Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Seminar for Applied Mathematics SAM

Scattered Manifold-Valued Data Approximation

P. Grohs and M. Sprecher and T. Yu

Research Report No. 2014-08
March 2014

Seminar für Angewandte Mathematik
Eidgenössische Technische Hochschule
CH-8092 Zürich
Switzerland

Funding SNF: 140635

Scattered Manifold-Valued Data Approximation *

Philipp Grohs Markus Sprecher Thomas Yu

March 5, 2014

Abstract

We consider the problem of approximating a smooth function f from an Euclidean domain to a manifold M by scattered samples $(f(\xi_i))_{i\in\mathcal{I}}$, where the data sites $(\xi_i)_{i\in\mathcal{I}}$ are assumed to be locally close but can otherwise be far apart points scattered throughout the domain. We introduce a natural approximant based on combining the moving least square method and the Karcher mean. We prove that the proposed approximant inherits the accuracy order and the smoothness from its linear counterpart. The analysis also tells us that the use of Karcher's mean (dependent on a Riemannian metric and the associated exponential map) is inessential and one can replace it by a more general notion of 'center of mass' based on a general retraction on the manifold. Consequently, we can substitute the Karcher mean by a more computationally efficient mean. We illustrate our work with numerical results which confirm our theoretical findings.

AMS Subject Classification (2010): primary 65D07 secondary 65D15, 53B20, 41AXX, 35RXX
Keywords: Riemannian data, manifold-valued function, approximation, scattered data, model reduction

1 Introduction

Let $f : \Omega \subset \mathbb{R}^d \to M$ with M a Riemannian manifold be an unknown smooth function and we only know its value at a set of distinct points $(\xi_i)_{i\in\mathcal{I}} \subset \Omega$. We are concerned with finding an approximant to f. Such an approximation problem for manifold-valued data arises in numerical geometric integration [16] and more recently in fast online methods for dimensionality reduced-order models [4]. In a Stanford press release associated with the aeronautical engineering publications [4, 2, 3], it is emphasized that being able to accurately and efficiently interpolate on manifolds is a key to fast online prediction of aerodynamic flutter, which, in turn, may help saving lives of pilots and passengers.

In the aforementioned references it was implicitly assumed that the data points $(f(\xi_i))_{i\in\mathcal{I}} \in M$ are close enough so that they can all be mapped to a single tangent space T_pM by the inverse exponential map log. In these previous works the base point $p \in M$ is typically one of $(f(\xi_i))_{i\in\mathcal{I}}$ and the choice can be quite arbitrary. In this setting, the problem simply reduces to a linear approximation problem on the tangent space T_pM. To approximate the value $f(x) \in M$, use any standard linear method (polynomial, spline, radial basis function etc.) to interpolate the values $(\log(p, (f(\xi_i)))_{i\in\mathcal{I}} \subset T_pM$ at the abscissa $(\xi_i)_{i\in\mathcal{I}}$, then evaluate the interpolant Q at the desired value x, and followed by applying the exponential map to get the approximation $f(x) \sim \exp(p, Q(x))$.

This 'push-interpolate-pull' technique only works when all the available data points $(f(\xi_i))_{i\in\mathcal{I}}$ fall within the injectivity radius of the point p, and the method only provides an approximation for $f(x)$ where x is near p. In this case the problem is local, and the topology of the manifold plays no role. One may then question what would be the difference if one uses the push-interpolate-pull approach but with the exponential map replaced by an arbitrary chart. With the exponential map, the push-interpolate-pull method respects the symmetry, if any, of the manifold and in some sense respects the metric of the manifold. However, it is not clear what is the practical advantage of the latter and if respecting symmetry is not the main concern, one is free to replace the exponential map by a retraction (see, e.g., [1, 18]) on the manifold.

*The research of PG and MS was supported by the Swiss National Fund under grant SNF 140635. The research of TY is supported in part by the National Science Foundation grant DMS 1115915.

The problem is more challenging when we have available data points $(f(\xi_i))_{i\in\mathcal{I}} \in M$ scattered at different parts of the manifold, and the manifold has a nontrivial topology. In this setting, we desire an approximation method with the following properties:

(i) The method is well-defined as long as the scattered data sites are reasonably close locally, but can otherwise be far apart globally.

(ii) The approximant should provide a decent approximation order when f is smooth. Since standard approximation theory tells us that for linear data there are methods which provide an accuracy of $O(h^m)$ when f is C^m smooth, and h is the meshwidth, it is natural to ask for a method for manifold-valued data with the same accuracy order.

(iii) The approximant itself should be smooth.

(iv) The approximant should be efficiently computable for any given x.

Such a method ought to 'act locally but think globally', meaning that the approximating value should only depend on the data $f(\xi_i)$ for ξ_i near x, but yet the approximant should be smooth and accurate for all x in the domain. This forces the method to be genuinely nonlinear.

In the shift-invariant setting, i.e. when $\Omega = \mathbb{R}^d$ and $(\xi_i)_{i\in\mathcal{I}} = h\mathbb{Z}^d$, the above problem was solved successfully first by a subdivision technique [23, 9] and more recently by a method based on combining a general quasi-interpolant with the Karcher mean [10]. Even more recently, the work [13] used a projection-based approach to generalize B-spline quasiinterpolation for regularly spaced data. In either case, it was shown that an approximation method for linear data can be suitably adapted to manifold-valued data without jeopardizing the accuracy or smoothness properties of the original linear method. This kind of (smoothness or approximation order) equivalence properties are analyzed by a method known as proximity analysis. It is also shown that in certain setups, the smoothness or approximation order equivalence can breakdown in unexpected ways [18, 23, 7, 13].

In this paper we provide a solution to the above problem in the multivariate scattered data setting. More precisely we combine the ideas of [10, 12] with the classical linear theory of scattered data approximation [22] to arrive at approximants which satisfy (i)-(iv) above.

We give a brief outline of this work. The following Section2 presents a brief overview of approximation results for scattered data approximation of Euclidean data. Then in Section 3 we present our generalization of the linear theory to the manifold-valued setting. This section also contains our main result regarding the approximation power of our nonlinear construction. It turns out that our scheme retains the optimal approximation rate as expected from the linear case. This is formalized in Theorem 3.5. In this theorem the dependence of the approximation rate on the geometry of M is made explicit and appears in form of norms of iterated covariant derivatives of the log function of M. To measure the smoothness of an M-valued function we utilize a smoothness descriptor introduced in [12] and which forms a natural generalization of Hölder norms to the manifold-valued setting. Our results also hold true for arbitrary choices of retractions. We discuss this extension in Subsection 3.3. Finally in Section 4 we present numerical experiments for the approximation of functions with values in the sphere and in the manifold of symmetric positive definite matrices. In all cases the approximation results derived in Section 3 are confirmed. We also examine an application to the interpolation of reduced order models and compare our method to the method introduced in [5] where it turns our that our method delivers superior approximation power.

2 Scattered Data Approximation in Linear Spaces

In this section we present classical results concerning the approximation of scattered data in Euclidean space. Our exposition mainly follows the monograph [22].

2.1 General Setup

We start by describing a general setup for scattered data approximation. Given a set $\Xi = (\xi_i)_{i\in\mathcal{I}} \subset \Omega \subset \mathbb{R}^d$ and associated basis functions $\Phi := (\varphi_i)_{i\in\mathcal{I}} \subset C(\Omega, \mathbb{R})$, where $C(\Omega, \mathbb{R})$ denotes the set of real-valued

continuous function on Ω. The linear quasi-interpolation procedure, applied to a continuous function $f : \Omega \to \mathbb{R}$, is defined as

$$\mathbf{Q}_{(\Xi,\Phi)} f(x) := \sum_{i \in \mathcal{I}} \varphi_i(x) f(\xi_i). \tag{1}$$

Essential for the approximation power of the operator $\mathbf{Q}_{(\Xi,\Phi)}$ is the property that polynomials up to a certain degree are reproduced:

Definition 2.1. *The pair (Ξ, Φ) reproduces polynomials of degree m if*

$$\sum_{i \in \mathcal{I}} \varphi_i(x) p(\xi_i) = p(x) \quad \text{for all } p \in \mathcal{P}_m(\mathbb{R}^d), \tag{2}$$

where $\mathcal{P}_m(\mathbb{R}^d)$ denotes the space of polynomials of (total) degree $\leq m$ on $\mathbb{R}^d$.

In particular, if (Ξ, Φ) reproduces polynomials of degree $m \geq 0$ then necessarily

$$\sum_{i \in \mathcal{I}} \varphi_i(x) = 1 \quad \text{for all } x \in \Omega. \tag{3}$$

Definition 2.2. *We define the local meshwidth $h_{(\Xi,\Phi)} : \Omega \to \mathbb{R}_{\geq 0}$ by*

$$h_{(\Xi,\Phi)}(x) := \sum_{i \in \mathcal{I}(x)} |\xi_i - x|,$$

where

$$\mathcal{I}(x) := \{i \in \mathcal{I} | \varphi_i(x) \neq 0\}$$

and $|\cdot|$ denotes the Euclidean norm.

The following example gives a simple procedure to approximate univariate data with polynomial reproduction degree 1.

Example 2.3. *Let $\Omega = [0, 1]$, $0 = \xi_1 < \xi_2 \cdots < \xi_n = 1$ and*

$$\varphi_i(x) := \begin{cases} \frac{x - \xi_{i-1}}{\xi_i - \xi_{i-1}} & \text{if } i > 1 \text{ and } \xi_{i-1} \leq x < \xi_i \\ \frac{\xi_{i+1} - x}{\xi_{i+1} - \xi_i} & \text{if } i < n \text{ and } \xi_i \leq x < \xi_{i+1} \\ 0 & \text{otherwise.} \end{cases}$$

These functions, also known as hat functions, satisfy the polynomial reproduction property with $m = 1$. The local meshwidth for $\xi_i < x < \xi_{i+1}$ is $h_{(\Xi,\Phi)}(x) = \xi_{i+1} - \xi_i$.

The following result gives a bound for the approximation error in terms of the local meshwidth and the polynomial reproduction degree. We will denote by $\overline{\Omega}$ the closure of Ω.

Theorem 2.4. *Assume that $\Omega \subset \mathbb{R}^d$ is a domain with smooth boundary and that (Ξ, Φ) reproduces polynomials of degree m. Then for all $1 \leq k < m + 1$ and $f \in C^k(\overline{\Omega}, \mathbb{R})$ there exists a generic constant $C > 0$, independent of f, Φ and Ξ, such that*

$$|f(x) - \mathbf{Q}_{(\Xi,\Phi)} f(x)| \leq C \sup_{i \in \mathcal{I}} |\varphi_i(x)| \|f\|_{C^k(\overline{\Omega},\mathbb{R})} h_{(\Xi,\Phi)}(x)^k \text{ for all } x \in \Omega.$$

Remark 2.5. *The theorem could be generalized to hold also for a larger class of domains Ω satisfying an interior cone condition, see [22].*

We omit a proof here as a general theorem will be proven in the next section. In contrast to other results, such as those in [22] Theorem 2.4 poses no restrictions on the data $\Xi = (\xi_i)_{i \in \mathcal{I}}$. In the following subsection we will show how the approximation results in [22] follow as a corollary to the previous Theorem 2.4.

2.2 Moving Least Squares

In this subsection, we will follow [22] and show how to construct a set of basis functions $\Phi^{\Xi,m,\delta} = (\varphi_i^{\Xi,m,\delta})_{i\in\mathcal{I}}$ with polynomial reproduction degree m for a set $\Xi = (\xi_i)_{i\in\mathcal{I}} \subset \Omega \subset \mathbb{R}^d$ which is (m,δ)-unisolvent.

Definition 2.6. *A set $\Xi = (\xi_i)_{i\in\mathcal{I}} \subset \Omega \subset \mathbb{R}^d$ is called (m,δ)-unisolvent if there exist no $x \in \Omega$ and $p \in \mathcal{P}_m(\mathbb{R}^d)$ with $p \neq 0$ and $p(\xi_i) = 0$ for all $\xi_i \in \Xi$ with $|\xi_i - x| \leq \delta$.*

Let $\alpha\colon [0,\infty) \to [0,1]$ with $\alpha(0) = 1$, $\alpha(z) = 0$ for $z \geq 1$ and $\alpha'(0) = \alpha'(1) = 0$, e.g. the Wendland function

$$\alpha(x) = \begin{cases} (1+4x)(1-x)^4 & 0 \leq x \leq 1 \\ 0 & x > 1. \end{cases}$$

We define

$$\varphi_j^{\Xi,m,\delta}(x) = \alpha_\delta(|x-\xi_j|)p_x(\xi_j)$$

where $\alpha_\delta(x) = \alpha(x/\delta)$ and $p_x \in \mathcal{P}_m(\mathbb{R}^d)$ is the unique solution of

$$\sum_{j\in\mathcal{I}} \alpha_\delta(|x-\xi_j|)p_x(\xi_j)p(\xi_j) = p(x) \qquad \forall\, p \in \mathcal{P}_m(\mathbb{R}^d).$$

It is an easy task to verify that p_x is well-defined and $(\Xi, \Phi^{\Xi,m,\delta})$ satisfies the polynomial reproduction property (2) of degree m.

Theorem 2.4 represents an approximation result for a single set of data sites Ξ and basis functions Φ. Let S be an index set. The approximation theory of the moving least squares method, as presented e.g. in [22] considers sets $(\Xi_j, \Phi^{\Xi_j,m,\delta_j})_{j\in S}$ of data sites. The theory in [22] requires $(\Xi_j,\delta_j)_{j\in S}$ to be quasi-uniform as defined below.

Definition 2.7. *A set $X = (\Xi_j,\delta_j)_{j\in S}$ of tuples is said to be quasi-uniform if there exists a constant $c > 0$, independent of $j \in S$ such that*

$$\delta_j \leq c\, q_{\Xi_j},$$

where q_Ξ is the separation distance defined by

$$q_\Xi \quad := \quad 1/2 \min_{\substack{i,j\in\mathcal{I} \\ i\neq j}} |\xi_i - \xi_j|. \tag{4}$$

Example 2.8. *Let $\Omega = [-1,1]$ and $\mathcal{I}_N = \{0,\dots,N-1\}$. For $i \in \mathcal{I}_N$ and $m \in \mathbb{N}$ we choose the node ξ_i^N uniformly at random in the interval $(-1+2i/N, -1+2(i+1)/N)$ and $\delta_N = 2(m+1)/N$. Then the set $\Xi_N = (\xi_i^N)_{i=0}^{N-1}$ is (m,δ_N)-unisolvent for all $N \in \mathbb{N}$ and $h_{(\Xi,\Phi)}(x) \leq 2(m+1)(2m+3)/N$ for all $x \in \Omega$. However $X = (\Xi_N, \Phi^{\Xi_N,m,\delta_N})_{N\in\mathbb{N}}$ is in general not quasiuniform.*

Example 2.9. *Let $\Omega = [-1,1]$ and $\mathcal{I}_N = \{0,\dots,N-1\}$. For $i \in \mathcal{I}_N$ and $m \in \mathbb{N}$ we choose the node ξ_i^N uniformly at random in the interval $[-1+(4i+1)/(2N), -1+(4i+3)/(2N)]$, $\delta_N = (2m+3)/N$. Then the set $\Xi_N = (\xi_i^N)_{i=0}^{N-1}$ is (m,δ_N)-unisolvent for all $N \in \mathbb{N}$, $h_{(\Xi,\Phi)}(x) \leq (2m+3)(2m+5)/N$ for all $x \in \Omega$ and $X = (\Xi_N, \Phi^{\Xi_N,m,\delta_N})_{N\in\mathbb{N}}$ is quasiuniform with $c = 2(2m+3)$.*

The following theorem is the main result of [22] regarding the approximation with moving least squares.

Theorem 2.10 ([22])**.** *Assume that $\Omega \subset \mathbb{R}^d$ is a domain with smooth boundary, $X = (\Xi_j,\delta_j)_{j\in S}$ is quasi-uniform, Ξ_j is (m,δ_j)-unisolvent for all $j \in S$ and $\Phi_j = \Phi^{\Xi_j,m,\delta_j}$ are the basis functions as defined above. Then for all $1 \leq k < m+1$ and $f \in C^k(\overline{\Omega},\mathbb{R})$ there exists a generic constant $C > 0$, independent of f, such that*

$$\|f - \mathbf{Q}_{(\Xi_j,\Phi_j)} f\|_{L^\infty(\Omega)} \leq C\|f\|_{C^k(\overline{\Omega},\mathbb{R})}\delta_j^k \text{ for all } x \in \Omega.$$

for all $j \in S$.

We show that Theorem 2.10 is a consequence of Theorem 2.4.

Proof. In the proof we use the letter C as a symbol for a generic constant whose value may change from equation to equation. Denote by $\mathcal{I}_j$ the index set associated with the sampling set $\Xi_j = (\xi_i^j)_{i\in\mathcal{I}_j}$. We first show that $\max_{x\in\Omega} |\mathcal{I}_j(x)| \leq C$ where C is independent of j. Note that

$$\mathcal{I}_j(x) \subseteq \{i \in \mathcal{I}_j | |x - \xi_i^j| \leq \delta_j\}.$$

Hence the pairwise disjoint balls with centers $(\xi_i^j)_{i\in\mathcal{I}_j(x)}$ and radius q_{Ξ_j} lie in the ball with center x and radius $\delta_j + q_{\Xi_j}$. As the fraction $(\delta_j + q_{\Xi_j})/q_{\Xi_j}$ is bounded from above by $(1+c)$, with c being the constant from Definition 2.7, we have that $\max_{x\in\Omega} |\mathcal{I}_j(x)|$ is bounded by the number of balls with radius 1 that fit into a ball of radius $(1 + c)$. Consequently

$$h_{(\Xi_j,\Phi_j)}(x) = \sum_{i\in\mathcal{I}_j(x)} |\xi_i^j - x| \leq C\delta_j$$

and hence by Theorem 2.4 we have, writing $\Phi_j = (\varphi_i^j)_{i\in\mathcal{I}_j}$,

$$\begin{aligned}
\|f - \mathbf{Q}_{(\Xi_j,\Phi_j)} f\|_{L^\infty(\Omega)} &= \sup_{x\in\Omega} |f(x) - \mathbf{Q}_{(\Xi_j,\Phi_j)} f(x)| \\
&\leq \sup_{x\in\Omega} C \sup_{i\in\mathcal{I}_j} |\varphi_i^j(x)| \|f\|_{C^k(\overline{\Omega},\mathbb{R})} h_{(\Xi_j,\Phi_j)}(x)^k \\
&\leq C \sup_{i\in\mathcal{I}_j} |\varphi_i^j(x)| \|f\|_{C^k(\overline{\Omega},\mathbb{R})} \delta_j^k.
\end{aligned}$$

Finally we note that, due to [22, Theorem 4.7 (2)] we have

$$\sup_{x\in\Omega,\, j\in S} \sup_{i\in\mathcal{I}_j(x)} |\varphi_i^j(x)| < \infty,$$

which yields the desired result. □

3 Scattered Data Approximation in Manifolds

The present section extends Theorem 2.4 and Theorem 2.10 to the case of manifold-valued functions $f : \Omega \to M$ with a Riemannian manifold M. First, in Subsection 3.1 we present a geometric generalization of the operator $\mathbf{Q}_{(\Xi,\Phi)}$ to the manifold-valued case. The construction is based on the idea to replace affine averages by a Riemannian mean as studied e.g. in [15]. Subsequently, in Subsection 3.2 we study the resulting approximation error. Finally, in Subsection 3.3 we extend our results to the case of more general notions of geometric average, induced by arbitrary retractions as studied e.g. in [11].

3.1 Definition of Riemannian Moving Least Squares

In this section we consider a Riemannian manifold M with distance metric $d : M \times M \to \mathbb{R}_{\geq 0}$ and metric tensor g, inducing a norm $|\cdot|_{g(p)}$ on the tangent space T_pM at $p \in M$.

Extending the classical theory which we briefly described in Section 2 we now aim to construct approximation operators for functions $f : \Omega \to M$. We follow the ideas of [19, 20, 21, 10, 12] where the sum in (1) is interpreted as a weighted mean of the points $(f(\xi_i))_{i\in\mathcal{I}(x)}$. Due to (3) this is justified.

In particular, for weights $\Gamma = (\gamma_i)_{i\in\mathcal{I}} \subset \mathbb{R}$ with $\sum_{i\in\mathcal{I}} \gamma_i = 1$ and points $\Pi = (p_i)_{i\in\mathcal{I}} \subset M$ we can define the Riemannian average

$$av_M(\Gamma, \Pi) := \operatorname*{argmin}_{p\in M} \sum_{i\in\mathcal{I}} \gamma_i d(p, p_i)^2. \tag{5}$$

One can show [15, 21] that $av_M(\Gamma, \Pi)$ is a well-defined operation, if the diameter of the set Π is small enough:

Theorem 3.1 ([21]). *Given a weight sequence $\Gamma = (\gamma_i)_{i\in\mathcal{I}} \subset \mathbb{R}$ with $\sum_{i\in\mathcal{I}} \gamma_i = 1$. Let $p_0 \in M$ and denote for $\rho > 0$ by B_ρ the geodesic ball of radius ρ around p_0. Then there exist $0 < \rho_1 \leq \rho_2 < \infty$, depending only on $\sum_{i\in\mathcal{I}} |\gamma_i|$ and the geometry of M such that for all points $\Pi = (p_i)_{i\in\mathcal{I}} \subset B_{\rho_1}$ the functional $\sum_{i\in\mathcal{I}} \gamma_i d(p, p_i)^2$ assumes a unique minimum in B_{ρ_2}.*

Whenever the assumptions of Theorem 3.1 hold true, the Riemannian average is uniquely determined by the first order condition

$$\sum_{i\in\mathcal{I}} \gamma_i \log\left(av_M(\Gamma, \Pi), p_i\right) = 0, \tag{6}$$

which could alternatively be taken as a definition of the weighted average in M, see also [21].

We can now define an M-valued analogue for Equation (1).

Definition 3.2. *Denoting $\Phi(x) := (\varphi_i(x))_{i\in\mathcal{I}} \subset \mathbb{R}$ and $f(\Xi) := (f(\xi_i))_{i\in\mathcal{I}} \subset M$ we define the nonlinear moving least squares approximant*

$$\mathbf{Q}^M_{(\Xi,\Phi)} f(x) := av_M(\Phi(x), f(\Xi)) \in M. \tag{7}$$

It is clear that in the linear case this definition coincides with (1). Furthermore it is easy to see that the smoothness of the basis functions Φ gets inherited by the approximation procedure $\mathbf{Q}^M_{(\Xi,\Phi)}$, see e.g. [20].

Remark 3.3. *We wish to emphasize that the approximation procedure as defined in Definition 3.2 is completely geometric in nature. In particular it is invariant under isometries of M. In mechanics this leads to the desirable property of objectivity.*

3.2 Approximation Error

We now wish to assess the approximation error of the nonlinear operator $\mathbf{Q}^M_{(\Xi,\Phi)}$ and generalize Theorem 2.4 to the M-valued case.

3.2.1 The Smoothness Descriptor of Manifold-Valued Functions

Two basic things need to be considered to that end. First we need to decide how we measure the error between the original function f and its approximation $\mathbf{Q}^M_{(\Xi,\Phi)} f$. This will be done pointwise using the geodesic distance $d\colon M \times M \to \mathbb{R}_{\geq 0}$ on M. Slightly more subtle is the question what is the right analogue to the term $\|f\|_{C^k(\overline{\Omega},\mathbb{R})}$ in the manifold-valued case? In [12] a so-called *smoothness descriptor* has been introduced to measure norms of derivatives of M-valued functions. Its definition requires the notion of covariant derivative in a Riemannian manifold. With $\frac{D}{dx^l}$ we denote the covariant partial derivative along f with respect to x^l. That is, given a function $f : \Omega \to M$ and a vector field $W : \Omega \to TM$ attached to f, i.e., $W(x) \in T_{f(x)}M$ for all $x \in \Omega$. Then in coordinates on M, the covariant derivative of W in x^l reads

$$\frac{D}{dx^l} W^r(x) := \frac{dW^r}{dx^l}(x) + \Gamma^r_{ij}(f(x)) \frac{df^i}{dx^l} W^j(x),$$

where we sum over repeated indices and denote with Γ^r_{ij} the Christoffel symbols associated to the metric of M, [6]. For iterated covariant derivatives we introduce the symbol $\mathcal{D}^{\vec{l}} f$ which means covariant partial differentiation along f with respect to the multi-index $\vec{l}$ in the sense that

$$\mathcal{D}^{\vec{l}} f := \frac{D}{dx^{l_k}} \cdots \frac{D}{dx^{l_2}} \frac{df}{dx^{l_1}}, \qquad \vec{l} \in \{1, \ldots, d\}^k,\ k \in \mathbb{N}_0. \tag{8}$$

Additionally we define $\mathcal{D}^{\vec{l}} f := 1$ (a constant function $\Omega \to \mathbb{R}$) if $\dim \vec{l} = 0$. Note that (8) differs from the usual multi-index notation, which cannot be used because covariant partial derivatives do not commute. The smoothness descriptor of an M-valued function is defined as follows.

Definition 3.4 (Smoothness Descriptor)**.** *For a function $f : \Omega \to M$, $k \geq 1$ we define the homogeneous k-th order smoothness descriptor*

$$\dot{\Theta}_{\infty,k,\Omega}(f) := \sum_{\substack{\vec{l}_j \in [d]^{m_j},\ j=1,\dots,k \\ \sum_{j=1}^k m_j = k}} \sup_{x\in\Omega} \prod_{j=1}^k \left| \mathcal{D}^{\vec{l}_j} f(x) \right|_{g(f(x))},$$

and its inhomogeneous version

$$\Theta_{\infty,k,\Omega}(f) := \sum_{l=1}^k \dot{\Theta}_{\infty,l,\Omega}(f).$$

The smoothness descriptor as defined above represents a geometric analogue of the classical notions of Hölder norms and seminorms. Note that, even in the Euclidean case (for instance $M = \mathbb{R}$) the expression $\dot{\Theta}_{\infty,k,\Omega}(f)$ is not equal to the Hölder seminorm $|f|_{C^k(\overline{\Omega},\mathbb{R})}$, as additional terms are present in the definition of $\dot{\Theta}_{\infty,k,\Omega}(f)$. But we have the implications

$$|f|_{C^k(\overline{\Omega},\mathbb{R})} < \infty \Leftrightarrow \dot{\Theta}_{\infty,k,\Omega}(f) < \infty.$$

In the proof of Theorem 3.5 it will become clear why the additional terms in $\dot{\Theta}_{\infty,k,\Omega}(f)$ are needed in the case of general M where, in contrast to $M = \mathbb{R}$, higher order covariant derivatives of the logarithm mapping need not vanish, compare also Remark 3.6 below.

3.2.2 Further Geometric Quantities

Coming back to the anticipated generalization of Theorem 2.4, we also aim to quantify exactly to which extent the approximation error depends on the geometry of M. To this end let $\log(p,\cdot) : M \to T_pM$ be the inverse of the exponential map at p. Denote by ∇_1, ∇_2 the covariant derivative of a bivariate function with respect to the first and second argument, respectively. In particular, for $l \in \mathbb{N}$ we will require the derivatives

$$\nabla_2^l \log(p,q) : (T_qM)^l \to T_pM$$

and their norms

$$\|\nabla_2^l \log(p,q)\| = \sup_{v_1,\dots,v_l \in T_qM} \frac{\left|\nabla_2^l \log(p,q)\,(v_1,\dots,v_l)\right|_{g(p)}}{\prod_{i=1}^l |v_i|_{g(q)}}. \tag{9}$$

its convex hull.

3.2.3 Main Approximation Result

Now we are ready to state and prove our main result. We write

$$\Omega_x := \operatorname{conv}(\{x\} \cup \{\xi_i : i \in \mathcal{I}(x)\})$$

for $x \in \Omega$ where we denote by $\operatorname{conv}(D)$ the convex hull of a set $D \subset \mathbb{R}^d$. Now our main result reads as follows.

Theorem 3.5. *Assume that $\Omega \subset \mathbb{R}^d$ is a domain with smooth boundary and (Ξ,Φ) reproduces polynomials of degree m. Then for all $1 \leq k < m+1$ and $f \in C^k(\overline{\Omega}, M)$ there exists a generic constant $C > 0$, independent of f, Ξ, Φ and the geometry of M such that*

$$d\left(\mathbf{Q}^M_{(\Xi,\Phi)} f(x), f(x)\right) \leq C\Theta_{\infty,k,\Omega}(f) \sup_{i\in\mathcal{I}(x)} |\varphi_i(x)| \sup_{1\leq r\leq k} \sup_{y\in\Omega_x\cap\Omega} \left\| \nabla_2^r \log\left(\mathbf{Q}^M_{(\Xi,\Phi)} f(x), f(y)\right)\right\| h_{(\Xi,\Phi)}(x)^k \tag{10}$$

for all $x \in \Omega$.

Proof. In the proof we use the letter C as a symbol for a generic constant whose value may change from equation to equation. Furthermore we will assume for simplicity that Ω is convex, the general case can be handled by smoothly extending the function f across the boundary of Ω. Note that if Ω is convex we have that $\Omega_x \subset \Omega$ for all $x \in \Omega$. The existence of a bounded extension operator for smooth domains is classical [14].

We shall use the balance law (6) which implies that we can write

$$\varepsilon(x) := \log\left(\mathbf{Q}^M_{(\Xi,\Phi)} f(x), f(x)\right) = \log\left(\mathbf{Q}^M_{(\Xi,\Phi)} f(x), f(x)\right) - \sum_{i\in\mathcal{I}(x)} \varphi_i(x) \log\left(\mathbf{Q}^M_{(\Xi,\Phi)} f(x), f(\xi_i)\right). \quad (11)$$

Now we consider the function $G : \Omega \times \Omega \to TM$ defined by

$$G(x,y) := \log\left(\mathbf{Q}^M_{(\Xi,\Phi)} f(x), f(y)\right) \in T_{\mathbf{Q}^M_{(\Xi,\Phi)} f(x)} M. \quad (12)$$

Since, for fixed $x \in \Omega$, the function G maps into a linear space we can perform a Taylor expansion of G in y around the point $(x,x) \in \Omega \times \Omega$ and obtain

$$G(x,y) = \sum_{\substack{\vec{l}\in\mathbb{N}^d \\ |\vec{l}\,|<k}} \frac{(y-x)^{\vec{l}}}{\vec{l}\,!} \partial_2^{\vec{l}} G(x,x) + \sum_{\substack{\vec{l}\in\mathbb{N}^d \\ |\vec{l}\,|=k}} R_{\vec{l}}(x,y)(y-x)^{\vec{l}}, \quad (13)$$

where for any $l = (l_1, \dots, l_d) \in \mathbb{N}^d$ and $z = (z_1, \dots, z_n) \in \mathbb{R}^d$ we define

$$|\vec{l}\,| := \sum_{i=1}^d l_i, \qquad z^{\vec{l}} := \prod_{i=1}^n (z_i)^{l_i}, \qquad \vec{l}\,! := \prod_{i=1}^n l_i!, \qquad \partial_2^{\vec{l}} G(x,z) := \frac{\partial^{|\vec{l}\,|} G(x,z)}{\prod_{i=1}^n (\partial z_i)^{l_i}}$$

and

$$R_{\vec{l}}(x,y) := \frac{|\vec{l}\,|}{\vec{l}\,!} \int_0^1 (1-t)^{|\vec{l}\,|-1} \partial_2^{\vec{l}} G(x, x+t(y-x)) dt. \quad (14)$$

We insert (13) into (11) and get the following expression for $\varepsilon(x)$:

$$\varepsilon(x) = G(x,x) - \sum_{i\in\mathcal{I}(x)} \varphi_i(x) \left(\sum_{\substack{\vec{l}\in\mathbb{N}^d \\ |\vec{l}\,|<k}} \frac{(\xi_i - x)^{\vec{l}}}{\vec{l}\,!} \partial_2^{\vec{l}} G(x,x) + \sum_{\substack{\vec{l}\in\mathbb{N}^d \\ |\vec{l}\,|=k}} R_{\vec{l}}(x,\xi_i)(\xi_i - x)^{\vec{l}} \right) \quad (15)$$

Exchanging summation order in (15) yields

$$\varepsilon(x) = \underbrace{G(x,x) - \sum_{\substack{\vec{l}\in\mathbb{N}^d \\ |\vec{l}\,|<k}} \sum_{i\in\mathcal{I}(x)} \varphi_i(x) \frac{(\xi_i - x)^{\vec{l}}}{\vec{l}\,!} \partial_2^{\vec{l}} G(x,x)}_{(I)} + \underbrace{\sum_{\substack{\vec{l}\in\mathbb{N}^d \\ |\vec{l}\,|=k}} \sum_{i\in\mathcal{I}(x)} \varphi_i(x) R_{\vec{l}}(x,\xi_i)(\xi_i - x)^{\vec{l}}}_{(II)}.$$

We will show that $(I) = 0$ and $(II) = O(h(x)^k)$ which implies our claim.

Let us start by showing that $(I) = 0$. As a first observation we note that, due to (3), we can write

$$(I) = \sum_{\substack{\vec{l}\in\mathbb{N}^d \\ 0<|\vec{l}\,|<k}} \underbrace{\sum_{i\in\mathcal{I}(x)} \varphi_i(x) \frac{(\xi_i - x)^{\vec{l}}}{\vec{l}\,!} \partial_2^{\vec{l}} G(x,x)}_{(I_{\vec{l}})}. \quad (16)$$

We claim that $(I_{\vec{l}}) = 0$ for all $\vec{l}\mathbb{N}^d$ with $|\vec{l}\,| < k$. Indeed, pick $x_* \in \Omega$ arbitrary. Then, by the polynomial reproduction property (2) and $k \le m+1$ we get

$$\sum_{i\in\mathcal{I}(x)} \varphi_i(x) \frac{(\xi_i - x_*)^{\vec{l}}}{\vec{l}\,!} \partial_2^{\vec{l}} G(x_*,x_*) = \partial_2^{\vec{l}} G(x_*,x_*) \frac{(x-x_*)^{\vec{l}}}{\vec{l}\,!} \quad \text{for all } x,\, x_* \in \Omega. \quad (17)$$

Setting $x_* = x$ in (17) yields

$$(I_{\vec{l}}) = \sum_{i\in\mathcal{I}(x)} \varphi_i(x)\frac{(\xi_i - x)^{\vec{l}}}{\vec{l}\,!}\partial_2^{\vec{l}}G(x,x) = \partial_2^{\vec{l}}G(x,x)\frac{(x-x)^{\vec{l}}}{\vec{l}\,!} = 0$$

which proves the desired claim.

We now move on to prove our second claim, namely that $(II) = O(h(x)^k)$. To this end we need to estimate, for any $\vec{l}\in\mathbb{N}^d$ with $|\vec{l}\,| = k$ the quantity

$$(II)_{\vec{l}} := \sum_{i\in\mathcal{I}(x)} \varphi_i(x) R_{\vec{l}}(x,\xi_i)(\xi_i - x)^{\vec{l}}. \tag{18}$$

To this end we consider, for fixed $\vec{l}$ and $i\in\mathcal{I}(x)$, the quantity $R_{\vec{l}}(x,\xi_i)$. Inserting Definition (12) and using the chain rule we obtain that

$$\left|\partial_2^{\vec{l}}G(x,y)\right|_{g\left(\mathbf{Q}^M_{(\Xi,\Phi)}f(x)\right)} \le C \sum_{\substack{1\le r\le k,\ \vec{l}_j\in[d]^{m_j}\\ \sum_{j=1}^r m_j = k}} \left|\nabla_2^r \log\left(\mathbf{Q}^M_{(\Xi,\Phi)}f(x), f(y)\right)\left(\mathcal{D}^{\vec{l}_1}f(y),\dots,\mathcal{D}^{\vec{l}_r}f(y)\right)\right|_{g\left(\mathbf{Q}^M_{(\Xi,\Phi)}f(x)\right)},$$

which can be estimated by

$$\begin{aligned}
\left|\partial_2^{\vec{l}}G(x,y)\right|_{g\left(\mathbf{Q}^M_{(\Xi,\Phi)}f(x)\right)} &\le C \sup_{1\le r\le k}\left\|\nabla_2^r \log\left(\mathbf{Q}^M_{(\Xi,\Phi)}f(x), f(y)\right)\right\| \\
&\qquad \sum_{\substack{1\le r\le k,\ \vec{l}_j\in[d]^{m_j}\\ \sum_{j=1}^r m_j = k}} \left|\mathcal{D}^{\vec{l}_1}f(y)\right|_{g(f(y))}\cdot\dots\cdot\left|\mathcal{D}^{\vec{l}_r}f(y)\right|_{g(f(y))} \\
&\le C \sup_{1\le r\le k}\left\|\nabla_2^r \log\left(\mathbf{Q}^M_{(\Xi,\Phi)}f(x), f(y)\right)\right\| \Theta_{\infty,k,\Omega}(f).
\end{aligned} \tag{19}$$

Inserting Estimate (19) into the definition (14) of $R_{\vec{l}}$ we get that

$$\begin{aligned}
\left|R_{\vec{l}}(x,\xi_i)\right|_{g\left(\mathbf{Q}^M_{(\Xi,\Phi)}f(x)\right)} &\le C\Theta_{\infty,k,\Omega}(f)\int_0^1 (1-t)^{k-1}\sup_{1\le r\le k}\left\|\nabla_2^r \log\left(\mathbf{Q}^M_{(\Xi,\Phi)}f(x), f(x+t(\xi_i-x))\right)\right\| dt \\
&\le C\Theta_{\infty,k,\Omega}(f)\sup_{1\le r\le k}\sup_{y\in\Omega_x}\left\|\nabla_2^r \log\left(\mathbf{Q}^M_{(\Xi,\Phi)}f(x), f(y)\right)\right\|.
\end{aligned} \tag{20}$$

Finally, putting (20) into (18) we get that

$$\begin{aligned}
\left|(II)_{\vec{l}}\right|_{g\left(\mathbf{Q}^M_{(\Xi,\Phi)}f(x)\right)} &\le C\Theta_{\infty,k,\Omega}(f)\sup_{1\le r\le k}\sup_{y\in\Omega_x}\left\|\nabla_2^r \log\left(\mathbf{Q}^M_{(\Xi,\Phi)}f(x), f(y)\right)\right\| \\
&\qquad \sum_{i\in\mathcal{I}(x)} |\varphi_i(x)||\xi_i - x|^{|\vec{l}\,|} \\
&\le C\Theta_{\infty,k,\Omega}(f)\sup_{1\le r\le k}\sup_{y\in\Omega_x}\left\|\nabla_2^r \log\left(\mathbf{Q}^M_{(\Xi,\Phi)}f(x), f(y)\right)\right\| \\
&\qquad \sup_{i\in\mathcal{I}(x)} |\varphi_i(x)| h_{(\Xi,\Phi)}(x)^k.
\end{aligned} \tag{21}$$

Summing up over all $|\vec{l}\,| = k$ we get the desired estimate. □

Two remarks are in order regarding Theorem 3.5.

Remark 3.6. *Clearly, in the linear case higher order (i.e. higher than order 1) derivatives of the logarithm mapping* $\log(p,q) = q - p$ *vanish. Using this fact it is easy to see that our proof of Theorem 3.5 reduces to Theorem 2.4 in the linear case.*

Remark 3.7. *The estimate in Theorem 3.5 completely separates the error contributions of f and of the geometry of M. We thus see that the only geometric quantity which influences the approximation consists of iterated covariant derivatives of the logarithm mapping.*

Using Theorem 3.5 we can now state and prove a geometric generalization to Wendland's main theorem on moving least squares approximation, e.g., Theorem 2.10.

Theorem 3.8. *Assume that $\Omega \subset \mathbb{R}^d$ is a domain with smooth boundary, $X = (\Xi_j, \delta_j)_{j\in S}$ is quasi-uniform, Ξ_j is (m, δ_j)-unisolvent for all $j \in S$ and $\Phi_j = \Phi^{\Xi_j,m,\delta_j}$ are the basis functions as defined in Section 2. Then for all $1 \leq k < m+1$ and $f \in C^k(\overline{\Omega}, M)$ there exists a generic constant $C > 0$, independent of f, such that*

$$\sup_{x\in\Omega} d\left(f(x), \mathbf{Q}^M_{(\Xi_j,\Phi_j)} f(x)\right) \leq C \sup_{1\leq r\leq k} \sup_{y\in\Omega_x\cap\Omega} \left\|\nabla_2^r \log\left(\mathbf{Q}^M_{(\Xi,\Phi)} f(x), f(y)\right)\right\| \Theta_{\infty,k,\Omega}(f)\delta_j^k$$

for all $j \in S$.

Proof. The proof proceeds exactly as the proof of Theorem 2.10, using Theorem 3.5 instead of Theorem 2.4. □

Our approximation operator $\mathbf{Q}^M_{(\Xi,\Phi)}$ satisfies $(i), (ii)$ and (iii) from the introduction.

3.3 Generalization to Retraction Pairs

The computation of the quasiinterpolant $\mathbf{Q}^M_{(\Xi,\Phi)} f(x)$ requires the efficient computation of the exponential and logarithm mapping of M. For many practical examples of M this is not an issue, however in certain cases (for instance the Stiefel manifold [11]) it is computationally expensive to compute the exponential or logarithm function of a given manifold. Then, alternative functions can sometimes be used. This idea is formalized by the concept of *retraction pairs*.

Definition 3.9 ([11], see also [1, 8])**.** *A pair (P, Q) of smooth functions*

$$P : TM \to M, \quad Q : M \times M \to TM$$

is called a retraction pair *if*

$$P(x, Q(x, y)) = y, \qquad \text{for all } x,\ y \in M, \text{ and } P(x, 0) = x, \quad \frac{d}{dv} P(x, v)\Big|_{v=0} = \mathrm{Id} \qquad \text{for all } x \in M.$$

In general P may only be defined locally around M, and Q around the diagonal of $M \times M$.

Example 3.10. *Certainly, the pair* $(\exp, \log)$ *satisfies the above assumptions [6], and therefore forms a retraction pair. Let $S^m = \{x \in \mathbb{R}^{m+1} | |x| = 1\}$ be the m-dimensional sphere. Here we can define a retraction pair (P, Q) by*

$$P(x, y) = \frac{x + y}{|x + y|} \text{ and } Q(x, y) = \frac{y}{\langle x, y\rangle} - x,$$

where $\langle\cdot,\cdot\rangle$ is the standard inner product. We refer to [1] for further examples of retraction pairs for several manifolds of practical interest.

Given a retraction pair (P, Q), we can construct generalized quasiinterpolants $\mathbf{Q}^{(P,Q)}_{(\Xi,\Phi)} f(x)$ based on the first order condition (6), which defines a geometric average based on (P, Q) via

$$\sum_{i\in\mathcal{I}} \gamma_i Q\left(av_{P,Q}(\Gamma, \Pi), p_i\right) = 0 \tag{22}$$

The results in [11] show that this expression is locally well-defined. The construction above allows us to define a geometric quasiinterpolant based on an arbitrary retraction pair as follows.

Definition 3.11. *Given a retraction pair (P, Q) and denoting $\Phi(x) := (\varphi_i(x))_{i\in\mathcal{I}} \subset \mathbb{R}$ and $f(\Xi) := (f(\xi_i))_{i\in\mathcal{I}} \subset M$ we define the nonlinear moving least squares approximant*

$$\mathbf{Q}^{(P,Q)}_{(\Xi,\Phi)} f(x) := av_{P,Q}(\Phi(x), f(\Xi)) \in M. \tag{23}$$

The following generalization of Theorem 3.5 holds true.

Theorem 3.12. *Assume that (P, Q) is a retraction pair. Assume further that (Ξ, Φ) reproduces polynomials of degree m. Then for all $1 \leq k < m+1$ and $f \in C^k(\overline{\Omega}, M)$ there exists a generic constant $C > 0$, independent of f, Ξ, Φ, P, Q and the geometry of M such that*

$$\left| Q\left(\mathbf{Q}^{(P,Q)}_{(\Xi,\Phi)} f(x), f(x)\right)\right|_{g(\mathbf{Q}^{(P,Q)}_{(\Xi,\Phi)} f(x))} \leq$$
$$C\Theta_{\infty,k,\Omega}(f) \sup_{i\in\mathcal{I}(x)} |\varphi_i(x)| \sup_{1\leq r\leq k} \sup_{y\in\Omega_x\cap\Omega} \left\| \nabla_2^r Q\left(\mathbf{Q}^{(P,Q)}_{(\Xi,\Phi)} f(x), f(y)\right)\right\| h_{(\Xi,\Phi)}(x)^k \quad (24)$$

for all $x \in \Omega$.

Proof. The proof is completely analogous to the proof of Theorem 3.5 with log replaced by Q. □

4 Numerical Examples

We demonstrate our theoretical findings for three different manifolds, the sphere S^m, the manifold of symmetric positive definite $k \times k$ matrices $SPD(k)$ and the manifold of invertible $k \times k$ matrices $GL(k)$. The main task will be the computation of the Riemannian averages. Having that we can define our approximations as in Definition 3.2. Throughout this section we assume that $\mathcal{I}$ is a finite set.

4.1 Interpolation of Sphere-Valued Functions

To find the Riemannian average $p \in S^m$ of points $(p_i)_{i\in\mathcal{I}} \subset S^m$ and weights $(\lambda_i)_{i\in\mathcal{I}} \subset \mathbb{R}$ with $\sum_{i\in\mathcal{I}} \lambda_i = 1$ we use the intrinsic version of Newton's method introduced in [1] on Equation (5). The metric $d\colon S^m \times S^m \to \mathbb{R}_+$ induced by the euclidean metric on $\mathbb{R}^{m+1}$ is

$$d(u, v) = \arccos(\langle u, v\rangle) \quad (25)$$

for all $u, v \in S^m$. Hence we have to find the minimum of the functional $J\colon S^m \to \mathbb{R}$ defined by

$$J(x) := \sum_{i\in\mathcal{I}} \lambda_i \arccos^2(\langle x, p_i\rangle). \quad (26)$$

The intrinsic version of Newton's method generates a sequence $(x_j)_{j>0} \subset S^m$ that converges to $p \in S^m$. The first element $x_1 \in S^m$ is a first guess. In Step $j > 0$ a quadratic approximation J_{x_j} of J at $x_j \in S^m$ is defined. The next element x_{j+1} is defined as the minimizer of J_{x_j}. The quadratic approximation J_{x_j} has the form

$$J_{x_j}(x) = J(x_j) + J^1_{x_j}(r(x)) + 1/2 J^2_{x_j}(r(x), r(x)) \quad (27)$$

where $r(x) = \log(x_j, x)$, $J^1_{x_j}\colon TS^m_{x_j} \to \mathbb{R}$ a linear form and $J^2_{x_j}\colon TS^m_{x_j} \times TS^m_{x_j} \to \mathbb{R}$ a symmetric bilinear form. These functions are uniquely determined by the condition

$$J(x) = J_{x_j}(x) + \mathcal{O}(|r(x)|^3).$$

In [1] it was shown that if the first guess $x_1 \in S^m$ is close enough to $p \in S^m$ then the sequence $(x_j)_{j=1}^{\infty}$ converges quadratically to $p \in S^m$. The minimizer $x_{j+1} \in S^m$ of J_{x_j} is the unique solution of

$$J^1_{x_j}(y) + J^2_{x_j}(r(x_{j+1}), y) = 0 \qquad \forall\, y \in TS^m_{x_j}. \quad (28)$$

This solution $x_j \in S^m$ can be found be choosing a basis of $TS^m_{x_j}$ and then solve the corresponding linear system of equations. Hence if we have the forms $J^1_{x_j}$ and $J^2_{x_j}$ we can compute the minimizer $x_{j+1} \in S^m$. A simple way to find $J^1_{x_j}$ and $J^2_{x_j}$ is to first extend the function $J\colon S^m \to \mathbb{R}$ to a function $\bar{J}$ defined on a set $U \subset \mathbb{R}^{m+1}$ and then to compute the derivatives of $\bar{J}$ at x_j. As $\arccos(x) = -i\, ln(x + i\sqrt{1-x^2})$ for all $x \in (-1, 1]$ where $i \in \mathbb{C}$ is the imaginary unit and $ln\colon \mathbb{C}\backslash\{x|x \leq 0\} \to \{z = x + yi | x \in \mathbb{R},\ -2\pi < y \leq 0\}$

is the main branch of the logarithm function we can extend $\arccos^2$ defined on $[-1,1]$ to a function $\alpha : [-1,\infty) \to \mathbb{R}$ defined by

$$\alpha(x) := \begin{cases} -ln^2(x + i\sqrt{1-x^2}) & \text{if } x > -1 \\ \pi^2 & \text{if } x = -1 \end{cases}.$$

The function α is analytic for all $x > -1$. The derivatives of α in the interval $(-1,1)$ are

$$\alpha'(x) = \frac{-2\arccos(x)}{\sqrt{1-x^2}} \qquad \text{and} \qquad \alpha''(x) = \frac{2+\alpha'(x)x}{1-x^2}. \tag{29}$$

The expressions get numerically unstable around $x \approx 1$. There, the series expansions

$$\alpha'(x) = -2 + \frac{2}{3}(x-1) + \mathcal{O}((x-1)^2), \tag{30}$$

$$\alpha''(x) = \frac{2}{3} - \frac{8}{15}(x-1) + \mathcal{O}((x-1)^2), \tag{31}$$

have to be used instead. The function $J\colon S^m \to \mathbb{R}$ can now be extended to a function $\bar{J}\colon U \subset \mathbb{R}^{m+1} \to \mathbb{R}$ defined by

$$\bar{J}(x) = \sum_{i \in \mathcal{I}} \lambda_i \alpha(\langle x, p_i \rangle)$$

where

$$U := \{x \in \mathbb{R}^{m+1} | \langle x, p_i \rangle \geq -1 \text{ for all } i \in \mathcal{I}\}.$$

We have

$$S^m \cap \mathrm{Int}(U) = S^m \backslash \{-p_i | i \in \mathcal{I}\}$$

where $\mathrm{Int}(U)$ denotes the interior of U. Hence whenever $x_j \neq -p_i$ for all $i \in \mathcal{I}$ we can do a Taylor expansion of $\bar{J}$ up to order 2 at x_j and obtain

$$\bar{J}(x) = \bar{J}(x_j) + d^1\bar{J}_{x_j}(x - x_j) + 1/2 d^2\bar{J}_{x_j}(x - x_j, x - x_j) + \mathcal{O}(|x - x_j|^3). \tag{32}$$

where

$$d^1\bar{J}_{x_j}(y) = \sum_{i \in \mathcal{I}} \lambda_i \alpha'\left(\langle x_j, p_i \rangle\right) \langle p_i, y \rangle \text{ and} \tag{33}$$

$$d^2\bar{J}_{x_j}(y,z) = \sum_{i \in \mathcal{I}} \lambda_i \alpha''\left(\langle x_j, p_i \rangle\right) \langle p_i, y \rangle \langle p_i, z \rangle. \tag{34}$$

The Taylor expansion of the exponential map exp up to order 2 is

$$\exp(x_j, s) = \cos(|s|)x_j + \frac{\sin(|s|)}{|s|}s \tag{35}$$

$$= x_j + s - \frac{|s|^2}{2}x_j + \mathcal{O}(|s|^3). \tag{36}$$

Let $r = \log(x_j, x)$. Replacing x by $\exp(x_j, r)$ in Equation (32) and using Equation (36) yields

$$\bar{J}(x) = \bar{J}(x_j) + d^1\bar{J}_{x_j}(r) + 1/2(d^2\bar{J}_{x_j}(r,r) - d^1\bar{J}_{x_j}(x_j)|r|^2) + \mathcal{O}(|r|^3). \tag{37}$$

Comparison of Equation (37) and (27) yields

$$J^1_{x_j}(y) = d^1\bar{J}_{x_j}(y) \tag{38}$$

$$J^2_{x_j}(y,y) = d^2\bar{J}_{x_j}(y,y) - d^1\bar{J}_{x_j}(x_j)|y|^2 \tag{39}$$

$$\tag{40}$$

for all $y \in TS^m_{x_j}$. In summary we arrive at Algorithm 1.

Algorithm 1: Find Riemannian Average on the Sphere

Input : $(p_i)_{i\in\mathcal{I}} \subset S^m$, $(\lambda_i)_{i\in\mathcal{I}} \subset \mathbb{R}$ with $\sum_{i\in\mathcal{I}} \lambda_i = 1$, tolerance *tol* and initial guess $x_1 \in S^m$.
Output: Riemannian average $av_{S^m}((p_i)_{i\in\mathcal{I}}, (\lambda_i)_{i\in\mathcal{I}})$

$x_0 = 0$, $j = 1$.
while $|x_j - x_{i-1}| < tol$ **do**
  compute $d^1\bar{J}_{x_j}$ and $d^2\bar{J}_{x_j}$ as in Equation (33) and (34).
  compute $J^1_{x_j}$ and $J^2_{x_j}$ as in Equation (38) and (39).
  choose a basis $y_1, \dots, y_m$ of $TS^m_{x_j}$.
  solve system of Equation $J^1_{x_j}(y_j) + J^2_{x_j}(r, y_j) = 0 \; \forall j = 1, \dots, m$ for r.
  compute $x_{j+1} = \exp(x_j, r)$.
  set $j = j + 1$.
end
return x_j

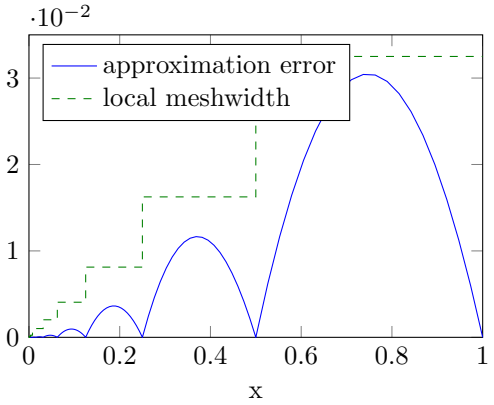

Figure 1: Approximation error for a sphere-valued function

Example 4.1. *We consider the function* $f\colon [0,1] \to S^m$ *defined by*

$$[0,1] \ni x \mapsto \frac{(1, x, x^2)}{|(1, x, x^2)|}$$

and the nodes $\{0\} \cup \{2^{-j} | j \in \{0, 1, \dots, 10\}\}$. *For interpolation we use the hat functions defined in Example 2.3. Figure 1 shows the error and the local meshwidth (multiplied by a constant). It illustrates that the error can be bounded by a constant times the local meshwidth as stated in Theorem 3.8.*

4.2 Interpolation of Functions With Symmetric Positive Definite Matrices As Values

For $k \in \mathbb{N}$ we denote the space of positive definite $k \times k$ matrices by $SPD(k)$. The following information on the geometry of $SPD(k)$ can be found in [17]. Let $S(k)$ be the space of symmetric $k \times k$ matrices. The tangent space of a point $P \in SPD(k)$ is equivalent to $S(k)$. The natural inner product on this space is

$$\langle A, B\rangle_P = tr(AP^{-1}BP^{-1})$$

where tr denotes the trace. The induced Riemannian metric is

$$d_{SPD}(X, Y) = \| \ln(X^{-1/2} Y X^{-1/2})\|_{tr}. \tag{41}$$

Here ln denotes the matrix logarithm and $\| \cdot \|_{tr}$ the trace norm, i.e. $\|X\|_{tr} = \sqrt{tr(XX^T)}$ where X^T denotes the transpose of X. The exponential map exp and logarithm map log on the manifold

of symmetric positive definite matrices are

$$\exp(P,Q) = P^{1/2}E(P^{-1/2}QP^{-1/2})P^{1/2}, \tag{42}$$
$$\log(P,Q) = P^{1/2}\ln(P^{-1/2}QP^{-1/2})P^{1/2}, \tag{43}$$

where $E(\cdot)$ denotes the matrix exponential.
To find the Riemannian average X of matrices $(X_i)_{i\in\mathcal{I}} \subset \mathbb{R}^{k\times k}$ and weights $(\lambda_i)_{i\in\mathcal{I}} \subset \mathbb{R}$ with $\sum_{i\in\mathcal{I}} \lambda_i = 1$ we solve Equation (6) which reduces in the case of positive definite matrices to

$$\sum_{i\in\mathcal{I}} \lambda_i \ln(X^{-1/2}X_iX^{-1/2}) = 0. \tag{44}$$

To solve this equation we use the standard Newton method. To do this we need to compute derivatives of the square root function, the matrix logarithm ln, the inverse function, the trace norm and compositions of these. To compute the derivative of the square root function observe that $Y^{k,l} := d\sqrt{X}/dx_{kl} \in \mathbb{R}^{k\times k}$ is the solution of the Lyapunov equation

$$Y^{k,l}\sqrt{X} + \sqrt{X}Y^{k,l} = E_{kl}, \tag{45}$$

where $E_{kl} \in \mathbb{R}^{n,n}$ is the matrix defined by $(E_{kl})_{mn} := \delta_{km}\delta_{ln}$. To compute the logarithm and its derivatives we use the identity

$$\ln(X) = 2^n \ln(X^{1/2^n}) \tag{46}$$

for all $X \in SPD(k)$ and the power series

$$\ln(I+X) = \sum_{i=1}^{\infty} (-1)^{i+1}\frac{X^i}{i} \tag{47}$$

for all $X \in \mathbb{R}^{k\times k}$ with the norm of its eigenvalues smaller than 1. With compositions of these two identities we are able to compute the logarithm of any SPD matrix up to a given tolerance. To compute the derivative of the matrix logarithm we need to be able to compute the derivative of the square root, which we already discussed (see Equation (45)), and the derivatives of powers of matrices which is

$$\frac{d(X^n)_{jk}}{dX_{lm}} = \sum_{i=0}^{n-1} (X^i)_{jl}(X^{n-i-1})_{mk}$$

for all $X \in \mathbb{R}^{N\times N}$ and $j,k,l,m \in \{1,\dots,N\}$. The derivative of the inverse function is

$$\frac{d(X^{-1})_{ij}}{dX_{kl}} = -(X^{-1})_{ik}(X^{-1})_{lj}$$

for all $X \in GL(N)$ and $i,j,k,l \in \{1,\dots,N\}$.

Algorithm 2: Find Riemannian Average on the Manifold of SPD-matrices

Input : $(X_i)_{i\in\mathcal{I}} \subset \mathbb{R}^{k\times k}$, $(\lambda_i)_{i\in\mathcal{I}} \subset \mathbb{R}$ with $\sum_{i\in\mathcal{I}} \lambda_i = 1$ and initial guess $Y_0 \in S^m$.
Output: Riemannian average $av_{SPD(k)}((X_i)_{i\in\mathcal{I}}, (\lambda_i)_{i\in\mathcal{I}})$

1 Use Newton's method to solve Equation (44) with initial value Y_0.

Example 4.2. *Let $\Omega = [0,1]$, $0 = \xi_1 < \xi_2 \dots < \xi_6 = 1$ and $f(x) = \cos(x\pi/2)A0 + \sin(x\pi/2)A1$, where $A0$ and $A1$ are randomly chosen SPD-matrices. For all $i \in \{1,2,\dots,6\}$ let $\varphi_i(x)$ be the hat functions from Example 2.3. In Figure 2 the error measured by the metric defined in Equation (41) and the local meshwidth (multiplied by a constant) is shown. It illustrates that the error can be bounded by a constant times the local meshwidth as stated in Theorem 3.8.*

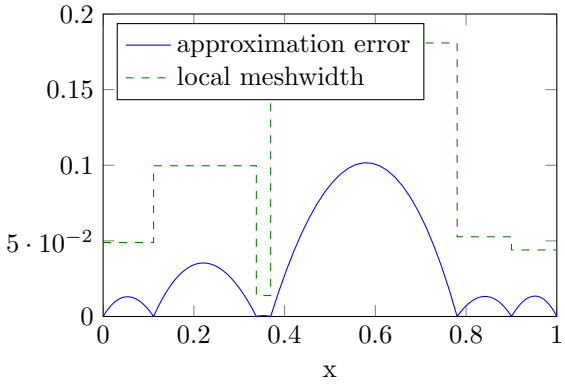

Figure 2: Approximation error for a SPD-valued function

4.3 Approximation of Reduced Order Models (ROMs)

We start by introducing linear time invariant systems. Then we present the ROM approximation method as well as an adaption of this method, based upon the theory in this paper, for approximating reduced order models (ROMs).

4.3.1 Linear time-invariant Systems

In a parameter dependent LTI (linear time-invariant) system as in [5] we assume that for each $x \in \Omega$ there exists a unique solution $z^x \colon [0,T] \to \mathbb{R}^q$ of

$$\frac{dw^x}{dt}(t) = A(x)w^x(t) + B(x)u(t), \tag{48}$$
$$z^x(t) = C(x)w^x(t) + D(x)u(t), \tag{49}$$

where $A \colon \Omega \to GL(n)$, with $GL(n)$ denoting the set of invertible matrices of size $n \times n$, $B \colon \Omega \to \mathbb{R}^{n\times p}$, $C \colon \Omega \to \mathbb{R}^{q\times n}$, $D \colon \Omega \to \mathbb{R}^{q\times p}$ and $u \colon [0,T] \to \mathbb{R}^p$. Typically z^x is an output functional of a dynamical system with control function u and $n \gg p, q$. Furthermore we assume that A, B, C and D are continuous. Define

$$M^{n,k} := \{U \in \mathbb{R}^{n\times k} | U^T U = I_k\},$$

where I_k is the $k \times k$ identity matrix. Let $U, V \colon \Omega \to M^{n,k}$ define test and trial bases, respectively. The state vector $w^x(t)$ will be approximated as a linear combination of column vectors of $V(x)$, i.e. $w^x(t) \approx V(x)w^{x,U,V}(t)$ where $w^{x,U,V}$ will be defined by substituting w^x by $V(x)w^{x,U,V}$ and multiplying Equation (48) from the left by $U(x)^T$. Hence we get the system of equations

$$U^T V(x)\frac{dw^{x,U,V}}{dt}(t) = U^T AV(x)w^{x,U,V}(t) + U^T B(x)u(t), \tag{50}$$
$$z^{x,U,V}(t) = CV(x)w^{x,U,V}(t) + D(x)u(t), \tag{51}$$

where all operations on matrix valued functions are defined in the natural way. Multiplying Equation (50) from the left by $(U^T V(x))^{-1}$ yields the new LTI system

$$\frac{dw^{x,U,V}}{dt}(t) = A^{U,V}(x)w^{x,U,V}(t) + B^{U,V}(x)u(t), \tag{52}$$
$$z^{x,U,V}(t) = C^V(x)w^{x,U,V}(t) + D(x)u(t), \tag{53}$$

where for all $x \in \Omega$

$$A^{U,V}(x) := (U^T V)^{-1}U^T AV(x) \in \mathbb{R}^{k\times k}, \tag{54}$$
$$B^{U,V}(x) := (U^T V)^{-1}U^T B(x) \in \mathbb{R}^{k\times p} \text{ and} \tag{55}$$
$$C^V(x) := CV(x) \in \mathbb{R}^{q\times k}. \tag{56}$$

The aim is that $z^{x,U,V}$ has the same properties and features as z. Furthermore k should be much smaller than n so that $z^{x,U,V}(t)$ can be computed (with the help of the matrices $A^{U,V}$, $B^{U,V}$, C^V and D) significantly faster than $z^x(t)$.
Let $O(k) := \{Q \in \mathbb{R}^{k\times k} | QQ^T = I_k\}$ be the space of orthogonal matrices. A coordinate transformation $\tilde{V}(x) = VQ(x)$, $\tilde{U}(x) = UQ(x)$ with an orthogonal matrix $Q(x) \in O(k)$ does transform the solution $z^{x,U,V}$ isometrically to $z^{x,\tilde{U},\tilde{V}} = Q^T(x)z^{x,U,V}$. We introduce the equivalence relation

$$(U, V) \sim (\tilde{U}, \tilde{V}) \qquad :\Leftrightarrow \qquad \exists Q \in O(k) \text{ s.t. } \tilde{U} = UQ \text{ and } \tilde{V} = VQ$$

on $(M^{n,k})^2$. Given a parameter dependent LTI-system and assume that for each choice of the parameter $x \in \Omega$ there exists a ROM with matrices $U(x), V(x), A^{U,V}(x), B^{U,V}(x), C^V(x)$. Furthermore assume that the map

$$\begin{aligned} f\colon \Omega &\to (M^{n,k})^2/\sim & (57)\\ x &\mapsto \Pi(U(x), V(x)), & (58)\end{aligned}$$

where $\Pi\colon (M^{n,k})^2 \to (M^{n,k})^2/\sim$ denotes the natural projection, is continuous. Note that U, V, $A^{U,V}$, $B^{U,V}$ and C^V do not have to be continuous. The problem of interpolating reduced-order models (ROMs) is: Given a reduction $(A^{U,V}(\xi_i), B^{U,V}(\xi_i), C^V(\xi_i))$ for several parameters ξ_i, $i \in \mathcal{I}$ find an approximation for a new parameter $x \in \Omega$ for a reduction $(A^{\tilde{U},\tilde{V}}(x), B^{\tilde{U},\tilde{V}}(x), C^{\tilde{V}}(x))$ with $(\tilde{U}(x), \tilde{V}(x)) \sim (U(x), V(x))$. As we are aiming for a fast method the running time should be independent of n. In addition to the matrices $(A^{U,V}(\xi_i), B^{U,V}(\xi_i), C^V(\xi_i))$ we can also use precomputed matrices of size $\ll n$. A key role for the algorithm in the next section will play the matrix-valued function P_{V,V_0} defined by

$$P_{V,V_0}(x) := V(x)^T V_0 \in \mathbb{R}^{k\times k}$$

for all $x \in \Omega$ where $V_0 \in \mathbb{R}^{n\times k}$. In [5] an $l_0 \in \mathcal{I}$ was chosen and V_0 was defined by $V_0 := V(\xi_{l_0})$. We will assume that $P_{V,V_0}(x) \in GL(k)$ for all $x \in \Omega$.

4.3.2 The ROM Approximation Method

We sketch the method proposed by Amsallem and Farhat in [5]. The algorithm is divided into two steps. In the first step we construct a continuous function $\tilde{f}^{U,V}\colon \Omega \to (M^{n,k})^2$ with $\tilde{f}^{U,V}(x) \sim (U(x), V(x))$ for all $x \in \Omega$, i.e. we choose a representant $\tilde{f}^{U,V}(x)$ of $f(x)$ for each $x \in \Omega$ where $\tilde{f}^{U,V}$ is continuous. We define

$$\begin{aligned} \tilde{f}^{U,V}_{V_0}\colon \Omega &\to (M^{n,k})^2 & (59)\\ x &\mapsto (\tilde{U}(x), \tilde{V}(x)) := (U(x)Q_{P_{V,V_0}}(x), V(x)Q_{P_{V,V_0}}(x)) & (60)\end{aligned}$$

where $Q_{P_{V,V_0}}(x) := \Pi_{O(k)}(P_{V,V_0}(x))$ with $\Pi_{O(k)}$ being the shortest point projection onto $O(k)$ defined below.

Definition 4.3. *We define the shortest point projection $\Pi_{O(k)}\colon GL(k) \to O(k)$ by*

$$\Pi_{O(k)}(X) := \underset{Y\in O(k)}{argmax}\ tr(X^TY).$$

To compute the shortest point projection we use the following well-known result based on the singular value decomposition (SVD).

Lemma 4.4. *Let (U, Σ, V) be the SVD of a matrix $X \in GL(k)$, i.e. $X = U\Sigma V^T$ with $U, V \in O(k)$ and Σ a diagonal matrix with positive real numbers on the diagonal. Then we have $\Pi_{O(k)}(X) = UV^T$.*

Proof. We have for $Y \in O(k)$

$$\operatorname{tr}(X^TY) = \operatorname{tr}(V\Sigma U^TY) = \operatorname{tr}(\Sigma U^TYV)$$

Let $Z = U^TYV$. As $U, Y, V \in O(k)$ we have $Z \in O(k)$. In particular the entries z_{ij} of Z satisfy $z_{ij} \leq 1$ Now we have

$$\operatorname{tr}(\Sigma Z) = \sum_{i=1}^{k} \sigma_i z_{ii} \leq \sum_{i=1}^{k} \sigma_i$$

with equality if and only if Z is the identity matrix I. Hence

$$U^T P_{O(k)}(X)V = I \Rightarrow \Pi_{O(k)}(X) = UV^T. \qquad \square$$

Lemma 4.5. *Let $R\colon \Omega \to O(k)$ then*

$$\tilde{f}_{V_0}^{UR,VR} = \tilde{f}_{V_0}^{U,V}$$

for all $x \in \Omega$.

Proof. We have

$$P_{VR,V_0} = (VR)^T V_0 = R^T P_{V,V_0},$$
$$Q_{P_{VR,V_0}} = \Pi_{O(k)}(P_{VR,V_0}) = \Pi_{O(k)}(R^T P_{V,V_0}) = R^T \Pi_{O(k)}(P_{V,V_0}) = R^T Q_{P_{V,V_0}},$$

and therefore

$$\begin{aligned}\tilde{f}_{V_0}^{UR,VR} &= (URQ_{P_{VR,V_0}}, VRQ_{P_{VR,V_0}}) = (URR^T Q_{P_{V,V_0}}, VRR^T Q_{P_{V,V_0}}) \\ &= (UQ_{P_{V,V_0}}, VQ_{P_{V,V_0}}) = \tilde{f}_{V_0}^{U,V}.\end{aligned}$$

$\square$

As $f\colon \Omega \to (M^{n,k})^2/\sim$ is continuous there exists a continuous function

$$\begin{aligned}\bar{f}\colon \Omega &\to (M^{n,k})^2 &(61)\\ x &\mapsto (\bar{U}(x), \bar{V}(x)) &(62)\end{aligned}$$

with $\Pi(\bar{f}) = f$. By Lemma 4.5 we have $\tilde{f}_{V_0}^{U,V} = \tilde{f}_{V_0}^{\bar{U},\bar{V}}$. Furthermore it is easy to check that $\tilde{f}_{V_0}^{\bar{U},\bar{V}}$ is continuous. The matrices of the reduced order model are

$$\begin{array}{rcccl} A^{\tilde{U},\tilde{V}} & = & (\tilde{U}^T\tilde{V})^{-1}\tilde{U}^T A\tilde{V} & = & Q_{P_{V,V_0}}^T A^{U,V} Q_{P_{V,V_0}}, \\ B^{\tilde{U},\tilde{V}} & = & (\tilde{U}^T\tilde{V})^{-1}\tilde{U}^T B & = & Q_{P_{V,V_0}}^T B^{U,V} \text{ and} \\ C^{\tilde{V}} & = & C\tilde{V} & = & C^V Q_{P_{V,V_0}}. \end{array}$$

The first step of the algorithm requires knowledge of the quantities

$$P_{V,V_0}(\xi_i) = V(\xi_i)^T V_0 \in \mathbb{R}^{k\times k}$$

for all $i \in \mathcal{I}$. The values

$$(\tilde{A}(\xi_i), \tilde{B}(\xi_i), \tilde{C}(\xi_i)) := (A^{\tilde{U},\tilde{V}}(\xi_i), B^{\tilde{U},\tilde{V}}(\xi_i), C^{\tilde{V}}(\xi_i))$$

are computed for all $i \in \mathcal{I}$.

Algorithm 3: Step 1 of ROM Approximation Algorithm

Input : $(\xi_i)_{i\in\mathcal{I}} \subset \Omega$, $(A_{\xi_i})_{i\in\mathcal{I}} \subset GL(k)$, $(B_{\xi_i})_{i\in\mathcal{I}} \subset \mathbb{R}^{k\times p}$, $(C_{\xi_i})_{i\in\mathcal{I}} \subset \mathbb{R}^{q\times k}$, and $P_{V,V_0}(\xi_i) \in \mathbb{R}^{k\times k}$ for all $i \in \mathcal{I}$

Output: $(\tilde{A}(\xi_i), \tilde{B}(\xi_i), \tilde{C}(\xi_i))$ for all $i \in \mathcal{I}$

1 Choose $V_0 \in M^{n,k}$.

2 Compute $(\tilde{A}(\xi_i), \tilde{B}(\xi_i), \tilde{C}(\xi_i))$ for all $i \in \mathcal{I}$.

In step 2 the data $\tilde{B}(x)$ and $\tilde{C}(x)$ is approximated with a method for linear spaces. Hence we can deal with each entry independently. The data $\tilde{A}(x)$ is approximated with respect to the space $GL(k)$. The exponential and logarithm map on $GL(k)$ are given by

$$\exp(X, Y) = E(Y)X \text{ and } \log(X, Y) = \ln(YX^{-1}),$$

Note that the matrix logarithm ln can be defined for all matrices with nonnegative eigenvalues and Equations (46) and (47) are still valid. An $i_0 \in \mathcal{I}$ is chosen and the data $\tilde{A}(\xi_i)$ is mapped by the logarithm map log with base point $\tilde{A}(\xi_{i_0})$ to the tangent space of this manifold at $\tilde{A}(\xi_{i_0})$. Then the new data is approximated with a method for linear spaces and finally mapped back to the manifold by the exponential map exp.

Algorithm 4: Step 2 of ROM Approximation Algorithm

Input : $(\tilde{A}(\xi_i), \tilde{B}(\xi_i), \tilde{C}(\xi_i))$ for all $i \in \mathcal{I}$ and $x \in \Omega$
Output: Approximation (A_x, B_x, C_x) for $(\tilde{A}(x), \tilde{B}(x), \tilde{C}(x))$

1 Interpolate each entry of the matrices $\tilde{B}(\xi_i), \tilde{C}(\xi_i)$, $i \in \mathcal{I}(x)$ independently to obtain B_x and C_x.
2 Choose $i_0 \in \mathcal{I}(x)$.
3 Compute $\bar{A}(\xi_i) = \log(\tilde{A}(\xi_{i_0}), \tilde{A}(\xi_i))$ for all $i \in \mathcal{I}(x)$.
4 Interpolate each entry of the matrices $\bar{A}(\xi_i)$, $i \in \mathcal{I}(x)$ independently to obtain $\bar{A}_x$.
5 Compute $A_x = \exp(\tilde{A}(\xi_{i_0}), \bar{A}_x)$.

4.3.3 Variation of the ROM Approximation Algorithm

We slightly change Step 2 of the algorithm described in Section 4.3.2. The data $\tilde{A}(x)$ is directly approximated on $GL(k)$. For this purpose we solve the Equation (6) which reduces in our case to

$$\sum_{i \in \mathcal{I}(x)} \varphi_i(x) \ln(\tilde{A}(\xi_i) X^{-1}) = 0 \tag{63}$$

for $X \in \mathbb{R}^{k \times k}$ by the standard Newton method. As starting value we choose $X_0 := \sum_{i \in \mathcal{I}(x)} \varphi_i(x) \ln(\tilde{A}(\xi_i))$.

Algorithm 5: Adaption of Step 2 of ROM Approximation Algorithm

Input : $(\tilde{A}(\xi_i), \tilde{B}(\xi_i), \tilde{C}(\xi_i))$ for all $i \in \mathcal{I}$ and $x \in \Omega$
Output: Approximation (A_x, B_x, C_x) for $(\tilde{A}(x), \tilde{B}(x), \tilde{C}(x))$

1 Interpolate each entry of the matrices $\tilde{B}(\xi_i), \tilde{C}(\xi_i)$, $i \in \mathcal{I}(x)$ independently to obtain B_x and C_x.
2 Interpolate the matrices $\tilde{A}(\xi_i)$, $i \in \mathcal{I}(x)$ on the space of non-singular matrices with Equation (63).

4.3.4 Numerical Experiments

In Section 5.1 of [5] a simple academic example where the ROM approximation method yields good results is shown. In the next example we have to use the variation of the algorithm to get reasonable approximations.

Example 4.6. *In this example we consider an interpolation of LTI-systems without a reduction. Hence we can set* $(\tilde{A}, \tilde{B}, \tilde{C}) = (A, B, C)$ *and omit step 1 of the ROM Approximation Algorithm. Let*

$$A(x) := \begin{pmatrix} \cos(h(x)) & \sin(h(x) \\ -\sin(h(x) & \cos(h(x)) \end{pmatrix},$$

where $h(x) = 4\sin(\pi x)$ *and* $\xi_i = \frac{i}{4}$ *for* $i \in \{-2, -1, 0, 1, 2\}$. *We choose* $i_0 = 0$ *and the hat functions* $\varphi_i(x)$ *from Example 2.3 as basis functions. The error for* A *in the trace norm for the ROM-Approximation and its adapted algorithm are illustrated in Figure 3. An illustration why the ROM approximation method has a large error is shown in Figure 4. As the matrices* $A(x)$ *are two dimensional rotation matrices we can see them as points on a circle. The Riemannian average of points* A_1 *and* A_2 *with weights* $\lambda_1 = \lambda_2 = 0.5$ *is* M. *However if we transform the points by the logarithm to the tangent space at* A_0, *interpolate on this tangent space and transform back by the exponential map we get a different point* $\tilde{M} \neq M$.

Example 4.7. *We consider the values* $n = 3$, $p = q = 1$ *and the matrices* $A(x) := A0 + xA1$, $B(x) := B$ *and* $C(x) := C$ *for all* $x \in [-1, 1]$ *where*

$$A_0 := \begin{pmatrix} 6 & 4 & 2 \\ 8 & 4 & 2 \\ 12 & 4 & 20 \end{pmatrix},\ A_1 := \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 3 & 2 & 1 \end{pmatrix},\ B := \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \text{ and } C := \begin{pmatrix} 1 & 2 & 3 \end{pmatrix}.$$

We set $U(x)$ *and* $V(x)$ *equal to the first* 2 *columns of the orthogonal matrices of the SVD of* $A(x)$. *We choose the data sites from Example 2.8 with* $m = 2$. *The* L^2*-norm of the error made in Step 2 of Algorithm 5 is shown in Figure 5. As we can see the convergence rate is as predicted by Theorem 3.8.*

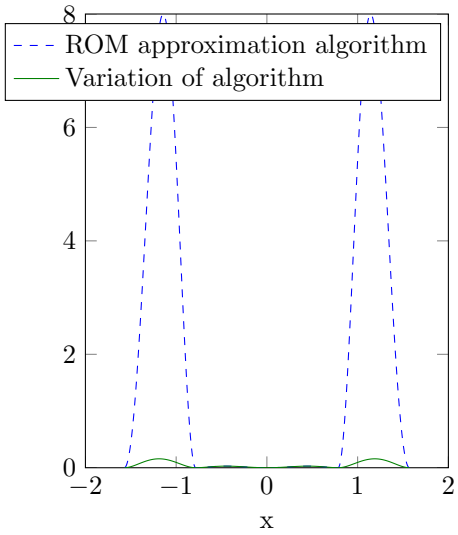

Figure 3: Error plot of ROM approximation algorithm and its adapted version

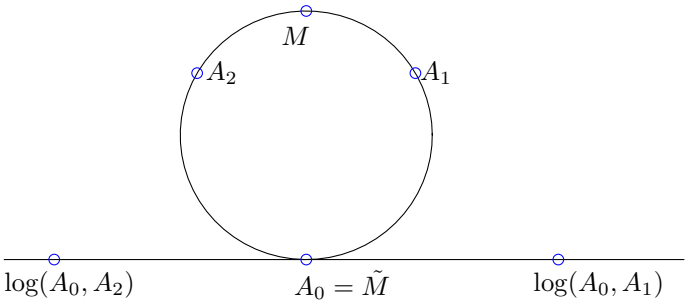

Figure 4: Illustration for the explanation of the large error in Figure 3

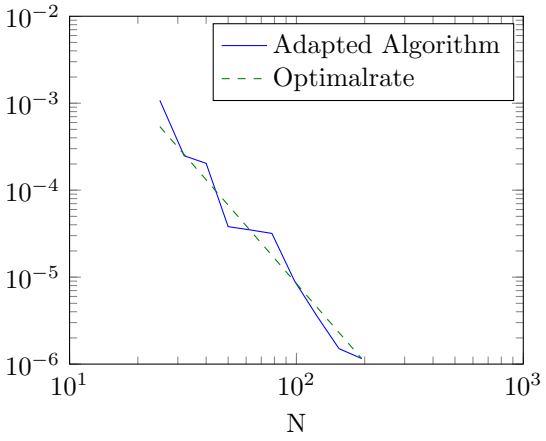

Figure 5: Convergence plot for error made in Algorithm 5

References

[1] P.A. Absil, R. Mahony, and R. Sepulchre. *Optimization algorithms on matrix manifolds*. Princeton University Press, 2009.

[2] D. Amsallem, J. Cortial, K. Carlberg, and C. Farhat. A method for interpolating on manifolds structural dynamics reduced-order models. *International Journal for Numerical Methods in Engineering*, 80(9):1241–1258, 2009.

[3] D. Amsallem, J. Cortial, and C. Farhat. Towards real-time computational-fluid-dynamics-based aeroelastic computations using a database of reduced-order information. *American Institute of Aeronautics and Astronautics Journal*, 48(9):2029–2037, 2010. doi: 10.2514/1.J050233.

[4] D. Amsallem and C. Farhat. Interpolation method for adapting reduced-order models and application to aeroelasticity. *American Institute of Aeronautics and Astronautics Journal*, 46(7):1803–1813, 2008. doi: 10.2514/1.35374.

[5] D. Amsallem and C. Farhat. An online method for interpolating linear parametric reduced-order models. *SIAM Journal on Scientific Computing*, 33(5):2169–2198, 2011.

[6] M. P. Do Carmo. *Riemannian geometry*. Birkhäuser Boston, 1992.

[7] T. Duchamp, G. Xie, and T. P.-Y. Yu. Single basepoint subdivision scheme for manifold-valued data: Time-symmetry without space-symmetry. To appear in Foundations of Computational Mathematics, available at `http://www.math.drexel.edu/~tyu`, 2011.

[8] P. Grohs. Smoothness equivalence properties of univariate subdivision schemes and their projection analogues. *Numerische Mathematik*, 113(2):163–180, 2009.

[9] P. Grohs. Approximation order from stability for nonlinear subdivision schemes. *Journal of Approximation Theory*, 162(5):1085–1094, 2010.

[10] P. Grohs. Quasi-interpolation in Riemannian manifolds. *IMA Journal of Numerical Analysis*, 2012.

[11] P. Grohs. Geometric multiscale decompositions of dynamic low-rank matrices. *Computer Aided Geometric Design*, 30(8):805–826, November 2013.

[12] P. Grohs, H. Hardering, and O. Sander. Optimal a priori discretization error bounds for geodesic finite elements. Technical Report 2013-16, Seminar for Applied Mathematics, ETH Zürich, Switzerland, 2013.

[13] P. Grohs and M. Sprecher. Projection-based quasiinterpolation in manifolds. Technical Report 2013-23, Seminar for Applied Mathematics, ETH Zürich, Switzerland, 2013.

[14] L. Hörmander. *The Analysis of Linear Partial Differential Operators: Vol.: 1.: Distribution Theory and Fourier Analysis*. Springer, 1983.

[15] H. Karcher. Mollifier smoothing and Riemannian center of mass. *Communications on Pure and Applied Mathematics*, 30:509–541, 1977.

[16] A. Marthinsen. Interpolation in Lie Groups. *SIAM Journal on Numerical Analysis*, 37(1):269–285, November 1999.

[17] M. Moakher and M. Zéraï. The Riemannian Geometry of the Space of Positive-Definite Matrices and Its Application to the Regularization of Positive-Definite Matrix-Valued Data. *Journal of Mathematical Imaging and Vision*, 40(2):171–187, June 2011.

[18] E. Nava-Yazdani and T. P.-Y. Yu. On Donoho's Log-Exp Subdivision Scheme: Choice of Retraction and Time-Symmetry. *Multiscale Modeling & Simulation*, 9(4):1801–1828, 2011.

[19] O. Sander. Geodesic finite elements for Cosserat rods. *International Journal for Numerical Methods in Engineering*, 82(13):1645–1670, 2010.

[20] O. Sander. Geodesic finite elements on simplicial grids. *International Journal for Numerical Methods in Engineering*, 92(12):999–1025, 2012.

[21] O. Sander. Geodesic finite elements of higher order. 2013. Institut für Geometrie und praktische Mathematik, Preprint 356, RWTH Aachen.

[22] H. Wendland. *Scattered Data Approximation*. Cambridge University Press, 2004.

[23] G. Xie and T. P.-Y. Yu. Approximation order equivalence properties of manifold-valued data subdivision schemes. *IMA Journal of Numerical Analysis*, 32(2):687–700, 2012.

Recent Research Reports

Nr.	Authors/Title

2013-48 R. Hiptmair and C. Jerez-Hanckes and C. Urzua
Optimal Operator Preconditioning for Boundary Elements on Open Curves

2013-49 P. Grohs and S. Hosseini
ε-Subgradient Algorithms for Locally Lipschitz Functions on Riemannian Manifolds

2013-50 A. Andersson and R. Kruse and S. Larsson
Duality in refined Watanabe-Sobolev spaces and weak approximations of SPDE

2014-01 M. Eigel and C.J. Gittelson and Ch. Schwab and E. Zander
A convergent adaptive stochastic Galerkin finite element method with quasi-optimal spatial meshes

2014-02 R. Kaeppeli and S. Mishra
Structure preserving schemes

2014-03 K. Grella
Sparse tensor phase space Galerkin approximation for radiative transport

2014-04 A. Hiltebrand and S. Mishra
Efficient preconditioners for a shock capturing space-time discontinuous Galerkin method for systems of conservation laws

2014-05 X. Claeys and R. Hiptmair and C. Jerez-Hanckes and S. Pintarelli
Novel Multi-Trace Boundary Integral Equations for Transmission Boundary Value Problems

2014-06 X. Claeys and R. Hiptmair
Integral Equations for Acoustic Scattering by Partially Impenetrable Composite Objects

2014-07 P. Grohs and S. Keiper and G. Kutyniok and M. Schaefer
Cartoon Approximation with α -Curvelets