *be manifolds with* $\dim(M)=n=\dim(\widehat{M})$ *equipped with covariant derivatives* ∇ *and* $\widehat{\nabla}$*, respectively. An intrinsic rolling of* (M,∇) *over* $(\widehat{M},\widehat{\nabla})$ *is a triple* $(v(t),\gamma(t),A(t))$*, where* $v\colon I\to M$ *is the rolling curve,* $\gamma\colon I\to\widehat{M}$ *is the development curve and* $A(t)\colon T_{v(t)}M\to T_{\gamma(t)}\widehat{M}$ *is a curve of linear isomorphisms such that for each* $t\in I$*, the following assertions are fulfilled:*

1. *No-Slip condition:* $\dot\gamma(t)=A(t)\dot v(t)$.
2. *No-Twist condition: A vector field* $Z\colon I\to TM$ *along* $v\colon I\to M$ *is parallel with respect to* ∇ *iff the vector field* $\widehat{Z}\colon I\to T\widehat{M}$ *given by* $\widehat{Z}(t)=A(t)Z(t)$ *is parallel along* $\gamma\colon I\to\widehat{M}$ *with respect to* $\widehat{\nabla}$.

In the sequel, we denote by G/H a reductive homogeneous space with fixed reductive decomposition $\mathfrak{g}=\mathfrak{h}\oplus\mathfrak{m}$, as usual.

Moreover, as in [41, Eq. (5.1)], we always endow $\mathfrak{m}$ with the covariant derivative $\nabla^{\mathfrak{m}}$ defined as follows. Let $V\colon\mathfrak{m}\ni v\mapsto(v,V_2(v))\in\mathfrak{m}\times\mathfrak{m}\cong T\mathfrak{m}$ and $W\colon\mathfrak{m}\ni v\mapsto(v,W_2(v))\in\mathfrak{m}\times\mathfrak{m}\cong T\mathfrak{m}$ denote smooth vector fields on $\mathfrak{m}$, where $V_2,W_2\colon\mathfrak{m}\to\mathfrak{m}$ are smooth maps. Then $\nabla^{\mathfrak{m}}$ evaluated at $(V,W)\in\Gamma^\infty(T\mathfrak{m})\times\Gamma^\infty(T\mathfrak{m})$ is given by

$$\nabla^{\mathfrak{m}}_V W\big|_v=\big(v,(T_vW_2)V_2(v)\big),\quad v\in\mathfrak{m}. \tag{3.19}$$

Rollings of $\mathfrak{m}$ over G/H equipped with an invariant covariant derivative ∇^α are characterized in [41]. In particular, the so-called kinematic equation from [41, Proposition 5.14] holds.

Proposition 3.7 *Let* $u\colon I\to\mathfrak{m}$ *be a curve and let* $(v,g,S)\colon I\ni t\mapsto(v(t),g(t),S(t))\in\mathfrak{m}\times G\times\mathrm{GL}(\mathfrak{m})$ *be a solution of the ODE*

$$\begin{aligned}\dot v(t)&=u(t),\\ \dot g(t)&=\big(T_e\ell_{g(t)}\circ S(t)\big)u(t),\\ \dot S(t)&=-\alpha(S(t)u(t),\cdot)\circ S(t).\end{aligned} \tag{3.20}$$

Then the triple $(v(t),\gamma(t),A(t))$ *is an intrinsic rolling of* $\mathfrak{m}$ *over* G/H *with respect to* ∇^α *whose rolling curve is given by* $v\colon I\to\mathfrak{m}$*. Moreover, the development curve is defined by* $\gamma\colon I\ni t\mapsto(\mathrm{pr}\circ g)(t)\in G/H$ *and the linear isomorphism* $A(t)\colon T_{v(t)}\mathfrak{m}\cong\mathfrak{m}\to T_{\gamma(t)}(G/H)$ *reads*

$$A(t)Z=\big(T_{g(t)}\,\mathrm{pr}\circ T_e\ell_{g(t)}\circ S(t)\big)Z,\quad Z\in\mathfrak{m}. \tag{3.21}$$

We refer to the ODE (3.20) from Proposition 3.7 as kinematic equation. By abuse of notation, an initial value problem associated with (3.20) is called kinematic equation, as well.

For the interpolation algorithm that we propose below, an explicit expression for a rolling of $\mathfrak{m}$ over G/H along a curve being the projection of a one-parameter subgroup in G that connects the initial point $x_0 \in G/H$ with the final point $x_k \in G/H$ is strongly desirable. If $x_0 = \mathrm{pr}(e)$ holds and a $\xi \in \mathfrak{g}$ with the property $x_k = \mathrm{pr}(\exp(\xi))$ is known, a closed form expression for the desired rolling of $\mathfrak{m}$ over G/H with respect to $\nabla^{\mathrm{can}2}$ is known. This is the next lemma, which is a reformulation of [41, Prop. 5.25].

Lemma 3.8 *Let $\xi \in \mathfrak{g}$. Then, the curve $(v, g, S)\colon I \to \mathfrak{m} \times G \times \mathrm{GL}(\mathfrak{m})$ given for $t \in I$ by*

$$\begin{aligned} v(t) &= \int_0^t \mathrm{Ad}_{\exp(s\xi_{\mathfrak{h}})}(\xi_{\mathfrak{m}})\, \mathrm{d}s, \\ g(t) &= \exp(t\xi)\exp(-t\xi_{\mathfrak{h}}), \\ S(t) &= \mathrm{id}_{\mathfrak{m}} \end{aligned} \tag{3.22}$$

defines the intrinsic rolling $(v(t), \gamma(t), A(t))$ of $\mathfrak{m}$ over G/H with respect to $\nabla^{\mathrm{can}2}$, where

$$\gamma\colon I \to G/H, \quad t \mapsto \gamma(t) = (\mathrm{pr} \circ g)(t) = \mathrm{pr}(\exp(t\xi)) \tag{3.23}$$

and

$$A(t)\colon T_{v(t)}\mathfrak{m} \cong \mathfrak{m} \ni Z \mapsto \big(T_{g(t)}\,\mathrm{pr} \circ T_e\ell_{g(t)}\big) Z \in T_{\gamma(t)}(G/H). \tag{3.24}$$

A more complicated expression is also available for a rolling with respect to $\nabla^{\mathrm{can}1}$ whose development curve is the projection of a one-parameter subgroup. This is the next lemma which is a reformulation of [41, Prop. 5.22].

Lemma 3.9 *Let $\xi \in \mathfrak{g}$. Then the curve $(v, g, S)\colon I \to \mathfrak{m} \times G \times \mathrm{GL}(\mathfrak{m})$ given for $t \in I$ by*

$$\begin{aligned} v(t) &= \int_0^t \exp\big(s\,\mathrm{pr}_{\mathfrak{m}} \circ \mathrm{ad}_{\xi_{\mathfrak{h}} + \frac{1}{2}\xi_{\mathfrak{m}}}\big)(\xi_{\mathfrak{m}})\, \mathrm{d}s \\ g(t) &= \exp(t\xi)\exp(-t\xi_{\mathfrak{h}}) \\ S(t) &= \mathrm{Ad}_{\exp(t\xi_{\mathfrak{h}})} \circ \exp\big(-t\,\mathrm{pr}_{\mathfrak{m}} \circ \mathrm{ad}_{\xi_{\mathfrak{h}} + \frac{1}{2}\xi_{\mathfrak{m}}}\big) \end{aligned} \tag{3.25}$$

defines the intrinsic rolling $(v(t), \gamma(t), A(t))$ of $\mathfrak{m}$ over G/H with respect to $\nabla^{\mathrm{can}1}$, where

$$\gamma\colon I \to G/H, \quad t \mapsto \gamma(t) = (\mathrm{pr} \circ g)(t) = \mathrm{pr}(\exp(t\xi)) \tag{3.26}$$

and

$$A(t)\colon T_{v(t)}\mathfrak{m} \cong \mathfrak{m} \ni Z \mapsto \big(T_{g(t)}\,\mathrm{pr} \circ T_e\ell_{g(t)} \circ S(t)\big) Z \in T_{\gamma(t)}(G/H). \tag{3.27}$$

Remark 3.10 The expressions obtained in Lemma 3.8 and Lemma 3.9, respectively, can be simplified if $\gamma\colon I \to G/H$ is the projection of a *horizontal* one-parameter subgroup in G, i.e. $\xi \in \mathfrak{m}$ holds. Using $\xi_{\mathfrak{h}} = 0$ and $\xi_{\mathfrak{m}} = \xi$ for $\xi \in \mathfrak{m}$, the curve $(v, g, S)\colon I \to \mathfrak{m} \times G \times \mathrm{GL}(\mathfrak{m})$ from Lemma 3.9 becomes for $t \in I$

$$\begin{aligned} v(t) &= t\xi \\ g(t) &= \exp(t\xi) \\ S(t) &= \exp\big(-\tfrac{1}{2}t\,\mathrm{pr}_{\mathfrak{m}} \circ \mathrm{ad}_{\xi}\big), \end{aligned} \tag{3.28}$$

see also [41, Cor. 5.24], while the expression from Lemma 3.8 simplifies to

$$\begin{aligned} v(t) &= t\xi, \\ g(t) &= \exp(t\xi) \\ S(t) &= \mathrm{id}_{\mathfrak{m}}\,. \end{aligned} \tag{3.29}$$

4 Interpolation via Intrinsic Rolling and Unwrapping

In this section, we come to the announced method to solve Problem 2.1 on a reductive homogeneous space. It is inspired by the rolling and unwrapping technique from [29, Sec. 5], see also [26, Sec. 21.3.3]. Although the algorithm proposed here is in some sense closely related to that technique, there are also some differences which will be clarified in Remark 4.6 below. We start by introducing some notations and assumptions.

Notation 4.1 Let G/H be a reductive homogeneous space with fixed reductive decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$ and let $\Phi\colon G \times N \to N$ be a smooth left-action of the Lie group G on a manifold N. We write for fixed $x \in N$

$$\Phi(\cdot, x)\colon G \to N, \quad g \mapsto \Phi(g, x) \tag{4.1}$$

and for fixed $g \in G$, we have the diffeomorphism

$$\Phi(g, \cdot) = \Phi_g\colon N \to N, \quad x \mapsto \Phi(g, x) = \Phi_g(x). \tag{4.2}$$

Moreover, let $\iota\colon G/H \to N$ be a G-equivariant embedding, i.e.

$$\Phi_g \circ \iota = \iota \circ \tau_g \tag{4.3}$$

holds for all $g \in G$, where $\tau_g\colon G/H \to G/H$ is defined in (3.12). We denote the image of ι by

$$M = \iota(G/H) \tag{4.4}$$

and set

$$x_0 = \iota(\mathrm{pr}(e)). \tag{4.5}$$

By viewing $\iota\colon G/H \to N$ as a diffeomorphism onto its image $M = \iota(G/H)$, one obtains

$$\iota \circ \tau_g = \Phi_g \circ \iota \iff \Phi_g = \iota \circ \tau_g \circ \iota^{-1} \tag{4.6}$$

for all $g \in G$.

The manifold $M = \iota(G/H)$ from Notation 4.1 is an embedded orbit of the G-action $\Phi\colon G \times N \to N$ through $x_0 \in N$ by [36, Thm. 6.4]. Clearly, the stabilizer subgroup of $x_0 \in M$ is H, i.e. $\mathrm{Stab}(x_0) = H$.

Moreover, the existence of a G-equivariant embedding, as in Notation 4.1, can be always assumed. Indeed, since τ defined in (3.11) is a transitive left action of G on G/H, it is always possible to choose $N = G/H$ and $\iota = \mathrm{id}_{G/H}\colon G/H \to G/H$.

Remark 4.2 In the sequel, we formulate Algorithm 1 to solve Problem 2.1 on a reductive homogeneous space G/H with a fixed reductive decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$. More precisely, by using Notation 4.1, we identify G/H with $M = \iota(G/H)$ via the embedding $\iota\colon G/H \to N$ viewed as diffeomorphism onto its image. Moreover, the linear isomorphism

$$T_e(\iota \circ \mathrm{pr})\big|_{\mathfrak{m}}\colon \mathfrak{m} \to T_{x_0}M, \quad v \mapsto \widehat{v} = T_e(\iota \circ \mathrm{pr})v \tag{4.7}$$

is used to identify $T_{x_0}M$ with $\mathfrak{m}$. By this approach, Algorithm 1 can be applied directly to a specific $M = \iota(G/H) \subseteq N$. In some examples, see Subsection 5.2 and Subsection 6.4 below, a method for solving Problem 2.1 on specific manifolds $M \subseteq N$ can be obtained that only uses the coordinates of the embedding space N, entirely suppressing the coset representation of G/H.

In order to provide a geometric interpretation of Algorithm 1, Step 2 below, we first explain in Lemma 4.3 below how a rolling $\big(\widehat{v}(t), \widehat{\gamma}(t), \widehat{A}(t)\big)$ of $T_{x_0}M$ over $M = \iota(G/H)$ can be constructed by means of a rolling $\big(v(t), \gamma(t), A(t)\big)$ of $\mathfrak{m}$ over G/H and the embedding $\iota\colon G/H \to N$.

Before that we recall the notion of pull-back covariant derivatives from [34, Lem. 4.37]. Let $\phi\colon M \to \widehat{M}$ be a diffeomorphism between smooth manifolds M and $\widehat{M}$ and let $\widehat{\nabla}$ be a covariant derivative on $\widehat{M}$. Then the pull-back $\nabla = \phi^*\widehat{\nabla}$ of $\widehat{\nabla}$ by ϕ defined by

$$\nabla_X Y \big(\phi^*\widehat{\nabla}\big)_X Y = (\phi^{-1})_*\big(\widehat{\nabla}_{\phi_* X}\phi_* Y\big), \quad X, Y \in \Gamma^\infty(TM) \tag{4.8}$$

is a covariant derivative on M, where $\phi_* X = T\phi \circ X \circ \phi^{-1} \in \Gamma^\infty\big(T\widehat{M}\big)$ is the well-known push-forward of $X \in \Gamma^\infty(TM)$, see e.g. [33, Chap. 8, p. 183]. Next, let $Z\colon I \to TM$ be a vector field along the curve $\gamma\colon I \to M$. Then, the covariant derivatives along curves γ and $\phi \circ \gamma$ defined by ∇ and $\widehat{\nabla}$ respectively, fulfill

$$T\phi \circ \nabla_{\dot{\gamma}(t)} Z\big|_t = \widehat{\nabla}_{\frac{\mathrm{d}}{\mathrm{d}t}(\phi\circ\gamma)(t)}(T\phi \circ Z)\big|_t, \tag{4.9}$$

for all $t \in I$, according to [34, Prop. 4.38 (a)].

Lemma 4.3 *Let $\big(v(t), \gamma(t), A(t)\big)$ be a rolling of $\big(\mathfrak{m}, \nabla^{\mathfrak{m}}\big)$ over $\big(G/H, \nabla^\alpha\big)$. Using (4.8) and Notation 4.1, define the covariant derivatives $\nabla^{T_{x_0}M}$ on $T_{x_0}M$ and $\nabla^{M,\alpha,\iota}$ on M by*

$$\nabla^{T_{x_0}M} = \Big(\big(T_e(\iota \circ \mathrm{pr})\big)^{-1}\Big)^* \nabla^{\mathfrak{m}} \quad \text{and} \quad \nabla^{M,\alpha,\iota} = \big(\iota^{-1}\big)^*\nabla^\alpha, \tag{4.10}$$

respectively. Also define the curves

$$\widehat{v}\colon I \ni t \mapsto \widehat{v}(t) = T_e(\iota \circ \mathrm{pr})v(t) \in T_{x_0}M \quad \text{and} \quad \widehat{\gamma}\colon I \ni t \mapsto \widehat{\gamma}(t) = (\iota \circ \gamma)(t) \in M \tag{4.11}$$

and let $\widehat{A}(t)\colon T_{\widehat{v}(t)}(T_{x_0}M) \cong T_{x_0}M \to T_{\widehat{\gamma}(t)}M$ be given by

$$\widehat{A}(t)Z = T_{\mathrm{pr}(g(t))}\iota \circ A(t) \circ \big(T_e(\iota \circ \mathrm{pr}\big|_{\mathfrak{m}})\big)^{-1} Z, \quad Z \in T_{x_0}M. \tag{4.12}$$

Then, the triple $\big(\widehat{v}(t), \widehat{\gamma}(t), \widehat{A}(t)\big)$ is a rolling of $\big(T_{x_0}M, \nabla^{T_{x_0}M}\big)$ over $\big(M, \nabla^{M,\alpha,\iota}\big)$ associated with the rolling $(v(t), \gamma(t), A(t))$.

Proof: Obviously, by the definition in (4.12), $\widehat{A}(t)\colon T_{x_0}M \cong T_{\widehat{v}(t)}(T_{x_0}M) \to T_{\widehat{\gamma}}M$ is a linear isomorphism. The no-slip condition for $(v(t), \gamma(t), A(t))$ yields for $t \in I$

$$\widehat{A}(t)\dot{\widehat{v}}(t) = \Big(T_{\mathrm{pr}(g(t))}\iota \circ A(t) \circ \big(T_e(\iota \circ \mathrm{pr}\big|_{\mathfrak{m}})\big)^{-1}\Big) T_e(\iota \circ \mathrm{pr})\dot{v}(t) = T_{\gamma(t)}\iota\dot{\gamma}(t) = \dot{\widehat{\gamma}}(t) \tag{4.13}$$

showing that $\big(\widehat{v}(t), \widehat{\gamma}(t), \widehat{A}(t)\big)$ satisfies the no-slip condition. Moreover, by exploiting (4.9) and using that $\big(v(t), \gamma(t), A(t)\big)$ satisfies the no-twist condition, it is straightforward to verify that $\big(\widehat{v}(t), \widehat{\gamma}(t), \widehat{A}(t)\big)$ fulfills the no-twist condition, as well. □

Algorithm 1 Interpolation on reductive homogeneous spaces

Input: $x_0, \dots, x_k \in M$ with $x_0 = \iota(\mathrm{pr}(e))$, initial velocity $\xi_0 \in T_{x_0}M$, final velocity $\xi_k \in T_{x_k}M$, instances of time $0 = t_0 < \cdots < t_k = T$, $\mathrm{Ad}(H)$-invariant bilinear map $\alpha\colon \mathfrak{m} \times \mathfrak{m} \to \mathfrak{m}$ corresponding to ∇^α.

1. Determine a solution $(v, g, S)\colon [0,T] \to \mathfrak{m} \times G \times \mathrm{GL}(\mathfrak{m})$ of the ODE (3.20) from Proposition 3.7 with some suitable control curve $u\colon [0,T] \to \mathfrak{m}$ such that $\iota(\gamma(0)) = x_0$, $\iota(\gamma(T)) = x_k$, $v(0) = 0$, $g(0) = e$ holds. Denote the associated intrinsic rolling of $\mathfrak{m}$ over G/H with respect to ∇^α from Proposition 3.7 by $(v(t), \gamma(t), A(t))$.

2. Define the curves $\widehat{v}\colon [0,T] \ni t \mapsto \widehat{v}(t) = T_e(\iota \circ \mathrm{pr})v(t) \in T_{x_0}M$ and $\widehat{\gamma}\colon [0,T] \ni t \mapsto \widehat{\gamma}(t) = (\iota \circ \gamma)(t) \in M$. Moreover, define for $t \in [0,T]$ the linear isomorphisms

$$\widehat{A}(t) = T_{\mathrm{pr}(g(t))}\iota \circ A(t) \circ \big(T_e(\iota \circ \mathrm{pr}\big|_{\mathfrak{m}})\big)^{-1}\colon T_{\widehat{v}(t)}T_{x_0}M \cong T_{x_0}M \to T_{\widehat{\gamma}(t)}M \tag{4.14}$$

$$\widehat{S}(t) = T_e(\iota \circ \mathrm{pr}) \circ S(t) \circ \big(T_e(\iota \circ \mathrm{pr})\big|_{\mathfrak{m}}\big)^{-1}\colon T_{x_0}M \to T_{x_0}M. \tag{4.15}$$

3. Unwrap the boundary data from M to $T_{x_0}M$ by defining

$$\begin{aligned} q_0 &= \widehat{v}(0) = 0, & q_k &= \widehat{v}(T) \\ \eta_0 &= \big(\widehat{A}(0)\big)^{-1}\xi_0, & \eta_k &= \big(\widehat{A}(T)\big)^{-1}\xi_k \end{aligned} \tag{4.16}$$

4. Let $\phi\colon U \subseteq M \to T_{x_0}M$ be a local diffeomorphism, where U is open with $x_0 \in U$, which satisfies

$$\phi(x_0) = 0 \quad \text{and} \quad T_{x_0}\phi = \mathrm{id}_{T_{x_0}M}, \tag{4.17}$$

where the identification $T_0(T_{x_0}M) \cong T_{x_0}M$ is used. Use ϕ to unwrap the remaining points $x_1, \dots, x_{k-1}$ by setting

$$q_i = \widehat{S}(t_i)^{-1}\big(\phi\big(\Phi_{g(t_i)^{-1}}(x_i)\big)\big) + \widehat{v}(t_i) \tag{4.18}$$

for $i \in \{1, \dots, k-1\}$.

5. Compute a $\mathscr{C}^2$-curve $y\colon [0,T] \to T_{x_0}M$ with $y(t_i) = q_i$ for all $i \in \{0, \dots, k\}$ and $\dot{y}(0) = \eta_0$ as well as $\dot{y}(T) = \eta_k$.

6. Define the curve $\beta\colon [0,T] \to M$ by wrapping $y\colon I \to T_{x_0}M$ back to the manifold by setting

$$\beta(t) = \Phi_{g(t)}\big(\phi^{-1}\big(\widehat{S}(t)\big(y(t) - \widehat{v}(t)\big)\big)\big) \tag{4.19}$$

for $t \in [0,T]$.

Output: The curve $\beta\colon [0,T] \to M$.

Notation 4.4 We say that Algorithm 1 is applied to M equipped with ∇^α to indicate that Algorithm 1 is applied to $M = \iota(G/H)$, where the $\mathrm{Ad}(H)$-invariant bilinear map $\alpha\colon \mathfrak{m} \times \mathfrak{m} \to \mathfrak{m}$ is chosen as the one corresponding to ∇^α by (3.16).

Remark 4.5 Let $U = \mathrm{dom}(\phi) \subset M$ and let $\mathrm{dom}(\phi^{-1}) \subseteq T_{x_0}M$ denote the domains of ϕ and ϕ^{-1}, respectively. Then, Algorithm 1 yields a well-defined curve $\beta\colon [0,T] \to M$ if the following assumptions are satisfied:

1. The points $\Phi_{g(t_i)^{-1}}(x_i)$ are in the domain $U = \mathrm{dom}(\phi)$ of ϕ for all $i \in \{1, \dots, k-1\}$.

2. One has $\widehat{S}(t)\big(y(t) - \widehat{v}(t)\big) \in \mathrm{dom}(\phi^{-1})$ for all $t \in [0, T]$.

Before we show that Algorithm 1 solves Problem 2.1, some remarks are in order.

Remark 4.6

1. Algorithm 1 yields a well-defined $\mathscr{C}^2$-curve solving Problem 2.1 on M provided that the assumptions from Notation 4.1 and Remark 4.5 are fulfilled. This is shown in Theorem 4.7 below.

2. Step 2 of Algorithm 1 can be interpreted as follows. By Lemma 4.3, the curves $\widehat{v}\colon [0, T] \to T_{x_0}M$ and $\widehat{\gamma}\colon [0, T] \to M$ as well as the curve of linear isomorphisms $\widehat{A}(t)\colon T_{\widehat{v}(t)}T_{x_0}M \cong T_{x_0}M \to T_{\widehat{\gamma}(t)}M$ form the triple $\big(\widehat{v}(t), \widehat{\gamma}(t), \widehat{A}(t)\big)$, which is exactly the rolling of $T_{x_0}M$ over M associated with the rolling $(v(t), \gamma(t), A(t))$ defined by Algorithm 1, Step 1.

3. Let G/H be endowed with an invariant Riemannian metric and assume that $\iota\colon G/H \to V$ is an isometric embedding into the Euclidean vector space V. Then, in general, the rolling and unwrapping technique from [26, 29] differs from Algorithm 1 applied to $M = \iota(G/H) \subseteq V$ with respect to ∇^α, even if ∇^α is chosen such that $(\iota^{-1})^*\nabla^\alpha = \nabla^{\mathrm{LC}}$ is the Levi-Civita covariant derivative on M. The method from [26, 29] depends on the normal part of the extrinsic rolling of M over an affine tangent space of M (with respect to ∇^{LC}) while Algorithm 1 only depends on an intrinsic rolling by construction. Nevertheless, for $\mathrm{SO}(n)$ the rolling and unwrapping technique from [26,29] yields a result that is similar to the output of Algorithm 1 applied to $\mathrm{SO}(n)$ equipped with ∇^{can1} as will be seen in Remark 5.9 below. Further discussions on the relation of the rolling and unwrapping technique from [26, 29] to Algorithm 1 are out of the scope of this text.

4. The interpolation curve obtained by Algorithm 1 depends on the choice of a local diffeomorphism $\phi\colon U \subseteq M \to T_{x_0}M$ satisfying some properties. A similar choice is required by the rolling and unwrapping technique from [26,29]. Moreover, in general, the result of Algorithm 1 depends on the choice of the rolling along a curve connecting the initial and final points x_0 and x_k, respectively. Similarly, one has to choose an extrinsic rolling along a curve connecting the initial and final points for applying the interpolation method from [26, 29].

5. In addition, Algorithm 1 depends on the choices of the $\mathrm{Ad}(H)$-invariant bilinear map $\alpha\colon \mathfrak{m} \times \mathfrak{m} \to \mathfrak{m}$. We illustrate this dependency in Section 5 on the example of matrix Lie groups.

6. Finally, we point out that Step 5 of Algorithm 1 can be solved efficiently in closed form by means of a suitable Euclidean cubic spline. This is also pointed out in [26,29]. In this context, we refer to [45, Sec. 2.4], where interpolation (in the vector space $\mathbb{R}$) by cubic splines is discussed.

Theorem 4.7 *Let $\beta\colon [0, T] \to M$ be defined by* (4.19) *in Algorithm 1, such that the assumptions of Notation 4.1 and Remark 4.5 are satisfied. Then $\beta\colon [0, T] \to M$ is a well-defined $\mathscr{C}^2$-curve that solves Problem 2.1 associated with the data $0 = t_0 < \cdots < t_k = T$, $x_0, \ldots, x_k \in M$, $\xi_0 \in T_{x_0}M$ and $\xi_k \in T_{x_k}M$.*

Proof: Obviously, the curve $\beta\colon [0,T] \to M$ is a $\mathscr{C}^2$-curve since it is a composition of $\mathscr{C}^2$-maps. We now compute for each $i \in \{1,\dots,k-1\}$

$$\begin{aligned}
\beta(t_i) &= \Phi_{g(t_i)}\Big(\phi^{-1}\Big(\widehat{S}(t_i)\big(y(t_i)-\widehat{v}(t_i)\big)\Big)\Big)\\
&= \Phi_{g(t_i)}\Big(\phi^{-1}\Big(\widehat{S}(t_i)\big(q_i-\widehat{v}(t_i)\big)\Big)\Big)\\
&\overset{(4.18)}{=} \Phi_{g(t_i)}\Big(\phi^{-1}\Big(\widehat{S}(t_i)\Big(\Big(\widehat{S}(t_i)^{-1}\big(\phi\big(\Phi_{g(t_i)^{-1}}(x_i)\big)\big)\Big)+\widehat{v}(t_i)\Big)-\widehat{v}(t_i)\Big)\Big)\\
&= \Phi_{g(t_i)}\Big(\phi^{-1}\Big(\widehat{S}(t_i)\widehat{S}(t_i)^{-1}\phi\big(\Phi_{g(t_i)^{-1}}(x_i)\big)\Big)\Big)\\
&= x_i.
\end{aligned} \tag{4.20}$$

Moreover, we obtain for $t=0$ due to $\widehat{v}(0)=0=q_0=y(0)$ and $g(0)=e$ by construction

$$\begin{aligned}
\beta(0) &= \Phi_{g(0)}\Big(\phi^{-1}\Big(\widehat{S}(0)\big(y(0)-\widehat{v}(0)\big)\Big)\Big)\\
&= \Phi_e\big(\phi^{-1}(0)\big)\\
&= \Phi_e(x_0)\\
&= x_0.
\end{aligned} \tag{4.21}$$

For $t=T$, one has $y(T)=q_k=\widehat{v}(T)$. This yields

$$\begin{aligned}
\beta(T) &= \Phi_{g(T)}\Big(\phi^{-1}\Big(\widehat{S}(T)\big(y(T)-\widehat{v}(T)\big)\Big)\Big)\\
&= \Phi_{g(T)}\Big(\phi^{-1}\Big(\widehat{S}(T)(0)\Big)\Big)\\
&= \Phi_{g(T)}\big(\phi^{-1}(0)\big)\\
&= \Phi_{g(T)}(x_0)\\
&= x_k,
\end{aligned} \tag{4.22}$$

where for the last equality we exploited $x_k = \widehat{\gamma}(T) = (\iota\circ \mathrm{pr})(g(T)) = (\iota \circ \mathrm{pr}\circ \ell_{g(T)})(e) = (\iota\circ\tau_{g(T)}\circ \mathrm{pr})(e) = \Phi_{g(T)}(x_0)$, where the last equality holds due to (4.3).

It remains to show that $\beta\colon [0,T]\to M$ fulfills the velocity boundary conditions. To this end, we need the following preparations. We note that

$$\begin{aligned}
&T_{\mathrm{pr}(e)}\tau_g \circ T_e \mathrm{pr}\big|_{\mathfrak{m}} = T_e(\tau_g\circ \mathrm{pr}\big|_{\mathfrak{m}}) = T_e(\mathrm{pr}\circ \ell_g)\big|_{\mathfrak{m}} = T_g\,\mathrm{pr}\circ T_e\ell_g\big|_{\mathfrak{m}}\\
\iff &T_{\mathrm{pr}(e)}\tau_g = T_g\,\mathrm{pr}\circ T_e\ell_g\circ\big(T_e\,\mathrm{pr}\big|_{\mathfrak{m}}\big)^{-1}
\end{aligned} \tag{4.23}$$

is fulfilled for all $g\in G$. In addition, we compute for $t\in[0,T]$

$$\begin{aligned}
\dot{\beta}(t) &= \tfrac{\mathrm{d}}{\mathrm{d}t}\left(\Phi_{g(t)}\Big(\phi^{-1}\Big(\widehat{S}(t)\big(y(t)-\widehat{v}(t)\big)\Big)\Big)\right)\\
&= T_{g(t)}\Phi\Big(\cdot,\phi^{-1}\Big(\widehat{S}(t)\big(y(t)-\widehat{v}(t)\big)\Big)\Big)\dot{g}(t)\\
&\quad + T_{\phi^{-1}\left(\widehat{S}(t)\left(y(t)-\widehat{v}(t)\right)\right)}\Phi\big(g(t),\cdot\big)\tfrac{\mathrm{d}}{\mathrm{d}t}\big(\phi^{-1}\big(\widehat{S}(t)\big(y(t)-\widehat{v}(t)\big)\big)\big)\\
&= T_{g(t)}\Phi\Big(\cdot,\phi^{-1}\Big(\widehat{S}(t)\big(y(t)-\widehat{v}(t)\big)\Big)\Big)\dot{g}(t)\\
&\quad + T_{\phi^{-1}\left(\widehat{S}(t)\left(y(t)-\widehat{v}(t)\right)\right)}\Phi\big(g(t),\cdot\big)\\
&\qquad \circ T_{\widehat{S}(t)\left(y(t)-\widehat{v}(t)\right)}\phi^{-1}\Big(\dot{\widehat{S}}(t)\big(y(t)-\widehat{v}(t)\big)+\widehat{S}(t)\big(\dot{y}(t)-\dot{\widehat{v}}(t)\big)\Big).
\end{aligned} \tag{4.24}$$

Recall that $\phi(x_0) = 0$ and $T_{x_0}\phi = \mathrm{id}_{T_{x_0}M}$ holds by assumptions, see Algorithm 1, Step 4. This implies $T_0\phi^{-1} = \big(T_{\phi^{-1}(0)}\phi\big)^{-1} = \big(\mathrm{id}_{T_{x_0}M}\big)^{-1} = \mathrm{id}_{T_{x_0}M}$. We continue with computing $\dot\beta(0)$ and $\dot\beta(T)$. Here we note that $y(0) = q_0 = \widehat{v}(0)$ and $y(T) = q_k = \widehat{v}(T)$ holds by Algorithm 1, Step 3 combined with Algorithm 1, Step 5. Hence the computation of $\dot\beta(0)$ and $\dot\beta(T)$ can be treated simultaneously by evaluating (4.24) at a fixed instance of time $t_* \in [0,T]$ with the property $y(t_*) = \widehat{v}(t_*)$. Using the definition $\widehat{v}(t) = T_e(\iota\circ\mathrm{pr})\big|_{\mathfrak{m}} v(t)$ for all $t \in [0,T]$ which implies

$$\dot{\widehat{v}}(t) = T_e(\iota\circ\mathrm{pr})\big|_{\mathfrak{m}}\dot v(t), \quad t \in [0,T] \tag{4.25}$$

and exploiting that $g\colon I \to G$ fulfills the ODE

$$\dot g(t) = \big(T_e\ell_{g(t)} \circ S(t)\big)\dot v(t), \quad t \in [0,T] \tag{4.26}$$

according to Proposition 3.7, one obtains by (4.24) due to the assumption $y(t_*) = \widehat{v}(t_*)$

$$\begin{aligned}
\dot\beta(t_*) &= T_{g(t_*)}\Phi\Big(\cdot, \phi^{-1}(0)\Big)\dot g(t_*) + T_{\phi^{-1}(0)}\Phi\big(g(t_*),\cdot\big)\circ T_0\phi^{-1}\Big(\widehat{S}(t_*)\big(\dot y(t_*) - \dot{\widehat{v}}(t_*)\big)\Big) \\
&= \tfrac{\mathrm{d}}{\mathrm{d}t}\Phi\big(g(t),x_0\big)\big|_{t=t_*} + T_{x_0}\Phi\big(g(t_*),\cdot\big)\circ\mathrm{id}_{T_{x_0}M}\Big(\widehat{S}(t_*)\big(\dot y(t_*) - \dot{\widehat{v}}(t_*)\big)\Big) \\
&= \tfrac{\mathrm{d}}{\mathrm{d}t}\Phi\big(g(t),\iota(\mathrm{pr}(e))\big)\big|_{t=t_*} + T_{\iota(\mathrm{pr}(e))}\Phi\big(g(t_*),\cdot\big)\Big(\widehat{S}(t_*)\big(\dot y(t_*) - \dot{\widehat{v}}(t_*)\big)\Big) \\
&\overset{(4.6)}{=} \tfrac{\mathrm{d}}{\mathrm{d}t}\iota\Big(\tau_{g(t)}\big(\mathrm{pr}(e)\big)\Big)\big|_{t=t_*} + T_{\iota(\mathrm{pr}(e))}\Big(\iota\circ\tau_{g(t_*)}\circ\iota^{-1}\Big)\Big(\widehat{S}(t_*)\big(\dot y(t_*) - \dot{\widehat{v}}(t_*)\big)\Big) \\
&\overset{(4.15)}{=} \tfrac{\mathrm{d}}{\mathrm{d}t}\iota\big(\mathrm{pr}(g(t))\big)\big|_{t=t_*} \\
&\quad + \Big(T_{\mathrm{pr}(g(t_*))}\iota\circ T_{\mathrm{pr}(e)}\tau_{g(t_*)}\circ T_{\iota(\mathrm{pr}(e))}\iota^{-1}\Big) \\
&\quad\circ\Big(T_e(\iota\circ\mathrm{pr})\circ S(t_*)\circ\big(T_e(\iota\circ\mathrm{pr})\big|_{\mathfrak{m}}\big)^{-1}\Big)\big(\dot y(t_*) - \dot{\widehat{v}}(t_*)\big) \\
&= \tfrac{\mathrm{d}}{\mathrm{d}t}\iota\big(\mathrm{pr}(g(t))\big)\big|_{t=t_*} \\
&\quad + \Big(T_{\mathrm{pr}(g(t_*))}\iota\circ T_{\mathrm{pr}(e)}\tau_{g(t_*)}\circ\big(T_{\mathrm{pr}(e)}\iota\big)^{-1}\Big) \\
&\quad\circ\Big(T_{\mathrm{pr}(e)}\iota\circ T_e\,\mathrm{pr}\circ S(t_*)\circ\big(T_e(\iota\circ\mathrm{pr})\big|_{\mathfrak{m}}\big)^{-1}\Big)\big(\dot y(t_*) - \dot{\widehat{v}}(t_*)\big) \\
&= \big(T_{\mathrm{pr}(g(t_*))}\iota\circ T_{g(t_*)}\,\mathrm{pr}\big)\dot g(t_*) \\
&\quad + \Big(T_{\mathrm{pr}(g(t_*))}\iota\circ\big(T_{\mathrm{pr}(e)}\tau_{g(t_*)}\circ T_e\,\mathrm{pr}\big)\circ S(t_*) \\
&\quad\circ\big(T_e(\iota\circ\mathrm{pr})\big|_{\mathfrak{m}}\big)^{-1}\Big)\big(\dot y(t_*) - \dot{\widehat{v}}(t_*)\big) \\
&\overset{(4.26),(4.25),(4.23)}{=} \big(T_{\mathrm{pr}(g(t_*))}\iota\circ T_{g(t_*)}\,\mathrm{pr}\circ T_e\ell_{g(t_*)}\circ S(t_*)\big)\dot v(t_*) \\
&\quad + \big(T_{\mathrm{pr}(g(t_*))}\iota\circ\big(T_{g(t_*)}\,\mathrm{pr}\circ T_e\ell_{g(t_*)}\big)\circ S(t_*)\circ\big(T_e(\iota\circ\mathrm{pr})\big|_{\mathfrak{m}}\big)^{-1}\dot y(t_*) \\
&\quad - \big(T_{\mathrm{pr}(g(t_*))}\iota\circ\big(T_{g(t_*)}\,\mathrm{pr}\circ T_e\ell_{g(t_*)}\big)\circ S(t_*)\circ\big(T_e(\iota\circ\mathrm{pr})\big|_{\mathfrak{m}}\big)^{-1} \\
&\quad\circ\big(T_e(\iota\circ\mathrm{pr}\big|_{\mathfrak{m}})\big)\dot v(t_*) \\
&= T_{\mathrm{pr}(g(t_*))}\iota\circ\Big(T_{g(t_*)}\,\mathrm{pr}\circ T_e\ell_{g(t_*)}\circ S(t_*)\Big)\circ\big(T_e(\iota\circ\mathrm{pr})\big|_{\mathfrak{m}}\big)^{-1}\dot y(t_*) \\
&\overset{(4.14)}{=} \widehat{A}(t_*)\dot y(t_*).
\end{aligned} \tag{4.27}$$

Since $y(0) = 0 = \widehat{v}(0)$, (4.27) holds for $t_* = 0$. So, plugging $t_* = 0$ into (4.27) yields

$$\dot\beta(0) = \widehat{A}(0)\dot y(0) = \widehat{A}(0)\eta_0 = \widehat{A}(0)\big(\big(\widehat{A}(0)\big)^{-1}\xi_0\big) = \xi_0. \tag{4.28}$$

Similarly, since $y(T) = q_k = \widehat{v}(T)$, plugging $t_* = T$ into (4.27) gives

$$\dot{\beta}(T) = \widehat{A}(T)\dot{y}(T) = \widehat{A}(T)\eta_k = \widehat{A}(T)\big(\big(\widehat{A}(T)\big)^{-1}\xi_k\big) = \xi_k. \tag{4.29}$$

This yields the desired result. □

Remark 4.8 Step 4 of Algorithm 1, see also the assumption in Remark 4.5, 1, can be weakened as follows: Unwrap the remaining points by setting for $i \in \{1, \ldots, k-1\}$

$$q_i = \widehat{S}(t_i)^{-1}(\widetilde{\eta}_i) + \widehat{v}(t_i), \tag{4.30}$$

where $\widetilde{\eta}_i \in T_{x_0}M$ denotes some tangent vector fulfilling

$$\phi^{-1}(\widetilde{\eta}_i) = \Phi_{g(t_i)^{-1}}(x_i). \tag{4.31}$$

Then, by the proof of Theorem 4.7, see in particular the calculation in (4.20), Algorithm 1 yields still a $\mathscr{C}^2$-curve solving the interpolation problem.

Remark 4.9 The proof of Theorem 4.7 reveals that a slight modification of Algorithm 1 can be used to obtain a $\mathscr{C}^k$-curve which solves Problem 2.1, where $k \geq 1$: The $\mathscr{C}^2$-curve $y\colon [0,T] \to T_{x_0}M$ in Step 5 of Algorithm 1 has to be replace by a suitable $\mathscr{C}^k$-curve.

5 Interpolation on Matrix Lie Groups

In this section, we show that Algorithm 1 can be used to solve Problem 2.1 on matrix Lie groups under some mild assumptions. Moreover, by considering the canonical invariant covariant derivative of first and second kind, we illustrate that the result of Algorithm 1 depends indeed on the choice of ∇^α.

5.1 Rolling Lie Groups Intrinsically

We consider a matrix Lie group $G \subseteq \mathrm{GL}(n)$, i.e. a closed subgroup G of $\mathrm{GL}(n)$. Clearly, one may identify G with the reductive homogeneous space $G \cong G/H$, where $H = \{e\}$ and the reductive decomposition is given by $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$, where $\mathfrak{h} = \{0\}$ and $\mathfrak{m} = \mathfrak{g}$. Then the canonical projection $\mathrm{pr}\colon G \to G/\{e\} \cong G$ becomes the identity map $\mathrm{id}_{\mathfrak{g}}$ by identifying $G/\{e\} \cong G$. In this section, the notations e and I_n are used both for the identity of a matrix Lie group $G \subseteq \mathrm{GL}(n)$ interchangeable.

Recall that G acts on itself by left-translation

$$\ell\colon G \times G \to G, \quad (g,h) \mapsto \ell_g(h) = gh. \tag{5.1}$$

Moreover, G as a matrix Lie group, acts on $\mathbb{R}^{n\times n}$ be its defining representation, namely by matrix multiplications from the left denoted by

$$\Phi\colon G \times \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}, \quad (g,X) \mapsto gX. \tag{5.2}$$

Next, let $g_0 \in G$ be fixed. We define a G-equivariant embedding fulfilling the requirements of Notation 4.1, namely

$$\iota\colon G \to \mathbb{R}^{n\times n}, \quad g \mapsto \iota(g) = gg_0 = r_{g_0}(g). \tag{5.3}$$

In fact, ι has the desired properties since in this case $\tau_g = \ell_g$ and

$$(\Phi_g \circ \iota)(h) = g\iota(h) = ghg_0 = \ell_g(h)g_0 = (\iota \circ \ell_g)(h) = (\iota \circ \tau_g)(h) \tag{5.4}$$

holds for all $g, h \in G$ and, moreover, $\iota(e) = g_0$ is satisfied.

5.1.1 Intrinsic Rollings with respect to ∇^{can1}

In order to apply Algorithm 1 to a matrix Lie group $G \subseteq \mathrm{GL}(n)$ endowed with ∇^{can1}, we recall in the next proposition a result on rolling Lie groups intrinsically with respect to ∇^{can1} from [41, Prop. 6.2], followed by a simplification for matrix Lie groups.

Proposition 5.1 *Let G be a Lie group. Moreover, let $u\colon I \to \mathfrak{g}$ be a given control curve and define $k, W\colon I \to G$ by the initial value problems*

$$\dot{k}(t) = \tfrac{1}{2} T_e \ell_{k(t)} u(t), \quad k(0) = e \quad \text{and} \quad \dot{W}(t) = -\tfrac{1}{2} T_e \ell_{W(t)} u(t), \quad W(0) = e, \tag{5.5}$$

respectively. Then, the curve $(v, g, S)\colon I \to \mathfrak{g} \times G \times \mathrm{GL}(\mathfrak{g})$ given by

$$\begin{aligned} v(t) &= \int_0^t u(s)\,\mathrm{d}s, \\ g(t) &= k(t) W(t)^{-1}, \\ S(t) &= \mathrm{Ad}_{W(t)}, \end{aligned} \tag{5.6}$$

for $t \in I$ defines an intrinsic rolling of $\mathfrak{g}$ over G with respect to ∇^{can1}, i.e. the triple $(v(t), g(t), A(t))$, where

$$A(t)\colon T_{v(t)}\mathfrak{g} \cong \mathfrak{g} \to T_{g(t)} G, \quad Z \mapsto A(t)Z = \big(T_e \ell_{g(t)} \circ \mathrm{Ad}_{W(t)}\big) Z, \tag{5.7}$$

is an intrinsic rolling of $\mathfrak{g}$ over G.

If G is a matrix Lie group, then (5.7) can be simplified to

$$A(t)Z = g(t) W(t) Z W(t)^{-1} = k(t) Z W(t)^{-1} \tag{5.8}$$

for all $Z \in \mathfrak{g}$.

Remark 5.2 Let G be equipped with a bi-invariant (pseudo-)Riemannian metric. Then, according to [41, Cor. 6.4], Proposition 5.1 applied to G yields rollings of $\mathfrak{g}$ over G with respect to the Levi-Civita covariant derivative ∇^{LC}.

For the discussion that follows, explicit expressions for rollings of $\mathfrak{g}$ over G with respect to ∇^{can1} are of particular interest. Such expressions are available if the development curves are one-parameter subgroups.

Corollary 5.3 *Let G be a Lie group and let $u\colon I \to \mathfrak{g}$ be constant, i.e. $u(t) = \xi$ for all $t \in I$ and some $\xi \in \mathfrak{g}$. Then $(v, g, S)\colon I \to \mathfrak{g} \times G \times \mathrm{GL}(\mathfrak{g})$ defines the rolling $(v(t), g(t), A(t))$ of $\mathfrak{g}$ over G with respect to ∇^{can1}, where*

$$v(t) = t\xi, \quad g(t) = \exp(t\xi), \qquad \text{and} \quad S(t) = \mathrm{Ad}_{\exp\left(-\frac{t}{2}\xi\right)} \tag{5.9}$$

as well as

$$A(t)\colon T_{v(t)}\mathfrak{g} \cong \mathfrak{g} \to T_{g(t)} G, \quad Z \mapsto \Big(T_e \ell_{\exp(t\xi)} \circ \mathrm{Ad}_{\exp\left(-\frac{t}{2}\xi\right)}\Big) Z \tag{5.10}$$

holds for all $t \in I$.

If $G \subseteq \mathrm{GL}(n)$ is a matrix Lie group, the expression for $A(t)$ simplifies to

$$A(t)Z = \exp(\tfrac{t}{2}\xi) Z \exp(\tfrac{t}{2}\xi) \tag{5.11}$$

for all $Z \in \mathfrak{g} \subseteq \mathfrak{gl}(n)$.

Proof: Using the notation from Proposition 5.1, we set $u(t) = \xi$ for all $t \in I$. Then, $v(t) = \int_0^t u(s)\,\mathrm{d}s = t\xi$ holds. Clearly, $k(t) = \exp(\frac{t}{2}\xi)$ and $W(t) = \exp(-\frac{t}{2}\xi)$ are the solutions of the initial value problems

$$\dot{k}(t) = \tfrac{1}{2} T_e \ell_{k(t)} \xi, \quad k(0) = e, \qquad \text{and} \qquad \dot{W}(t) = -\tfrac{1}{2} T_e \ell_{W(t)} \xi, \quad W(0) = e, \tag{5.12}$$

respectively. The desired result is obtained by Proposition 5.1 due to

$$g(t) = k(t) W(t)^{-1} = \exp(\tfrac{t}{2}\xi) \big(\exp(-\tfrac{t}{2}\xi) \big)^{-1} = \exp(\tfrac{t}{2}\xi) \exp(\tfrac{t}{2}\xi) = \exp(t\xi) \tag{5.13}$$

and

$$A(t) Z = \big(T_e \ell_{\exp(t\xi)} \circ \mathrm{Ad}_{\exp\left(-\frac{t}{2}\xi\right)} \big) Z = \exp(\tfrac{t}{2}\xi) Z \exp(\tfrac{t}{2}\xi) \tag{5.14}$$

for all $Z \in \mathfrak{g}$ by a straightforward computation, where the last equality holds if G is a matrix Lie group. □

5.1.2 Intrinsic Rollings with respect to $\nabla^{\mathrm{can}2}$

We also apply Algorithm 1 to G equipped with $\nabla^{\mathrm{can}2}$ below. Therefore intrinsic rollings of $\mathfrak{g}$ over G with respect to $\nabla^{\mathrm{can}2}$ are considered in the next proposition.

Proposition 5.4 *Let G be a Lie group and let $u\colon I \to \mathfrak{g}$ be some control curve. Moreover, let $(v, g, S)\colon I \ni t \mapsto (v(t), g(t), \mathrm{id}_{\mathfrak{g}}) \in \mathfrak{g} \times G \times \mathrm{GL}(\mathfrak{g})$, where*

$$v(t) = \int_0^t u(s)\,\mathrm{d}s, \quad t \in I, \tag{5.15}$$

and $g\colon I \to G$ is defined by the initial value problem

$$\dot{g}(t) = T_e \ell_{g(t)} u(t), \quad g(0) = e. \tag{5.16}$$

Then $(v(t), g(t), A(t))$ defines an intrinsic rolling of $\mathfrak{g}$ over G with respect to $\nabla^{\mathrm{can}2}$, where

$$A(t)\colon T_{v(t)}\mathfrak{g} \cong \mathfrak{g} \to T_{g(t)} G, \quad Z \mapsto T_e \ell_{g(t)} Z. \tag{5.17}$$

If G is a matrix Lie group, (5.17) *simplifies to $A(t)Z = g(t)Z$.*

Proof: According to Definition 3.5, 2, $\nabla^{\mathrm{can}2}$ corresponds to the $\mathrm{Ad}(\{e\})$-invariant bilinear map $\alpha\colon \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ defined by $\alpha(X, Y) = 0$ for all $X, Y \in \mathfrak{g}$. Now it is straightforward to verify that the curve $(v, g, S)\colon I \ni t \mapsto (v(t), g(t), \mathrm{id}_{\mathfrak{g}}) \in \mathfrak{g} \times G \times \mathrm{GL}(\mathfrak{g})$ fulfills the kinematic equation (3.20) from Proposition 3.7 with $\alpha = 0$. This yields the desired result. □

The special case of Proposition 5.4, where the development curve is a one-parameter subgroup, is also of interest. Thus we state the next corollary.

Corollary 5.5 *Let G be a Lie group and let $u\colon I \to \mathfrak{g}$ be constant, i.e. $u(t) = \xi$ holds for all $t \in I$ and some $\xi \in \mathfrak{g}$. Then $(v, g, S)\colon I \ni t \mapsto (v(t), g(t), S(t)) = (t\xi, \exp(t\xi), \mathrm{id}_{\mathfrak{g}}) \in \mathfrak{g} \times G \times \mathrm{GL}(\mathfrak{g})$ defines the rolling $(v(t), g(t), A(t))$ of $\mathfrak{g}$ over G with respect to $\nabla^{\mathrm{can}2}$, where*

$$A(t)\colon T_{v(t)}\mathfrak{g} \cong \mathfrak{g} \to T_{g(t)} G, \quad Z \mapsto (T_e \ell_{\exp(t\xi)}) Z. \tag{5.18}$$

holds for all $t \in I$. If G is a matrix Lie group, (5.18) *simplifies to*

$$A(t) Z = \exp(t\xi) Z \tag{5.19}$$

for all $Z \in \mathfrak{g} \subseteq \mathfrak{gl}(n)$.

Proof: The desired result follows by specifying $u\colon I \to \mathfrak{g}$ in Proposition 5.4 as $u(t) = \xi$ for all $t \in I$. □

5.2 Interpolation on Matrix Lie Groups via Intrinsic Rolling and Unwrapping

Next we specify Algorithm 1 to a matrix Lie group $G \subseteq \mathrm{GL}(n)$, where we make the following assumptions and choices.

Notation 5.6 Let $g_0 \in G$ be fixed. We assume that G is embedded into $\mathbb{R}^{n\times n}$ via

$$\iota\colon G \to \mathbb{R}^{n\times n}, \quad g \mapsto \iota(g) = gg_0 = r_{g_0}(g), \tag{5.20}$$

see also (5.3). Moreover, we assume that there exists some $\xi \in \mathfrak{g}$ such that $\exp(T\xi)g_0 = g_k$ holds, where $T > 0$ is a fixed real number included here such that the notation coincides with the one used in the algorithms below.

In general, for two given points $g_0, g_k \in G$, the existence of a $\xi \in \mathfrak{g}$ with $\exp(T\xi)g_0 = g_k$ is not ensured. Nevertheless, if the exponential map $\exp\colon \mathfrak{g} \to G$ is surjective, one can always find some $\xi \in \mathfrak{g}$ with $g_0 \exp(T\xi) = g_k$. Next we note that

$$T_e(\iota \circ \mathrm{pr})Z = T_e\iota \circ T_e \operatorname{id}_G Z = T_e r_{g_0} Z = Zg_0 \tag{5.21}$$

holds for all $Z \in \mathfrak{g}$ due to $\mathrm{pr} = \mathrm{id}_G\colon G \to G$. This yields

$$\big(T_e(\iota \circ \mathrm{pr})\big)^{-1} Z = Zg_0^{-1}. \tag{5.22}$$

It is well-known that the exponential map $\exp\colon \mathfrak{g} \to G$ is a local diffeomorphism around $0 \in \mathfrak{g}$. We denote its inverse by $\log\colon U \subseteq G \to \mathfrak{g}$ which is defined on a suitable open neighbourhood of $e \in G$. This map is used to construct a local diffeomorphism ϕ with the properties required by Algorithm 1, Step 4. To this end, we set

$$\phi\colon r_{g_0}(U) \subseteq G \to T_{g_0}G, \quad g \mapsto \big(T_e r_{g_0} \circ \log \circ (r_{g_0})^{-1}\big)(g) = \log\big(gg_0^{-1}\big)g_0. \tag{5.23}$$

Then, ϕ has the desired properties (4.17) due to $\phi(g_0) = \log(e)g_0 = 0g_0 = 0$ and

$$T_{g_0}\phi = T_e r_{g_0} \circ T_e \log \circ T_{g_0}(r_{g_0})^{-1} = T_e r_{g_0} \circ \mathrm{id}_{\mathfrak{g}} \circ \big(T_e r_{g_0}\big)^{-1} = \mathrm{id}_{T_{g_0}G}. \tag{5.24}$$

Its inverse $\phi^{-1}\colon V \subseteq T_{g_0}G \to G$ defined on some suitable open neighbourhood of $0 \in T_{g_0}G$ is given by

$$\phi^{-1}(Z) = \big(T_e r_{g_0} \circ \log \circ (r_{g_0})^{-1}\big)^{-1} Z = r_{g_0} \circ \log^{-1} \circ \big(T_e r_{g_0}\big)^{-1} Z = \exp\big(Zg_0^{-1}\big)g_0 \tag{5.25}$$

for all $Z \in T_{g_0}G$.

Remark 5.7 If the map $\log\colon U \subseteq G \to \mathfrak{g}$ used for the construction of ϕ in (5.23) is replaced by another suitable local diffeomorphism $\varphi\colon \widehat{U} \subseteq G \to \mathfrak{g}$ defined on an open neighbourhood $\widehat{U}$ of $e \in G$ satisfying $\varphi(e) = 0$ and $T_e\varphi = \mathrm{id}_{\mathfrak{g}}$, the map $\widehat{\phi} = T_e r_{g_0} \circ \varphi \circ (r_{g_0})^{-1}$ also satisfies the properties required by Algorithm 1, Step 4. However, for simplicity, we restrict to the local diffeomorphism ϕ defined in (5.23).

Remark 5.8 In this subsection we focus on matrix Lie groups to obtain rather simple expressions. Nevertheless, since Proposition 5.1 and Corollary 5.3 as well as Proposition 5.4 and Corollary 5.5 include expressions for rollings which are valid for general Lie groups, in principle, Algorithm 2 and Algorithm 3 can be adapted to general Lie groups.

5.2.1 Interpolation via Intrinsic Rollings with respect to $\nabla^{\mathrm{can}1}$

Using Notation 5.6, we apply Algorithm 1 to a matrix Lie group $G \subseteq \mathrm{GL}(n)$ equipped with $\nabla^{\mathrm{can}1}$. This yields Algorithm 2. Afterwards, in Lemma 5.10 below, we show that Algorithm 2 is indeed a special case of Algorithm 1.

Algorithm 2 Interpolation on matrix Lie groups equipped with $\nabla^{\mathrm{can}1}$

Input: $g_0, \dots, g_k \in G \subseteq \mathrm{GL}(n)$, initial velocity $\xi_0 \in T_{g_0}G$, final velocity $\xi_k \in T_{g_k}G$, instances of time $0 = t_0 < \dots < t_k = T$.

1. Compute $\xi \in \mathfrak{g}$ such that $\iota(\exp(T\xi)) = \exp(T\xi)g_0 = g_k$ holds.
2. Unwrap the boundary data to $T_{g_0}G$ by defining
$$\begin{aligned} q_0 &= 0, \quad && q_k = T\xi g_0, \\ \eta_0 &= \xi_0, \quad && \eta_k = \exp(-\tfrac{T}{2}\xi)\big(\xi_k g_0^{-1}\big)\exp\big(-\tfrac{T}{2}\xi\big)g_0. \end{aligned} \tag{5.26}$$
3. Unwrap the remaining data for $i \in \{1, \dots, k-1\}$ by setting
$$q_i = \exp\big(\tfrac{t_i}{2}\xi\big)\Big(\log\Big(\big(\exp(-t_i\xi)g_i\big)g_0^{-1}\Big)\Big)\exp\big(-\tfrac{t_i}{2}\xi\big)g_0 + t_i\xi g_0. \tag{5.27}$$
4. Compute a $\mathscr{C}^2$-curve $y\colon [0,T] \to T_{g_0}G$ fulfilling $y(t_i) = q_i$ for all $i \in \{0, \dots, k\}$ and $\dot{y}(0) = \eta_0$ as well as $\dot{y}(T) = \eta_k$.
5. Define the curve $\beta\colon [0,T] \to G$ by setting
$$\beta(t) = \exp\big(\tfrac{t}{2}\xi\big)\exp\Big(y(t)g_0^{-1} - t\xi\Big)\exp\big(\tfrac{t}{2}\xi\big)g_0. \tag{5.28}$$

Output: The curve $\beta\colon [0,T] \to G$.

Remark 5.9 When Algorithm 2 is applied to $G = \mathrm{SO}(n)$, the curve $\beta\colon [0,T] \to \mathrm{SO}(n)$ given by (5.28) is of a form similar to the curve in [29, Eq. (5.11)], obtained by the rolling and unwrapping technique from [26, 29] applied to $\mathrm{SO}(n)$.

Lemma 5.10 *Algorithm 2 is a special case of Algorithm 1 applied to the matrix Lie group G equipped with $\nabla^{\mathrm{can}1}$, where the choices from Notation 5.6 are used.*

Proof: We only need to verify that the steps of Algorithm 2 are obtained by specializing the corresponding steps of Algorithm 1. This is done by the following computations, where the assumptions and notations from Algorithm 2 are used:

1. Clearly, Algorithm 2, Step 1 combined with Corollary 5.3 can be seen as a special case of Algorithm 1, Step 1 applied to G equipped with $\nabla^{\mathrm{can}1}$. In particular, the curve $(v, g, S)\colon [0,T] \to \mathfrak{g} \times G \times \mathrm{GL}(\mathfrak{g})$ associated with the rolling of $\mathfrak{g}$ over G whose development curve is given by $g\colon [0,T] \ni t \mapsto \exp(t\xi) \in G$ reads
$$v(t) = t\xi, \quad g(t) = \exp(t\xi) \quad \text{and} \quad S(t) = \mathrm{Ad}_{\exp\left(-\frac{t}{2}\xi\right)} \tag{5.29}$$
for all $t \in [0,T]$, see also (5.9) in Corollary 5.3.

2. Algorithm 2, Step 2 is a special case of Algorithm 1, Step 2 and Algorithm 1, Step 3 combined in one step. Note that

$$\widehat{v}(t) = (T_e\iota)v(t) = t\xi g_0 \quad \text{and} \quad \widehat{\gamma}(t) = (\iota \circ g)(t) = \exp(t\xi)g_0 \tag{5.30}$$

holds for all $t \in [0,T]$. Moreover, since $\mathrm{pr} = \mathrm{id}_G$ and $(T_e\iota)^{-1}Z = r_{g_0^{-1}}(Z)$ for all $Z \in T_{g_0}G$, $\widehat{S}(t)$ defined in (4.15) of Algorithm 1 simplifies to

$$\begin{aligned} \widehat{S}(t)Z &= \big(T_e\iota \circ S(t) \circ (T_e\iota)^{-1}\big)Z \\ &= \big(\operatorname{Ad}_{\exp(-\frac{t}{2}\xi)}(Zg_0^{-1})\big)g_0 \\ &= \exp\big(-\tfrac{t}{2}\xi\big)\big(Zg_0^{-1}\big)\exp\big(\tfrac{t}{2}\xi\big)g_0, \end{aligned} \tag{5.31}$$

and $\widehat{A}(t)$ defined in (4.14) of Algorithm 1 specializes to

$$\begin{aligned} \widehat{A}(t)Z &= T_{g(t)}\iota \circ A(t) \circ (T_e\iota)^{-1}Z \\ &= (T_{g(t)}\iota)\exp\big(\tfrac{t}{2}\xi\big)\big(Zg_0^{-1}\big)\exp\big(\tfrac{t}{2}\xi\big) \\ &= \exp\big(\tfrac{t}{2}\xi\big)\big(Zg_0^{-1}\big)\exp\big(\tfrac{t}{2}\xi\big)g_0, \end{aligned} \tag{5.32}$$

where $A(t)$ is given by (5.11). For $Z \in T_{\widehat{\gamma}(t)}G$, its inverse is given by

$$\widehat{A}(t)^{-1}Z = \big((T_e\iota) \circ A(t)^{-1} \circ (T_{g(t)}\iota)^{-1}\big)Z = \exp\big(-\tfrac{t}{2}\xi\big)\big(Zg_0^{-1}\big)\exp\big(-\tfrac{t}{2}\xi\big)g_0. \tag{5.33}$$

By (5.30) and (5.33) the definition of the boundary data in (5.26) is a special case of Algorithm 1, Step 3.

3. Algorithm 2, Step 3 is a special case of the corresponding step in Algorithm 1. Define $\phi\colon U \subseteq G \to T_{g_0}G$ by $g \mapsto \log\big(gg_0^{-1}\big)g_0$. Then, by the discussion at the beginning of Subsection 5.2, ϕ satisfies the requirements of the local diffeomorphism from Algorithm 1, Step 4. Next, using $\widehat{v}(t) = t\xi g_0$ and

$$\widehat{S}(t)^{-1}Z = \Big(T_e\iota \circ \Big(\operatorname{Ad}_{\exp\left(-\frac{t}{2}\xi\right)}\Big)^{-1} \circ (T_e\iota)^{-1}\Big)Z = \Big(\operatorname{Ad}_{\exp(\frac{t}{2}\xi)}\big(Zg_0^{-1}\big)\Big)g_0, \tag{5.34}$$

for $Z \in T_{g_0}G$, one obtains for $i \in \{1,\dots,k-1\}$

$$\begin{aligned} q_i &= \widehat{S}(t_i)^{-1}\big(\phi\big(g(t_i)^{-1}g_i\big)\big) + \widehat{v}(t_i) \\ &= \Big(\operatorname{Ad}_{\exp\left(\frac{t_i}{2}\xi\right)}\Big(\phi\big(\exp(t_i\xi)^{-1}g_i\big)g_0^{-1}\Big)\Big)g_0 + t_i\xi g_0 \\ &= \Big(\operatorname{Ad}_{\exp\left(\frac{t_i}{2}\xi\right)}\Big(\Big(\log\Big(\big(\exp(t_i\xi)^{-1}g_i\big)g_0^{-1}\Big)g_0\Big)g_0^{-1}\Big)\Big)g_0 + t_i\xi g_0 \\ &= \Big(\operatorname{Ad}_{\exp\left(\frac{t_i}{2}\xi\right)}\Big(\log\big(\big(\exp(t_i\xi)^{-1}g_i\big)g_0^{-1}\big)\Big)\Big)g_0 + t_i\xi g_0 \\ &= \exp\big(\tfrac{t_i}{2}\xi\big)\log\Big(\exp(-t_i\xi)g_ig_0^{-1}\Big)\exp\big(-\tfrac{t_i}{2}\xi\big)g_0 + t_i\xi g_0. \end{aligned} \tag{5.35}$$

4. Obviously, Algorithm 2, Step 4 corresponds to Algorithm 1, Step 5.

5. Algorithm 2, Step 5 is a special case of the corresponding step in Algorithm 1. Using

$\widehat{v}(t) = t\xi g_0$, and recalling $g(t) = \exp(t\xi)$, we compute

$$\begin{aligned}\beta(t) &= \Phi_{g(t)}\big(\phi^{-1}\big(\widehat{S}(t)\big(y(t) - \widehat{v}(t)\big)\big)\big)\\ &\overset{(5.25),(5.31)}{=} \Phi_{g(t)}\Big(\exp\Big(\Big(\exp\big(-\tfrac{t}{2}\xi\big)\big(\big(y(t) - \widehat{v}(t)\big)g_0^{-1}\big)\exp\big(\tfrac{t}{2}\xi\big)g_0\Big)g_0^{-1}\Big)g_0\Big)\\ &\overset{(5.2)}{=} \exp(t\xi)\Big(\exp(-\tfrac{t}{2}\xi)\exp\Big(y(t)g_0^{-1} - t\xi\Big)\big(\exp(-\tfrac{t}{2}\xi)\big)^{-1}g_0\Big)\\ &= \exp\big(\tfrac{t}{2}\xi\big)\exp\Big(y(t)g_0^{-1} - t\xi\Big)\exp\big(\tfrac{t}{2}\xi\big)g_0\end{aligned} \tag{5.36}$$

as desired. □

5.2.2 Interpolation via Intrinsic Rollings with respect to $\nabla^{\mathrm{can}2}$

In order to illustrate that Algorithm 1 depends on the choice of the covariant derivative, we now apply it to a matrix Lie group equipped with $\nabla^{\mathrm{can}2}$. This yields Algorithm 3 below.

Algorithm 3 Interpolation on matrix Lie groups equipped with $\nabla^{\mathrm{can}2}$

Input: $g_0, \ldots, g_k \in G \subseteq \mathrm{GL}(n)$, initial velocity $\xi_0 \in T_{g_0}G$, final velocity $\xi_k \in T_{g_k}G$, instances of time $0 = t_0 < \cdots < t_k = T$.

1. Compute $\xi \in \mathfrak{g}$ such that $\iota(\exp(T\xi)) = \exp(T\xi)g_0 = g_k$ holds.
2. Unwrap the boundary data to $T_{g_0}G$ by defining

$$\begin{aligned} q_0 &= 0, & q_k &= T\xi g_0,\\ \eta_0 &= \xi_0, & \eta_k &= \exp(-T\xi)\xi_k. \end{aligned} \tag{5.37}$$

3. Unwrap the remaining data for $i \in \{1, \ldots, k-1\}$ by setting

$$q_i = \log\big(\exp(-t_i\xi)g_i g_0^{-1}\big)g_0 + t_i\xi g_0. \tag{5.38}$$

4. Compute a $\mathscr{C}^2$-curve $y\colon [0,T] \to T_{g_0}G$ satisfying $y(t_i) = q_i$ for all $i \in \{0, \ldots, k\}$ and $\dot{y}(0) = \eta_0$ as well as $\dot{y}(T) = \eta_k$.
5. Define the curve $\beta\colon [0,T] \to G$ by setting

$$\beta(t) = \exp(t\xi)\exp\big(y(t)g_0^{-1} - t\xi\big)g_0. \tag{5.39}$$

Output: The curve $\beta\colon [0,T] \to G$.

Lemma 5.11 *Algorithm 3 is a special case of Algorithm 1 applied to the matrix Lie group G with the choices from Notation 5.6, where G is equipped with $\nabla^{\mathrm{can}2}$.*

Proof: Proceeding analogously to the proof of Lemma 5.10, we verify that the steps of Algorithm 3 are obtained by specializing the corresponding steps of Algorithm 1. In detail, using the assumptions and notations from Algorithm 3, we obtain:

1. Algorithm 3, Step 1 is a special case of Algorithm 1, Step 1 applied to G equipped with $\nabla^{\mathrm{can}2}$. Indeed, by Corollary 5.5, the curve $(v, g, S)\colon [0,T] \ni t \mapsto (t\xi, \exp(t\xi), \mathrm{id}_{\mathfrak{g}}) \in \mathfrak{g} \times G \times \mathrm{GL}(\mathfrak{g})$ is associated with the rolling $\big(v(t), g(t), A(t)\big) = \big(t\xi, \exp(t\xi), T_e\ell_{\exp(t\xi)}\big)$ of $\mathfrak{g}$ over G whose development curve is given by $g\colon [0,T] \ni t \mapsto \exp(t\xi) \in G$.

2. Algorithm 3, Step 2 is a special case of Algorithm 1, Step 2 and Algorithm 1, Step 3 combined in one step. Note that

$$\widehat{v}(t) = (T_e\iota)v(t) = t\xi g_0 \quad \text{and} \quad \widehat{\gamma}(t) = (\iota \circ g)(t) = \exp(t\xi)g_0 \tag{5.40}$$

holds for all $t \in [0, T]$. Moreover, $\widehat{S}(t)$ defined in (4.15) of Algorithm 1 simplifies to

$$\widehat{S}(t) = \big(T_e\iota \circ S(t) \circ (T_e\iota)^{-1}\big) = \big(T_e\iota \circ \mathrm{id}_{\mathfrak{g}} \circ (T_e\iota)^{-1}\big) = \mathrm{id}_{T_{g_0}G} \tag{5.41}$$

and $\widehat{A}(t)$ defined in (4.14) of Algorithm 1 specializes to

$$\widehat{A}(t)Z = T_{g(t)}\iota \circ A(t) \circ \big(T_e(\iota \circ \mathrm{pr}\,\big|_{\mathfrak{m}})\big)^{-1} Z = \exp(t\xi)Z, \quad Z \in T_{g_0}G, \tag{5.42}$$

where $A(t)$ is given by (5.19) and $\mathrm{pr} = \mathrm{id}_G$ as well as $(T_e\iota)^{-1}(Z) = Zg_0^{-1}$ are used. Clearly, its inverse is given by $\widehat{A}(t)^{-1}Z = \exp(-t\xi)Z$ for all $Z \in T_{g_0}G$. By this expression and (5.42), the definition of the boundary data in (5.37) is a special case of Algorithm 1, Step 3.

3. Algorithm 2, Step 3 is a special case of the corresponding step in Algorithm 1. Define $\phi\colon U \subseteq G \to T_{g_0}G$ by $g \mapsto \log\big(gg_0^{-1}\big)g_0$. By the discussion at the beginning of Subsection 5.2, ϕ satisfies the requirements of the local diffeomorphism from Algorithm 1, Step 4. Next, using $\widehat{v}(t) = t\xi g_0$ and $\widehat{S}(t)^{-1} = \mathrm{id}_{T_{g_0}G}$, one obtains for $i \in \{1, \dots, k-1\}$

$$q_i = \widehat{S}(t_i)^{-1}\big(\phi\big(g(t_i)^{-1}g_i\big)\big) + \widehat{v}(t_i) = \log\Big(\big(\exp(-t_i\xi)g_i\big)g_0^{-1}\Big)g_0 + t_i\xi g_0. \tag{5.43}$$

4. Obviously, Algorithm 2, Step 4 corresponds to Algorithm 1, Step 5.

5. Algorithm 2, Step 5 is a special case of the corresponding step in Algorithm 1. Using (5.41) and $\widehat{v}(t) = t\xi g_0$ as well as $\phi^{-1}(Z) = \exp\big(Zg_0^{-1}\big)g_0$ according to (5.25), we obtain

$$\beta(t) = \Phi_{g(t)}\big(\phi^{-1}\big(\widehat{S}(t)\big(y(t) - \widehat{v}(t)\big)\big)\big) = \exp\big(t\xi\big)\exp\Big(y(t)g_0^{-1} - t\xi\Big)g_0 \tag{5.44}$$

as desired. □

We already mentioned the Lie group SO(n) in Remark 5.9 as an example to which Algorithm 2 can be applied. Next we briefly discuss the manifold SE(3) equipped with two different Lie group structures.

Example 5.12 1. The special Euclidean group SE(n) can be viewed as the matrix Lie group

$$\mathrm{SE}(n) = \Big\{ \big[\begin{smallmatrix} R & b \\ 0 & 1 \end{smallmatrix}\big] \,\Big|\, R \in \mathrm{SO}(n),\ b \in \mathbb{R}^n \Big\} \subseteq \mathrm{GL}(n+1), \tag{5.45}$$

whose Lie algebra is given by

$$\mathfrak{se}(n) = \Big\{ \big[\begin{smallmatrix} \Omega & v \\ 0 & 0 \end{smallmatrix}\big] \,\Big|\, \Omega \in \mathfrak{so}(n),\ v \in \mathbb{R}^n \Big\} \subseteq \mathfrak{gl}(n+1), \tag{5.46}$$

see e.g. [18, Sec. 2.6]. It is well-known that the exponential $\exp\colon \mathfrak{se}(n) \to \mathrm{SE}(n)$ is surjective [18, Thm. 2.12]. Obviously, Algorithm 1 can be applied to solve Problem 2.1 on SE(n) with Algorithm 2 and Algorithm 3 being particular instances. For details on evaluating the exponential and logarithm numerically, we refer to [23, Chap. 11-12], see also [10, Sec. 5]. We mention that the case $n = 3$ is of interest for various applications and, moreover, interpolation problems on SE(3) are addressed in several works in the literature, see for example [13, 51] and references therein.

2. In some applications, see e.g. [13], it is convenient to endow SE(3) with a Lie group structure different from that in 1. As in [13], we write $\mathrm{PCG}(3) = \mathrm{SO}(3) \times \mathbb{R}^3 = \{(R, b) \mid R \in \mathrm{SO}(3),\ b \in \mathbb{R}^3\}$ for the direct product of SO(3) and $\mathbb{R}^3$, the so-called pose change group. Obviously, as manifolds $\mathrm{PCG}(3) \cong \mathrm{SE}(3)$ are diffeomorphic. However, PCG(3) is not isomorphic to SE(3) as a Lie group. Clearly, Algorithm 1 can be applied to the reductive homogeneous spaces $\mathrm{PCG}(3) \cong \mathrm{PCG}(3)/\{e\}$, too. Moreover, it is straightforward to verify that the map

$$\mathrm{PCG}(3) \ni (R, b) \mapsto \begin{bmatrix} R & 0 & 0 \\ 0 & I_3 & b \\ 0 & 0 & 1 \end{bmatrix} \in \mathrm{GL}(7) \subseteq \mathbb{R}^{7\times 7} \tag{5.47}$$

is an isomorphism of Lie groups onto its image turning PCG(3) into a matrix Lie group. Thus Algorithm 2 and Algorithm 3 can be applied to PCG(3), as well.

Next we briefly discuss Problem 2.2 on a matrix Lie group G. This leads to Algorithm 4 below, where Algorithm 2 or Algorithm 3 might be an essential building block for solving the sub-problem in Step 2 of Algorithm 4.

Algorithm 4 Interpolation on matrix Lie groups (for solving Problem 2.2)

Input: $g_0, \dots, g_k \in G$, velocities $\xi_0 \in T_{x_0}G, \dots, \xi_k \in T_{x_k}G$, instances of time $0 = t_0 < \dots < t_k = T$.

1. For $i = 0, \dots, k-1$ do:
   Compute a $\mathscr{C}^1$-curve $\beta_i \colon [0, t_{i+1} - t_i] \to G$ satisfying $\beta_i(0) = g_i$, $\beta_i(t_{i+1} - t_i) = g_{i+1}$ and $\dot{\beta}_i(0) = \xi_i$ as well as $\dot{\beta}_i(t_{i+1} - t_i) = \xi_{i+1}$.
2. Define $\beta \colon [0, T] \to M$ for $t \in [0, T]$ by

$$\beta\big|_{[t_i, t_{i+1})}(t) = \beta_i(t - t_i) \quad, i \in \{0, \dots, k-2\} \quad \text{and} \quad \beta\big|_{[t_{k-1}, T]}(t) = \beta_{k-1}(t - t_{k-1}) \tag{5.48}$$

Output: The curve $\beta \colon [0, T] \to G$.

Assuming that the interpolation problems arising in Step 1 of Algorithm 4 can be solved by Algorithm 2 or Algorithm 3, a specific algorithm for solving Problem 2.2 on G is obtained. More generally, the next lemma shows that Algorithm 4 yields a valid solution of Problem 2.2 on G independently of the particular method for solving Algorithm 4, Step 1.

Lemma 5.13 *Let $g_0, \dots, g_k \in G$ and $\xi_0 \in T_{g_0}G, \dots, \xi_k \in T_{g_k}G$ as well as $0 = t_0 < \dots < t_k = T$. Then Algorithm 4 solves Problem 2.2 on G associated with the given data.*

PROOF: Obviously, $\beta \colon [0, T] \to G$ is well-defined. Let $i \in \{1, \dots, k-1\}$. Then

$$\lim_{t \to t_i^-} \beta(t) = \beta_{i-1}(t_i - t_{i-1}) = g_i \quad \text{and} \quad \lim_{t \to t_i^+} \beta_i(t) = \beta_i(t_i - t_i) = g_i \tag{5.49}$$

holds, proving that β is continuous. Thus β is in fact a piecewise $\mathscr{C}^1$-curve. Moreover, one has

$$\lim_{t \to t_i^-} \dot{\beta}(t) = \dot{\beta}_{i-1}(t_i - t_{i-1}) = \xi_i \quad \text{and} \quad \lim_{t \to t_i^+} \dot{\beta}_i(t) = \dot{\beta}_i(t_i - t_i) = \xi_i \tag{5.50}$$

showing that $\beta \colon [0, T] \to G$ is a $\mathscr{C}^1$-curve. In addition, $\beta(0) = \beta_0(0) = g_0$ and $\beta(T) = \beta_{k-1}(T - t_{k-1}) = g_k$, as well as $\dot{\beta}(0) = \dot{\beta}_0(0) = \xi_0$ and $\dot{\beta}(T) = \dot{\beta}_k(T - t_{k-1}) = \xi_k$ holds. Together with (5.49) and (5.50), this yields $\beta(t_i) = g_i$ and $\dot{\beta}(t_i) = \xi_i$, for all $i \in \{0, \dots, k\}$, as desired. □

6 Interpolation on Stiefel Manifolds

In this section, we apply Algorithm 1 to the Stiefel manifold $\mathrm{St}_{n,k}$, where we put an emphasis on obtaining an efficient algorithm for the case $k \ll n$. To this end, we first recall some facts on Stiefel manifolds and the so-called quasi-geodesics on them. Afterwards, Algorithm 1 is adapted to $\mathrm{St}_{n,k}$. However, we start with motivating the choice of the reductive decomposition and our focus on quasi-geodesics, see Subsection 6.2 and Subsection 6.3 below.

Remark 6.1 In [41, Sec. 6.2.1] and [42, Sec. 5.1], the Stiefel manifold $\mathrm{St}_{n,k}$ equipped with a fixed α-metric from [27] is identified with a normal naturally reductive homogeneous space. Moreover, using the closed-form expressions for intrinsic rollings of $\mathrm{St}_{n,k}$ with respect to ∇^{LC} defined by an arbitrary α-metric from [41, Sec. 6.2.2] and [42, Sec. 5.4], Algorithm 1 could be applied to $\mathrm{St}_{n,k}$, viewed as (normal naturally) reductive homogeneous space G/H equipped with ∇^{LC}, where $G = \mathrm{O}(n) \times \mathrm{O}(k)$, $H \cong \mathrm{O}(n-k) \times \mathrm{O}(k)$, and the reductive decomposition $\mathfrak{so}(n) \times \mathfrak{so}(k) = \mathfrak{h} \oplus \mathfrak{m}$ is orthogonal with respect to a specific scalar product. However, this would require the computation of exponentials of elements in $\mathfrak{so}(\mathfrak{m})$, where $\mathfrak{m}$ is of dimension $\dim(\mathfrak{m}) = \dim(\mathrm{St}_{n,k}) = nk - k(k-1)/2$. Clearly, this is not tractable from a computational point of view for “big” n.

In addition, the explicit expressions for extrinsic rollings of Stiefel manifolds equipped with the Euclidean metric from [42, Sec. 5.4] would allow for applying the rolling and unwrapping technique from [26, 29] to $\mathrm{St}_{n,k}$. This approach is not tractable for “big” n, as well, since it also requires computations of exponentials of elements in $\mathfrak{so}(\mathfrak{m})$.

To overcome this difficulty, we apply Algorithm 1 to $\mathrm{St}_{n,k}$ identified with a reductive homogeneous space equipped with $\nabla^{\mathrm{can}2}$, where the reductive decomposition is chosen such that the so-called quasi-geodesics are projections of *horizontal* one-parameter subgroups in $\mathrm{O}(n) \times \mathrm{O}(k)$. Ultimately, this yields Algorithm 6 below. Here our choice of the covariant derivative $\nabla^{\mathrm{can}2}$ is motivated by the simple form of the curve $S\colon I \to \mathrm{GL}(\mathfrak{m})$ given by the kinematic equation from Lemma 3.8, see also (3.29) in Remark 3.10.

6.1 Stiefel Manifolds

We start with recalling some basic facts on the Stiefel manifold, viewed as embedded submanifold of $\mathbb{R}^{n\times k}$ and as homogeneous space. Here we mainly follow [27]. The Stiefel manifold is the embedded submanifold

$$\mathrm{St}_{n,k} = \{X \in \mathbb{R}^{n\times k} \mid X^\top X = I_k\} \subseteq \mathbb{R}^{n\times k}, \tag{6.1}$$

where $1 \leq k \leq n$. The characterization of the tangent space of $\mathrm{St}_{n,k}$ at $X \in \mathrm{St}_{n,k}$ given by

$$T_X \mathrm{St}_{n,k} = \{V \in \mathbb{R}^{n\times k} \mid X^\top V = -V^\top X\} \tag{6.2}$$

is well-known. Next we consider the $G = (\mathrm{O}(n) \times \mathrm{O}(k))$-action

$$\Phi\colon (\mathrm{O}(n) \times \mathrm{O}(k)) \times \mathbb{R}^{n\times k} \to \mathbb{R}^{n\times k}, \quad ((R,\theta), X) \mapsto RX\theta^\top \tag{6.3}$$

from the left which restricts to the action on the Stiefel manifold

$$\Phi\colon (\mathrm{O}(n) \times \mathrm{O}(k)) \times \mathrm{St}_{n,k} \to \mathrm{St}_{n,k}, \quad ((R,\theta), X) \mapsto RX\theta^\top \tag{6.4}$$

which is known to be transitive.

Let $X \in \mathrm{St}_{n,k}$ be fixed. Then $\mathrm{St}_{n,k}$ can be identified with the homogeneous space $G/H_X = (\mathrm{O}(n) \times \mathrm{O}(k))/H_X$, where

$$H_X = \mathrm{Stab}(X) = \{(R,\theta) \in \mathrm{O}(n) \times \mathrm{O}(k) \mid \Phi((R,\theta),X) = X\} \subseteq \mathrm{O}(n) \times \mathrm{O}(k) \tag{6.5}$$

is the stabilizer subgroup of X under the transitive action from (6.4). Here we refer to [36, Thm. 6.4] and [33, Thm. 21.18] for more details on Lie group actions and their orbits. In particular, the following diagram

$$\begin{array}{ccc} \mathrm{O}(n) \times \mathrm{O}(k) & & \\ \Big\downarrow \mathrm{pr}_X & \searrow^{\Phi(\cdot,X)} & \\ (\mathrm{O}(n) \times \mathrm{O}(k))/H_X & \xrightarrow{\iota_X} & \mathrm{St}_{n,k} \end{array} \tag{6.6}$$

commutes, where

$$\mathrm{pr}_X \colon \mathrm{O}(n) \times \mathrm{O}(k) \to (\mathrm{O}(n) \times \mathrm{O}(k))/H_X, \quad (R,\theta) \mapsto (R,\theta) \cdot H_X \tag{6.7}$$

is the canonical projection and

$$\iota_X \colon (\mathrm{O}(n) \times \mathrm{O}(k))/H_X \ni (R,\theta) \cdot H_X \mapsto \Phi_{(R,\theta)}(X) = RX\theta^\top \in \mathrm{St}_{n,k} \tag{6.8}$$

is a G-equivariant diffeomorphism. Here $(R,\theta) \cdot H_X = \mathrm{pr}_X(R,\theta)$ denotes the coset in $(\mathrm{O}(n) \times \mathrm{O}(k))/H_X$ defined by $(R,\theta) \in G = \mathrm{O}(n) \times \mathrm{O}(k)$, as usual. The tangent map of $\iota_X \circ \mathrm{pr}_X = \Phi(\cdot, X) \colon \mathrm{O}(n) \times \mathrm{O}(k) \to \mathrm{St}_{n,k}$ at $(I_n, I_k) \in \mathrm{O}(n) \times \mathrm{O}(k)$ is given by

$$T_{(I_n,I_k)}(\iota_X \circ \mathrm{pr}_X) \colon \mathfrak{so}(n) \times \mathfrak{so}(k) \to T_X\mathrm{St}_{n,k}, \quad (\Omega, \Psi) \mapsto \Omega X - X\Psi. \tag{6.9}$$

Moreover, by [27], the stabilizer subgroup $H_X = \mathrm{Stab}(X)$ is isomorphic to the Lie group $H_X \cong \mathrm{O}(n-k) \times \mathrm{O}(k)$ and the Lie algebra $\mathfrak{h}_X$ of H_X, viewed as a Lie subalgebra of $\mathfrak{so}(n) \times \mathfrak{so}(k)$ is given by

$$\mathfrak{h}_X = \ker\big(T_{(I_n,I_k)}(\iota_X \circ \mathrm{pr}_X)\big) = \{(\Omega,\Psi) \in \mathfrak{so}(n) \times \mathfrak{so}(k) \mid \Omega X - X\Psi = 0\}. \tag{6.10}$$

Notation 6.2 From now on, we assume $1 \leq k < n$ as we are mainly interested in the case $k \ll n$. Moreover, note that $\mathrm{St}_{n,n} = \mathrm{O}(n)$ has two connected components. For solving Problem 2.1 on one of its connected component, i.e. $\mathrm{SO}(n) \subseteq \mathrm{O}(n)$ or $R_0\mathrm{SO}(n) \subseteq \mathrm{O}(n)$ for some $R_0 \in \mathrm{O}(n)$ with $\det(R_0) = -1$, we refer to Section 5.

6.2 A Reductive Decomposition

The Stiefel manifold can be equipped with a one-parameter family of metrics introduced in [27]. In that work, the Lie group $\mathrm{O}(n)\times\mathrm{O}(k)$ is equipped with a bi-invariant metric which depends on a parameter $\alpha \in \mathbb{R}\setminus\{-1, 0\}$. Afterwards, this family of metrics on $\mathrm{O}(n)\times\mathrm{O}(k)$ is used to construct a one-parameter family of metrics on $\mathrm{St}_{n,k} \cong (\mathrm{O}(n) \times \mathrm{O}(k))/H_X$ such that $\iota_X \circ \mathrm{pr}_X \colon \mathrm{O}(n)\times\mathrm{O}(k) \to \mathrm{St}_{n,k}$ becomes a pseudo-Riemannian submersion. Moreover, in [27, Sec. 4.2], it is observed that the limit $\alpha \to \infty$ yields the decomposition of $\mathfrak{so}(n) \times \mathfrak{so}(k)$ corresponding to the so-called quasi-geodesic horizontal distribution on $\mathrm{O}(n) \times \mathrm{O}(k)$ from [30, Sec. 3.2.1]. This decomposition is the reductive decomposition of $\mathfrak{so}(n) \times \mathfrak{so}(k)$ which is of interest in this text. To discuss it in some more detail, we first consider a map $\mathrm{pr}_{\mathfrak{m}_X} \colon \mathfrak{so}(n)\times\mathfrak{so}(k) \to \mathfrak{so}(n)\times\mathfrak{so}(k)$ which is implicitly defined in [27, Lem. 3.2] by taking the limit $\alpha \to \infty$.

Definition 6.3 *Let $X \in \mathrm{St}_{n,k}$ be arbitrary. Define the map*

$$\mathrm{pr}_{\mathfrak{m}_X} : \mathfrak{so}(n) \times \mathfrak{so}(k) \to \mathfrak{so}(n) \times \mathfrak{so}(k), \quad (\Omega, \Psi) \mapsto (\Omega^{\perp_X}, \Psi^{\perp_X}) \tag{6.11}$$

by

$$\begin{aligned} \Omega^{\perp_X} &= XX^\top \Omega + \Omega XX^\top - 2XX^\top \Omega XX^\top, \\ \Psi^{\perp_X} &= \Psi - X^\top \Omega X, \end{aligned} \tag{6.12}$$

for all $(\Omega, \Psi) \in \mathfrak{so}(n) \times \mathfrak{so}(k)$.

The subscript of $\mathrm{pr}_{\mathfrak{m}_X}$ in Definition 6.3 coincides with the notation for its image, see Definition 6.6 below.

According to [27, Sec. 4.3], at the point $I_{n,k} = \left[\begin{smallmatrix} I_k \\ 0 \end{smallmatrix}\right] \in \mathrm{St}_{n,k}$, one has the decomposition of

$$(\Omega, \Psi) = \left(\begin{bmatrix} A & -B^\top \\ B & C \end{bmatrix}, \Psi \right) \in \mathfrak{so}(n) \times \mathfrak{so}(k), \quad A \in \mathfrak{so}(n), C \in \mathfrak{so}(n-k), B \in \mathbb{R}^{(n-k)\times k} \tag{6.13}$$

given by

$$(\Omega, \Psi) = \left(\left[\begin{smallmatrix} A & 0 \\ 0 & C \end{smallmatrix}\right], A \right) + \left(\begin{bmatrix} 0 & -B^\top \\ B & 0 \end{bmatrix}, \Psi - A \right) \tag{6.14}$$

with the first summand being in $\mathfrak{h}_{I_{n,k}}$ and the second summand being in $\mathfrak{m}_{I_{n,k}}$. Here $\mathfrak{h}_{I_{n,k}}$ and $\mathfrak{m}_{I_{n,k}}$ are parameterized by

$$\mathfrak{h}_{I_{n,k}} = \left\{ \left(\left[\begin{smallmatrix} C & 0 \\ 0 & D \end{smallmatrix}\right], C \right) \,\middle|\, C \in \mathfrak{so}(k), D \in \mathfrak{so}(n-k) \right\} \tag{6.15}$$

and

$$\mathfrak{m}_{I_{n,k}} = \left\{ \left(\begin{bmatrix} 0 & -B^\top \\ B & 0 \end{bmatrix}, A \right) \,\middle|\, B \in \mathbb{R}^{(n-k)\times k}, A \in \mathfrak{so}(k) \right\}. \tag{6.16}$$

In other words,

$$\mathrm{pr}_{\mathfrak{m}_{I_{n,k}}} \left(\begin{bmatrix} A & -B^\top \\ B & C \end{bmatrix}, \Psi \right) = \left(\begin{bmatrix} 0 & -B^\top \\ B & 0 \end{bmatrix}, \Psi - A \right) \tag{6.17}$$

holds by (6.14).

Remark 6.4 The subspace $\mathfrak{m}_{I_{n,k}} \subseteq \mathfrak{so}(n) \times \mathfrak{so}(k)$ is used in [30, Sec. 3.2.1] to define the "quasi-geodesic horizontal distribution" on $\mathrm{O}(n) \times \mathrm{O}(k)$ by left-translation, i.e. by setting

$$\mathrm{Hor}\big(\mathrm{O}(n) \times \mathrm{O}(k)\big)_{(R,\theta)} = T_e \ell_{(R,\theta)} \mathfrak{m}_{I_{n,k}} = (R, \theta)\mathfrak{m}_{I_{n,k}} \subseteq T_{(R,\theta)}(\mathrm{O}(n) \times \mathrm{O}(k)) \tag{6.18}$$

for all $(R, \theta) \in \mathrm{O}(n) \times \mathrm{O}(k)$.

Lemma 6.5 *The decomposition $\mathfrak{so}(n) \times \mathfrak{so}(k) = \mathfrak{h}_{I_{n,k}} \oplus \mathfrak{m}_{I_{n,k}}$ is reductive.*

PROOF: Clearly, $\mathfrak{h}_{I_{n,k}} \oplus \mathfrak{m}_{I_{n,k}} = \mathfrak{so}(n) \times \mathfrak{so}(k)$ is a direct sum. This is observed for example in [30, Sec. 3.2.1] and [27, Sec. 3]. Next let

$$h = \left(\begin{bmatrix} R_1 & 0 \\ 0 & R_2 \end{bmatrix}, R_1 \right) \in H_{I_{n,k}} \quad \text{and} \quad \xi = \left(\begin{bmatrix} 0 & -B^\top \\ B & 0 \end{bmatrix}, A \right) \in \mathfrak{m}_{I_{n,k}}, \tag{6.19}$$

i.e. $R_1 \in \mathrm{O}(k)$, $R_2 \in \mathrm{O}(n-k)$, $B \in \mathbb{R}^{(n-k)\times k}$ and $A \in \mathfrak{so}(k)$. We calculate

$$\begin{aligned} \mathrm{Ad}_h(\xi) &= \left(\begin{bmatrix} R_1 & 0 \\ 0 & R_2 \end{bmatrix}, R_1 \right) \left(\begin{bmatrix} 0 & -B^\top \\ B & 0 \end{bmatrix}, A \right) \left(\begin{bmatrix} R_1 & 0 \\ 0 & R_2 \end{bmatrix}^\top, R_1^\top \right) \\ &= \left(\begin{bmatrix} 0 & -R_1 B^\top R_2^\top \\ R_2 B R_1^\top & 0 \end{bmatrix}, R_1 A R_1^\top \right) \in \mathfrak{m}_{I_{n,k}} \end{aligned} \tag{6.20}$$

proving the inclusion $\mathrm{Ad}_h(\mathfrak{m}_{I_{n,k}}) \subseteq \mathfrak{m}_{I_{n,k}}$ for all $h \in H_{I_{n,k}}$. This yields the desired result.□

Next we consider the more general situation, where $X \in \mathrm{St}_{n,k}$ is arbitrary.

Definition 6.6 *Let* $X \in \mathrm{St}_{n,k}$ *and define*

$$\mathfrak{m}_X = \operatorname{im}\big(\operatorname{pr}_{\mathfrak{m}_X}\big) \subseteq \mathfrak{so}(n) \times \mathfrak{so}(k). \tag{6.21}$$

Clearly, for each $X \in \mathrm{St}_{n,k}$, $\mathfrak{m}_X$ is a subspace of $\mathfrak{so}(n) \times \mathfrak{so}(k)$.

Lemma 6.7 *Let* $X \in \mathrm{St}_{n,k}$. *Then*

$$\mathfrak{so}(n) \times \mathfrak{so}(k) = \mathfrak{h}_X \oplus \mathfrak{m}_X \tag{6.22}$$

is a reductive decomposition turning $(\mathrm{O}(n)\times\mathrm{O}(k))/H_X$ *into a reductive homogeneous space.*

Proof: The decomposition $\mathfrak{so}(n) \times \mathfrak{so}(k) = \mathfrak{h}_{I_{n,k}} \oplus \mathfrak{m}_{I_{n,k}}$ is a reductive decomposition by Lemma 6.5. Since $\mathrm{O}(n) \times \mathrm{O}(k)$ acts transitively on $\mathrm{St}_{n,k}$, each $X \in \mathrm{St}_{n,k}$ can be written as $X = \Phi_{(R,\theta)}(I_{n,k}) = RI_{n,k}\theta^\top$ for some $(R, \theta) \in \mathrm{O}(n) \times \mathrm{O}(k)$. One obtains for $(\Omega, \Psi) \in \mathfrak{h}_{I_{n,k}} = \ker\big(T_{(I_n,I_k)}\Phi(\cdot, I_{n,k})\big)$

$$\begin{aligned} T_{(I_n,I_k)}\Phi(\cdot, X)\big(\operatorname{Ad}_{(R,\theta)}(\Omega, \Psi)\big) &= (R\Omega R^\top)X - X(\theta\Psi\theta^\top) \\ &= (R\Omega R^\top)(RI_{n,k}\theta^\top) - (RI_{n,k}\theta^\top)(\theta\Psi\theta^\top) \\ &= R(\Omega I_{n,k} - I_{n,k}\Psi)\theta^\top \\ &= 0. \end{aligned} \tag{6.23}$$

Thus $\operatorname{Ad}_{(R,\theta)}(\mathfrak{h}_{I_{n,k}}) \subseteq \ker\big(T_{(I_n,I_k)}\Phi(\cdot, X)\big) = \mathfrak{h}_X$ follows. By counting dimensions, this inclusion is in fact an equality, i.e. $\operatorname{Ad}_{(R,\theta)}(\mathfrak{h}_{I_{n,k}}) = \mathfrak{h}_X$. Next, let $(\Omega, \Psi) \in \mathfrak{so}(n) \times \mathfrak{so}(k)$. By a straightforward calculation, one obtains

$$\operatorname{pr}_{\mathfrak{m}_{RX\theta^\top}}\big(\operatorname{Ad}_{(R,\theta)}(\Omega, \Psi)\big) = \operatorname{Ad}_{(R,\theta)}\big(\operatorname{pr}_{\mathfrak{m}_X}(\Omega, \Psi)\big). \tag{6.24}$$

For $X = RI_{n,k}\theta^\top$, Equation (6.24) yields

$$\mathfrak{m}_X = \mathfrak{m}_{RI_{n,k}\theta^\top} = \operatorname{Ad}_{(R,\theta)}\big(\operatorname{im}(\operatorname{pr}_{\mathfrak{m}_{I_{n,k}}})\big) = \operatorname{Ad}_{(R,\theta)}\big(\mathfrak{m}_{I_{n,k}}\big). \tag{6.25}$$

Since $\operatorname{Ad}_{(R,\theta)} \colon \mathfrak{so}(n) \times \mathfrak{so}(k) \to \mathfrak{so}(n) \times \mathfrak{so}(k)$ is a linear isomorphism, we obtain the direct sum

$$\mathfrak{so}(n) \times \mathfrak{so}(k) = \operatorname{Ad}_{(R,\theta)}(\mathfrak{h}_{I_{n,k}}) \oplus \operatorname{Ad}_{(R,\theta)}(\mathfrak{m}_{I_{n,k}}) = \mathfrak{h}_X \oplus \mathfrak{m}_X, \tag{6.26}$$

where we used that $\mathfrak{so}(n) \times \mathfrak{so}(k) = \mathfrak{h}_{I_{n,k}} \oplus \mathfrak{m}_{I_{n,k}}$ holds. It remains to show that this decomposition turns $(\mathrm{O}(n) \times \mathrm{O}(k))/H_X$ into a reductive homogeneous space. To this end, let $(\widetilde{R}, \widetilde{\theta}) \in H_X$, i.e. $\Phi_{(\widetilde{R},\widetilde{\theta})}(X) = \widetilde{R}X\widetilde{\theta}^\top = X$ is satisfied. By (6.24), we obtain

$$\operatorname{Ad}_{(\widetilde{R},\widetilde{\theta})}(\mathfrak{m}_X) \overset{(6.24)}{=} \mathfrak{m}_{\widetilde{R}X\widetilde{\theta}^\top} = \mathfrak{m}_X \tag{6.27}$$

as desired. □

Notation 6.8 From now on, if the point $X \in \mathrm{St}_{n,k}$ is clear by the context, we suppress the subscript X in the notation, i.e. we write $H = H_X$ and denote $\mathfrak{h}_X$ and $\mathfrak{m}_X$ by $\mathfrak{h}$ and $\mathfrak{m}$, respectively.

Lemma 6.9 *Let $X \in \mathrm{St}_{n,k}$. The inverse of the linear isomorphism*

$$T_{(I_n,I_k)}(\iota_X \circ \mathrm{pr}_X)\big|_{\mathfrak{m}} = T_{(I_n,I_k)}\Phi(\cdot, X)\big|_{\mathfrak{m}} \colon \mathfrak{m} \ni (\Omega, \Psi) \mapsto \Omega X - X\Psi \in T_X\mathrm{St}_{n,k} \tag{6.28}$$

is given by

$$\big(T_{(I_n,I_k)}(\iota_X \circ \mathrm{pr}_X)\big|_{\mathfrak{m}}\big)^{-1} \colon T_X\mathrm{St}_{n,k} \ni V \mapsto \big(VX^\top - XV^\top + 2XV^\top XX^\top, -X^\top V\big) \in \mathfrak{m}. \tag{6.29}$$

Proof: This statement can be viewed as a reformulation of [32, Prop. 5], see also [27, Prop. 3] after taking the limit $\alpha \to \infty$. Alternatively, one can prove for $V \in T_X\mathrm{St}_{n,k}$

$$\big(T_{(I_n,I_k)}(\iota_X \circ \mathrm{pr}_X)\big|_{\mathfrak{m}} \circ \big(T_{(I_n,I_k)}(\iota_X \circ \mathrm{pr}_X)\big|_{\mathfrak{m}}\big)^{-1}\big)V = V \tag{6.30}$$

and

$$\mathrm{pr}_{\mathfrak{m}}\left(\big(T_{(I_n,I_k)}(\iota_X \circ \mathrm{pr}_X)\big|_{\mathfrak{m}}\big)^{-1}V\right) = \big(T_{(I_n,I_k)}(\iota_X \circ \mathrm{pr}_X)\big|_{\mathfrak{m}}\big)^{-1}V \tag{6.31}$$

as well as

$$\big(T_{(I_n,I_k)}(\iota_X \circ \mathrm{pr}_X)\big|_{\mathfrak{m}}\big)^{-1}\big) \circ \big(T_{(I_n,I_k)}(\iota_X \circ \mathrm{pr}_X)\big|_{\mathfrak{m}}(\xi) = \xi \tag{6.32}$$

for all $\xi \in \mathfrak{m}$ by straightforward computations. □

6.3 Quasi-Geodesics on Stiefel Manifolds

Next we consider special curves on $\mathrm{St}_{n,k}$, the so-called quasi-geodesics introduced in [32] for solving interpolation problems on Stiefel manifolds. An alternative formula which is more efficient from a computational point of view has been derived in [6]. These curves are also discussed in [27, 30].

Before we continue, we point out that the notations $\exp(A)$ and e^A are both used for the exponential of a matrix $A \in \mathbb{R}^{n \times n}$.

By [27, Sec. 5.3], the quasi-geodesic through $X \in \mathrm{St}_{n,k}$ with initial velocity $V \in T_X\mathrm{St}_{n,k}$ is given by

$$\gamma \colon \mathbb{R} \to \mathrm{St}_{n,k}, \quad t \mapsto \mathrm{e}^{t(VX^\top - XV^\top + 2XV^\top XX^\top)} X \mathrm{e}^{tX^\top V}. \tag{6.33}$$

Clearly, (6.33) is in accordance with the definition of a quasi-geodesic in [30, Sec. 3.3.3], see also [32, Prop. 6] combined with [32, Prop. 5]. In those references, a quasi-geodesic is defined as a projection of a one-parameter subgroup in $\mathrm{O}(n) \times \mathrm{O}(k)$ which is horizontal with respect to the quasi-geodesic distribution considered in Remark 6.4. Indeed, defining the curve

$$g \colon \mathbb{R} \ni t \mapsto \big(\mathrm{e}^{t(VX^\top - XV^\top + 2XV^\top XX^\top)}, \mathrm{e}^{-tX^\top V}\big) \in \mathrm{O}(n) \times \mathrm{O}(k), \tag{6.34}$$

which is horizontal with respect to the distribution from Remark 6.4, one has $\gamma = \iota_X \circ \mathrm{pr}_X \circ g$ as desired.

Next we list some more properties of quasi-geodesics.

Remark 6.10 Let $\gamma \colon I \to \mathrm{St}_{n,k}$ be the quasi-geodesic from (6.33), where $X \in \mathrm{St}_{n,k}$ and $V \in T_X\mathrm{St}_{n,k}$ are fixed. Then the following assertions are fulfilled:

1. According to [27, Sec. 5.3], γ is a solution of the initial value problem

$$\begin{aligned} &\ddot{\gamma}(t) = -2(\dot{\gamma}(t)\dot{\gamma}(t)^\top \gamma(t)) - \gamma(t)\big(2(\gamma(t)^\top \dot{\gamma}(t))^2 + \dot{\gamma}(t)^\top \dot{\gamma}(t)\big), \\ &\gamma(0) = X, \quad \dot{\gamma}(0) = V. \end{aligned} \tag{6.35}$$

2. The curve $\mathbb{R} \ni t \mapsto (\mathrm{pr}_X \circ g)(t) \in (\mathrm{O}(n) \times \mathrm{O}(k))/H_X$ being the projection of the horizontal one-parameter subgroup g from (6.34) is a geodesic on $(\mathrm{O}(n) \times \mathrm{O}(k))/H_X$ with respect to the canonical invariant covariant derivative of second kind $\nabla^{\mathrm{can}2}$. Hence $\gamma\colon \mathbb{R} \to \mathrm{St}_{n,k}$ can be considered as a geodesic with respect to $\nabla^{\mathrm{can}2}$ by identifying $\mathrm{St}_{n,k} \cong (\mathrm{O}(n) \times \mathrm{O}(k))/H_X$ via ι_X. To be more precise, the quasi-geodesic $\gamma\colon \mathbb{R} \to \mathrm{St}_{n,k}$ is the geodesic through $\gamma(0) = X$ with initial velocity $\dot{\gamma}(0) = V$ with respect the pull-back covariant derivative $\big(\iota_X^{-1}\big)^* \nabla^{\mathrm{can}2}$, see (4.8) above, where this definition is recalled.

3. Let $X \in \mathrm{St}_{n,k}$. By [32, Prop. 6], the map

$$\mathrm{Exp}_X \colon T_X\mathrm{St}_{n,k} \to \mathrm{St}_{n,k}, \quad V \mapsto \mathrm{e}^{VX^\top - XV^\top + 2XV^\top XX^\top} X \mathrm{e}^{X^\top V} = \gamma(1), \tag{6.36}$$

where γ is the quasi-geodesic that satisfies $\gamma(0) = X$ and $\dot{\gamma}(0) = V$, yields the retraction

$$T\mathrm{St}_{n,k} \ni (X, V) \mapsto \mathrm{Exp}_X(V) \in \mathrm{St}_{n,k} \tag{6.37}$$

in the sense of [2, Sec. 3], see also [32, Def. 4]. For more details on retractions we refer to [1, Chap. 4] in the context of optimization and also to [3] for applications to numerical integration of ODEs on manifolds. In particular, the map (6.36) is a local diffeomorphism on some suitable open neighbourhood of $0 \in T_X\mathrm{St}_{n,k}$ which satisfies $\mathrm{Exp}_X(0) = X$ and $T_0\mathrm{Exp}_X = \mathrm{id}_{T_X\mathrm{St}_{n,k}}$ up to the identification $T_0(T_X\mathrm{St}_{n,k}) \cong T_X\mathrm{St}_{n,k}$.

4. In view of Claim 2, the map $\mathrm{Exp}_X \colon T_X\mathrm{St}_{n,k} \to \mathrm{St}_{n,k}$ from (6.36) is not only some retraction but the exponential map associated with the covariant derivative $\big(\iota_X^{-1}\big)^* \nabla^{\mathrm{can}2}$ in the sense of [48, Sec. 15.1] since it fulfills $\mathrm{Exp}_X(V) = \gamma(1)$, where $\gamma\colon \mathbb{R} \to \mathrm{St}_{n,k}$ is the geodesic with respect to $\big(\iota_X^{-1}\big)^* \nabla^{\mathrm{can}2}$ through $\gamma(0) = X$ with initial velocity $\dot{\gamma}(0) = V$. This justifies to denote it by "Exp_X".

5. By Proposition 6.19 below, the map defined in (6.36) is surjective for all $X \in \mathrm{St}_{n,k}$. Moreover, a method for computing some $V \in T_X\mathrm{St}_{n,k}$ with $\mathrm{Exp}_X(V) = Y$ for an *arbitrary* $Y \in \mathrm{St}_{n,k}$ is given by Algorithm 5 below.

We now take a closer look at quasi-geodesics in order to apply Algorithm 1 to $\mathrm{St}_{n,k}$. Here we focus on formulas that are efficient from a computational point of view for $k \ll n$. The following discussion is inspired by the work [6]. In the sequel, some of the results in [6] are recalled and generalized.

We start by deriving an alternative formula for the expression

$$\mathrm{e}^{t(VX^\top - XV^\top + 2XV^\top XX^\top)} Y, \tag{6.38}$$

where $X \in \mathrm{St}_{n,k}$, $V \in T_X\mathrm{St}_{n,k}$, $t \in \mathbb{R}$ and $Y \in \mathbb{R}^{n \times k}$. As preparation, we state the following lemmas concerning some probably well-known facts on the singular value decomposition (SVD). Here we use the terminology concerning "thin" and "compact" SVDs from [7, Sec. 3.4].

Notation 6.11 Let $k \leq n$ and let $A \in \mathbb{R}^{n \times k}$ be a matrix of rank $r = \mathrm{rank}(A)$.

1. A decomposition $A = Q\Sigma S^\top$, where $Q \in \mathrm{St}_{n,r}$, $S \in \mathrm{St}_{k,r}$ and $\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_r) \in \mathbb{R}^{r \times r}$ is diagonal, is called a thin SVD of A.

2. A decomposition $A = \widehat{Q}\widehat{\Sigma}\widehat{S}^\top$, where $\widehat{Q} \in \mathrm{St}_{n,k}$, $\widehat{S} \in \mathrm{O}(k)$ and $\widehat{\Sigma} = \mathrm{diag}(\widehat{\sigma}_1, \ldots, \widehat{\sigma}_k) \in \mathbb{R}^{k \times k}$ is diagonal, is called a compact SVD of A.

Neither 1 nor 2 imposes any ordering or non-negative assumption on the diagonal entries of the matrices $\Sigma \in \mathbb{R}^{r\times r}$ and $\widehat{\Sigma} \in \mathbb{R}^{k\times k}$. If we assume that the singular values are arranged in a specific way or are non-negative, this will be indicated explicitly. Clearly, if $\operatorname{rank}(A) = k$, a thin SVD of A is also a compact SVD of A and vice versa.

Lemma 6.12 *Let $X \in \mathrm{St}_{n,k}$, $Y \in \mathbb{R}^{n\times k}$ and let $(I_n - XX^\top)Y = \widehat{Q}\widehat{\Sigma}\widehat{S}^\top$ be a compact SVD, where $\widehat{\Sigma} = \operatorname{diag}(\widehat{\sigma}_1, \ldots, \widehat{\sigma}_k) \in \mathbb{R}^{k\times k}$. If $\widehat{\sigma}_j \neq 0$ holds for some $j \in \{1, \ldots, k\}$, then $(X^\top \widehat{Q})_{ij} = 0$ is fulfilled for all $i \in \{1, \ldots, k\}$. In particular, $(I_n - XX^\top)\widehat{Q} = \widehat{Q}$ holds if $(I_n - XX^\top)Y$ has full rank.*

PROOF: We obtain for $X \in \mathrm{St}_{n,k}$ by $X^\top X = I_k$ and $\widehat{S}^\top \widehat{S} = I_k$

$$0 = X^\top (I_n - XX^\top)Y = X^\top(\widehat{Q}\widehat{\Sigma}\widehat{S}^\top) \iff X^\top \widehat{Q}\widehat{\Sigma} = 0\widehat{S} = 0. \tag{6.39}$$

The equality on the right hand side of (6.39) can be rewritten component-wise for $i, j \in \{1, \ldots, k\}$ by

$$(X^\top \widehat{Q}\widehat{\Sigma})_{ij} = \sum_{\ell=1}^{k} (X^\top \widehat{Q})_{i\ell}\widehat{\Sigma}_{\ell j} = (X^\top \widehat{Q})_{ij}\widehat{\sigma}_j = 0. \tag{6.40}$$

Thus $\widehat{\sigma}_j \neq 0$ implies $(X^\top \widehat{Q})_{ij} = 0$. This shows the first claim.

We now assume $\operatorname{rank}\big((I_n - XX^\top)Y\big) = k$. Then $\widehat{\sigma}_j \neq 0$ holds for all $j \in \{1, \ldots, k\}$ implying $(X^\top \widehat{Q}) = 0$ due to (6.40). In addition, we obtain $(I_n - XX^\top)\widehat{Q} = \widehat{Q} - 0 = \widehat{Q}$.□

Lemma 6.13 *Let $X \in \mathrm{St}_{n,k}$, $Y \in \mathbb{R}^{n\times k}$ and let $(I_n - XX^\top)Y = Q\Sigma S^\top$ be a thin SVD. Then $X^\top Q = 0 \in \mathbb{R}^{k\times r}$ and $(I_n - XX^\top)Q = Q$ holds, where $r = \operatorname{rank}\big((I_n - XX^\top)Y\big)$.*

PROOF: By $X^\top X = I_k$, one calculates $0 = X^\top(I_n - XX^\top)Y = X^\top(Q\Sigma S^\top)$. Since Σ is invertible and $S^\top S = I_r$ holds by assumption, we obtain

$$X^\top Q\Sigma S^\top = 0 \implies X^\top Q\Sigma S^\top S = X^\top Q\Sigma = 0 \implies X^\top Q = 0 \tag{6.41}$$

as desired. Clearly, $X^\top Q = 0$ yields $(I_n - XX^\top)Q = Q - 0 = Q$. □

After this preparation, we derive an alternative expression for (6.38).

Lemma 6.14 *Let $X \in \mathrm{St}_{n,k}$, $V \in T_X\mathrm{St}_{n,k}$, $t \in \mathbb{R}$ and let $Y \in \mathbb{R}^{n\times k}$. Let $Q\Sigma S^\top = (I_n - XX^\top)V$ be a thin SVD of $(I_n - XX^\top)V$. Then*

$$\begin{aligned}
&\mathrm{e}^{t(VX^\top - XV^\top + 2XV^\top XX^\top)}Y \\
&= X\Big(S\cos(t\Sigma)S^\top X^\top Y + (I_k - SS^\top)X^\top Y - S\sin(t\Sigma)Q^\top Y\Big) \\
&\quad + Q\Big(\sin(t\Sigma)S^\top X^\top Y + \cos(t\Sigma)Q^\top Y + Q^\top Y\Big)
\end{aligned} \tag{6.42}$$

holds.

PROOF: Let $X_\perp \in \mathrm{St}_{n,n-k}$ such that $[X \mid X_\perp] \in \mathrm{O}(n)$ holds. Clearly $X^\top X_\perp = 0$. Using ideas from [54, Sec. 3.1], we set

$$B = X_\perp^\top (I_n - XX^\top)V = X_\perp^\top V \in \mathbb{R}^{(n-k)\times k} \tag{6.43}$$

and calculate

$$\begin{aligned}
&\mathrm{e}^{t(VX^\top - XV^\top + 2XV^\top XX^\top)}Y \\
&= [X \mid X_\perp][X \mid X_\perp]^\top \mathrm{e}^{t(VX^\top - XV^\top) + 2XV^\top XX^\top}[X \mid X_\perp][X \mid X_\perp]^\top Y \\
&= [X \mid X_\perp]\exp\left(t\begin{bmatrix} 0 & -B^\top \\ B & 0\end{bmatrix}\right)\begin{bmatrix} X^\top Y \\ X_\perp^\top Y\end{bmatrix}.
\end{aligned} \tag{6.44}$$

Next, we take a closer look at the matrix exponential

$$\exp\Big(t\begin{bmatrix} 0 & -B^\top \\ B & 0 \end{bmatrix}\Big). \tag{6.45}$$

To this end, we use ideas from [7, Sec. 3.4]. Let $Q\Sigma S^\top = (I_n - XX^\top)V$ be a thin SVD. In particular, we have $Q \in \mathrm{St}_{n,r}$, $\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_r)$ and $S \in \mathrm{St}_{k,r}$, where $r = \operatorname{rank}\big((I_n - XX^\top)V\big) \leq k$. Moreover, we define

$$W = X_\perp^\top Q \in \mathbb{R}^{(n-k)\times r}. \tag{6.46}$$

Using $X^\top Q = 0$ by Lemma 6.13, we obtain

$$W^\top W = Q^\top X_\perp X_\perp^\top Q = Q^\top (I_n - XX^\top)Q = Q^\top Q - Q^\top XX^\top Q = I_r, \tag{6.47}$$

i.e. $W \in \mathrm{St}_{n-k,r}$ is fulfilled. Moreover, let $W_\perp \in \mathrm{St}_{n-k,n-k-r}$ be a matrix such that $[W \mid W_\perp] \in \mathrm{O}(n-k)$ holds and let $S_\perp \in \mathrm{St}_{k,k-r}$ be a matrix fulfilling $[S \mid S_\perp] \in \mathrm{O}(k)$. By this notation, see also [18, Sec. 23.9] for a similar computation, we obtain

$$\begin{bmatrix} S & S_\perp & 0 & 0 \\ 0 & 0 & W & W_\perp \end{bmatrix}\begin{bmatrix} 0 & 0 & -t\Sigma & 0 \\ 0 & 0 & 0 & 0 \\ t\Sigma & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}\begin{bmatrix} S^\top & 0 \\ S_\perp^\top & 0 \\ 0 & W^\top \\ 0 & W_\perp^\top \end{bmatrix} = \begin{bmatrix} 0 & -tS\Sigma W^\top \\ tW\Sigma S^\top & 0 \end{bmatrix} = \begin{bmatrix} 0 & -tB^\top \\ tB & 0 \end{bmatrix}, \tag{6.48}$$

where we used $B = X_\perp^\top (I_n - XX^\top)V = X_\perp^\top Q\Sigma S^\top = W\Sigma S^\top$. Using (6.48), we calculate

$$\begin{aligned} \exp\Big(t\begin{bmatrix} 0 & -B^\top \\ B & 0 \end{bmatrix}\Big) &= \begin{bmatrix} S & S_\perp & 0 & 0 \\ 0 & 0 & W & W_\perp \end{bmatrix}\begin{bmatrix} \cos(t\Sigma) & 0 & -\sin(t\Sigma) & 0 \\ 0 & I_{k-r} & 0 & 0 \\ \sin(t\Sigma) & 0 & \cos(t\Sigma) & 0 \\ 0 & 0 & 0 & I_{n-k-r} \end{bmatrix}\begin{bmatrix} S^\top & 0 \\ S_\perp^\top & 0 \\ 0 & W^\top \\ 0 & W_\perp^\top \end{bmatrix} \\ &= \begin{bmatrix} S\cos(t\Sigma)S^\top + S_\perp S_\perp^\top & -S\sin(t\Sigma)W^\top \\ W\sin(t\Sigma)S^\top & W\cos(t\Sigma)W^\top + W_\perp W_\perp^\top \end{bmatrix} \end{aligned} \tag{6.49}$$

as in the proof of [7, Prop. 3.3]. Next we derive an alternative expression for the last line of (6.44). To this end, we note that

$$X_\perp W = X_\perp (X_\perp^\top Q) = (I_n - XX^\top)Q = Q \tag{6.50}$$

holds by Lemma 6.13, and consequently $W^\top X_\perp^\top = (X_\perp W)^\top = Q^\top$ is fulfilled. Using (6.50) and $SS_\perp^\top = I_k - SS^\top$, we obtain by (6.49)

$$\begin{aligned} &[X \mid X_\perp]\exp\Big(t\begin{bmatrix} 0 & -B^\top \\ B & 0 \end{bmatrix}\Big)\begin{bmatrix} X^\top Y \\ X_\perp^\top Y \end{bmatrix} \\ &= [X \mid X_\perp]\begin{bmatrix} S\cos(t\Sigma)S^\top + S_\perp S_\perp^\top & -S\sin(t\Sigma)W^\top \\ W\sin(t\Sigma)S^\top & W\cos(t\Sigma)W^\top + W_\perp W_\perp^\top \end{bmatrix}\begin{bmatrix} X^\top Y \\ X_\perp^\top Y \end{bmatrix} \\ &= [X \mid X_\perp]\begin{bmatrix} S\cos(t\Sigma)S^\top X^\top Y + S_\perp S_\perp^\top X^\top Y - S\sin(t\Sigma)W^\top X_\perp^\top Y \\ W\sin(t\Sigma)S^\top X^\top Y + W\cos(t\Sigma)W^\top X_\perp^\top Y + W_\perp W_\perp^\top X_\perp^\top Y \end{bmatrix} \\ &= X\Big(S\cos(t\Sigma)S^\top X^\top Y - S\sin(t\Sigma)(W^\top X_\perp^\top)Y + S_\perp S_\perp^\top X^\top Y\Big) \\ &\quad + (X_\perp W)\sin(t\Sigma)S^\top X^\top Y + (X_\perp W)\cos(t\Sigma)(W^\top X_\perp^\top)Y \\ &\quad + ((X_\perp W_\perp)(W_\perp^\top X_\perp^\top))Y \\ &= X\Big(S\cos(t\Sigma)S^\top X^\top Y + S_\perp S_\perp^\top X^\top Y - S\sin(t\Sigma)Q^\top Y\Big) \\ &\quad + Q\sin(t\Sigma)S^\top X^\top Y + Q\cos(t\Sigma)Q^\top Y + QQ^\top Y\Big) \\ &= X\Big(S\cos(t\Sigma)S^\top X^\top Y + (I_k - SS^\top)X^\top Y - S\sin(t\Sigma)Q^\top Y\Big) \\ &\quad + Q\Big(\sin(t\Sigma)S^\top X^\top Y + \cos(t\Sigma)Q^\top Y + Q^\top Y\Big). \end{aligned} \tag{6.51}$$

Comparing (6.51) with (6.44) yields the desired result. □

Lemma 6.14 implies the next corollary on quasi-geodesics. We point out that the first and third expression in (6.52) below already appeared in [27, Sec. 5.3] and [6, Prop. 1], respectively.

Corollary 6.15 *Let $X \in \mathrm{St}_{n,k}$ and $V \in T_X\mathrm{St}_{n,k}$. Moreover, let $Q\Sigma S^\top = (I_n - XX^\top)V$ be a thin SVD and let $\widehat{Q}\widehat{\Sigma}\widehat{S}^\top = (I_n - XX^\top)V$ be a compact SVD. Then the quasi-geodesic $\gamma\colon \mathbb{R} \to \mathrm{St}_{n,k}$ starting at $\gamma(0) = X$ with initial velocity $\dot{\gamma}(0) = V$ is given by*

$$\begin{aligned}\gamma(t) &= \mathrm{e}^{t(VX^\top - XV^\top + 2XV^\top XX^\top)} X \mathrm{e}^{tX^\top V} \\ &= \big(\big(XS\cos(t\Sigma)S^\top + Q\sin(t\Sigma)\big)S^\top + X(I_k - SS^\top)\big)\exp(tX^\top V) \\ &= \big(X\widehat{S}\cos(t\widehat{\Sigma}) + \widehat{Q}\sin(t\widehat{\Sigma})\big)\widehat{S}^\top \exp(tX^\top V)\end{aligned} \tag{6.52}$$

for $t \in \mathbb{R}$.

Proof: The first equality in (6.52) already appeared in [27, Sec. 5.3]. The second equality follows by Lemma 6.14 and Lemma 6.13. The third equality of (6.52) coincides with the expression from [6, Prop. 1]. But we give next an alternative proof using the second equality. Let $r = \operatorname{rank}\big((I_n - XX^\top)V\big)$. Without loss of generality, we assume $\widehat{Q} = [Q \mid Q_\perp] \in \mathrm{St}_{n,k}$, $\widehat{\Sigma} = \operatorname{diag}(\Sigma, 0_{k-r}) \in \mathbb{R}^{k\times k}$ and $\widehat{S} = [S \mid S_\perp] \in \mathrm{O}(k)$ for some suitable $Q_\perp \in \mathrm{St}_{n,k-r}$ and $S_\perp \in \mathrm{St}_{k,k-r}$. By this notation and using $(I_k - SS^\top) = S_\perp S_\perp^\top$, we compute

$$\begin{aligned}&\big(X\widehat{S}\cos(t\widehat{\Sigma}) + \widehat{Q}\sin(t\widehat{\Sigma})\big)\widehat{S}^\top \\ &= X[S \mid S_\perp]\begin{bmatrix}\cos(t\Sigma) & 0 \\ 0 & I_{k-r}\end{bmatrix}\begin{bmatrix}S^\top \\ S_\perp^\top\end{bmatrix} + [Q \mid Q_\perp]\begin{bmatrix}\sin(t\Sigma) & 0 \\ 0 & 0_{k-r}\end{bmatrix}\begin{bmatrix}S^\top \\ S_\perp^\top\end{bmatrix} \\ &= X(S\cos(t\Sigma)S^\top + S_\perp S_\perp^\top) + Q\sin(t\Sigma)S^\top \\ &= XS\cos(t\Sigma)S^\top + Q\sin(t\Sigma)S^\top + X(I_k - SS^\top).\end{aligned}$$

This yields the desired result. □

Corollary 6.16 *Let $X \in \mathrm{St}_{n,k}$ and $V \in T_X\mathrm{St}_{n,k}$. Define the curve*

$$g = (R, \theta)\colon \mathbb{R} \to \mathrm{O}(n) \times \mathrm{O}(k), \quad t \mapsto \big(\mathrm{e}^{t(VX^\top - XV^\top + 2XV^\top XX^\top)}, \mathrm{e}^{-tX^\top V}\big). \tag{6.53}$$

Moreover, let $Y \in \mathbb{R}^{n\times k}$ and let $Q\Sigma S^\top = (I_n - XX^\top)V$ be a thin SVD. Then

$$\begin{aligned}\Phi_{g(t)}(Y) = \Bigg(&X\Big(S\cos(t\Sigma)S^\top X^\top Y + (I_k - SS^\top)X^\top Y - S\sin(t\Sigma)Q^\top Y\Big) \\ &+ Q\Big(\sin(t\Sigma)S^\top X^\top Y + \cos(t\Sigma)Q^\top Y + Q^\top Y\Big)\Bigg)\mathrm{e}^{tX^\top V}\end{aligned} \tag{6.54}$$

holds for all $t \in \mathbb{R}$.

Proof: Clearly, one obtains

$$\Phi_{g(t)}(Y) = \mathrm{e}^{t(VX^\top - XV^\top + 2XV^\top XX^\top)} Y \mathrm{e}^{tX^\top V} \tag{6.55}$$

by the definition of $\Phi\colon (\mathrm{O}(n) \times \mathrm{O}(k)) \times \mathbb{R}^{n\times k} \ni ((R,\theta), X) \mapsto RX\theta^\top \in \mathbb{R}^{n\times k}$ in (6.3) by exploiting $X^\top V = -V^\top X$ for $X \in \mathrm{St}_{n,k}$ and $V \in T_X\mathrm{St}_{n,k}$. Applying Lemma 6.14 to (6.55) yields the desired result. □

Moreover, we need the "logarithm" associated with the quasi-geodesics starting at a point $X \in \mathrm{St}_{n,k}$, i.e. the inverse of the map $\mathrm{Exp}_X \colon T_X\mathrm{St}_{n,k} \to \mathrm{St}_{n,k}$ from Remark 6.10, 3. A procedure to compute such a "logarithm" is proposed in Algorithm 5 below.

Remark 6.17 Algorithm 5, is obtained by modifying [6, Alg. 1, Prop. 2]. In more detail, we adapt [6, Alg. 1, Prop. 2] such that for *all* $Y \in \mathrm{St}_{n,k}$, an element $V \in T_X\mathrm{St}_{n,k}$ with $\mathrm{Exp}_X(V) = Y$ is returned by Algorithm 5. This is in contrast to [6, Alg. 1]. In particular, that algorithm does not work for pairs $(X, Y) \in \mathrm{St}_{n,k}$, where $Y = XR$ for any $R \in \mathrm{O}(k)$ with $\det(R) = -1$. In this context, we also mention that the method for computing a quasi-geodesic joining $X \in \mathrm{St}_{n,k}$ with $Y \in \mathrm{St}_{n,k}$ from [32] imposes for fixed X some constraints on $Y \in \mathrm{St}_{n,k}$, see in particular [32, Thm. 7] and [32, Re. 9].

Algorithm 5 Logarithm of the quasi-geodesic exponential

Input: $(X, Y) \in \mathrm{St}_{n,k} \times \mathrm{St}_{n,k}$.

1. Compute a SVD: $\widetilde{Q}\widetilde{D}\widetilde{R}^\top = Y^\top X$.
2. Define $T = \operatorname{diag}\big(1, \ldots, 1, \det(\widetilde{Q}\widetilde{R}^\top)\big) \in \big(\mathrm{O}(1)\big)^k \subseteq \mathrm{O}(k)$.
3. Set $R = \widetilde{Q}T\widetilde{R}^\top$.
4. Set $Y_* = YR$.
5. Define $A = \log(R^\top)$.
6. Compute a compact SVD: $\widehat{Q}\widehat{D}\widehat{S}^\top = (I_n - XX^\top)Y_*$.
7. If $T = I_k$
   (a) Define $\widehat{\Sigma} = \arcsin(\widehat{D})$.
   (b) Set $\widetilde{V} = \widehat{Q}\widehat{\Sigma}\widehat{S}^\top$.
   Else
   (c) Define $\widehat{\Sigma} = \arcsin\big((I_k - \widetilde{D}^2)^{1/2}\big) - \operatorname{diag}(0, \ldots, 0, \pi)$.
   (d) Set $\widetilde{V} = (\widehat{Q}\widehat{S}^\top \widetilde{R}T)\widehat{\Sigma}\widetilde{R}^\top$.

Output: $V = \widetilde{V} + XA$ (initial velocity of the quasi-geodesic starting at X at $t = 0$ and reaching Y at $t = 1$.)

In Step 5 of Algorithm 5, the definition $A = \log(R^\top)$ is understood to be the principal logarithm of $R^\top$ if $R^\top$ has no eigenvalue equal to -1. Otherwise, some matrix $A \in \mathfrak{so}(k)$ with $\exp(A) = R^\top \in \mathrm{SO}(k)$ is denoted by $\log(R^\top)$.

Notation 6.18 Let $(X, Y) \in \mathrm{St}_{n,k} \times \mathrm{St}_{n,k}$ and let $V \in \mathbb{R}^{n \times k}$ be an output of Algorithm 5 applied to (X, Y). Then we write

$$V = \mathrm{Log}_X(Y). \tag{6.56}$$

This is a slight abuse of notation since there can be some ambiguities. In particular, the results of Algorithm 5, Step 3, and Algorithm 5, Step 5 might be not uniquely determined. Nevertheless, restricted to a suitable open neighbourhood U of X, $\mathrm{Log}_X \colon U \to \mathrm{Log}_X(U) \subseteq T_X\mathrm{St}_{n,k}$ is in fact the inverse of Exp_X, see Proposition 6.19, Claim 2, below.

Proposition 6.19 *Let* $(X,Y) \in \mathrm{St}_{n,k} \times \mathrm{St}_{n,k}$.

1. *Let* $V = \mathrm{Log}_X(Y)$ *be defined by Algorithm 5 applied to* (X,Y). *Then* $V \in T_X\mathrm{St}_{n,k}$ *holds and*
$$Y = \mathrm{Exp}_X(V) = \mathrm{e}^{VX^\top - XV^\top + 2XV^\top XX^\top} X \mathrm{e}^{X^\top V} \tag{6.57}$$
is fulfilled, i.e. $\mathrm{Exp}_X\big(\mathrm{Log}_X(Y)\big) = Y$. *In particular,* $\mathrm{Exp}_X \colon T_X\mathrm{St}_{n,k} \to \mathrm{St}_{n,k}$ *is surjective.*

2. *Assume that* $V \in T_X\mathrm{St}_{n,k}$ *satisfies the following properties:*
   (a) $\|(I_n - XX^\top)V\|_2 < \frac{\pi}{2}$.
   (b) $\|X^\top V\|_2 < \pi$.

   Then $\mathrm{Log}_X\big(\mathrm{Exp}_X(V)\big) = V$ *holds.*

PROOF: We adapt the proof of [6, Prop. 2], and also use ideas from [7, Thm. 5.4] and [7, Thm. 5.5].

Following Algorithm 5, Step 1, let $Y^\top X = \widetilde{Q}\widetilde{D}\widetilde{R}^\top$ be a SVD of $Y^\top X$, where $\widetilde{D} = \mathrm{diag}(\widetilde{\sigma}_1, \ldots, \widetilde{\sigma}_k) \in \mathbb{R}^{k\times k}$ with $\widetilde{\sigma}_1 \geq \cdots \geq \widetilde{\sigma}_k \geq 0$ which is well-known to be non-unique. In more detail, following [24, Thm. 3.1.1'], we write $\widetilde{D} = \mathrm{diag}(\widetilde{D}_1, \ldots, \widetilde{D}_s)$, where $\widetilde{D}_i = \widetilde{\sigma}_i I_{k_i} \in \mathbb{R}^{k_i \times k_i}$ and $k_i \in \mathbb{N}_0$ is the multiplicity of the singular value $\widetilde{\sigma}_i$. Then, every other SVD of $Y^\top X$ with singular values arranged in descending order is given by
$$Y^\top X = \big(\widetilde{Q}\,\mathrm{diag}(\widetilde{R}_1, \ldots, \widetilde{R}_{s-1}, \widetilde{B}_s)\big)\widetilde{D}\big(\widetilde{R}\,\mathrm{diag}(\widetilde{R}_1, \ldots, \widetilde{R}_{s-1}, \widetilde{W}_s)\big)^\top, \tag{6.58}$$
where $\widetilde{R}_i \in \mathrm{O}(k_i)$ for $i \in \{1, \ldots, s-1\}$ and $\widetilde{B}_s, \widetilde{W}_s \in \mathrm{O}(k_s)$. Here $\widetilde{B}_s = \widetilde{W}_s$ holds if $\widetilde{D}_s \neq 0$. Nevertheless, despite of this ambiguity, we will prove in the sequel that Algorithm 5 yields always a valid result. Next define
$$T = \mathrm{diag}\big(1, \ldots, 1, \det(\widetilde{Q}\widetilde{R}^\top)\big) \in \big(\mathrm{O}(1)\big)^k \subseteq \mathrm{O}(k) \tag{6.59}$$
following Algorithm 5, Step 2. Moreover, by Step 3 of Algorithm 5, we have the definitions
$$R = \widetilde{Q}T\widetilde{R}^\top \quad \text{and} \quad Y_* = YR. \tag{6.60}$$
Using (6.60) and $Y^\top X = \widetilde{Q}\widetilde{D}\widetilde{R}^\top$, we compute
$$X^\top Y_* = X^\top Y R = (Y^\top X)^\top R = (\widetilde{R}\widetilde{D}\widetilde{Q}^\top)(\widetilde{Q}T\widetilde{R}^\top) = \widetilde{R}(\widetilde{D}T)\widetilde{R}^\top \tag{6.61}$$
This yields
$$Y_* = XX^\top Y_* + (I_n - XX^\top)Y_* = X\widetilde{R}(\widetilde{D}T)\widetilde{R}^\top + (I_n - XX^\top)Y_*. \tag{6.62}$$
Setting $L = (I_n - XX^\top)Y_*$, we obtain by $Y_*^\top Y_* = I_k$ and (6.61)
$$L^\top L = Y_*^\top (I_n - XX^\top) Y_* = \widetilde{R}\big(I_k - \widetilde{D}^2\big)\widetilde{R}^\top. \tag{6.63}$$
Thus $\widetilde{R} \in \mathrm{O}(k)$ diagonalizes $L^\top L$. Clearly, the eigenvalues of $L^\top L$ are in the interval $[0,1]$. Thus the matrix $\bar{D} = \big(I_k - \widetilde{D}^2\big)^{1/2}$ is a diagonal matrix whose diagonal entries are the singular values of L.

We now consider the cases $T = I_k$ and $T \neq I_k$ separately. Our goal is to show that both cases, i.e. Step 7b and Step 7d of Algorithm 5 yield a $\widetilde{V} \in T_X\mathrm{St}_{n,k}$ satisfying $\mathrm{Exp}_X(\widetilde{V}) = Y_*$.

1. We start with $T = I_k$. Let $L = \widehat{Q}\widehat{D}\widehat{S}^\top$ be a compact SVD computed in Algorithm 5, Step 6, where we assume, without loss of generality, that the singular values, i.e. the diagonal entries of $\widehat{D}$, are arranged in descending order. Thus, by well-known properties of the SVD, there is a matrix $P \in \mathrm{O}(k)$ such that $\widehat{D} = P\bar{D}P^\top$ holds. Next we write $\widehat{D} = \operatorname{diag}\big(\widehat{D}_1, \ldots, \widehat{D}_s\big) = \operatorname{diag}\big(I_{k_s} - \widetilde{D}_s^2, \ldots, I_{k_1} - \widetilde{D}_1^2\big)$. Then, see e.g. [24, Thm. 3.1.1'], any other compact SVD of L with singular values arranged in descending order is given by

$$L = \big(\widehat{Q}\operatorname{diag}\big(\widehat{R}_1, \ldots, \widehat{R}_{s-1}, \widehat{B}_s\big)\big)\widehat{D}\big(\widehat{S}\operatorname{diag}\big(\widehat{R}_1, \ldots, \widehat{R}_{s-1}, \widehat{W}_s\big)\big)^\top, \tag{6.64}$$

where $\widehat{R}_i \in \mathrm{O}(k_{s+1-i})$ for $i \in \{1, \ldots, s-1\}$ and $\widehat{B}_s, \widehat{W}_s \in \mathrm{O}(k_1)$. Here $\widehat{B}_s = \widehat{W}_s$ holds if $\widehat{D}_s = (1 - \widetilde{D}_1^2) \neq 0$. Moreover, one has

$$P\big(\widetilde{R}(L^\top L)\widetilde{R}^\top\big)P^\top = P\big(I_k - \widetilde{D}^2\big)P^\top = P\bar{D}^2P^\top = \widehat{D}^2 \tag{6.65}$$

by (6.63) showing that $P\widetilde{R} \in \mathrm{O}(k)$ diagonalizes $L^\top L$. Hence, by well-known properties of the compact SVD, $\widehat{S} \in \mathrm{O}(k)$ defined by the compact SVD $L = \widehat{Q}\widehat{D}\widehat{S}^\top$ is given by $\widehat{S} = \widetilde{R}P^\top$ up to the ambiguity described in (6.64). Furthermore, by using $\bar{D}^2 = (I_k - \widetilde{D}^2)$, one verifies that

$$Y^\top X = (\widetilde{Q}P)\widetilde{D}(\widetilde{R}P)^\top = \widetilde{Q}(P\widetilde{D}P^\top)\widetilde{R}^\top \tag{6.66}$$

is fulfilled, where $P\widetilde{D}P^\top \in \mathbb{R}^{k\times k}$ is diagonal. Thus $Y^\top X = (\widetilde{Q}P)\widetilde{D}(\widetilde{R}P)^\top$ is also a SVD of $Y^\top X$, where the diagonal entries of $P\widetilde{D}P^\top$ are arranged in ascending order. Moreover, the matrix $R \in \mathrm{O}(k)$ from Algorithm 5, Step 3 is not affected by the choice of $P \in \mathrm{O}(k)$ for $T = I_k$ due to

$$(\widetilde{Q}P)I_k(\widetilde{R}P)^\top = \widetilde{Q}PP^\top\widetilde{R}^\top = \widetilde{Q}\widetilde{R}^\top = R \tag{6.67}$$

in accordance with (6.60). Next define $\widehat{\Sigma} = \arcsin(\widehat{D})$. Then $\widehat{D} = \sin(\widehat{\Sigma})$ and $\cos(\widehat{\Sigma}) = (I_k - \sin(\widehat{\Sigma})^2)^{1/2} = (I_k - \widehat{D}^2)^{1/2}$ holds. By (6.62) and $T = I_k$, this yields

$$Y_* = X\widehat{S}\cos(\widehat{\Sigma})\widehat{S}^\top + \widehat{Q}\sin(\widehat{\Sigma})\widehat{S}^\top. \tag{6.68}$$

By $\arcsin(0) = 0$, the ambiguity of the compact SVD $L = \widehat{Q}\widehat{D}\widehat{S}^\top$ described in (6.64), does not affect the definition of Y_* in (6.68). Next, define $\widetilde{V} = \widehat{Q}\widehat{\Sigma}\widehat{S}^\top$. Using again $\arcsin(0) = 0$, the calculation in (6.40) in Lemma 6.12 reveals that $X^\top(\widehat{Q}\widehat{\Sigma}) = 0$ holds, implying $X^\top\widetilde{V} = X^\top(\widehat{Q}\widehat{\Sigma}\widehat{S}^\top) = 0 \in \mathfrak{so}(k)$. Thus $\widetilde{V} \in T_X\mathrm{St}_{n,k}$ is satisfied. Now (6.68) combined with Corollary 6.15 yields

$$\operatorname{Exp}_X(\widetilde{V}) = \big(X\widehat{S}\cos(\widehat{\Sigma})\widehat{S}^\top + \widehat{Q}\sin(\widehat{\Sigma})\widehat{S}^\top\big)\mathrm{e}^{X^\top\widetilde{V}} = Y_*\mathrm{e}^0 = Y_*, \tag{6.69}$$

because of $X^\top\widetilde{V} = 0$.

2. Next we treat the case $T \neq I_k$. As above, let $L = \widehat{Q}\widehat{D}\widehat{S}^\top$ be a compact SVD from Algorithm 5, Step 6. Clearly, one has $\widehat{S}^\top L^\top L\widehat{S} = \widehat{D}^2$, i.e. $\widehat{S} \in \mathrm{O}(k)$ diagonalizes $L^\top L$. Combining this identity with (6.63) yields

$$\widehat{D}^2 = \widehat{S}^\top\big(L^\top L\big)\widehat{S} = \widehat{S}^\top\big(\widetilde{R}(I_k - \widetilde{D}^2)\widetilde{R}^\top\big)\widehat{S} = \big(\widehat{S}^\top\widetilde{R}\big)\big(I_k - \widetilde{D}^2\big)\big(\widehat{S}^\top\widetilde{R}\big)^\top. \tag{6.70}$$

Obviously, writing $\bar{D} = (I_k - \widetilde{D}^2)^{1/2}$, we obtain by (6.70)

$$\widehat{D} = (\widehat{S}^\top\widetilde{R})\bar{D}(\widehat{S}^\top\widetilde{R})^\top. \tag{6.71}$$

By using (6.71), the factorization

$$L = \widehat{Q}\widehat{D}\widehat{S}^\top = \widehat{Q}\big((\widehat{S}^\top \widetilde{R})\bar{D}(\widehat{S}^\top \widetilde{R})^\top\big)\widehat{S}^\top = \big(\widehat{Q}\widehat{S}^\top \widetilde{R}\big)\bar{D}\widetilde{R}^\top \tag{6.72}$$

is also a compact SVD of L because $\bar{D} \in \mathbb{R}^{k\times k}$ is diagonal and $\widehat{Q}\widehat{S}^\top \widetilde{R} \in \mathrm{St}_{n,k}$ as well as $\widetilde{R} \in \mathrm{O}(k)$ holds. We now set $\widetilde{\Sigma} = \arcsin(\bar{D})$. Similar to 1, i.e. for the case $T = I_k$, we now obtain $\sin(\widetilde{\Sigma}) = \bar{D}$ because of $\bar{D}_{ii} \in [0, 1]$ for $i \in \{1, \ldots, k\}$. Moreover, we calculate

$$\cos(\widehat{\Sigma}) = (I_k - \sin(\widehat{\Sigma})^2)^{1/2} = \big(I_k - \bar{D}^2\big)^{1/2} = \big(I_k - (I_k - \widetilde{D}^2)\big)^{1/2} = \widetilde{D}. \tag{6.73}$$

Next, following Algorithm 5, Step 7c, we define

$$\widehat{\Sigma} = \arcsin\big(\bar{D}\big) - \operatorname{diag}(0, \ldots, 0, \pi) = \widetilde{\Sigma} - \operatorname{diag}(0, \ldots, 0, \pi) \in \mathbb{R}^{k\times k}. \tag{6.74}$$

Then, by using $\sin(x - \pi) = -\sin(x)$ and $\cos(x - \pi) = -\cos(x)$ for all $x \in \mathbb{R}$, the definition $T = \operatorname{diag}(1, \ldots, 1, -1) \in \mathrm{O}(k)$ implies

$$\sin(\widehat{\Sigma}) = T\bar{D} \quad \text{and} \quad \cos\big(\widehat{\Sigma}\big) = \cos(\widetilde{\Sigma})T \overset{(6.73)}{=} \widetilde{D}T. \tag{6.75}$$

Combining (6.62) with (6.72) and (6.75) yields

$$\begin{aligned} Y_* &= X\widetilde{R}(\widetilde{D}T)\widetilde{R}^\top + \big(\widehat{Q}\widehat{S}^\top \widetilde{R}\big)(TT)\bar{D}\widetilde{R}^\top \\ &= X\widetilde{R}\cos(\widehat{\Sigma})\widetilde{R}^\top + \big(\widehat{Q}\widehat{S}^\top \widetilde{R}T\big)\sin(\widehat{\Sigma})\widetilde{R}^\top. \end{aligned} \tag{6.76}$$

We now define $\widetilde{V} = \big(\widehat{Q}\widehat{S}^\top \widetilde{R}T\big)\widehat{\Sigma}\widetilde{R}^\top$. This factorization is a compact SVD of $\widetilde{V}$ with $\widehat{Q}\widehat{S}^\top \widetilde{R}T \in \mathrm{St}_{n,k}$, $\widetilde{R} \in \mathrm{O}(k)$ and a diagonal matrix $\widehat{\Sigma}$ which is allowed to have a negative diagonal entry. By $\arcsin(0) = 0$, the calculation in (6.40) in Lemma 6.12 reveals that $X^\top\big(\big(\widehat{Q}\widehat{S}^\top \widetilde{R}T\big)\widehat{\Sigma}\big) = 0$ holds, implying $X^\top \widetilde{V} = X^\top\big(\big(\widehat{Q}\widehat{S}^\top \widetilde{R}T\big)\widehat{\Sigma}\widetilde{R}^\top\big) = 0 \in \mathfrak{so}(k)$. Thus $\widetilde{V} \in T_X\mathrm{St}_{n,k}$ is satisfied. Now (6.76) combined with Corollary 6.15 and using $X^\top\widetilde{V} = 0$ yields

$$\operatorname{Exp}_X(\widetilde{V}) = \Big(X\widetilde{R}\cos(\widehat{\Sigma})\widetilde{R}^\top + \big(\widehat{Q}\widehat{S}^\top \widetilde{R}T\big)\sin(\widehat{\Sigma})\widetilde{R}^\top\Big)\mathrm{e}^{X^\top\widetilde{V}} = Y_*\mathrm{e}^0 = Y_* \tag{6.77}$$

as desired.

Next we observe that R defined in (6.60) is always an element in $\mathrm{SO}(k)$ since $\det(R) = \det(\widetilde{Q})\det(T)\det(\widetilde{R}^\top) = 1$ holds by the definition of T in (6.59). Hence there exists an $A \in \mathfrak{so}(k)$ with $\mathrm{e}^A = R^\top$ which is obtained by Algorithm 5, Step 5. Setting $V = \widetilde{V} + XA$ yields $X^\top V = X^\top(\widetilde{V} + XA) = A$ due to $X^\top\widetilde{V} = 0$. By Corollary 6.15 combined with (6.60) and (6.69) if $T = I_k$ or (6.77) for $T \neq I_k$, we now obtain

$$\operatorname{Exp}_X(V) = \operatorname{Exp}_X(\widetilde{V})\mathrm{e}^{X^\top V} = Y_*\mathrm{e}^A = YRR^\top = Y \tag{6.78}$$

as desired. Since $Y \in \mathrm{St}_{n,k}$ was arbitrary, the surjectivity of $\operatorname{Exp}_X \colon T_X\mathrm{St}_{n,k} \to \mathrm{St}_{n,k}$ is proven, as well.

It remains to prove Claim 2. Let $V \in T_X\mathrm{St}_{n,k}$ satisfying Assumption 2a and Assumption 2b. Denote by $\bar{Q}\bar{\Sigma}\bar{S}^\top = (I_n - XX^\top)V$ a compact SVD. Here the matrices are decorated with a bar to distinguish them from the matrices defined by Algorithm 5. Then Condition 2a yields $0 \leq \bar{\Sigma}_{ii} < \pi/2$ for all $i \in \{1, \ldots, k\}$, see e.g. [23, Appendix B.7]. We define

$$Y = \operatorname{Exp}_X(V) = \big(X\bar{S}\cos(\bar{\Sigma})\bar{S}^\top + \bar{Q}\sin(\bar{\Sigma})\bar{S}^\top\big)\mathrm{e}^{X^\top V}, \tag{6.79}$$

where Corollary 6.15 is used to obtain the second equality. Using $X^\top \bar{Q}\sin(\bar{\Sigma}) = 0$ due to $\sin(0) = 0$ by the calculation in (6.40) in Lemma 6.12, we obtain

$$Y^\top X = \mathrm{e}^{-X^\top V}\big(\bar{S}\sin(\bar{\Sigma})\bar{Q}^\top + \bar{S}\cos(\bar{\Sigma})\bar{S}^\top X^\top)\big)X = \mathrm{e}^{-X^\top V}\bar{S}\cos(\bar{\Sigma})\bar{S}^\top. \tag{6.80}$$

Defining $\widetilde{Q} = \mathrm{e}^{-X^\top V}\bar{S}$, $\widetilde{D} = \cos(\bar{\Sigma})$, $\widetilde{R} = \bar{S}$, and exploiting $\cos(\bar{\Sigma})_{ii} > 0$ for all $i \in \{1,\dots,k\}$ due to $0 \leq \bar{\Sigma}_{ii} < \pi/2$, the decomposition $Y^\top X = \widetilde{Q}\widetilde{D}\widetilde{R}$ is an SVD of $Y^\top X$ because of (6.80). Furthermore, we have

$$\det(\widetilde{Q}\widetilde{R}^\top) = \det(\mathrm{e}^{-X^\top V}\bar{S}\bar{S}^\top) = \det(\mathrm{e}^{-X^\top V}) = 1 \tag{6.81}$$

showing that Step 2 of Algorithm 5 yields $T = I_k$. Thus

$$R = \widetilde{Q}T\widetilde{R}^\top = \mathrm{e}^{-X^\top V}\widetilde{S}I_k\widetilde{S}^\top = \mathrm{e}^{-X^\top V} \tag{6.82}$$

holds by Algorithm 5, Step 3. Moreover, by $\widetilde{D}_{ii} = \cos(\bar{\Sigma})_{ii} > 0$, the matrix R is uniquely determined by (6.82) because the ambiguity of the SVD $Y^\top X = \widetilde{Q}\widetilde{D}\widetilde{R}^\top$ from (6.58) does not affect R defined in Algorithm 5, Step 3 if $T = I_k$ and $\operatorname{rank}(Y^\top X) = k$. In fact, we have

$$\big(\widetilde{Q}\operatorname{diag}(\widetilde{R}_1,\dots,\widetilde{R}_{s-1},\widetilde{B}_s)\big)I_k\big(\widetilde{R}\operatorname{diag}(\widetilde{R}_1,\dots,\widetilde{R}_{s-1},\widetilde{B}_s)\big)^\top = \widetilde{Q}\widetilde{R}^\top = R, \tag{6.83}$$

where the notation from (6.58) is used. This implies

$$Y_* = YR = X\bar{S}\cos(\bar{\Sigma})\bar{S}^\top + \bar{Q}\sin(\bar{\Sigma})\bar{S}^\top \tag{6.84}$$

by (6.79). Using again $X^\top\bar{Q}\sin(\bar{\Sigma}) = 0$ due to $\sin(0) = 0$, we obtain

$$(I_n - XX^\top)Y_* = \bar{Q}\sin(\bar{\Sigma})\bar{S}^\top. \tag{6.85}$$

Moreover, $0 \leq \sin(\bar{\Sigma})_{ii} < 1$ holds because of $0 \leq \bar{\Sigma}_{ii} < \pi/2$ showing that $(I_n - XX^\top)Y_* = \bar{Q}\sin(\bar{\Sigma})\bar{S}^\top$ is a compact SVD by (6.85). Thus a SVD from Step 6 of Algorithm 5, i.e. $\widehat{Q}\widehat{D}\widehat{S}^\top = (I_n - XX^\top)Y_*$, is given by $\widehat{Q} = \bar{Q}$, $\widehat{D} = \sin(\bar{\Sigma})$ and $\widehat{S} = \bar{S}$. In addition, by $0 \leq \sin(\bar{\Sigma})_{ii} < 1$, Algorithm 5, Step 7a yields $\widehat{\Sigma} = \arcsin(\widehat{D}) = \arcsin(\sin(\bar{\Sigma})) = \bar{\Sigma}$. Moreover, Algorithm 5, Step 5 yields $A = \log(R^\top) = \log(\mathrm{e}^{X^\top V}) = X^\top V$ by $\|X^\top V\|_2 < \pi$, see e.g. [23, Problem 1.39]. Finally, we compute

$$\widehat{Q}\widehat{\Sigma}\widehat{S}^\top + XA = \bar{Q}\bar{\Sigma}\bar{S}^\top + XX^\top V = (I_n - XX^\top)V + XX^\top V = V \tag{6.86}$$

as desired. □

6.4 Interpolation on Stiefel Manifolds via Intrinsic Rolling and Unwrapping

After the preparation in the previous subsections, we now adapt Algorithm 1 to the Stiefel manifold $\mathrm{St}_{n,k}$. In this subsection, we replace the symbol k in Algorithm 1 by ℓ, i.e. the number of points is denoted by $(\ell + 1) \in \mathbb{N}$, since the symbol k occurs in our notation for the Stiefel manifold $\mathrm{St}_{n,k}$. Using this new notation, we state Algorithm 6 for solving Problem 2.1 on $\mathrm{St}_{n,k}$.

Algorithm 6 Interpolation on Stiefel Manifolds

Input: $X_0, \dots, X_\ell \in \mathrm{St}_{n,k}$, initial velocity $V_0 \in T_{X_0}\mathrm{St}_{n,k}$, final velocity $V_\ell \in T_{X_\ell}\mathrm{St}_{n,k}$, instances of time $0 = t_0 < \dots < t_\ell = T$.

1. Compute $V = \frac{1}{T}\mathrm{Log}_{X_0}(X_\ell)$ by Algorithm 5 and define the curve
$$g\colon \mathbb{R} \mapsto \mathrm{O}(n) \times \mathrm{O}(k), \quad t \mapsto \big(\mathrm{e}^{t(VX_0^\top - X_0V^\top + 2X_0V^\top X_0X_0^\top)}, \mathrm{e}^{-tX_0^\top V}\big). \tag{6.87}$$
2. Unwrap the boundary data to $T_{X_0}\mathrm{St}_{n,k}$ by defining
$$\begin{aligned} q_0 &= 0, & q_\ell &= TV, \\ \eta_0 &= \Phi_{g(0)}(V_0) = V_0, & \eta_\ell &= \Phi_{g(-T)}(V_\ell), \end{aligned} \tag{6.88}$$
where (6.54) is used to compute $\eta_\ell \in T_{X_0}\mathrm{St}_{n,k}$.
3. Unwrap the remaining data by defining for $i \in \{1, \dots, \ell - 1\}$
$$q_i = \mathrm{Log}_{X_0}\big(\Phi_{g(-t_i)}(X_i)\big) + t_iV, \tag{6.89}$$
where $\Phi_{g(-t_i)}(X_i)$ is computed by (6.54) and Log_{X_0} is evaluated by Algorithm 5.
4. Compute a $\mathscr{C}^2$-curve $y\colon [0,T] \to T_{X_0}\mathrm{St}_{n,k}$ with $y(t_i) = q_i$ for all $i \in \{0, \dots, \ell\}$ and $\dot{y}(0) = V_0$ as well as $\dot{y}(T) = V_\ell$.
5. Define the curve $\beta\colon [0,T] \to \mathrm{St}_{n,k}$ via wrapping $y\colon I \to T_{X_0}\mathrm{St}_{n,k}$ back to $\mathrm{St}_{n,k}$ by setting
$$\beta(t) = \Phi_{g(t)}\big(\mathrm{Exp}_{X_0}\big(y(t) - tV\big)\big) \tag{6.90}$$
for $t \in [0,T]$. Here $\Phi_{g(t)}(\cdot)$ is computed by (6.54) and $\mathrm{Exp}_{X_0}\colon T_{X_0}\mathrm{St}_{n,k} \to \mathrm{St}_{n,k}$ is evaluated by using the third expression in (6.52).

Output: The curve $\beta\colon [0,T] \to \mathrm{St}_{n,k}$.

Remark 6.20 In Algorithm 6, Step 1, Algorithm 5 is invoked to compute $V = \frac{1}{T}\mathrm{Log}_{X_0}(X_\ell)$. For evaluating (6.54) in Step 2 and Step 3 of Algorithm 6, a thin SVD of $(I_n - X_0X_0^\top)V$ is needed. Instead of computing again this thin SVD from scratch, the results of Algorithm 5 can be reused. Indeed, if T defined in Algorithm 5, Step 2 satisfies $T = I_k$, the matrices $\widehat{Q} \in \mathrm{St}_{n,k}$, $\widehat{S} \in \mathrm{O}(k)$ and $\widehat{\Sigma} \in \mathbb{R}^{k\times k}$ from Algorithm 5, Step 7b can be used to construct a thin SVD of $(I_n - X_0X_0^\top)V \in \mathbb{R}^{n\times k}$. Similarly, if $T \neq I_k$, the matrices $\widehat{Q}\widehat{S}^\top\widetilde{R}T \in \mathrm{St}_{n,k}$, $\widehat{\Sigma} \in \mathbb{R}^{k\times k}$ and $\widetilde{R} \in \mathrm{O}(k)$ from Algorithm 5, Step 7d can be used to obtain a thin SVD of $(I_n - X_0X_0^\top)V \in \mathbb{R}^{n\times k}$.

Next we show that Algorithm 6 is obtained by applying Algorithm 1 to the Stiefel manifold $\mathrm{St}_{n,k}$ and, consequently, it yields a solution of Problem 2.1 on $\mathrm{St}_{n,k}$.

Proposition 6.21 *Let $X_0, \dots, X_\ell \in \mathrm{St}_{n,k}$, let $0 = t_0 < \dots < t_\ell = T$ and let $V_0 \in T_{X_0}\mathrm{St}_{n,k}$ as well as $V_\ell \in T_{X_\ell}\mathrm{St}_{n,k}$ be given. Then, the curve $\beta\colon [0,T] \to \mathrm{St}_{n,k}$ defined by Algorithm 6 is a $\mathscr{C}^2$-curve solving Problem 2.1 on $\mathrm{St}_{n,k}$ associated with the given data.*

Proof: We show that Algorithm 6 is a special case of Algorithm 1 applied to $\mathrm{St}_{n,k}$. Then the assertion follows by Theorem 4.7.

To this end, we identify the Stiefel manifold $\mathrm{St}_{n,k}$ with the reductive homogeneous space $(\mathrm{O}(n) \times \mathrm{O}(k))/H_{X_0}$ via the $(\mathrm{O}(n) \times \mathrm{O}(k))$-equivariant diffeomorphism

$$\iota_{X_0} \colon (\mathrm{O}(n) \times \mathrm{O}(k))/H_{X_0} \ni (R, \theta) \cdot H_{X_0} \mapsto R X_0 \theta^\top \in \mathrm{St}_{n,k}, \tag{6.91}$$

where we chose the reductive decomposition $\mathfrak{so}(n) \times \mathfrak{so}(k) = \mathfrak{h}_{X_0} \oplus \mathfrak{m}_{X_0} = \mathfrak{h} \oplus \mathfrak{m}$ from Lemma 6.7. Moreover, $(\mathrm{O}(n) \times \mathrm{O}(k))/H_{X_0}$ is equipped with $\nabla^{\mathrm{can}2}$.

By using these identifications and choices, Algorithm 6 is an application of Algorithm 1 to $\mathrm{St}_{n,k}$. In order to prove this, we proceed with the following steps:

1. Algorithm 1, Step 1 and Algorithm 1, Step 2 are combined in Algorithm 6, Step 1. Indeed, let $V \in T_{X_0}\mathrm{St}_{n,k}$ be defined by Algorithm 6, Step 1 and let

$$\xi = \big(T_{(I_n,I_k)}(\iota_{X_0} \circ \mathrm{pr})\big|_{\mathfrak{m}}\big)^{-1} V = \big(V X^\top - X V^\top + 2 X V^\top X X^\top, -X^\top V\big) \in \mathfrak{m}. \tag{6.92}$$

Define the curve

$$v \colon [0, T] \ni t \mapsto v(t) = t\xi \in \mathfrak{m} \tag{6.93}$$

which clearly fulfills

$$T_{(I_n,I_k)}(\iota_{X_0} \circ \mathrm{pr}_{X_0}) v(t) = tV \tag{6.94}$$

by Lemma 6.9. Moreover, let $g \colon [0, T] \to \mathrm{O}(n) \times \mathrm{O}(k)$ be given by (6.87). Obviously, $g(t) = \exp(t\xi)$ holds for all $t \in [0, T]$. Then, by Lemma 3.8, the curve

$$(v, g, S) \colon [0, T] \ni t \mapsto (v(t), g(t), \mathrm{id}_{\mathfrak{m}}) \in \mathfrak{m} \times (\mathrm{O}(n) \times \mathrm{O}(k)) \times \mathrm{GL}(\mathfrak{m}) \tag{6.95}$$

defines the intrinsic rolling $(v(t), \gamma(t), A(t))$ of $\mathfrak{m}$ over $(\mathrm{O}(n) \times \mathrm{O}(k))/H_{X_0}$ with respect to $\nabla^{\mathrm{can}2}$, where

$$\gamma = \mathrm{pr}_{X_0} \circ g \colon I \to (\mathrm{O}(n) \times \mathrm{O}(k))/H_{X_0}, \quad t \mapsto \mathrm{pr}_{X_0}(\exp(t\xi)) \tag{6.96}$$

and $A(t) \colon T_{v(t)}\mathfrak{m} \cong \mathfrak{m} \to T_{\mathrm{pr}(g(t))}((\mathrm{O}(n) \times \mathrm{O}(k))/H_{X_0})$ is given by

$$A(t)Z = (T_{g(t)} \mathrm{pr}_{X_0} \circ T_e \ell_{g(t)}) Z, \quad Z \in \mathfrak{m}. \tag{6.97}$$

Using (6.91), we simplify the curve of linear isomorphisms

$$\widehat{A}(t) \colon T_{X_0}\mathrm{St}_{n,k} \to T_{\iota_{X_0}(\gamma(t))}\mathrm{St}_{n,k} \tag{6.98}$$

defined in (4.14) in Algorithm 1, Step 2. By using the chain-rule and exploiting $\tau_g \circ \mathrm{pr}_{X_0} = \mathrm{pr}_{X_0} \circ \ell_g$ as well as $\iota_{X_0} \circ \tau_g = \Phi_g \circ \iota_{X_0}$ for all $g \in \mathrm{O}(n) \times \mathrm{O}(k)$, we compute for $Z \in T_{X_0}\mathrm{St}_{n,k}$

$$\begin{aligned}
&\widehat{A}(t)Z \\
&= T_{\mathrm{pr}_{X_0}(g(t))}\iota_{X_0} \circ A(t) \circ \big(T_{(I_n,I_k)}(\iota_{X_0} \circ \mathrm{pr}_{X_0})\big|_{\mathfrak{m}}\big)^{-1} Z \\
&= T_{\mathrm{pr}_{X_0}(g(t))}\iota_{X_0} \circ \Big(T_{g(t)} \mathrm{pr}_{X_0} \circ T_{(I_n,I_k)} \ell_{g(t)}\Big) \circ \big(T_{(I_n,I_k)}(\iota_{X_0} \circ \mathrm{pr}_{X_0})\big|_{\mathfrak{m}}\big)^{-1} Z \\
&= T_{(I_n,I_k)}(\iota_{X_0} \circ \mathrm{pr}_{X_0} \circ \ell_{g(t)}) \circ \big(T_{(I_n,I_k)}(\iota_{X_0} \circ \mathrm{pr}_{X_0})\big|_{\mathfrak{m}}\big)^{-1} Z \\
&= T_{(I_n,I_k)}(\iota_{X_0} \circ \tau_{g(t)} \circ \mathrm{pr}_{X_0}) \circ \big(T_{(I_n,I_k)}(\iota_{X_0} \circ \mathrm{pr}_{X_0})\big|_{\mathfrak{m}}\big)^{-1} Z \\
&= T_{(I_n,I_k)}(\Phi_{g(t)} \circ \iota_{X_0} \circ \mathrm{pr}_{X_0}) \circ \big(T_{(I_n,I_k)}(\iota_{X_0} \circ \mathrm{pr}_{X_0})\big|_{\mathfrak{m}}\big)^{-1} Z \\
&= T_{X_0}\Phi_{g(t)} \circ T_{(I_n,I_k)}(\iota_{X_0} \circ \mathrm{pr}_{X_0}) \circ \big(T_{(I_n,I_k)}(\iota_{X_0} \circ \mathrm{pr}_{X_0})\big|_{\mathfrak{m}}\big)^{-1} Z \\
&= T_{X_0}\Phi_{g(t)}(Z) \\
&= \Phi_{g(t)}(Z),
\end{aligned} \tag{6.99}$$

where we have used the fact that $\Phi_{g(t)}\colon \mathbb{R}^{n\times k} \ni Z \to \Phi_{g(t)}(Z) \in \mathbb{R}^{n\times k}$ is a linear map to obtain the last equality.

Moreover, we have $S(t) = \mathrm{id}_{\mathfrak{m}}$ for all $t \in [0,T]$ by (6.95) implying that the map $\widehat{S}(t)\colon T_{X_0}\mathrm{St}_{n,k} \to T_{X_0}\mathrm{St}_{n,k}$ defined in (4.15) in Algorithm 1, Step 2 simplifies to

$$\widehat{S}(t) = T_{(I_n,I_k)}(\iota_{X_0} \circ \mathrm{pr}_{X_0}) \circ S(t) \circ \big(T_{(I_n,I_k)}(\iota_{X_0} \circ \mathrm{pr}_{X_0})\big|_{\mathfrak{m}}\big)^{-1} = \mathrm{id}_{T_{X_0}\mathrm{St}_{n,k}}\,. \tag{6.100}$$

2. Obviously, Algorithm 6, Step 2 corresponds to Algorithm 1, Step 3 applied to $\mathrm{St}_{n,k}$.

3. By $\widehat{S}(t) = \mathrm{id}_{T_{X_0}\mathrm{St}_{n,k}}$ for all $t \in [0,T]$, Algorithm 6, Step 3 corresponds to Algorithm 1, Step 4. Here the map ϕ is chosen as $\mathrm{Log}_{X_0}\colon \mathrm{St}_{n,k} \to T_{X_0}\mathrm{St}_{n,k}$ in Algorithm 6, Step 3. This map has the desired properties by Remark 6.10, 3. due to $\mathrm{Log}_{X_0} = (\mathrm{Exp}_{X_0})^{-1}$ in a suitable open neighbourhood $U \subseteq \mathrm{St}_{n,k}$ containing $X_0 \in \mathrm{St}_{n,k}$.

4. Clearly, Algorithm 6, Step 4 corresponds to Algorithm 1, Step 5.

5. Using again $\widehat{S}(t) = \mathrm{id}_{T_{X_0}\mathrm{St}_{n,k}}$ for all $t \in [0,T]$ and using $(\mathrm{Log}_{X_0})^{-1} = \mathrm{Exp}_{X_0}$, one verifies that Algorithm 6, Step 5 corresponds to Algorithm 1, Step 6.

Hence Algorithm 6 is a special case of Algorithm 1 applied to $\mathrm{St}_{n,k}$. This yields the desired result by Theorem 4.7. □

Remark 6.22 Algorithm 6 always yields a valid solution of Problem 2.1 on $\mathrm{St}_{n,k}$, i.e. no restriction on the given data $X_0, \dots, X_\ell$ and $V_0 \in T_{X_0}\mathrm{St}_{n,k}$ as well as $V_\ell \in T_{X_\ell}\mathrm{St}_{n,k}$ needs to be imposed for Algorithm 6 to yield some $\mathscr{C}^2$-curve solving Problem 2.1. Indeed, by Proposition 6.19, Algorithm 5 yields for all $(X,Y) \in \mathrm{St}_{n,k} \times \mathrm{St}_{n,k}$ some $V \in T_X\mathrm{St}_{n,k}$ such that $Y = \mathrm{Exp}_X(V)$ holds. Thus the assertion follows by Remark 4.8.

Remark 6.23 Proposition 6.21 shows that Algorithm 6 is obtained by applying Algorithm 1 to $\mathrm{St}_{n,k}$, where the local diffeomorphism ϕ in Algorithm 1, Step 4 is chosen as $\mathrm{Log}_{X_0}\colon U \subseteq \mathrm{St}_{n,k} \to T_{X_0}\mathrm{St}_{n,k}$. In principal, other choices of ϕ are possible. For instance, let $R\colon T\mathrm{St}_{n,k} \to \mathrm{St}_{n,k}$ be a retraction. Then one may set $\phi^{-1} = R_{X_0}\colon U \subseteq T_{X_0}\mathrm{St}_{n,k} \to \mathrm{St}_{n,k}$ and $\phi = (R_{X_0})^{-1}\colon \widetilde{U} \subseteq \mathrm{St}_{n,k} \to T_{X_0}\mathrm{St}_{n,k}$, where U and $\widetilde{U}$ are a suitable open neighbourhoods of $0 \in T_{X_0}\mathrm{St}_{n,k}$ and $X_0 \in \mathrm{St}_{n,k}$, respectively. Discussing specific choices for R is out of the scope of this text but we refer to [1, Chap. 4] for some examples of retractions on $\mathrm{St}_{n,k}$.

Moreover, Algorithm 6 can be used as an essential building block for an algorithm for solving Problem 2.2 on $\mathrm{St}_{n,k}$ which is computationally efficient for $k \ll n$.

Algorithm 7 Interpolation on Stiefel Manifolds (for solving Problem 2.2)

Input: $X_0, \dots, X_\ell \in \mathrm{St}_{n,k}$, velocities $V_0 \in T_{X_0}\mathrm{St}_{n,k}, \dots, V_\ell \in T_{X_\ell}\mathrm{St}_{n,k}$, instances of time $0 = t_0 < \dots < t_\ell = T$.

1. For $i = 0, \dots, \ell-1$ do:
   Compute a $\mathscr{C}^2$-curve $\beta_i\colon [0, t_{i+1} - t_i] \to \mathrm{St}_{n,k}$ satisfying $\beta_i(0) = X_i$, $\beta_i(t_{i+1} - t_i) = X_{i+1}$ and $\dot{\beta}_i(0) = V_i$ as well as $\dot{\beta}_i(t_{i+1} - t_i) = V_{i+1}$ by Algorithm 6.

2. Define $\beta\colon [0,T] \to M$ for $t \in [0,T]$ by

$$\beta\big|_{[t_i,t_{i+1})}(t) = \beta_i(t - t_i) \quad , i \in \{0, \dots, \ell-2\} \quad \text{and} \quad \beta\big|_{[t_{\ell-1},T]}(t) = \beta_{\ell-1}(t - t_{\ell-1}) \tag{6.101}$$

Output: The curve $\beta\colon [0,T] \to \mathrm{St}_{n,k}$.

Proposition 6.24 *Let* $X_0, \dots, X_\ell \in \mathrm{St}_{n,k}$*,* $V_0 \in T_{X_0}\mathrm{St}_{n,k}, \dots, V_\ell \in T_{X_\ell}\mathrm{St}_{n,k}$ *and* $0 = t_0 < \dots < t_\ell = T$*. Then, the curve* $\beta\colon [0,T] \to \mathrm{St}_{n,k}$ *defined by Algorithm 7 is a* $\mathscr{C}^1$*-curve solving Problem 2.2 on* $\mathrm{St}_{n,k}$ *associated with the given data.*

Proof: The assertion is shown analogously to the proof of Lemma 5.13. □

7 Some Geometric Properties of the Interpolating Curves

In this section, we investigate some geometric properties of the curves obtained by Algorithm 2, Algorithm 3 and Algorithm 6. To be more precise, we provide the relevant formulas for evaluating the functionals $L, E\colon \mathscr{C}^2([0,T], M) \to \mathbb{R}$ defined by

$$\begin{aligned} L(\beta) &= \int_0^T \|\dot\beta(t)\|_{\beta(t)} \,\mathrm{d}t \\ E(\beta) &= \tfrac{1}{2}\int_0^T \langle \dot\beta(t), \dot\beta(t)\rangle_{\beta(t)} \,\mathrm{d}t \end{aligned} \tag{7.1}$$

numerically, where β is computed by Algorithm 2, Algorithm 3 or Algorithm 6 and $\langle\cdot,\cdot\rangle$ is an appropriated (pseudo-)Riemannian metric on M. Here M is a matrix Lie group or a Stiefel manifold. Moreover, under the same assumptions, we briefly comment on the evaluation of the acceleration and jerk, defined by

$$\begin{aligned} A(\beta) &= \tfrac{1}{2}\int_0^T \langle \nabla_{\dot\beta(t)}\dot\beta(t), \nabla_{\dot\beta(t)}\dot\beta(t)\rangle \,\mathrm{d}t \\ J(\beta) &= \tfrac{1}{2}\int_0^T \langle \nabla_{\dot\beta(t)}\nabla_{\dot\beta(t)}\dot\beta(t), \nabla_{\dot\beta(t)}\nabla_{\dot\beta(t)}\dot\beta(t)\rangle \,\mathrm{d}t, \end{aligned} \tag{7.2}$$

respectively, where ∇ is a suitable covariant derivative on M. If $\nabla = \nabla^{\mathrm{LC}}$ in (7.2), the critical curves of the functional A and J are exactly the (geometric) Riemannian cubic and Riemannian quintic polynomial splines, see [9] for details.

Before we continue, we recall that the tangent map of the matrix exponential at $A \in \mathfrak{gl}(n) = \mathbb{R}^{n\times n}$ evaluated at $B \in T_A\mathfrak{gl}(n) \cong \mathbb{R}^{n\times n}$ is given by

$$(T_A \exp)B = \Big(T_{I_n}\ell_{\exp(A)} \circ \int_0^1 \mathrm{e}^{-\operatorname{ad}_{sA}} \,\mathrm{d}s\Big)(B) = \exp(A)\Big(\sum_{j=0}^\infty \tfrac{(-1)^j}{(j+1)!}(\operatorname{ad}_A)^j(B)\Big), \tag{7.3}$$

see e.g. [16, Sec. 1.5] and also [18, Sec. 3.2] for the second identity. It can be computed by using the formula

$$\exp\left(\left[\begin{smallmatrix} A & B \\ 0 & A \end{smallmatrix}\right]\right) = \left[\begin{smallmatrix} \exp(A) & (T_A\exp)B \\ 0 & \exp(A) \end{smallmatrix}\right], \tag{7.4}$$

see e.g. [23, Sec. 10.6].

7.1 Interpolating Curves on Lie Groups

For computing $L(\beta)$ and $E(\beta)$, where $\beta\colon [0,T] \to G$ is obtained by Algorithm 2 or Algorithm 3, an expression for $\dot\beta\colon [0,T] \to TG$ is required.

Lemma 7.1 *Let* $G \subseteq \mathrm{GL}(n)$ *be a matrix Lie group.*

1. *Let* $\beta\colon [0,T] \to G$ *be generated by Algorithm 2, i.e.*

$$\beta(t) = \exp\left(\tfrac{t}{2}\xi\right)\exp\Big(y(t)g_0^{-1} - t\xi\Big)\exp\left(\tfrac{t}{2}\xi\right)g_0. \tag{7.5}$$

Define the curve $q\colon [0,T] \to G$ *by* $q(t) = \exp\big(y(t)g_0^{-1} - t\xi\big)$*, so that*

$$\begin{aligned}\dot\beta(t) &= \tfrac{\xi}{2}\exp\big(\tfrac{t}{2}\xi\big)q(t)\exp\big(\tfrac{t}{2}\xi\big)g_0 + \exp\big(\tfrac{t}{2}\xi\big)\dot q(t)\exp\big(\tfrac{t}{2}\xi\big)g_0 \\ &\quad + \exp\big(\tfrac{t}{2}\xi\big)q(t)\exp\big(\tfrac{t}{2}\xi\big)\tfrac{\xi}{2}g_0,\end{aligned} \tag{7.6}$$

holds for all $t \in [0,T]$*, where*

$$\dot q(t) = \big(T_{(y(t)g_0^{-1}-t\xi)}\exp\big)(\dot y(t)g_0^{-1} - \xi). \tag{7.7}$$

2. *Let* $\beta\colon [0,T] \to G$ *be generated by Algorithm 3, i.e.*

$$\beta(t) = \exp(t\xi)\exp\big(y(t)g_0^{-1} - t\xi\big)g_0 \tag{7.8}$$

Define the curve $q\colon [0,T] \to G$ *as in Claim 1. Then*

$$\dot\beta(t) = \xi\exp(t\xi)q(t) + \exp(t\xi)\dot q(t) = \xi\beta(t) + \exp(t\xi)\dot q(t), \tag{7.9}$$

holds for all $t \in [0,T]$*, where an expression for* $\dot q(t)$ *is given by* (7.7).

PROOF: The assertions follow by some straightforward computations. □

Let $\beta\colon [0,T] \to G$ be given by Algorithm 2 or Algorithm 3. Since the definition of $E, L\colon \mathscr{C}^2([0,T],G) \to \mathbb{R}$ assumes that a (pseudo-)Riemannian metric on G is chosen, once that is fixed on G, L and E can be evaluated, at least numerically, using Lemma 7.1. A possible, non-canonical choice, exploiting that $G \subseteq \mathrm{GL}(n)$ is a matrix Lie group, is given by the metric induced by the Frobenius scalar product, namely

$$\langle \xi_1, \xi_2\rangle_g = \operatorname{tr}\big(\xi_1^\top \xi_2\big), \tag{7.10}$$

where $g \in G \subseteq \mathrm{GL}(n)$ and $\xi_1, \xi_2 \in T_gG \subseteq T_g\mathrm{GL}(n) \cong \mathbb{R}^{n\times n}$.

Remark 7.2 Let $\beta\colon [0,T] \to G$ be given by Algorithm 2 or Algorithm 3. Since β is given in closed form, in principle, it is possible to compute $A(\beta)$ and $J(\beta)$ defined in (7.2), where for the latter, β is assumed to be a piecewise $\mathscr{C}^3$-curve. Here G is equipped with an invariant covariant derive ∇^α. In this context, one may use the following expression for the associated covariant derivative of a vector field $Z\colon I \to G$ along the curve $\beta\colon I \to G$. Define

$$z\colon I \ni t \mapsto (T_e\ell_{\beta(t)})^{-1}Z(t) \in \mathfrak{g} \quad \text{and} \quad x\colon I \ni t \mapsto (T_e\ell_{\beta(t)})^{-1}\dot\beta(t) \in \mathfrak{g}. \tag{7.11}$$

By applying [39, Prop. 4.26] to $G \cong G/\{e\}$, one obtains for $t \in I$

$$\nabla^\alpha_{\dot\beta(t)} Z\big|_t = T_e\ell_{\beta(t)}\big(\dot z(t) + \alpha\big(x(t), z(t)\big)\big). \tag{7.12}$$

7.2 Interpolating Curves on Stiefel Manifolds

We now consider $L(\beta)$ and $E(\beta)$ on $\mathrm{St}_{n,k}$, where $\mathrm{St}_{n,k}$ is equipped with an α-metric from [27]. For $\nu = -\frac{2\alpha+1}{\alpha+1}$, the α-metric on $\mathrm{St}_{n,k}$ is given by

$$\langle V, W\rangle^\nu_X = 2\operatorname{tr}\big(V^\top W\big) + \nu\operatorname{tr}\big(V^\top X X^\top W\big), \quad X \in \mathrm{St}_{n,k}, \quad V, W \in T_X\mathrm{St}_{n,k} \tag{7.13}$$

according to [27, Eq. (43)]. This metric includes the Euclidean metric, see e.g. [1], scaled by the factor of 2 (corresponding to $\nu = 0$) and the so-called canonical metric also scaled by the factor of 2 (corresponding to $\nu = -1$), see e.g. [17, Eq. (2.39)], as special cases.

In order to evaluate $L(\beta)$ and $E(\beta)$, an expression for $\dot\beta$ is required.

Lemma 7.3 *Let $\beta\colon [0,T]\to \mathrm{St}_{n,k}$ be given by Algorithm 6, i.e.*

$$\beta(t) = \Phi_{g(t)}\big(\mathrm{Exp}_{X_0}\big(y(t)-tV\big)\big), \tag{7.14}$$

where

$$g(t) = \big(\mathrm{e}^{t(VX_0^\top - X_0V^\top + 2X_0V^\top X_0X_0^\top)}, \mathrm{e}^{-tX_0^\top V}\big) = \big(\mathrm{e}^{t\Omega_{X_0}(V)}, \mathrm{e}^{-t\Psi_{X_0}(V)}\big) \tag{7.15}$$

and define

$$Q\colon [0,T]\to \mathrm{St}_{n,k}, \quad t\mapsto Q(t) = \mathrm{Exp}_{X_0}(y(t)-tV). \tag{7.16}$$

Then

$$\dot\beta(t) = \Omega_{X_0}(V)\beta(t) + \mathrm{e}^{t\Omega_{X_0}(V)}\dot Q(t)\mathrm{e}^{t\Psi_{X_0}(V)} + \beta(t)\Psi_{X_0}(V) \tag{7.17}$$

holds for all $t\in[0,T]$, where

$$\dot Q(t) = (T_{(y(t)-tV)}\mathrm{Exp}_{X_0})(\dot y(t) - V). \tag{7.18}$$

Proof: Using $\beta(t) = \mathrm{e}^{t\Omega_{X_0}(V)}Q(t)\mathrm{e}^{t\Psi_{X_0}(V)}$, the desired result is obtained by a straightforward computation. □

Remark 7.4 For computing $\dot\beta(t)$, a method for evaluating (7.18) is needed. To this end, an expression for tangent map of Exp_{X_0} at $V\in T_{X_0}\mathrm{St}_{n,k}$ evaluated at $W\in T_V(T_{X_0}\mathrm{St}_{n,k})\cong T_{X_0}\mathrm{St}_{n,k}$ is desirable. By using the notation from (7.15), the chain rule and (7.3) as well as (7.4), one obtains

$$\begin{aligned}
T_V\mathrm{Exp}_{X_0}W &= \big(T_{\Omega_{X_0}(V)}\exp\Omega_{X_0}(W)\big)X_0\mathrm{e}^{\Psi_{X_0}(V)} \\
&\quad + \mathrm{e}^{\Omega_{X_0}(V)}X_0\big(T_{\Psi_{X_0}(V)}\exp\Psi_{X_0}(W)\big) \\
&= \mathrm{e}^{\Omega_{X_0}(V)}\Big(\int_0^1 \mathrm{e}^{-s\,\mathrm{ad}_{\Omega_{X_0}(V)}}\Omega_{X_0}(W)\,\mathrm{d}s\Big)X_0\mathrm{e}^{\Psi_{X_0}(V)} \\
&\quad + \mathrm{e}^{\Omega_{X_0}(V)}X_0\mathrm{e}^{\Psi_{X_0}(V)}\Big(\int_0^1 \mathrm{e}^{-s\,\mathrm{ad}_{\Psi_{X_0}(V)}}\Psi_{X_0}(W)\,\mathrm{d}s\Big) \\
&= [I_n \mid 0]\exp\Big(\begin{bmatrix}\Omega_{X_0}(V) & \Omega_{X_0}(W)\\ 0 & \Omega_{X_0}(V)\end{bmatrix}\Big)\begin{bmatrix}0\\ X_0\end{bmatrix}\mathrm{e}^{\Psi_{X_0}(V)} \\
&\quad + \mathrm{e}^{\Psi_{X_0}(V)}X_0[I_k\mid 0]\exp\Big(\begin{bmatrix}\Psi_{X_0}(V) & \Psi_{X_0}(W)\\ 0 & \Psi_{X_0}(V)\end{bmatrix}\Big)\begin{bmatrix}0\\ I_k\end{bmatrix}.
\end{aligned} \tag{7.19}$$

Moreover, assuming $\operatorname{rank}\big((I_n - X_0X_0^\top)V\big) = k$, one can adapt the formulas from [7, Sec. 3.5], see also [52, Sec. 4.1], to evaluate $T_V\mathrm{Exp}_{X_0}W$ numerically in an efficient way for $k\ll n$. We are not aware of a similar approach to evaluate $T_V\mathrm{Exp}_{X_0}W$ for arbitrary $V,W\in T_{X_0}\mathrm{St}_{n,k}$ not fulfilling the rank constraint mentioned above. Nevertheless, as soon as a method for evaluating (7.18) is available, the functionals $E(\beta)$ and $L(\beta)$ can be approximated, at least, numerically, where $\mathrm{St}_{n,k}$ is equipped with an α-metric given by (7.13).

Remark 7.5 Similar to Remark 7.2, in principle, one can compute $A(\beta)$ and $J(\beta)$ defined in (7.2), where β is obtained by Algorithm 6 and $\mathrm{St}_{n,k}$ is equipped with an α-metric. Here the following formula for the Levi-Civita covariant derivative with respect to an α-metric of a vector field $Z\colon I\to T\mathrm{St}_{n,k}$ along the curve $\beta\colon I\to\mathrm{St}_{n,k}$ might be helpful. As a consequence of [40, Cor. 6.10], one has

$$\begin{aligned}
\nabla^{\mathrm{LC}}_{\dot\beta(t)}Z\big|_t &= \dot Z(t) - \tfrac{\nu}{2}\big(\dot\beta(t)Z(t)^\top + Z(t)\dot\beta(t)^\top\big)\beta(t) \\
&\quad - \tfrac{\nu}{2}\beta(t)\big(\beta(t)^\top\dot\beta(t)\beta(t)^\top Z(t) + \beta(t)^\top Z(t)\beta(t)^\top\dot\beta(t)\big) \\
&\quad + \tfrac{1}{2}\beta(t)\big(\dot\beta(t)^\top Z(t) + Z(t)^\top\dot\beta(t)\big)
\end{aligned} \tag{7.20}$$

where $\nu = -\frac{2\alpha+1}{\alpha+1}$, see [40, Re. 3.5].

8 Conclusion

We presented a new algorithm for solving Problem 2.1 on a general reductive homogeneous space G/H with a fixed reductive decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{m}$. This algorithm, which builds on intrinsic rollings of $\mathfrak{m}$ over G/H with respect to some invariant covariant derivative, can be used to obtain closed-form solutions of Problem 2.1. We illustrated that the proposed method can be applied to particular reductive homogeneous spaces in a straightforward way on the example of matrix Lie groups. Moreover, this method is applied to Stiefel manifolds. By choosing a suitable reductive decomposition and an appropriated invariant covariant derivative, an algorithm for solving Problem 2.1 on the Stiefel manifold $\mathrm{St}_{n,k}$ is obtained that is also efficient from a computational point of view for $k \ll n$.

Acknowledgments

The first author has been supported by the German Federal Ministry of Education and Research (BMBF-Projekt, funding numbers: 05M20WWA and 05M20WBA, Verbundprojekt 05M2020 - DyCA). The second author has been supported by the German Federal Ministry of Education and Research (BMBF-Projekt 05M20WWA: Verbundprojekt 05M2020 - DyCA). The work of the last author has been supported by Fundação para a Ciência e Tecnologia (FCT) under the project UIDB/00048/2020 (`https://doi.org/10.54499/UIDB/00048/2020`).

References

[1] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization algorithms on matrix manifolds*, Princeton University Press, Princeton, NJ, 2008.

[2] R. L. Adler, J.-P. Dedieu, J. Y. Margulies, M. Martens, and M. Shub, *Newton's method on Riemannian manifolds and a geometric model for the human spine*, IMA J. Numer. Anal., 22 (2002), pp. 359–390.

[3] M. Barbero-Liñán and D. M. de Diego, *Retraction maps: A seed of geometric integrators*, Foundations of Computational Mathematics, 23 (2023), pp. 1335–1380.

[4] J. Batista, K. Krakowski, L. Machado, P. Martins, and F. Silva Leite, *Multi-source domain adaptation using C^1-smooth subspaces interpolation*, in 2016 IEEE International Conference on Image Processing (ICIP), 2016, pp. 2846–2850.

[5] J. Batista, K. Krakowski, and F. Silva Leite, *Exploring quasi-geodesics on Stiefel manifolds in order to smooth interpolate between domains*, in 2017 IEEE 56th Annual Conference on Decision and Control (CDC), 2017, pp. 6395–6402.

[6] T. Bendokat and R. Zimmermann, *Efficient quasi-geodesics on the Stiefel manifold*, in Geometric Science of Information, F. Nielsen and F. Barbaresco, eds., Cham, 2021, Springer International Publishing, pp. 763–771.

[7] T. Bendokat, R. Zimmermann, and P.-A. Absil, *A Grassmann manifold handbook: basic geometry and computational aspects*, Adv. Comput. Math., 50 (2024), pp. Paper No. 6, 51.

[8] A. Bloch, M. Camarinha, and L. Colombo, *Variational obstacle avoidance problem on Riemannian manifolds*, in 2017 IEEE 56th Annual Conference on Decision and Control (CDC), 2017, pp. 145–150.

[9] M. Camarinha, F. Silva Leite, and P. E. Crouch, *High-Order Splines on Riemannian Manifolds*, Proceedings of the Steklov Institute of Mathematics, 321 (2023), pp. 158–178.

[10] J. R. Cardoso and F. Silva Leite, *Exponentials of skew-symmetric matrices and logarithms of orthogonal matrices*, J. Comput. Appl. Math., 233 (2010), pp. 2867–2875.

[11] R. Caseiro, J. F. Henriques, P. Martins, and J. Batista, *Beyond the shortest path: Unsupervised domain adaptation by sampling subspaces along the spline flow*, in 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, pp. 3846–3854.

[12] R. Caseiro, P. Martins, J. F. Henriques, F. Silva Leite, and J. Batista, *Rolling Riemannian manifolds to solve the multi-class classification problem*, in 2013 IEEE Conference on Computer Vision and Pattern Recognition, 2013, pp. 41–48.

[13] G. S. Chirikjian, R. Mahony, S. Ruan, and J. Trumpf, *Pose Changes From a Different Point of View*, Journal of Mechanisms and Robotics, 10 (2018), p. 021008.

[14] P. Crouch and F. Silva Leite, *The de Dasteljau algorithm on lie groups and spheres*, Journal of Dynamical and Control Systems, 5 (1999), pp. 397–429.

[15] P. Crouch and F. Silva Leite, *Rolling motions of pseudo-orthogonal groups*, in 2012 IEEE 51st IEEE Conference on Decision and Control (CDC), 2012, pp. 7485–7491.

[16] J. J. Duistermaat and J. A. C. Kolk, *Lie groups*, Universitext, Springer-Verlag, Berlin, 2000.

[17] A. Edelman, T. A. Arias, and S. T. Smith, *The geometry of algorithms with orthogonality constraints*, SIAM J. Matrix Anal. Appl., 20 (1999), pp. 303–353.

[18] J. Gallier and J. Quaintance, *Differential geometry and Lie groups—a computational perspective*, vol. 12 of Geometry and Computing, Springer, Cham, 2020.

[19] J. R. Goodman, *Path Planning on Riemannian Manifolds with Applications to Quadrotor Load Transportation*, dissertation, Universidad Autónoma de Madrid, Madrid, 2023.

[20] E. Grong, *Controllability of rolling without twisting or slipping in higher dimensions*, SIAM Journal on Control and Optimization, 50 (2012), pp. 2462–2485.

[21] M. Hanik, *Geometric Data Analysis: Advancements of the Statistical Methodology and Applications*, dissertation, Freien Universität Berlin, Berlin, 2023.

[22] M. Hanik, E. Nava-Yazdani, and C. von Tycowicz, *De Casteljau's algorithm in geometric data analysis: Theory and application*, 2024. https://arxiv.org/abs/2402.07550.

[23] N. J. Higham, *Functions of matrices*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2008.

[24] R. A. Horn and C. R. Johnson, *Topics in matrix analysis*, Cambridge University Press, Cambridge, 1991.

[25] K. Hüper, M. Kleinsteuber, and F. Silva Leite, *Rolling Stiefel manifolds*, Internat. J. Systems Sci., 39 (2008), pp. 881–887.

[26] K. Hüper, K. A. Krakowski, and F. Silva Leite, *Rolling Maps and Nonlinear Data*, Springer International Publishing, Cham, 2020, pp. 577–610.

[27] K. Hüper, I. Markina, and F. Silva Leite, *A Lagrangian approach to extremal curves on Stiefel manifolds*, J. Geom. Mech., 13 (2021), pp. 55–72.

[28] K. Hüper, Y. Shen, and F. Silva Leite, *Geometric splines and interpolation on S^2: Numerical experiments*, in Proceedings of the 45th IEEE Conference on Decision and Control, 2006, pp. 6403–6407.

[29] K. Hüper and F. Silva Leite, *On the geometry of rolling and interpolation curves on S^n, SO_n, and Grassmann manifolds*, J. Dyn. Control Syst., 13 (2007), pp. 467–502.

[30] V. Jurdjevic, I. Markina, and F. Silva Leite, *Extremal curves on Stiefel and Grassmann manifolds*, J. Geom. Anal., 30 (2020), pp. 3948–3978.

[31] P. Kokkonen, *Étude du modèle des variétés roulantes et de sa commandabilité*, PhD thesis, Université Paris Sud - Paris XI, Paris, 2012.

[32] K. A. Krakowski, L. Machado, F. Silva Leite, and J. Batista, *A modified Casteljau algorithm to solve interpolation problems on Stiefel manifolds*, Journal of Computational and Applied Mathematics, 311 (2017), pp. 84–99.

[33] J. M. Lee, *Introduction to smooth manifolds*, vol. 218 of Graduate Texts in Mathematics, Springer, New York, second ed., 2013.

[34] ——, *Introduction to Riemannian manifolds*, vol. 176 of Graduate Texts in Mathematics, Springer, Cham, 2018.

[35] L. Machado, F. Pina, and F. Silva Leite, *Rolling maps for the essential manifold*, in Dynamics, Games and Science, J.-P. Bourguignon, R. Jeltsch, A. A. Pinto, and M. Viana, eds., Cham, 2015, Springer International Publishing, pp. 399–415.

[36] P. W. Michor, *Topics in differential geometry*, vol. 93 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2008.

[37] K. Nomizu, *Invariant affine connections on homogeneous spaces*, Amer. J. Math., 76 (1954), pp. 33–65.

[38] B. O'Neill, *Semi-Riemannian geometry*, vol. 103 of Pure and Applied Mathematics, Academic Press, Inc., New York, 1983. With applications to relativity.

[39] M. Schlarb, *Covariant derivatives on homogeneous spaces – horizontal lifts and parallel transport*, 2023. `https://arxiv.org/abs/2308.07089`.

[40] ——, *A multi-parameter family of metrics on Stiefel manifolds and applications*, Journal of Geometric Mechanics, 15 (2023), pp. 147–187.

[41] ——, *Rolling reductive homogeneous spaces*, 2023. `https://arxiv.org/abs/2308.08221`.

[42] M. Schlarb, K. Hüper, I. Markina, and F. Silva Leite, *Rolling Stiefel manifolds equipped with α-metrics*, Mathematics, 11 (2023).

[43] R. W. Sharpe, *Differential geometry*, vol. 166 of Graduate Texts in Mathematics, Springer-Verlag, New York, 1997.

[44] Y. Shen, K. Huper, and F. Leite, *Smooth interpolation of orientation by rolling and wrapping for robot motion planning*, in Proceedings 2006 IEEE International Conference on Robotics and Automation, 2006. ICRA 2006., 2006, pp. 113–118.

[45] J. Stoer and R. Bulirsch, *Introduction to numerical analysis*, vol. 12 of Texts in Applied Mathematics, Springer-Verlag, New York, third ed., 2002. Translated from the German by R. Bartels, W. Gautschi and C. Witzgall.

[46] T. Szwagier and X. Pennec, *Rethinking the Riemannian logarithm on flag manifolds as an orthogonal alignment problem*, in Geometric Science of Information, F. Nielsen and F. Barbaresco, eds., Cham, 2023, Springer Nature Switzerland, pp. 375–383.

[47] A. Séguin and D. Kressner, *Hermite interpolation with retractions on manifolds*, 2024. `https://arxiv.org/abs/2212.12259`.

[48] L. W. Tu, *Differential geometry*, vol. 275 of Graduate Texts in Mathematics, Springer, Cham, 2017. Connections, curvature, and characteristic classes.

[49] R. Vemulapalli and R. Chellappa, *Rolling rotations for recognizing human actions from 3d skeletal data*, in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 4471–4479.

[50] K. Ye, K. S.-W. Wong, and L.-H. Lim, *Optimization on flag manifolds*, Mathematical Programming, 194 (2022), pp. 621–660.

[51] M. Zefran, V. Kumar, and C. Croke, *On the generation of smooth three-dimensional rigid body motions*, IEEE Transactions on Robotics and Automation, 14 (1998), pp. 576–589.

[52] R. Zimmermann, *Hermite interpolation and data processing errors on Riemannian matrix manifolds*, SIAM Journal on Scientific Computing, 42 (2020), pp. A2593–A2619.

[53] R. Zimmermann, *Manifold interpolation and model reduction*, in Volume 1 System- and Data-Driven Methods and Algorithms, 2021, pp. 229–274.

[54] R. Zimmermann and K. Hüper, *Computing the Riemannian logarithm on the Stiefel manifold: Metrics, methods, and performance*, SIAM Journal on Matrix Analysis and Applications, 43 (2022), pp. 953–980.