Eidgenössische Technische Hochschule Zürich
Ecole polytechnique fédérale de Zurich
Politecnico federale di Zurigo
Swiss Federal Institute of Technology Zurich

Finite elements of arbitrary order and quasiinterpolation for data in Riemannian manifolds

P. Grohs

Research Report No. 2011-56
September 2011

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

Finite Elements of Arbitrary Order and Quasiinterpolation for Data in Riemannian Manifolds *

P. Grohs

September 5, 2011

Abstract

We consider quasiinterpolation operators for functions assuming their values in a Riemannian manifold. We construct such operators from corresponding linear quasiinterpolation operators by replacing affine averages with the Riemannian center of mass. As a main result we show that the approximation rate of such a nonlinear operator is the same as for the linear operator it has been derived from. In order to formulate this result in an intrinsic way we use the Sasaki metric to compare the derivatives of the function to be approximated with the derivatives of the nonlinear approximant. Numerical experiments confirm our theoretical findings.

AMS Subject Classification: primary 41A25, 65DXX secondary 53BXX

1 Introduction

A fundamental problem in computational science is the handling of massive amounts of data. In addition to the sheer mass of data to be processed, in recent times many modern sensing mechanisms produce data which is of nonstandard type, with data points assuming their value in nonlinear geometries. Examples include

- Deformation Tensors, where the data points consist of elements of the Cartan-Hadamard space of positive definite symmetric matrices. The data is modeled as a function $\mathbb{R}^3 \to \mathrm{SPD}(3)$. Such data arises for instance in Diffusion Tensor MRI in medical imaging [6] or strain and stress measurement in materials science.
- Positions of rigid bodies, where the data points consist of elements of the Lie group of rigid body motions. The data is modeled as a function $\mathbb{R} \to \mathrm{SE}(3)$. Data of this type arises for instance in kinematics or motion design [30], or Cosserat rod modeling [24].
- Orientations, where the data points consist of elements of the Lie group of orthogonal matrices. The data is modeled as a function $\mathbb{R} \to \mathrm{SO}(3)$. Orientation-valued data arrays arise for instance as 'black box' recordings of the orientation of aircrafts, varying with time [23].
- Subspaces, where the data points consist of elements of a Grassmanian manifold. The data is modeled as a function $\mathbb{R} \to \mathrm{G}(k,n)$. These data types can arise for instance in array signal processing [23].
- Orthogonal matrices with positive determinants also occur in isospectral-flow problems [17].
- Functions with values in direct products of the Lie groups $\mathrm{O}(3)$, $\mathrm{SO}(3)$ have been used in the modeling of human motion data [18].

*This research was supported by the European Research Council under grant ERC AdG 247277. The results of this paper have been presented at the Geometry Workshop in Obergurgl on June 22 2011.

This (incomplete) list suggests that it is of eminent interest to develop useful computational and theoretical tools capable of processing manifold-valued functions.

In linear problems, data is usually approximated by finite element spaces [1], both when the function to be approximated is given explicitly as well as implicitly as a solution of an optimization problem. Among the key properties which determine the usefulness of such a construction is a high rate of approximation of the finite element spaces.

It is the aim of the present paper to define nonlinear finite element manifolds which naturally generalize the linear theory and to prove that the associated approximation properties are exactly of the same order as for corresponding linear constructions. We do this by constructing explicitly a simple (nonlinear) projection operator onto these finite element manifolds which is shown to behave optimally.

More specifically, the objective of the present paper is to present a complete extension of the theory of quasiinterpolation, [3, 5], to the nonlinear case. We will start with a linear quasiinterpolation scheme and, by replacing affine averages with the Riemannian center of mass [20], wind up with an intrinsic approximation procedure for manifold-valued functions. Our main result is that the linear properties completely carry over to the nonlinear case.

Even stating such a result is nontrivial since it is at first glance not clear how to compare between a differential of a function with the differential of an approximant which both assume their values in the (iterated) tangent bundle of the manifold. We solve this problem by utilizing the so-called Sasaki metric [26] which is a canonical Riemannian structure defined on the tangent bundle of a Riemannian manifold. Using this formulation we are able to give a complete and intrinsic extension of the linear theory, see Theorem 3.8 below.

To give a flavor of our main result consider a linear quasiinterpolation operator

$$f \mapsto \bar{Q}^h f(\cdot) := \sum_{j\in\mathbb{Z}} f(hj)\Phi(h^{-1}\cdot - j),$$

where Φ might be for instance an affine combination of the integer translates of the fundamental cardinal cubic B-spline function as in Example 2.5 below (of course higher orders than cubic and different functions are possible).

The well-known linear theory establishes results regarding the decay of the approximation error

$$\left\| \left(\frac{d}{dx}\right)^l (f - \bar{Q}^h f) \right\|_\infty$$

with the stepwidth h tending to zero. This approximation error in h is related to a few simple properties of Φ such as smoothness or polynomial exactness, see Section 2.1. The quasiinterpolation operator can thus be used as a projection onto the finite element space spanned by cubic B-splines with meshwidth h and to study the asymptotic approximation properties of these spaces.

Our main idea is to regard the expression for $\bar{Q}^h f$ as a weighted average of the samples $f(hj)$. By replacing affine averages with the Riemannian center of mass [20] we arrive at a definition of an associated quasiinterpolation operator $Q^h f$ for functions f with values in a differentiable manifold M, see Section 2.2.

In order to study the approximation errors of Q^h we need to be able to compare derivatives $d^l f$ with $d^l Q^h f$, both taking their values in the iterated tangent bundle $T^l M$ of M. This is achieved by using the so-called Sasaki metric on $T^l M$ which is described in Section 3.2. With the geodesic distance $\mathfrak{s}_l$ induced by the Sasaki metric our main result is that the approximation error

$$\mathfrak{s}_l \left(d^l f, d^l Q^h f\right)$$

behaves in the exact same way as the corresponding linear error, see Theorem 3.8 for a more precise statement.

As an application we construct nonlinear (finite element) approximation manifolds (see Definition 2.7 below) with a prescribed number of degrees of freedom and determine their approximation properties, see Theorem 3.9 below.

Besides approximating an explicitly given function, a main motivation for our construction is the potential use of these approximation manifolds to solve manifold-valued optimization problems in a finite element like fashion, compare [24, 25]. Theorem 3.9 opens up the door for a theoretical analysis of such algorithms.

We would like to note that our construction can easily be adapted to the multivariate case and also over simplicial grids (compare [25]). The study of the associated approximation properties will be the subject of forthcoming work.

1.1 Previous Work

There exists by now a substantial body of previous work related to nonlinear data types of which we only mention some examples. In the paper [23] a manifold-valued wavelet transform has been derived, its theoretical properties are investigated in [14, 13, 15]. The idea to use nonlinear subdivision schemes for the approximation of manifold-valued data has been investigated in [30, 12, 31, 34, 32, 28, 11, 33]. Among these we would like to single out [33] where a different construction of quasiinterpolants is presented and analyzed numerically. The already mentioned papers [24, 25] present a construction of first order geodesic finite element spaces much in the spirit of our (higher order) construction. Related to these constructions, various optimization problems are studied. In [18] a Riemannian framework is introduced for the modeling of human biomechanics. The paper [6] analyzes computational methods in tensor image processing based on geometric constructions. In [16] variational splines on manifolds are introduced. The paper [21] is concerned with the study of Riemannian cubics. Finally, we would like to mention [17], where Runge-Kutta methods are extended to to the case of Lie-valued (and more general) ODEs.

1.2 Outline

The outline is as follows. In Section 2 we will lay out the necessary background from linear approximation theory and also describe the nonlinear setup we will work with. After that, in Section 3 we will first show that it suffices to study our approximation problems in a chart, and then we will discuss the Sasaki metric which will allow us to formulate our results in an intrinsic fashion. Section 4 contains the proof of our main result and forms the main technical part of this paper. Finally, in Section 5 we present some numerical computations which confirm our theory. We also include an appendix containing some auxiliary results which will be needed in the course of the proof.

1.3 Notation

For a function $f : \mathbb{R} \to \mathbb{R}^d$ we use the usual terminology $\|f\|_\infty := \sup_x |f(x)|$ with $|\cdot|$ the maximum norm on $\mathbb{R}^d$. The space of continuous functions with values in a manifold M is denoted by $C^0(\mathbb{R}, M)$. If M is a Euclidean space we simply write C^0. Furthermore, we use the symbol l_∞ for the Banach space of bounded, real valued sequences on $\mathbb{Z}$. We use boldface notation for multiindices $\mathbf{j} \in \mathbb{Z}^k$ and denote $|\mathbf{j}|_1 := j_1 + \cdots + j_k$. Distance metrics on Riemannian manifolds are usually expressed by fraktur letters, for instance $\mathfrak{d}$. The space of polynomials of degree $\leq m$ on $\mathbb{R}$ shall be denoted by Π_m. For a set A, the symbol χ_A denotes its indicator function. The expression $\lfloor \alpha \rfloor$ denotes the largest integer which is smaller or equal $\alpha \in \mathbb{R}$. We use the usual notation $A = \mathcal{O}(B)$ or $A \lesssim B$ to indicate that the quantity A is bounded by a constant times B.

2 Preliminaries

2.1 Linear Theory

We start by reviewing some well-known facts from linear quasiinterpolation theory. For more information we refer to [3, 5, 7]. In the following definition we introduce the smoothness spaces C^α which we will work with.

Definition 2.1. *For a function $f : \mathbb{R} \to \mathbb{R}^d$ we define $\Delta_h^l f$, the l-th order forward difference with stepwidth $h > 0$ inductively by*

$$\Delta_h^0 f := f, \quad \textit{and} \quad \Delta_h^l f(x) := \Delta_h^{l-1}(x+h) - \Delta_h^{l-1}(x).$$

We say that $f \in C^\alpha$, $\alpha > 0$ if we have

$$\left\| \Delta_h \left(\frac{d}{dx}\right)^{\lfloor\alpha\rfloor} f \right\|_\infty = \mathcal{O}\left(h^{\alpha-\lfloor\alpha\rfloor}\right).$$

Linear finite element spaces on a regular partition of an interval are usually constructed from a function $\Phi : \mathbb{R} \to \mathbb{R}$ satisfying the following requirements:

$$\sum_{j\in\mathbb{Z}} \Phi(\cdot - j) = 1 \qquad \text{[resolution of identity]} \tag{1}$$

$$\exists N \in \mathbb{R}: \ \operatorname{supp} \Phi \subset [-N, N] \qquad \text{[locality]} \tag{2}$$

$$\sum_{j\in\mathbb{Z}} p(i)\Phi(\cdot - j) = p(\cdot), \ \forall p \in \Pi_{m-1} \qquad \text{[polynomial exactness]} \tag{3}$$

$$\Phi \in C^s \qquad \text{[smoothness]} \tag{4}$$

Given such a function, we can define the *linear finite element spaces*

$$V_h(\Phi) := \left\{ \sum_{j\in\mathbb{Z}} a(j)\Phi\left(h^{-1}\cdot - j\right) : \ (a(j))_{j\in\mathbb{Z}} \in l_\infty \right\}.$$

In many applications it is necessary to approximate a given function f by the finite element spaces $V_h(\Phi)$. This can be done either implicitly, when f is given as a solution of an operator equation, or explicitly, when f is explicitly given. One of the very fundamental results in approximation theory is that the approximation rate of the finite element spaces for a given function f is exactly governed by the smoothness of f:

Theorem 2.2 (Jackson Theorem)**.** *Assume that $f \in C^\alpha$ with $\alpha < m$. Then for $l < \alpha$ and $l < s$ we have*

$$\inf_{g\in V_h(\Phi)} \left\| \left(\frac{d}{dx}\right)^l (f - g) \right\|_\infty = \mathcal{O}(h^{\alpha-l}). \tag{5}$$

In fact, more can be said: under the assumptions $(1 - 4)$, a quasioptimal approximant $g \in V_h(\Phi)$ can be constructed explicitly.

Definition 2.3. *For Φ satisfying (1 – 4) we define the* linear quasiinterpolation operator $\bar{Q}^h : C^0 \to V_h(\Phi)$ *defined via*

$$f(\cdot) \mapsto \bar{Q}^h f(\cdot) := \sum_{j\in\mathbb{Z}} f(hj)\Phi(h^{-1}\cdot - j) \in V_h(\Phi). \tag{6}$$

Now we can state the following stronger form of Theorem 2.2.

Theorem 2.4. *Assume that $f \in C^\alpha$ with $\alpha < m$. Then for $l < \alpha$ and $l < s$ we have*

$$\left\| \left(\frac{d}{dx}\right)^l (f - \bar{Q}^h f) \right\|_\infty = \mathcal{O}(h^{\alpha-l}). \tag{7}$$

Example 2.5 (Cubic B-spline quasiinterpolation)**.** *A classical example of a quasiinterpolation operator can be constructed from the cardinal cubic B-spline function $B_3(\cdot)$, [3]. Even though the choice $\Phi = B_3$ would only lead to the low degree polynomial reproduction of 1, it is possible to preprocess the sampling data of f and arrive at a fourth order quasiinterpolation scheme*

$$\bar{Q}^h f(\cdot) := \sum_{j\in\mathbb{Z}} B_3(h^{-1}\cdot - j)\left(-\frac{1}{6}f(h(j-1)) + \frac{4}{3}f(hj) - \frac{1}{6}f(h(j+1))\right),$$

which falls into our definition by putting

$$\Phi(\cdot) = -\frac{1}{6}B_3(\cdot - 1) + \frac{4}{3}B_3(\cdot) - \frac{1}{6}B_3(\cdot + 1)$$

which can be shown to satisfy (3) with $m = 4$ [3]. The linear approximation spaces are given by piecewise cubic polynomials, C^2 at the knots.

Clearly, Theorem 2.4 implies Theorem 2.2, for a proof of Theorem 2.4 we refer to [3]. We would like to close this section by noting that the backbone of any approximation method or finite element solver is given by the validity of a Jackson theorem [1].

2.2 Nonlinear Approximation

Having presented a brief introduction the the construction of regular finite element spaces for functions $f : \mathbb{R} \to \mathbb{R}^d$, we would now like to ask whether such a construction can be meaningfully extended to the case of manifold-valued functions $f : \mathbb{R} \to M$, M being a d-dimensional differentiable manifold. More specifically, our goal is to construct a quasiinterpolation operator operating on M-valued functions such that an equivalent to Theorem 2.4 holds. The approach we will take is to regard the sum

$$\sum_{j\in\mathbb{Z}} f(hj)\Phi(h^{-1}\cdot -j)$$

in (6) as a weighted average of the points $(f(hj))_{j\in\mathbb{Z}}$ with weights given by $\left(\Phi(h^{-1}\cdot -j)\right)_{j\in\mathbb{Z}}$. By (1) this is justified. The key insight is that weighted averages also exist in Riemannian manifolds, i.e. a pair (M, g) with M a differentiable manifold and g a metric tensor field on M, see [8] for more information on Riemannian geometry.

Definition 2.6. *Assume that (M, g) is a Riemannian manifold with induced metric $\mathfrak{d}$. For points $(p(j))_{j\in J}$ and weights $(w(j))_{j\in J}$ the Riemannian center of mass*

$$x^* = av_M\left((p(j))_{j\in J}, (w(j))_{j\in J}\right)$$

is defined as

$$x^* = argmin_{x\in M} \sum_{j\in J} w(j)\mathfrak{d}(p(j), x)^2.$$

It can be shown that locally the Riemannian center of mass exists and is unique. Furthermore, locally, it is characterized by the first order equilibrium condition

$$\sum_{j\in J} w(j) \log_M(x^*, p(j)) = 0, \tag{8}$$

see e.g. [20]. Here, $\log_M$ denotes the logarithm mapping of M, e.g., the local inverse of the exponential function $\exp_M$ of M, [8].

The natural idea is now to replace the affine weighted average in (6) with the Riemannian average in M. This leads to the following definition.

Definition 2.7. *We define the* nonlinear finite element manifolds

$$V_h^M(\Phi) := \left\{ av_M\left((a(j))_{j\in\mathbb{Z}}, (\Phi(h^{-1}\cdot -j))_{j\in\mathbb{Z}}\right) : \text{ whenever the average is well-defined } \right\}.$$

It is our interest to study the approximation properties of these nonlinear finite element manifolds, that is instead of approximating with linear spaces we are approximating with nonlinear manifolds with the same number of degrees of freedom. We will do this by considering explicit projection operators onto these manifolds in terms of nonlinear quasiinterpolation as defined below.

Definition 2.8. *Define the* nonlinear quasiinterpolation operator

$$Q^h : C^0(\mathbb{R}, M) \to C^0(\mathbb{R}, M)$$

by

$$Q^h f(\cdot) := av_M\left((f(hj))_{j\in\mathbb{Z}}, (\Phi(h^{-1}\cdot -j)_{j\in\mathbb{Z}}\right). \tag{9}$$

Remark 2.9. *Note that for h sufficiently small and $f \in C^\alpha$ for any $\alpha > 0$ the expression $Q^h f$ is always defined. This is due to the locality of Φ and the fact that all sampling points used in the computation of $Q^h f(x)$ lie in a set of arbitrarily small diameter. Therefore, by local well-definedness of the Riemannian* *center of mass, the average leading to $Q^h f(x)$ exists. For this reason we will ignore the issue of well-definedness in the sequel and tacitly assume that the sampling width h is sufficiently small.*

We would also like to remark that for several examples of practical interest, the manifold M possesses particular structural properties which make the Riemannian averages defined for any *initial data. One such example is the manifold $SPD(n)$ of symmetric positive definite $n \times n$ matrices arising e.g. in Diffusion Tensor MRI in medical imaging [22].*

Remark 2.10. *Karcher [20] constructs a similar, continuous mollifying procedure for approximating manifold-valued functions (which may also have a general manifold as domain of definition). In [20, Page 521] he writes*

> *The approximation of higher derivatives of f, if they exist, by the corresponding derivatives of* [the approximant] *is not clear to me. [...]*

For the case of functions defined on $\mathbb{R}$ and the semidiscrete mollifying operation defined by the Q^h operator we are able to settle this question. We believe that our approach could be suitable for clarifying this question also in the general case.

The idea to replace affine averages by the Riemannian center of mass is not new. In [12, 29] it has been applied to study smoothness and approximation properties of manifold-valued subdivision schemes. The very recent and interesting work [24, 25] constructs first order finite element spaces essentially in the same way as we do. These spaces are then used to solve manifold-valued optimization problems arising e.g. in Cosserat rod modeling in a very natural fashion. We expect our present results to be relevant in this direction. Finally we would like to mention [22] where Riemannian methods are used for the filtering, denoising and statistical analysis of Diffusion Tensor MRI data.

The central question to ask is whether the approximation properties which are known in the linear case (e.g., Theorem 2.4) also hold for the nonlinear quasiinterpolation operators. Answering this question affirmatively will be the main theme of the present paper.

In order to pose the question of approximability of the derivatives of f correctly, it is necessary to give a canonical notion of difference between elements of the tangent bundles $T^l M$. In Section 3 below we address this problem and present an appropriate construction, namely the so-called Sasaki metric [26]. Then, in Section 4 we show that indeed the linear approximation results also hold in the same form for manifold-valued quasiinterpolation.

3 Localization and Natural Metrics

The present section discusses the effect of conducting our computations in a chart. After that we introduce the concept of the Sasaki metric on tangent bundles of Riemannian manifolds. This allows us to state our final result Theorem 3.8 in an intrinsic fasion.

3.1 Computations in Charts

We now consider a chart $\gamma : M \to \mathbb{R}^d$ and its induced chart $(\gamma, d\gamma) : TM \to \mathbb{R}^{2d}$. With respect to this chart, using (8), we can write

$$\gamma \circ Q^h f(x) = u\left(\gamma \circ Q^h f(x), \sum_{j\in\mathbb{Z}} \Phi(h^{-1}x - j) v\left(\gamma \circ Q^h f(x), \gamma \circ f(hj)\right)\right), \tag{10}$$

with

$$u(x,y) := \gamma \circ \exp_M\left(\gamma^{-1}(x), (d\gamma)^{-1}\big|_x y\right) \quad \text{and} \quad v(x,y) := d\gamma\big|_{\gamma^{-1}x}\left(\log_M\left(\gamma^{-1}x, \gamma^{-1}y\right)\right), \quad x,y \in \mathbb{R}^d,$$

whenever defined.

Our main result is that in any chart γ we have the following approximation rate:

Theorem 3.1. *Assume that $f \in C^\alpha$ with $\alpha < m$ and $\alpha < s$. Then for $l < \alpha$ and any chart γ we have*

$$\left\| \left(\frac{d}{dx}\right)^l \left(\gamma \circ f - \gamma \circ Q^h f\right) \right\|_\infty = \mathcal{O}(h^{\alpha - l}). \tag{11}$$

The proof follows in Section 4.

In particular, the previous theorem easily implies that for *any* Riemannian metric g_l on $T^l M$ with induced distance $\mathfrak{d}_l$ (meaning that $\mathfrak{d}_l$ measures the *geodesic distance* between two points, see [8]) we have the following result

Theorem 3.2. *Assume that $f \in C^\alpha$ with $\alpha < m$ and $\alpha < s$. Then for $l < \alpha$ and any Riemannian metric g_l on the vector bundle $T^l M$ with induced distance metric $\mathfrak{d}_l$, we have the estimate*

$$\sup_x \mathfrak{d}_l \left(d^l f, d^l Q^h f\right) = \mathcal{O}(h^{\alpha - l}), \tag{12}$$

where $d^l f : \mathbb{R} \to T^l M$ denotes the total differential of order l [8]. The implicit constant is uniform for data values in a compact set.

Proof. Any chart γ induces the chart $d^l \gamma : T^l M \to \mathbb{R}^{2^l d}$. With respect to this latter chart, by Theorem 3.1 we have

$$\left\| d^l \gamma \circ \left(d^l Q^h f(\cdot)\right) - d^l \gamma \circ \left(d^l f(\cdot)\right) \right\|_\infty = \mathcal{O}(h^{\alpha - l}).$$

Since the inverse $(d^l \gamma)^{-1}$ of the chart $d^l \gamma$ is smooth and in particular Lipschitz with respect to the metric $\mathfrak{d}_l$, we arrive at the estimate

$$\begin{aligned} \sup_x \mathfrak{d}_l \left(d^l f(x), d^l Q^h f(x)\right) &= \sup_x \mathfrak{d}_l \left((d^l \gamma)^{-1} \circ d^l \gamma \circ \left(d^l f(x)\right), (d^l \gamma)^{-1} \circ d^l \gamma \circ \left(d^l Q^h f(x)\right)\right) \\ &= \mathcal{O}\left(\left\| d^l \gamma \circ \left(d^l Q^h f(\cdot)\right) - d^l \gamma \circ \left(d^l f(\cdot)\right) \right\|_\infty\right) = \mathcal{O}(h^{\alpha - l}) \end{aligned}$$

which proves the assertion for data contained in the domain of definition for any single chart γ. If the data values are contained in a compact set, finitely many charts cover this set and therefore the uniformity of the implied constant follows. □

3.2 Sasaki Metric

Theorem 3.2 above states that for *any* Riemannian metric defined on the tangent bundle $T^l M$ we can show an approximation theorem as strong as for the linear case. It is a well-known fact that any differentiable manifold admits a Riemannian metric [27]. However, it would be nice to be able to single out one specific Riemannian metric on $T^l M$ within which we can measure the approximation error between $d^l f$ and its approximation $d^l Q^h f$. The present section describes such a canonical metric, the so-called Sasaki metric which has been introduced in [26].

3.2.1 Construction

Before we can describe the Sasaki metric we need some preliminary facts from Riemannian geometry of tangent bundles, see [35] for more information. Assume we are given a Riemannian manifold (M, g), where g is a symmetric $(0, 2)$-tensor field on M. Consider its tangent bundle TM which carries a natural manifold structure [8]. For $p \in M$ we denote $T_p M$ the tangent space attached to p and likewise we denote $T_{(p,u)} TM$ the tangent space attached to the tangent bundle of M at the tangent vector $(p, u) \in TM$. We want to find a suitable metric tensor g_T on the manifold TM, meaning that g_T acts on pairs of tangent vectors of TM. With $\pi : TM \to M$ denoting the bundle projection, we define the *vertical subspace*

$$\mathcal{V}_{(p,u)} := \ker \left(d\pi|_{(p,u)}\right)$$

of $T_{(p,u)} TM$ at $(p, u) \in TM$. The terminology 'vertical' stems from the fact that a vertical vector leaves the basepoint p stationary and only moves along the fiber $\pi^{-1}\{p\} \subset TM$. Without going into detail, we would like to mention that the vertical subspace $\mathcal{V}_{(p,u)}$ can be complemented by the so-called *horizontal subspace*

$$\mathcal{H}_{(p,u)} := \ker\left(K_{(p,u)}\right),$$

where $K_{(p,u)} : T_{(p,u)}TM \to T_pM$ is defined as $(c(t), U(t)) \in TM,\ (c(0), U(0)) = (p, u) \mapsto \nabla_{c'}U(0)$, where ∇ denotes the Levi-Civita connection of (M, g) ($\mathcal{H}_{(p,u)}$ consists of tangent vectors of parallel vector fields). We have

$$T_{(p,u)}TM = \mathcal{H}_{(p,u)} \oplus \mathcal{V}_{(p,u)},$$

see also [8, 9].

Definition 3.3. *Let $X \in T_pM$. Then the* horizontal lift *of X at $(p, u) \in TM$ is the unique vector $X^h \in \mathcal{H}_{(p,u)}$ such that $d\pi_{(p,u)}(X^h) = X$. The* vertical lift *X^v of X at (p, u) is the unique vector $X^v \in \mathcal{V}_{(p,u)}$ such that $K_{(p,u)}(X^v) = X$.*

Each $Z \in T_{(p,u)}TM$ can be uniquely expressed as $Z = X^h + Y^v$ for $X, Y \in T_pM$. Based on this decomposition we can now define natural metrics on TM [19].

Definition 3.4. *A metric g_T on TM is called* natural *if*

$$g_T(X^h, Y^h) = g(X, Y) \quad \textit{and} \quad g_T(X^h, Y^v) = 0 \quad \textit{for all vector fields } X, Y \textit{ on } M.$$

Several different natural metrics for TM can be defined by specifying conditions on $g_T(X^v, Y^v)$. For instance, the so-called *Cheeger-Gromoll metric* is uniquely defined by setting

$$g_T^{CG}(X^v, Y^v)_{(p,u)} = \frac{1}{1 + g(X, Y)^2}\left(g(X, Y) + g(X, u)g(Y, u)\right),$$

see [2]. We will focus on the simpler *Sasaki metric.*

Definition 3.5. *The Sasaki metric is the unique natural metric on TM which satisfies*

$$g_T^S(X^v, Y^v) = g(X, Y) \quad \textit{for all vector fields } X, Y \textit{ on } M.$$

By iterating this construction, a Sasaki metric can be defined on the iterated tangent bundles T^lM, $l \in \mathbb{Z}_+$.

The geodesic distance metric *on T^lM induced by the Riemannian Sasaki metric will be denoted by $\mathfrak{s}_l$ – it gives a natural notion of (local) distance $\mathfrak{s}_l(x, y)$ between points $x, y \in T^lM$.*

For a list of various properties satisfied by the Sasaki metric we refer to the original paper [26].

3.2.2 Intrinsic Approximation Results

Now that we have singled out a canonical metric on the tangent bundles T^lM we can define smoothness properties for M-valued functions intrinsically as follows.

Definition 3.6. *A M-valued function $f : \mathbb{R} \to M$ is in $C^\alpha(\mathbb{R}, M)$ if*

$$\mathfrak{s}_{\lfloor\alpha\rfloor}\left(d^{\lfloor\alpha\rfloor}f(\cdot + h), d^{\lfloor\alpha\rfloor}f(\cdot)\right) = \mathcal{O}(h^{\alpha - \lfloor\alpha\rfloor})$$

with the implicit constant uniform in x for data values $f(x)$ in a compact set.

Remark 3.7. *It is easy to see that $f \in C^\alpha(\mathbb{R}, M)$ if and only if $\gamma \circ f \in C^\alpha$ for all charts γ.*

We can finally state our main theorem.

Theorem 3.8. *Assume that $f \in C^\alpha(\mathbb{R}, M)$ with $\alpha < m$ and $\alpha < s$. Then for $l < \alpha$ we have the estimate*

$$\mathfrak{s}_l\left(d^lf, d^lQ^hf\right) = \mathcal{O}(h^{\alpha - l}). \tag{13}$$

The implicit constant is uniform for data values $f(x)$ in a compact set.

Proof. The proof is a direct consequence of Theorem 3.2. □

This theorem immediately implies the following approximation result for the approximation manifold defined in Definition 2.7.

Theorem 3.9. *Assume that $f \in C^\alpha(\mathbb{R}, M)$ with $\alpha < m$ and $\alpha < s$. Then for $l < \alpha$ we have the estimate*

$$\inf_{g \in V_h^M(\Phi)} \mathfrak{s}_l \left(d^l f, d^l g\right) = \mathcal{O}(h^{\alpha - l}). \tag{14}$$

The implicit constant is uniform for data values $f(x)$ in a compact set.

4 Proof of Theorem 3.1

From now on we will simply write $Q^h f(\cdot)$ in place of $\gamma \circ Q^h f(\cdot)$ with the understanding that we are actually computing in a chart. The only property we shall use is the representation (10) and the smoothness of the functions u, v.

In a chart the linear quasiinterpolation operator can be written as

$$\bar{Q}^h f(x) = \sum_{j \in \mathbb{Z}} \Phi(h^{-1}x - j) u \left(Q^h f(x), v\left(Q^h f(x), f(hj)\right)\right). \tag{15}$$

The main idea is to compare the nonlinear approximation operator $Q^h f$ with its linear counterpart $\bar{Q}^h f$ and to use known properties of the linear operator to show the desired result. Estimating the error between linear and nonlinear quasiinterpolation will make up the bulk of this section, culminating in the proof of Theorem 3.1 in Section 4.2 below.

We now define an important asymptotic quantity which will be frequently used in the sequel.

Definition 4.1. *For $f : \mathbb{R} \to \mathbb{R}^d$, $h > 0$, $l \in \mathbb{N}$, we define the quantity*

$$\Omega_{l,h}(f) := \sum_{|\mathbf{r}|_1 = l} \prod_p \|\Delta_h^{r_p} f\|_\infty. \tag{16}$$

We collect some simple properties of the quantities $\Omega_{l,h}(f)$.

Lemma 4.2. *Assume that $f \in C^\alpha$ with $\alpha > l$. Then*

$$\Omega_{l,h}(f) = \mathcal{O}(h^l). \tag{17}$$

Proof. Assume that $f \in C^\alpha$. Then for any $l' < \alpha$ we have

$$\|\Delta_h^{l'} f\|_\infty = \mathcal{O}(h^{l'}).$$

It follows that

$$\Omega_{l,h}(f) = \sum_{|\mathbf{r}|_1 = l} \prod_p \|\Delta_h^{r_p} f\|_\infty = \mathcal{O}\left(\sum_{|\mathbf{r}|_1 = l} \prod_p h^{r_p}\right) = \mathcal{O}(h^l).$$

□

Lemma 4.3. *We have*

$$\Omega_{l_1,h}(f)\Omega_{l_2,h}(f) = \mathcal{O}\left(\Omega_{l_1+l_2,h}(f)\right). \tag{18}$$

Proof. The proof follows simply by noting that all terms occurring on the left hand side of (18) also occur on the right hand side. □

4.1 Proximity Inequality

Our basic strategy will be to compare the linear representation with the nonlinear one in order to infer properties for the nonlinear operator from properties of the linear one.

In the sequel we will make use of the following definition.

Definition 4.4. *For an arbitrary vector-valued sequence $p : \mathbb{Z} \to \mathbb{R}^d$ we define the l-th forward difference via*

$$\delta^l p(j) := \sum_{i=0}^{l} (-1)^{l-i} \binom{l}{i} p(j+i). \tag{19}$$

In addition we shall use the following notation. For a vector $v \in \mathbb{R}^d$ we define the *concatenation*

$$[v]^k := (v, \dots, v) \in \mathbb{R}^{kd}.$$

Furthermore, for a function $f(x)$ and a multiindex $\mathbf{l} \in \mathbb{Z}_+^k$ we denote

$$\boldsymbol{\Delta}_h^{\mathbf{l}}[f]^k(\mathbf{x}) := \left(\Delta_h^{l_1} f(x_1), \dots, \Delta_h^{l_k} f(x_k)\right), \quad \mathbf{x} \in \mathbb{R}^k,$$

and similarly for a sequence $p(j)$

$$\delta^{\mathbf{l}}[p]^k(\mathbf{j}) := \left(\delta^{l_1} p(j_1), \dots, \delta^{l_k} p(j_k)\right), \quad \mathbf{j} \in \mathbb{Z}^k.$$

We will use Taylor expansion of the smooth function u in the second component, namely

$$u(y,z) = \sum_{k=0}^{M} \frac{1}{k!} d_2^{(k)} u\big|_{(y,0)} \left([z]^k\right) + R^{(M+1)}(y,z) \left([z]^{M+1}\right), \tag{20}$$

with $R^{(M+1)}$ a $(M+1)$-multilinear form, smoothly depending on y, z, and $d_2^{(k)}$ denoting the k-th total derivative in the second coordinate. Therefore, Taylor expanding (10) up to order M we obtain

$$\begin{aligned} Q^h f(x) = \sum_{k=0}^{M} \frac{1}{k!} d_2^{(k)} u\big|_{(Q^h f(x),0)} \left(\left[\sum_{j\in\mathbb{Z}} \Phi(h^{-1}x - j) v\left(Q^h f(x), f(hj)\right) \right]^k \right) + \\ R^{(M+1)} \left(Q^h f(x), \sum_{j\in\mathbb{Z}} \Phi(h^{-1}x - j) v\left(Q^h f(x), f(hj)\right) \right) \left(\left[\sum_{j\in\mathbb{Z}} \Phi(h^{-1}x - j) v\left(Q^h f(x), f(hj)\right) \right]^{M+1} \right). \end{aligned} \tag{21}$$

Taylor expanding (15) up to order M we obtain

$$\begin{aligned} \bar{Q}^h f(x) = \sum_{k=0}^{M} \sum_{j\in\mathbb{Z}} \Phi(h^{-1}x - j) \frac{1}{k!} d_2^{(k)} u\big|_{(Q^h f(x),0)} \left(\left[v\left(Q^h f(x), f(hj)\right)\right]^k \right) + \\ \sum_{j\in\mathbb{Z}} \Phi(h^{-1}x - j) R^{(M+1)} \left(Q^h f(x), v\left(Q^h f(x), f(hj)\right)\right) \left(\left[v\left(Q^h f(x), f(hj)\right)\right]^{M+1} \right). \end{aligned} \tag{22}$$

Our first goal is to gather useful estimates for the differences

$$\begin{aligned} F_k := d_2^{(k)} u\big|_{(Q^h f(x),0)} \left(\left[\sum_{j\in\mathbb{Z}} \Phi(h^{-1}x - j) v\left(Q^h f(x), f(hj)\right) \right]^k \right) - \\ \sum_{j\in\mathbb{Z}} \Phi(h^{-1}x - j) d_2^{(k)} u\big|_{(Q^h f(x),0)} \left(\left[v\left(Q^h f(x), f(hj)\right)\right]^k \right), \end{aligned} \tag{23}$$

namely the following lemma. The proof is based on a combinatorial argument exploiting the polynomial reproduction property (3) of Φ. Similar arguments have been used in previous work in different contexts [10, 32, 4].

Lemma 4.5. *For $l < s$ we have*

$$|\Delta_{h'}^{l} F_k| = \mathcal{O}\left(\Omega_{m,h}(f)\sum_{i=0}^{l}(h^{-1}h')^{l-i}\Omega_{i,h'}\left(Q^h f\right)\right). \tag{24}$$

Proof. Define $\bar{j} := \lfloor h^{-1}x - N \rfloor$. Then every index j occurring in the summation formula for F_k satisfies $j \geq \bar{j}$ and $j - \bar{j} = \mathcal{O}(1)$, due to the support properties of Φ. By the definition of the forward differences we can represent each vector $v\left(Q^h f(x), f(hj)\right)$ by

$$v\left(Q^h f(x), f(hj)\right) = \sum_{i=0}^{j-\bar{j}}\binom{j-\bar{j}}{i}\delta^i v\left(Q^h f(x), f(h\bar{j})\right). \tag{25}$$

Inserting (25) into the definition of F_k we get $F_k = F_k^1 + F_k^2$, where

$$F_k^1 := \sum_{i_1,\dots,i_k\in\mathbb{Z}_+}\sum_{j_1,\dots,j_k\in\mathbb{Z}}\prod_r \Phi(h^{-1}x - j_r)\binom{j_r-\bar{j}}{i_r}$$
$$d_2^{(k)}u\big|_{(Q^h f(x),0)}\left(\delta^{i_1}v\left(Q^h f(x), f(h\bar{j})\right),\dots,\delta^{i_k}v\left(Q^h f(x), f(h\bar{j})\right)\right),$$

and

$$F_k^2 := \sum_{i_1,\dots,i_k\in\mathbb{Z}_+}\sum_{j\in\mathbb{Z}}\Phi(h^{-1}\cdot - j)\prod_r\binom{j-\bar{j}}{i_r}$$
$$d_2^{(k)}u\big|_{(Q^h f(x),0)}\left(\delta^{i_1}v\left(Q^h f(x), f(h\bar{j})\right),\dots,\delta^{i_k}v\left(Q^h f(x), f(h\bar{j})\right)\right).$$

Let us now fix $(i_1,\dots,i_k)$ with $i_1+\dots+i_k < m$. Then the functions $p_r := p_r(j_r) = \binom{j_r-\bar{j}}{i_r}$, $p(j) := \prod_{r=1}^{k}\binom{j-\bar{j}}{i_r}$ are polynomials of total degree $< m$. Therefore, by polynomial reproduction of the function Φ of order $m-1$ we get

$$\sum_{j_1,\dots,j_k\in\mathbb{Z}}\prod_r \Phi(h^{-1}x - j_r)\binom{j_r-\bar{j}}{i_1} = \sum_{j\in\mathbb{Z}}\Phi(h^{-1}x - j)\prod_r\binom{j-\bar{j}}{i_r} = p(h^{-1}x),$$

and therefore we only need to consider terms with $i_1+\dots+i_k \geq m$ in the summation formula for $F_k = F_k^1 - F_2^k$. In summary, F_k can be expressed as a finite linear combination of terms of the form

$$c(x)d_2^{(k)}u\big|_{(Q^h f(x),0)}\left(\delta^{i_1}v\left(Q^h f(x), f(h\bar{j})\right),\dots,\delta^{i_k}v\left(Q^h f(x), f(h\bar{j})\right)\right),$$

with

$$i_1+\dots+i_k \geq m$$

and

$$c(x) := \prod_{r=1}^{k}\Phi(h^{-1}x - j_1)p_r(j_r) - \chi_{j_1=\dots=j_k}(j_1,\dots,j_k)\Phi(h^{-1}x - j_1)p(j_1). \tag{26}$$

We need to bound $\Delta_{h'}^{l}F_k$ and by the above discussion it suffices to get a bound on

$$\Delta_{h'}^{l}c(x)d_2^{(k)}u\big|_{(Q^h f(x),0)}\left(\delta^{i_1}v\left(Q^h f(x), f(h\bar{j})\right),\dots,\delta^{i_k}v\left(Q^h f(x), f(h\bar{j})\right)\right) \tag{27}$$

with $i_1+\dots+i_k \geq m$. The expression (27) can be rewritten as a finite linear combination of terms of the form

$$\left(\Delta_{h'}^{l_1}c(x_1)\right)\left(\Delta_{h'}^{l_2}d_2^{(k)}u\big|_{(Q^h f(x_2),0)}\right)\left(\Delta_{h'}^{l_3}\delta^{i_1}v\left(Q^h f(x_3), f(h\bar{j})\right),\dots,\Delta_{h'}^{l_{k+2}}\delta^{i_k}v\left(Q^h f(x_{k+2}), f(h\bar{j})\right)\right),$$

with $l_1+\dots+l_{k+2} = l$. By (26) and the smoothness of Φ we can estimate

$$\Delta_{h'}^{l_1}c(x) = \mathcal{O}\left(h^{-l_1}(h')^{l_1}\right). \tag{28}$$

Further, due to the smoothness of u and Lemma A.2 we can estimate

$$\left\| \Delta_{h'}^{l_2} d_2^{(k)} u \big|_{(Q^h f(x_2),0)} \right\| = \mathcal{O}\left(\Omega_{l_2,h'}(Q^h f)\right). \tag{29}$$

Finally, by Lemma A.3 we have for general l', i' that

$$\Delta_{h'}^{l'} \delta^{i'} v\left(Q^h f(x), f(h\bar{j})\right) = \mathcal{O}\left(\Omega_{l',h'}(Q^h f)\Omega_{i',h}(f)\right). \tag{30}$$

By putting estimates (28), (29) and (30) into (27) we finally arrive at the desired result. $\square$

We next treat the remainder terms in the Taylor representations.

Lemma 4.6. *For $M > 0$ and $l < s$ we have*

$$\Delta_{h'}^{l} R^{(M+1)}\left(Q^h f(x), \sum_{j\in\mathbb{Z}} \Phi(h^{-1}x - j) v\left(Q^h f(x), f(hj)\right)\right)\left(\left[\sum_{j\in\mathbb{Z}} \Phi(h^{-1}x - j) v\left(Q^h f(x), f(hj)\right)\right]^{M+1}\right)$$
$$= \|\Delta_h f\|_\infty^{M-l} \mathcal{O}\left(\sum_{i=0}^{l} (h^{-1}h')^{l-i} \Omega_{i,h'}\left(Q^h f\right)\right) \tag{31}$$

and

$$\Delta_{h'}^{l} \sum_{j\in\mathbb{Z}} \Phi(h^{-1}x - j) R^{(M+1)}\left(Q^h f(x), v\left(Q^h f(x), f(hj)\right)\right)\left(\left[v\left(Q^h f(x), f(hj)\right)\right]^{M+1}\right)$$
$$= \|\Delta_h f\|_\infty^{M-l} \mathcal{O}\left(\sum_{i=0}^{l} (h^{-1}h')^{l-i} \Omega_{i,h'}\left(Q^h f\right)\right). \tag{32}$$

Proof. We start with (31). The proof goes by iteratively rewriting the divided differences in order to arrive at simpler expressions. First, note that the left-hand-side of (31) can be expressed as a linear combination of terms of the form

$$\left(\Delta_{h'}^{l_1} R^{(M+1)}\left(Q^h f(x_1), \sum_{j\in\mathbb{Z}} \Phi(h^{-1}x_1 - j) v\left(Q^h f(x_1), f(hj)\right)\right)\right)$$
$$\left(\boldsymbol{\Delta}_{h'}^{\mathbf{l_2}} \left[\sum_{j\in\mathbb{Z}} \Phi(h^{-1}\cdot - j) v\left(Q^h f(\cdot), f(hj)\right)\right]^{M+1} (\mathbf{x_2})\right),$$

with $l_1 + |\mathbf{l_2}|_1 = l$. Due to the smoothness of u and (48) we can estimate

$$\left\| \Delta_{h'}^{l_1} R^{(M+1)}\left(Q^h f(x), \sum_{j\in\mathbb{Z}} \Phi(h^{-1}x - j) v\left(Q^h f(x), f(hj)\right)\right) \right\| \leq$$
$$\sum_{r_1+r_2=l_1} \Omega_{r_1,h'}(Q^h f) \Omega_{r_2,h'}\left(\sum_{j\in\mathbb{Z}} \Phi(h^{-1}\cdot - j) v\left(Q^h f(\cdot), f(hj)\right)\right).$$

Due to the smoothness of $R^{(M+1)}$ and Equation (48), this expression can further be estimated by

$$\sum_{r_1+r_2=l_1} \Omega_{r_1,h'}\left(Q^h f\right) \sup_j \Omega_{r_2,h'}\left(\Phi(h^{-1}\cdot - j) v\left(Q^h f(\cdot), f(hj)\right)\right). \tag{33}$$

For the second term in this product we employ the following estimate for $r \geq 0$:

$$\left\|\Delta_{h'}^{r} \Phi(h^{-1}\cdot - j) v\left(Q^h f(\cdot), f(hj)\right)\right\|_\infty \lesssim \sum_{t_1+t_2=r} \Omega_{t_1,h'}\left(\Phi(h^{-1}\cdot - j)\right) \Omega_{t_2,h'}\left(v\left(Q^h f(\cdot), f(hj)\right)\right)$$

which, by the smoothness of v and (48), can be bounded by

$$\sum_{t_1+t_2=r} \Omega_{t_1,h'}\left(\Phi(h^{-1}\cdot -j)\right)\Omega_{t_2,h'}\left(Q^h f(\cdot)\right). \tag{34}$$

Due to the smoothness of Φ, this expression can be estimated by

$$\sum_{t_1+t_2=r} h^{-t_1}(h')^{t_1}\Omega_{t_2,h'}\left(Q^h f(\cdot)\right). \tag{35}$$

Combining (33) and (35) and using Lemma 4.3, we obtain

$$\left\|\Delta_{h'}^{l_1} R^{(M+1)}\left(Q^h f(x), \sum_{j\in\mathbb{Z}} \Phi(h^{-1}x-j)v\left(Q^h f(x), f(hj)\right)\right)\right\| = \mathcal{O}\left(\sum_{i\le l_1}(h^{-1}h')^{l_1-i}\Omega_{i,h'}\left(Q^h f\right)\right). \tag{36}$$

Now we go on to estimate entries in the vector

$$\mathbf{\Delta}_{h'}^{\mathbf{l_2}}\left[\sum_{j\in\mathbb{Z}}\Phi(h^{-1}\cdot -j)v\left(Q^h f(x), f(hj)\right)\right]^{M+1}.$$

since $|\mathbf{l_2}|_1 \le l$, we can assume that only the first l entries of $\mathbf{l_2}$ are nonzero. The other $M+1-l$ entries can be bounded by $\mathcal{O}\left(\|\Delta_h f\|_\infty\right)$ by Lemma A.1. For the other entries we can use the estimate (35) with r replaced by $(l_2)_i$. Combining this estimate with (36) finally yields the desired estimate to show (31). The proof for (32) works analogously. □

Putting together the two previous estimates we arrive at the following general result.

Corollary 4.7. *We have the proximity inequality*

$$\left\|\Delta_{h'}^{l}\left(Q^h f(\cdot) - \bar{Q}^h f(\cdot)\right)\right\|_\infty = \mathcal{O}\left(\Omega_{m,h}(f)\sum_{i=0}^{l}(h^{-1}h')^{l-i}\Omega_{i,h'}\left(Q^h f\right)\right). \tag{37}$$

Proof. If we disregard the residue terms in the Taylor expansions of $Q^h f$ and $\bar{Q}^h f$, then the estimate follows directly from Lemma 4.5. The residual terms are handled by Lemma 4.6 and by choosing $M \ge m+1+l$. □

4.2 Main Theorem

We are almost in a position to give a proof of our main Theorem 3.1. First we need the following lemma which states that the quasiinterpolants $Q^h f$ are uniformly smooth, independent of h.

Lemma 4.8. *Assume that $f \in C^\alpha$ with $\alpha < m$ and $\alpha < s$. Then for $l \le \alpha$ we have*

$$\left\|\Delta_{h'}^{l} Q^h f\right\|_\infty = \mathcal{O}\left((h')^l\right), \tag{38}$$

the implicit constant being independent of h.

Proof. We perform induction on l, the case $l=0$ being trivial. Let us now assume that for all $l' < l$ we have an inequality of the form (38). We can estimate

$$\left\|\Delta_{h'}^{l} Q^h f\right\|_\infty \le \left\|\Delta_{h'}^{l}\left(Q^h f - \bar{Q}^h f\right)\right\|_\infty + \left\|\Delta_{h'}^{l}\bar{Q}^h f\right\|_\infty$$

which, by Corollary 4.7 can be bounded by

$$C\left(\sum_{|\mathbf{s}|_1=m}\prod_q \|\Delta_h^{s_q} f\|_\infty \sum_{|\mathbf{r}|_1\le l}(h^{-1}h')^{l-|\mathbf{r}|_1}\prod_p \|\Delta_{h'}^{r_p} Q^h f\|_\infty\right) + \left\|\Delta_{h'}^{l}\bar{Q}^h f\right\|_\infty.$$

By the smoothness of f, Lemma 4.2, and the induction hypothesis, we can bound the above quantity by

$$Ch^{\alpha-l}(h')^l + Ch^{\alpha}\|\Delta_{h'}^l Q^h f\|_\infty + \left\|\Delta_{h'}^l \bar{Q}^h f\right\|_\infty,$$

with another constant C.

Utilizing the fact that $\left\|\Delta_{h'}^l \bar{Q}^h f\right\|_\infty = \mathcal{O}\left((h')^l\right)$, we arrive at the estimate

$$\left\|\Delta_{h'}^l Q^h f\right\|_\infty \leq Ch^{\alpha-l}(h')^l + Ch^{\alpha}\|\Delta_{h'}^l Q^h f\|_\infty + \bar{C}(h')^l.$$

Now let h be small enough such that $Ch^\alpha \leq \frac{1}{2}$. Then we have

$$\frac{1}{2}\left\|\Delta_{h'}^l Q^h f\right\|_\infty \leq Ch^{\alpha-l}(h')^l + \bar{C}(h')^l,$$

and this shows the desired assertion. □

We can finally conclude the proof of Theorem 3.1.

Proof of Theorem 3.1. The idea is to use the linear theory which states that

$$\left\|\left(\frac{d}{dx}\right)^l \left(f - \bar{Q}^h f\right)\right\|_\infty = \mathcal{O}(h^{\alpha-l}),$$

together with a suitable estimate for the difference between the linear and the nonlinear approximation procedure, namely we will show that

$$\left\|\left(\frac{d}{dx}\right)^l \left(Q^h f - \bar{Q}^h f\right)\right\|_\infty = \mathcal{O}(h^{\alpha-l}). \tag{39}$$

In order to arrive at (39) we observe that the expression to be estimated can be written as

$$\left\|\left(\frac{d}{dx}\right)^l \left(Q^h f - \bar{Q}^h f\right)\right\|_\infty = \lim_{h'\to 0}(h')^{-l}\left\|\Delta_{h'}^l\left(Q^h f - \bar{Q}^h f\right)\right\|_\infty. \tag{40}$$

By corollary 4.7, the right hand side in (40) can be estimated by a constant times

$$(h')^{-l}\Omega_{m,h}(f)\sum_{i=0}^{l}(h^{-1}h')^{l-i}\Omega_{i,h'}\left(Q^h f\right). \tag{41}$$

By Lemma 4.8 we can estimate

$$\|\Delta_{h'}^r Q^h f\|_\infty = \mathcal{O}((h')^r), \quad r = 0, \ldots, l,$$

and hence

$$\Omega_{i,h'}\left(Q^h f\right) = \mathcal{O}\left((h')^i\right) \tag{42}$$

with an implicit constant independent of h, h'. Furthermore, due to the fact that $f \in C^\alpha$ and Lemma 4.2 we have

$$\Omega_{m,h}(f) = \mathcal{O}(h^\alpha). \tag{43}$$

Putting (42) and (43) into (41), we arrive at the estimate

$$(h')^{-l}\left\|\Delta_{h'}^l\left(Q^h f - \bar{Q}^h f\right)\right\|_\infty = \mathcal{O}(h^{\alpha-l}),$$

with the implicit constant independent of h'. By (40), this implies that

$$\left\|\left(\frac{d}{dx}\right)^l \left(Q^h f - \bar{Q}^h f\right)\right\|_\infty = \mathcal{O}(h^{\alpha-l}).$$

This proves (39) and hence the theorem. □

5 Numerical Experiments

In the present section we conduct some simple numerical experiments which confirm our theoretical findings. We will confine ourselves to the case $M = \mathrm{SO}(2)$, the manifold of orthogonal 2×2 matrices with positive determinant. This is a compact Riemannian manifold and also a Lie group. Its tangent bundle TM is given by $\mathrm{SO}(2) \times \mathfrak{so}_2$, where $\mathfrak{so}_2$ is the Lie algebra of 2×2 skew-symmetric matrices.

The exponential function of M is defined by the matrix exponential

$$\exp_M(p, q) := p\left(\sum_{i=0}^{\infty} \frac{q^i}{i!}\right), \quad (p, q) \in \mathrm{SO}(2) \times \mathfrak{so}_2.$$

its (local) inverse is given by

$$\log_M(p, q) := \sum_{i=0} (-1)^i \frac{(qp^{-1} - I)^i}{i}, \quad p, q \in \mathrm{SO}(2),$$

the usual matrix logarithm.

We compute the Riemannian center of mass

$$\mathrm{av}_M((p(j))_j, (w(j))_j)$$

via the fixed point iteration

$$x_{n+1} = \exp_M\left(x_n, \sum_j w(j) \log_M(x_n, p(j))\right).$$

In [20] it is shown that this iterative procedure converges linearly to the center of mass.

Example 5.1 (Quasiinterpolation with cubic B-splines: smooth data)**.** *In this example we set*

$$\Phi(x) := -\frac{1}{6}B_3(x-1) + \frac{4}{3}B_3(x) - \frac{1}{6}B_3(x+1),$$

with B_3 the cardinal cubic B-spline function [3]. In the linear case this gives a well known quasiinterpolation scheme with polynomial reproduction $m = 4$ and smoothness $s = 2$. We study the approximation rate for the smooth $SO(2)$-valued function

$$f(x) = \begin{pmatrix} \cos(\sin(2x)) & -\sin(\sin(2x)) \\ \sin(\sin(2x)) & \cos(\sin(2x)) \end{pmatrix} \tag{44}$$

and its first two derivatives, see Figure 1. The error is measured in the Frobenius norm. Our experiment confirms the approximation rates predicted by the theory.

Example 5.2 (Quasiinterpolation with cubic B-splines: nonsmooth data)**.** *In this example we study the same approximation procedure as in Example 5.1, this time with the nonsmooth function $f : [0, 1] \to SO(2)$ defined via*

$$f(x) = \begin{pmatrix} \cos(|x - .5|^{1/2}) & -\sin(|x - .5|^{1/2}) \\ \sin(|x - .5|^{1/2}) & \cos(|x - .5|^{1/2}) \end{pmatrix}. \tag{45}$$

The function f is only in $C^{1/2}$ and therefore we only expect an approximation rate of $\frac{1}{2}$ which is observed in Figure 2.

Example 5.3 (High order approximation)**.** *In our final example we consider quasiinterpolation with a quintic B-spline function. With B_5 the cardinal B-spline function of degree 5 we put*

$$\Phi(\cdot) = \frac{13}{240}B_5(x-2) - \frac{7}{15}B_5(x-1) + \frac{73}{40}B_5(x) - \frac{7}{15}B_5(x+1) + \frac{13}{240}B_5(x+2).$$

It can be shown that this function satisfies the assumptions (1–4) with $m = 6$ and $s = 4$. It follows that the approximation error

$$f - Q^h f$$

is expected to be of order h^6. This is confirmed by the numerical experiment in Figure 3, where the smooth function f from Example 5.1 is approximated numerically.

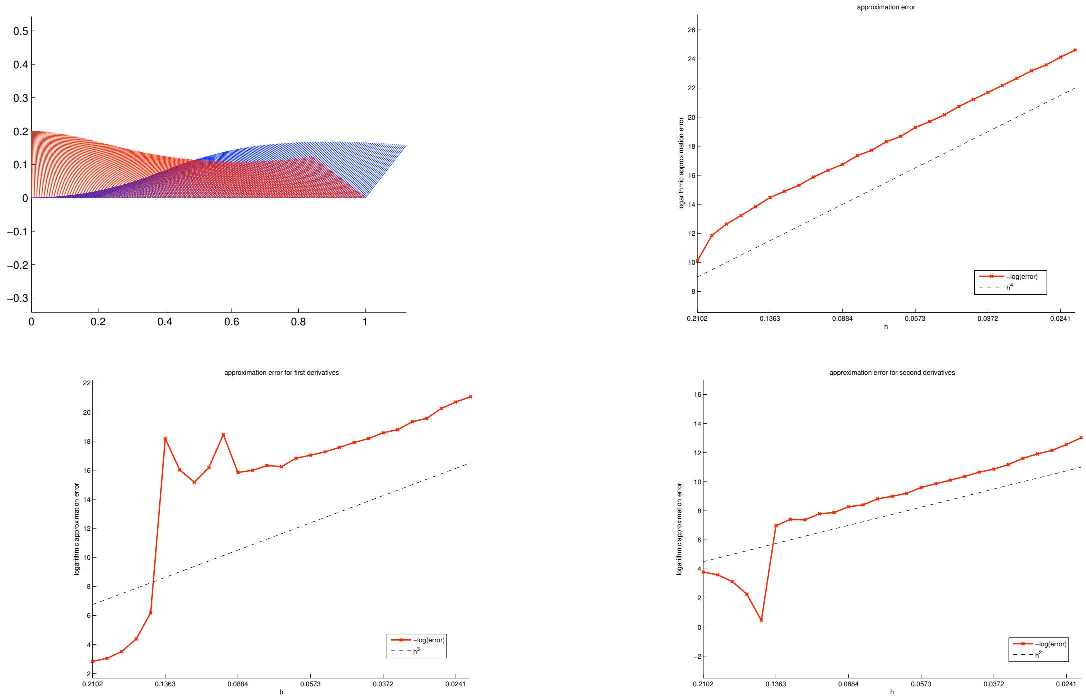

Figure 1: In this example the function $f(x) : [0,1] \to \mathrm{SO}(2)$ given by (44) is approximated by nonlinear cubic B-spline quasiinterpolation. Top left: $f(x)$. Top right: Approximation error. Bottom left: Approximation error of first derivatives. Bottom right: Approximation error of second derivatives.

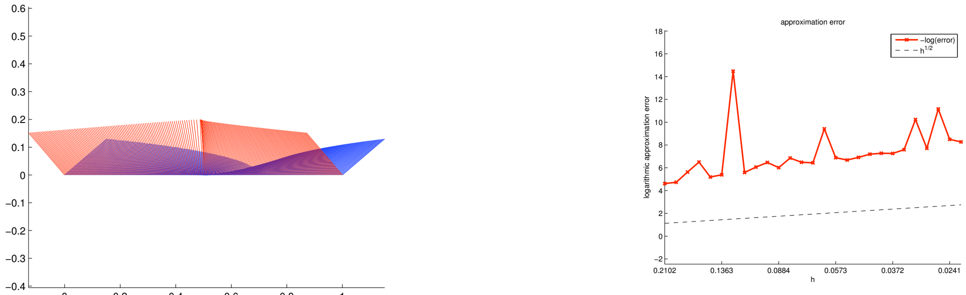

Figure 2: In this example the function $f(x) : [0,1] \to \mathrm{SO}(2)$ given by (45) is approximated by nonlinear cubic B-spline quasiinterpolation. Top left: $f(x)$. Top right: Approximation error.

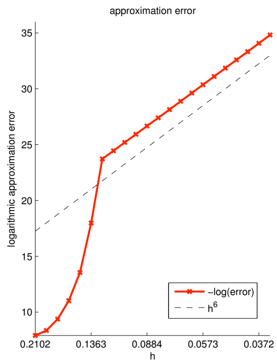

Figure 3: Approximation rate of the smooth function f given in (44) by quintic B-spline quasiinterpolation. The plot suggests the expected approximation order h^6.

References

[1] D. Braess. *Finite Elements*. Cambridge University Press, 2007.

[2] J. Cheeger and D. Gromoll. On the structure of complete manifolds of nonnegative curvature. *Annals of Mathematics*, 96(3):413–443, 1972.

[3] C. Chui. *An Introduction to Wavelets*. Academic Press, 1992.

[4] I. Daubechies, O. Runborg, and W. Sweldens. Normal multiresolution approximation of curves. *Constructive Approximation*, 20(3):399–463, 2004.

[5] C. de Boor and G. Fix. Spline approximation by quasiinterpolants. *Journal of Approximation Theory*, 8:19–45, 1973.

[6] M. D. Denis Le Bihan, J. Mangin, C. Poupon, C. Clark, S. Pappata, N. Molko, and H. Chabriat. Diffusion tensor imaging: concepts and applications. *Journal of Magnetic Resonance Imaging*, 13:534–546, 2001.

[7] R. DeVore and G. Lorentz. *Constructive Approximation*. Springer, 1993.

[8] M. DoCarmo. *Riemannian Geometry*. Birkhäuser, 1992.

[9] P. Dombrowski. On the geometry of the tangent bundle. *Journal für die reine und angewandte Mathematik (Crelles Journal)*, 1962(210):73–88, 1962.

[10] P. Grohs. Smoothness of interpolatory multivariate subdivision in Lie groups. *IMA Journal of Numerical Analysis*, 29:760–772, 2009.

[11] P. Grohs. Approximation order from stability of nonlinear subdivision schemes. *Journal of Approximation Theory*, 162:1085–1094, 2010.

[12] P. Grohs. A general proximity analysis of nonlinear subdivision schemes. *SIAM Journal on Mathematical Analysis*, 42:729–750, 2010.

[13] P. Grohs. Stability of manifold-valued subdivision schemes and multiscale transformations. *Constructive Approximation*, 32:569–596, 2010.

[14] P. Grohs and J. Wallner. Interpolatory wavelets for manifold-valued data. *Applied and Computational Harmonic Analysis*, 27(3):325–333, 2009.

[15] S. Harizanov and P. Oswald. Stability of nonlinear subdivision and multiscale transforms. *Constructive Approximation*, 31:359–393, 2010.

[16] M. Hofer and H. Pottmann. Energy-minimizing splines in manifolds. In *ACM Transactions on Graphics (Proceedings SIGGRAPH)*, volume 23, pages 284–293. ACM, 2004.

[17] A. Iserles, H. Munthe-Kaas, S. Nørsett, and A. Zanna. Lie-group methods. *Acta Numerica 2000*, 9(1):215–365, 2000.

[18] V. Ivancevic. Symplectic rotational geometry in human biomechanics. *SIAM Review*, 46(3):455–474, 2004.

[19] E. Kappos. Natural metrics on tangent bundles. Master's thesis, Lund University, Sweden, 2001.

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

[21] L. Noakes, G. Heinzinger, and B. Paden. Cubic splines on curved spaces. *IMA Journal of Mathematical Control and Information*, 6(4):465, 1989.

[22] X. Pennec, P. Fillard, and N. Ayache. A Riemannian framework for tensor computing. *International Journal of Computer Vision*, 66(1):41–66, 2006.

[23] I. U. Rahman, I. Drori, V. C. Stodden, D. Donoho, and P. Schröder. Multiscale representations for manifold-valued data. *Multiscale modeling and Simulation*, 4(4):1201–1232, 2006.

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

[25] O. Sander. Geodesic finite elements on simplicial grids. 2011. preprint.

[26] S. Sasaki. On the differential geometry of tangent bundles of Riemannian manifolds. *Tohoku Mathematical Journal*, 10(3):338–354, 1958.

[27] M. Spivak. *A comprehensive introduction to differential geometry*, volume 1. Publish or Perish, 1979.

[28] J. Wallner and N. Dyn. Convergence and C^1 analysis of subdivision schemes on manifolds by proximity. *Computer Aided Geometric Design*, 22:593–622, 2005.

[29] J. Wallner, E. Navayazdani, and A. Weinmann. Convergence and smoothness analysis of subdivision rules in riemannian and symmetric spaces. *Advances in Computational Mathematics*, 2010.

[30] J. Wallner and H. Pottmann. Intrinsic subdivision with smooth limits for graphics and animation. *ACM Transactions on Graphics*, 25:356–274, 2006.

[31] A. Weinmann. Nonlinear subdivision schemes on irregular meshes. *Constructive Approximation*, 31:395–415, 2010.

[32] G. Xie and T. Yu. Smoothness equivalence properties of interpolatory Lie group subdivision schemes. *IMA Journal of Numerical Analysis*, 2008.

[33] G. Xie and T. P.-Y. Yu. Approximation order equivalence properties of manifold-valued data subdivision schemes. Manuscript, submitted for publication, April 2008.

[34] G. Xie and T. P.-Y. Yu. Smoothness equivalence properties of general manifold-valued data subdivision schemes. *Multiscale Modeling and Simulation*, 7(3):1073–1100, 2008.

[35] K. Yano and S. Ishihara. *Tangent and Cotangent Bundles: Differential Geometry*, volume 16. Marcel Dekker, 1973.

A Appendix

Lemma A.1. *We have*

$$\chi_{[-N,N]}(h^{-1}x-i)v(Q^h(x),f(hj)) = \mathcal{O}\left(\|\Delta_h f(\cdot)\|_\infty\right).$$

Proof. Let us assume for simplicity that M is embedded in Euclidean space. First note that

$$\log_M(x,y) = \log_M(x,x) + \mathcal{O}(|x-y|) = \mathcal{O}(|x-y|) \tag{46}$$

since $v(x,x)=0$. By [20, 1.5.1] it holds that

$$|f(hj)-Q^h f(x)| \lesssim \mathfrak{d}\left(f(hj),Q^h f(x)\right) \lesssim \left|\sum_{l\in\mathbb{Z}} \Phi(h^{-1}x-l)\log_M(f(hl),f(hi))\right|. \tag{47}$$

This expression is only nonzero if

$$h^{-1}x-l\in[-N,N].$$

Further, by assumption we have

$$h^{-1}x-i\in[-N,N],$$

which implies that $|hl-hi|\lesssim h$. Therefore, by (46) and (47) we get

$$\chi_{[-N,N]}(h^{-1}x-i)|f(hj)-Q^h f(x)| = \mathcal{O}\left(\|\Delta_h f\|_\infty\right).$$

The final estimate follows from noting that

$$\left|v\left(f(hj),Q^h f(x)\right)\right| \lesssim |f(hj)-Q^h f(x)|,$$

which can be shown in the same fashion as (46). □

Lemma A.2. *Assume that $g:\mathbb{R}^k\to\mathbb{R}$ is a smooth function and $\mathbf{f}=(f_1,\dots,f_k):\mathbb{R}\to\mathbb{R}^k$. Then*

$$\Delta^l_{h'} g\left(\mathbf{f}(\cdot)\right) = \mathcal{O}\left(\sum_{|\mathbf{r}|_1=l}\prod_q \Omega_{r_q,h'}(f_q)\right) \tag{48}$$

Proof. We use Taylor expansion of g at $(f_1(x),\dots,f_k(x))$:

$$\begin{aligned} g\left(\mathbf{f}(x+y)\right) = \sum_{i=1}^{l-1}\sum_{\mathbf{s}\in\{1,\dots,k\}^i} d^i_{\mathbf{s}} g\big|_{(\mathbf{f}(x))}\left(f_{s_1}(x+y)-f_{s_1}(x),\dots,f_{s_i}(x+y)-f_{s_i}(x)\right)\\ + \sum_{\mathbf{s}\in\{1,\dots,k\}^l} d^l_{\mathbf{s}} g\big|_{\xi(\mathbf{f}(x+y))}\left(f_{s_1}(x+y)-f_{s_1}(x),\dots,f_{s_i}(x+y)-f_{s_i}(x)\right). \end{aligned}$$

Here the expression $d^i_{\mathbf{s}}g$ denotes the partial derivative $\partial_{s_1}\dots\partial_{s_i}g$. The last term in the above expression is clearly of order

$$\mathcal{O}\left(\sum_{\mathbf{s}\in\{1,\dots,k\}^l}\prod_q \|\Delta_{h'} f_{s_q}\|^l_\infty\right) \quad \text{for} \quad |y|\le lh',$$

and therefore we can estimate

$$\begin{aligned} \Delta^l_{h'} g\left(\mathbf{f}(x)\right) = \sum_{i=1}^{l-1}\sum_{\mathbf{s}\in\{1,\dots,k\}^i} \Delta^l_{h'} d^i_{\mathbf{s}} g\big|_{\mathbf{f}}\left(f_{s_1}(x+y)-f_{s_1}(x),\dots,f_{s_i}(x+y)-f_{s_i}(x)\right)\\ + \mathcal{O}\left(\sum_{\mathbf{s}\in\{1,\dots,k\}^l}\prod_q \|\Delta_{h'} f_{s_q}\|^l_\infty\right). \end{aligned} \tag{49}$$

Each of the terms

$$\Delta_{h'}^{l} d_{\mathbf{s}}^{i} g\big|_{\mathbf{f}(\mathbf{x})} \left(f_{s_1}(x+y) - f_{s_1}(x), \ldots, f_{s_i}(x+y) - f_{s_i}(x)\right)$$

can be expressed as a finite linear combination of terms of the form

$$d_{\mathbf{s}}^{i} g\big|_{\mathbf{f}(x)} \left(\Delta_{h'}^{l_1} f_{s_1}(x_1 + \cdot) - f_{s_1}(x_1), \ldots, \Delta_{h'}^{l_i} f_{s_i}(x_i + \cdot) - f_{s_i}(x_i)\right)$$

with $l_1 + \cdots + l_i = l$ and some $(x_1, \ldots, x_i) \in \mathbb{R}^i$. This gives the final bound. $\square$

Lemma A.3. *We have the estimate*

$$\Delta_{h'}^{l_1} \delta^{l_2} v\left(Q^h(x), f(hj)\right) = \mathcal{O}\left(\Omega_{l_1,h'}(Q^h f) \Omega_{l_2,h}(f)\right). \tag{50}$$

Proof. Again we use Taylor expansion of the bivariate function $v(x,y)$ together with arguments akin to the proof of Lemma A.2. Put $\bar{j} = \lfloor h^{-1}x - N \rfloor$. Then $h\bar{j} - hj = \mathcal{O}(h)$ for all j such that $v\left(Q^h f(x), f(hj)\right) \neq 0$. We have

$$v\left(Q^h f(x), f(hj)\right) = \sum_{i=0}^{l_2-1} d_2^{(i)} v\big|_{(Q^h f(x), f(h\bar{j}))} \left([f(hj) - f(h\bar{j})]^i\right) \\ + R^{(l_2)}\left(Q^h f(x), f(hj)\right)\left([f(hj) - f(h\bar{j})]^i\right).$$

For any i we can write

$$\delta^{l_2} \Delta_{h'}^{l_1} d_2^{(i)} v\big|_{(Q^h f(x), f(h\bar{j}))} \left([f(hj) - f(h\bar{j})]^i\right)$$

as a finite linear combination of terms of the form

$$\Delta_{h'}^{l_1} d_2^{(i)} v\big|_{(Q^h f(x), f(h\bar{j}))} \left(\delta^{\mathbf{s}} [f(hj) - f(h\bar{j})]^i\right)$$

with $|\mathbf{s}|_1 = l_2$. This expression can be bounded by

$$\left\|\Delta_{h'}^{l_1} d_2^{(i)} v\big|_{(Q^h f(x), f(h\bar{j}))}\right\| \sum_{|\mathbf{s}|_1 = l_2} \prod_q \|\Delta_h^{s_q} f\|_\infty.$$

Furthermore, by Lemma A.2 we have

$$\left\|\Delta_{h'}^{l_1} d_2^{(i)} v\big|_{(Q^h f(x), f(h\bar{j}))}\right\| = \mathcal{O}\left(\sum_{|\mathbf{r}|_1 = l_1} \prod_p \|\Delta^{r_p} Q^h f\|_\infty\right).$$

To handle the residual term we note that we can bound the expression

$$\delta^{l_2} \Delta_{h'}^{l_1} R^{(l_2)}\left(Q^h f(x), f(hj)\right)\left([f(hj) - f(h\bar{j})]^{l_2}\right)$$

by terms of the form

$$\Delta_{h'}^{l_1} R^{(l_2)}\left(Q^h f(x), f(hj)\right)\left([f(hj) - f(h\bar{j})]^{l_2}\right),$$

which can in turn be bounded by

$$\left\|\Delta_{h'}^{l_1} R^{(l_2)}\left(Q^h f(x), f(hj)\right)\right\| |f(hj) - f(h\bar{j})|^{l_2} = \mathcal{O}\left(\sum_{|\mathbf{r}|_1 = l_1} \prod_p \|\Delta^{r_p} Q^h f\|_\infty \|\Delta_h f\|_\infty^{l_2}\right).$$

We have used Lemma A.2 and the fact that $hj - h\bar{j} = \mathcal{O}(h)$. Summing up these estimates gives the desired result. $\square$

Research Reports

No.	Authors/Title
11-56	*P. Grohs* Finite elements of arbitrary order and quasiinterpolation for data in Riemannian manifolds
11-55	*P. Grohs* Bandlimited shearlet-type frames with nice duals
11-54	*A. Kunoth and Ch. Schwab* Analytic regularity and GPC approximation for control problems constrained by linear parametric elliptic and parabolic PDEs
11-53	*Ch. Schwab and S.A. Tokareva* High order approximation of probabilistic shock profiles in hyperbolic conservation laws with uncertain initial data
11-52	*F.Y. Kuo, Ch. Schwab and I.H. Sloan* Quasi-Monte Carlo finite element methods for a class of elliptic partial differential equations with random coefficients
11-51	*A. Chernov and Ch. Schwab* First order k-th moment finite element analysis of nonlinear operator equations with stochastic data
11-50	*E. Fonn* Shearlet Galerkin for transport equations: implementation and stability
11-49	*M.J. Castro, U.S. Fjordholm, S. Mishra and C. Parés* Entropy conservative and entropy stable schemes for non-conservative hyperbolic systems
11-48	*D. Kressner and Ch. Tobler* Preconditioned low-rank methods for high-dimensional elliptic PDE eigenvalue problems
11-47	*H. Heumann and R. Hiptmair* Convergence of lowest order semi-Lagrangian schemes
11-46	*K. Grella and Ch. Schwab* Sparse discrete ordinates method in radiative transfer
11-45	*X. Claeys and R. Hiptmair* Boundary integral formulation of the first kind for acoustic scattering by composite structures
11-44	*A. Chkifa, A. Cohen, R. DeVore and Ch. Schwab* Sparse adaptive Taylor approximation algorithms for parametric and stochastic elliptic PDEs