SMOOTHNESS EQUIVALENCE PROPERTIES OF MANIFOLD-VALUED DATA SUBDIVISION SCHEMES BASED ON THE PROJECTION APPROACH

*GANG XIE * AND THOMAS P.-Y. YU †*

Abstract:

Interpolation of manifold-valued data is a fundamental problem which has applications in many fields. The linear subdivision method is an efficient and well-studied method for interpolating or approximating real-valued data in a multiresolution fashion. A natural way to apply a linear subdivision scheme $\overline{S}$ to interpolate manifold-valued data is to first embed the manifold at hand to an Euclidean space and construct a projection operator P that maps points from the ambient space to a closest point on the embedded surface, and then consider the *nonlinear* subdivision operator $S := P \circ \overline{S}$. When applied to symmetric spaces such as S^{n-1}, $SO(n)$, $SL(n)$, $SE(n)$, $G(n,k)$ the projection method can also be carried out in such a way that the resulted schemes enjoy natural coordinate invariance properties and robust numerical implementations.

Despite such nice features, the mathematical analysis of such nonlinear subdivision schemes is at its infancy. In this article, we attack the so-called Smoothness Equivalence Conjecture, which asserts that the smoothness property of S is exactly the same as that of $\overline{S}$. We show that in the cases of S^{n-1}, $SO(n)$ and related manifolds, we have a proximity condition of the form:

$$\left|(S-\overline{S})y\right|_\infty \lesssim \sum_{i=1}^{p-1} |\Delta^i y|_\infty |\Delta^{p-i} y|_\infty,$$

where p is the accuracy order of $\overline{S}$. Armed with this proximity condition and other known approximation theoretic results, we can establish the result that the Hölder smoothness exponent of S is always as high as that of $\overline{S}$ – no matter how high the latter is.

Acknowledgments. We thank Tom Duchamp for many stimulating discussions. The work of this research was partially supported by the National Science Foundation grants CCF 9984501 (CAREER Award) and DMS 0512673.

Keywords. Linear subdivision schemes, Nonlinear subdivision schemes, Riemannian manifold, Lie groups, Sphere, $SO(n)$, Smoothness, Approximation order, Closest Point Projection, Interpolation, Singular Value Decomposition

Mathematics Subject Classification. 26A15, 26A16, 26A18, 41A05, 42C40

*Department of Mathematics, Drexel University, 3141 Chestnut Street, 206 Korman Center, Philadelphia, PA 19104, U.S.A.. Email: `gangxie2006@gmail.com`

† Department of Mathematics, Drexel University, 3141 Chestnut Street, 206 Korman Center, Philadelphia, PA 19104, U.S.A.. Email: `yut@drexel.edu`. URL: `http://www.math.drexel.edu/~tyu`.

1. Introduction. Given a smooth manifold $\mathcal{M}$ and a sequence of points $p_i \in \mathcal{M}$ on the manifold (called a control polygon), a fundamental problem is to find a smooth curve that either interpolates or approximates the control polygon. A typical application in robotics and computer vision is rigid body motion interpolation, in which the manifold is the Lie group $\mathcal{M} = SE(3)$ [2]. In the curve design problem on the sphere described in [27], we have $\mathcal{M} = \mathbb{S}^2$. In the 'grand tour' method for visualizing p-dimensional data points based on ortho-projection of the points to 2 dimensional subspaces [1], we have the Grassmannian manifold $\mathcal{M} = G(p, 2)$. Time series that take values on a Grassmannian manifold $G(n, k)$ also arise in array signal processing. There is also a lot of interest in numerical analysis for interpolation in Lie groups, especially in the geometric methods for the numerical solution of ODEs, see, e.g., [20, 18, 17].

Some, if not many, of the existing methods for the manifold interpolation problem appear to be quite specific to the manifold at hand. On the other hand, in certain applications it is quite desirable if the curve can be given in various level-of-details, in this case subdivision based methods would be very attractive.

In recent papers [30, 24, 31], various general subdivision methods are proposed which can be used to subdivide data that takes values on a manifold (or data that obeys some other forms of nonlinear constraints.) A common feature of these methods is that they are all based on adapting a linear subdivision scheme to a manifold, resulting in nonlinear algorithms that are easy to implement but difficult to analyze. Since in each case the method is based on taking a *linear* subdivision scheme and applying it to subdivision of data obeying *nonlinear* constraints, we refer to such a method as a **linearization** method. Here, we discuss three classes of linearization methods:

- **Tangent plane approach**: In this approach, the basic operation is that for any given $p \in \mathcal{M}$, there is a map f_p with inverse f_p^{-1} that map points in a neighborhood of p back and forth between the manifold and the tangent plane $T_p\mathcal{M}$. Then a subdivided point is obtained from a group of points in the coarser scale by first mapping these points to the tangent plane based at one of these points using the corresponding map f_p, then applying the linear subdivision rule in $T_p\mathcal{M}$ (a linear space) and mapping the subdivided point in the tangent plane back to the manifold using f_p^{-1}. The specific f_p and f_p^{-1} proposed in [24] are the logarithmic and exponential maps (in either the setting of Riemannian manifold or Lie group, see, e.g., [3, 11]); we refer to this linearization method as the **Log-Exp scheme**.
- **Factorization-geodesic scheme**: The mask of any linear subdivision scheme can be factorized (in a non-unique way) into a number of two point weighted averages [30, Theorem 1]. On any Riemannian manifold, the notion of geodesic is well-defined, so one can perform subdivision based on replacing the weighted averages by 'weighted geodesic averages'. See [30] for details.
- **Projection approach**: In this approach $\mathcal{M}$ is supposed to be immersed or embedded in an Euclidean space $\mathbb{R}^n$; as such, we view $\mathcal{M}$ as a subset of $\mathbb{R}^n$. Also a projection operator that maps points in a neighborhood of $\mathcal{M}$ onto $\mathcal{M}$ is chosen. A natural example of such an operator would be the one that maps a $p \in \mathbb{R}^n$ to a point in $\mathcal{M}$ *closest* to p. We refer to this as the **closest point projection scheme**. A subdivision step is based on first applying the linear subdivision rule in $\mathbb{R}^n$, resulting in points that are typically not in $\mathcal{M}$, then followed by applying the projection operator to force these subdivided points back to $\mathcal{M}$.

Any of the above linearization methods can be used in conjunction with an arbitrary linear subdivision scheme.

The Log-Exp scheme and the factorization-geodesic schemes are both based on geodesics and are *intrinsic* in nature, whereas the projection approach is *extrinsic* in nature: In the projection approach, one may have two isometric embeddings of a given Riemannian manifold $(\mathcal{M}, g)$ to two Euclidean spaces, and the (closest point, say) projection scheme may result in two different curves on $\mathcal{M}$ for the same set of initial points on $\mathcal{M}$. From a practical point of view, the projection method is probably the most natural to use for manifolds with a "natural" embedding [1] into an Euclidean space, e.g. $\mathbb{S}^n \hookrightarrow \mathbb{R}^{n+1}$, $SO(n) \hookrightarrow SL(n) \hookrightarrow GL(n) \hookrightarrow \mathbb{R}^{n\times n}$, or $SE(n) \hookrightarrow GL(n+1) \hookrightarrow \mathbb{R}^{(n+1)\times(n+1)}$.

1.1. Linear subdivision schemes. We recall in this section some of the basic definitions, notions and notations related to linear subdivision schemes. We keep the exposition to the minimum as there are plenty of references on this topic, see, for example, [4, 14, 25, 6, 7] and the references therein.

[1]If G is a group that acts transitiviely on $\mathcal{M}$, some authors call an embedding $\Phi : \mathcal{M} \to \mathbb{R}^n$ *natural* if there is a smooth group homomorphism $\rho : G \to SE(n)$ such that $\phi(g \cdot x) = \rho(g) \cdot \Phi(x)$ for all $x \in \mathcal{M}$ and $g \in G$.

In the simplest setting, a linear stationary subdivision scheme is defined by a linear operator S on $\ell(\mathbb{Z}) := \{x \mid x : \mathbb{Z} \to \mathbb{R}\}$ of the form

$$(Sx)(2k) = \sum_{i\in\mathbb{Z}} x(i)\, a_e(k-i), \quad (Sx)(2k+1) = \sum_{i\in\mathbb{Z}} x(i)\, a_o(k-i), \tag{1.1}$$

where a_e and a_o are two finitely supported real-valued sequences such that $\sum_i a_e(i) = \sum_i a_o(i) = 1$. This operator is usually written by analysts in the following more compact form: $(Sx)(k) = \sum_{i\in\mathbb{Z}} x(i)\, \mathbf{a}(k-2i)$. Here, $\mathbf{a}$ is called the **mask** of the subdivision scheme and can be easily assembled from the a_o and a_e above.

If $a_e = \delta_0$ (the Kronecker sequence), then we say S is **interpolatory**. Interpolatory subdivision schemes are first studied in [12, 9].

A subdivision operator S is meant to be *iterated*. Moreover, for any initial sequence $v : \mathbb{Z} \to \mathbb{R}$, one is supposed to visualize $S^j v$ as a function on the grid $2^{-j}\mathbb{Z}$, as opposed to a function on $\mathbb{Z}$ as our mathematical notation may unduely suggest. We say S is **convergent** if for any $v \in \ell(\mathbb{Z})$, the sequence $f_j := \sum_{k\in\mathbb{Z}} v_{j,k} 1_{[2^{-j}k,\, 2^{-j}(k+1))}$, $j = 0, 1, 2, \ldots$, $v_j := S^j v$, converges uniformly on compact sets to a limit function; we denote this (necessarily unique) limit function by $S^\infty v$. For a convergent subdivision operator S, we define its critical Hölder smoothness (a.k.a. L^∞-Lipschitz smoothness) by

$$s_\infty(S) := \inf_{v\in\ell^\infty} \sup\{\alpha : S^\infty v \in \mathrm{Lip}\,\alpha\}. \tag{1.2}$$

While a subdivision scheme S as defined above operates on scalars, one can apply it in a componentwise fashion to m-vectors; this, in particular, gives a practical curve drawing algorithm with input being a coarse control polygon in $\mathbb{R}^m$ ($m = 2$ or 3). When we later write Sy where y is a sequence of m-vectors, this is to be interpreted as the componentwise application of S to y.

For any (finite or infinite) sequence of m-vectors $y = (y_i)_{i\in\mathcal{I}}$, $y_i \in \mathbb{R}^m$, $\mathcal{I} = \mathbb{Z}$ or $\{1, \ldots, n\}$, we write

$$|y|_\infty := \sup_{i\in\mathcal{I}} \|y_i\|_2$$

and define its difference sequences by $(\Delta y)_i := y_i - y_{i-1}$, $(\Delta^k y)_i := (\Delta^{k-1} y)_i - (\Delta^{k-1} y)_{i-1}$, $k > 1$, where i ranges through the appropriate set of indices when $\mathcal{I}$ is finite. We denote by $\ell^\infty(\mathbb{Z} \to \mathbb{R}^m)$, or simply ℓ^∞ when there is no source of confusion, the space of all sequences $y : \mathbb{Z} \to \mathbb{R}^m$ such that $|y|_\infty$ is finite.

An interpolatory subdivision scheme S is said to have approximation order p ($\in \mathbb{Z}_+$) if for any bounded C^p function $f : \mathbb{R} \to \mathbb{R}$, the interpolant of f on the grid $h\mathbb{Z}$ defined by

$$f_h := (S^\infty v)(\cdot/h), \quad v_k = f(kh), \tag{1.3}$$

satisfies

$$\|f - f_h\|_\infty = O(h^p), \quad h \to 0.$$

For an interpolatory scheme S, it is not hard to show that, based on simple twists of the arguments already presented by Dubuc [12], that S has approximation order p if and only if S reproduces the polynomial space Π_{p-1}, i.e. $S(p\,|_{\mathbb{Z}}) = p\,|_{\frac{1}{2}\mathbb{Z}}$ for all $p \in \Pi_{p-1}$. This condition imposes a set of linear conditions on the mask of S.

If $s_\infty(S) > p$, then S must have approximation order $p+1$; the converse is far from the truth: an interpolatory scheme can have an arbitrarily high approximation order but an arbitrarily low smoothness.

1.2. Smoothness equivalence. Intuitively, both the Log-Exp scheme and the factorization-geodesic scheme try to use geodesics to mimic weighted averages (1.1) on a manifold locally. Hence, one may expect that the nonlinearities presented in the resulted schemes are rather weak *at fine subdivision levels.* This seems to be true in the sense that the nonlinearity at fine scales is too weak to affect either the smoothness or the approximation order. We discuss smoothness in this paper and approximation order in the companion paper [35].

From many computational experiments both the Log-Exp scheme and the factorization-geodesic scheme are observed empirically to produce limit paths on the manifold with Hölder regularity exactly the same as that produced by the underlying linear scheme [24, 36]. This is the so-called **smoothness equivalence** property, and is conjectured to hold for both the Log-Exp scheme and the factorization-geodesic scheme when applied in conjunction to any linear subdivision scheme on any C^∞ Riemannian manifold.

The smoothness equivalence conjecture is true for both Log-Exp and geodesic-factorization if $\mathcal{M}$ is an Abelian Lie group (e.g., the n-torus $\mathbb{T}^n$) viewed as a Riemannian manifold with a bi-invariant metric. (Caution: For $\mathbb{T}^n$, we are referring to the so-called "flat torus"; when $n = 2$, it is diffeomorphic, but not isometric, to the bagel-like torus as drawn in $\mathbb{R}^3$ with the induced Riemannian metric from $\mathbb{R}^3$. We still believe that the conjecture is true for the "bagel-like torus", just that the proof is not going to be as trivial.)

The goal of this paper is to study the smoothness equivalence properties of the closest point projection scheme, which is *not* based on geodesics. Our empirical experiments suggest that the corresponding smoothness equivalence conjecture for closest point projection scheme also holds true. We note that, even in the case of the circle $\mathbb{S}^1$ ($\cong \mathbb{T}^1$) the proof is, unlike that for Log-Exp or geodesic-factorization, already not trivial.

1.3. Warming up on the circle. As a warmup, we consider here the case when $\mathcal{M} = \mathbb{S}^1 = \{x \in \mathbb{R}^2 : \|x\|_2 = 1\}$ and the linear subdivision schemes are Deslauriers-Dubuc's $2L$-point interpolatory subdivision schemes. Here the metric on $\mathbb{S}^1$ is induced by the standard Euclidean metric of $\mathbb{R}^2$. In this case the geodesic distance is just arc length on the circle, and the Log-Exp and geodesic-factorization schemes are essentially the same. If $p_{0,k} = [\cos(\theta_{0,k}), \sin(\theta_{0,k})]^T \in \mathbb{S}^1$, and assume for simplicity that all $\theta_{0,k} \in (-\delta, \delta)$ for a not so big δ, then either scheme generates the subdivided points $p_{j,k} = [\cos(\theta_{j,k}), \sin(\theta_{j,k})]^T$ based on *linearly* subdividing the *angles* $\theta_{j,k}$ using the linear rule:

$$\theta_{j+1,2k} = \theta_{j,k}, \quad \theta_{j+1,2k+1} = \textstyle\sum_{i=-L+1}^{L} w_i \theta_{j,k+i}. \tag{1.4}$$

Here the w_i's are the mask entries of the $2L$-point Deslauriers-Dubuc scheme; e.g. $w_0 = w_1 = 9/16$, $w_{-1} = w_2 = -1/16$ if $L = 2$. In this case, the Log-Exp scheme produces a limit function $f : \mathbb{R} \to \mathbb{S}^1$ of the form $f(t) = [\cos(\theta(t)), \sin(\theta(t))]^T$ where $\theta(t)$ is the scalar-valued limit function obtained by applying the Deslauriers-Dubuc scheme to the initial *scalar* data $(\theta_{0,k})_{k\in\mathbb{Z}}$. So smoothness equivalence between the linear subdivision scheme and the associated "nonlinear" scheme is readily clear.

On the other hand, the nonlinear subdivision scheme based on closest point projection gives: $p_{j+1,2k} = p_{j,k}$, $p_{j+1,2k+1} = \sum_{i=-L+1}^{L} w_i p_{j,k+i} / \| \sum_{i=-L+1}^{L} w_i p_{j,k+i}\|_2$, or

$$\theta_{j+1,2k} = \theta_{j,k}, \quad \theta_{j+1,2k+1} = \arctan \frac{\sum_{i=-L+1}^{L} w_i \sin(\theta_{j,k+i})}{\sum_{i=-L+1}^{L} w_i \cos(\theta_{j,k+i})}.$$

The smoothness equivalence property is far less obvious in this case. Using Taylor's theorem, one can show that

$$\arctan \frac{\sum_{i=-L+1}^{L} w_i \sin(\theta_{j,k+i})}{\sum_{i=-L+1}^{L} w_i \cos(\theta_{j,k+i})} = \sum_{i=-L+1}^{L} w_i \theta_{j,k+i} + O\left(|\Delta\theta_j|_\infty^3\right). \tag{1.5}$$

Using the proximity and perturbation arguments detailed in this article, the above estimate allows us to conclude that the closest point projection scheme on $\mathbb{S}^1$ shares the same Hölder regularity of the underlying linear subdivision scheme *if the linear scheme has a critical Hölder regularity* $\leqslant 3$. However, when $L \geqslant 4$, the Deslauriers-Dubuc schemes have critical Hölder regularity > 3, and the above estimate is insufficient for proving smoothness equivalence. We can still use (1.5) to conclude that the nonlinear scheme has critical Hölder smoothness no less than 3, but we can no longer conclude that the nonlinear scheme is as smooth as the linear scheme.

We note that (1.5) is a kind of proximity condition between a linear and a nonlinear scheme, similar to, but not the same, as the proximity results in [30]. Using either a general result in [30] or the first part of Theorem 3.7, we have:

$$\frac{\sum_{i=-L+1}^{L} w_i p_{j,k+i}}{\| \sum_{i=-L+1}^{L} w_i p_{j,k+i}\|} = \sum_{i=-L+1}^{L} w_i p_{j,k+i} + O\left(|\Delta p_j|_\infty^2\right). \tag{1.6}$$

This is yet another proximity condition. While (1.6) may seem more natural than (1.5), the former is not as powerful as the latter: (1.6) only allows us to conclude smoothness equivalence when the linear scheme has smoothness $\leqslant 2$.

Regardless, neither (1.6) nor (1.5) is nearly good enough to prove the smoothness equivalence conjecture of the $\mathbb{S}^1$-closest point projection scheme based on Deslauriers-Dubuc schemes for an *arbitrary* L, as it is well-known that the smoothness of Deslauriers-Dubuc schemes grows unboundedly as $L \to \infty$.

1.4. Contributions and Organization of this paper. This paper aims to attack the *arbitrary degree* smoothness equivalence conjectures as described in previous sections. We prove that smoothness equivalence holds when the closest point projection scheme is applied in conjunction with an *arbitrary* linear interpolatory subdivision schemes to data that takes values on $\mathbb{S}^n$, $SO(n)$ and their direct products.

These results stand in contrast to the recent low degree smoothness equivalence results in [30, 29, 36]. Related low degree smoothness equivalence results of nonlinear subdivision schemes arising from other applications can be found in [33, 32, 8].

The idea – and limitation – of [30, 36] are that by bounding the nonlinearity using a quadratic term, one can show smoothness equivalence when – and only when – the linear scheme has smoothness $\leqslant 2$. Wallner's preprint [29] shows how one may go from "$\leqslant 2$" to "$\leqslant 3$" for certain linear schemes used in conjunction with the factorization-geodesic and the projection schemes.

The key discovery reported in this paper is that the polynomial reproducibility property of a linear interpolatory subdivision scheme $\overline{S}$ can lend itself to a useful 'factorization' of the nonlinearity (Section 3.1): by bounding the nonlinearity of the $\mathbb{S}^n$-closest point projection scheme expressed in *specific* quadratic terms that involve the interpolatory subdivision mask (see (3.7)-(3.9)), we can somehow lift the quadratic proximity bound $|(S - \overline{S})y|_\infty = O(|\Delta y|_\infty^2)$ into a higher order proximity bound

$$(*) \qquad |(S - \overline{S})y|_\infty \leqslant B_p \sum_{i=1}^{p-1} |\Delta^i y|_\infty |\Delta^{p-i} y|_\infty$$

when $\overline{S}$ reproduces Π_{p-1}.

Once $(*)$ is proved, we have a strong enough proximity condition that allows us, when combined with the perturbation theorem [8, Theorem 3.3] and a 'bootstrapping' argument that relies on the interpolatory property of the subdivision scheme, to prove the desired arbitrary degree smoothness equivalence result. See Section 3.2.

In Section 3.3, we show how to relax the closest point projection operator to a near-closest projection operation without jeopardizing the arbitrary degree smoothness equivalence property.

In Section 4, we show how to extend the main smoothness equivalence result to $SO(n)$, $SE(n)$ and related manifolds.

We discuss in Section 5 possible extensions of our results.

In Section 2, we streamline some of the general proximity results in [30, 29]. The proofs of a number of lemmas and theorems are recorded in the Appendix.

2. General Proximity Results. The following theorem is a restatement of [29, Theorem 2]. It is an extension of [30, Theorem 2].

THEOREM 2.1. *Let $T_1, T_2 : \ell^\infty \to \ell^\infty$. Suppose that there exist $\delta, A > 0, \mu \in (0,1)$ and $\alpha > 1$ such that for any $p \in \ell^\infty$ with $|\Delta p|_\infty < \delta$, we have*

$$|\Delta T_1^j p|_\infty \leqslant \mu^j |\Delta p|_\infty, \quad \forall j \in \mathbb{N}, \qquad (2.1)$$
$$|T_1 p - T_2 p|_\infty \leqslant A |\Delta p|_\infty^\alpha.$$

Then for any $\epsilon > 0$ there exists $\delta' > 0$ such that when $|\Delta p|_\infty < \delta'$,

$$|\Delta T_2^j p|_\infty \leqslant (\mu + \epsilon)^j |\Delta p|_\infty, \quad \forall j \in \mathbb{N}.$$

Our goal is to extend the above theorem to Theorem 2.4. For this purpose, we need the next two lemmas.

LEMMA 2.2. *Let $T_1, T_2 : \ell^\infty \to \ell^\infty$. Suppose that there exist $C, \delta, A > 0$ and $\alpha > 1$ such that for all $p \in \ell^\infty$ with $|\Delta p|_\infty < \delta$, we have*

$$|\Delta T_1 p|_\infty \leqslant C |\Delta p|_\infty, \qquad (2.2)$$
$$|T_1 p - T_2 p|_\infty \leqslant A |\Delta p|_\infty^\alpha. \qquad (2.3)$$

Then there exist $C' > 0$ and $\delta' > 0$ such that when $|\Delta p|_\infty < \delta'$,

$$|\Delta T_2 p|_\infty \leqslant C'|\Delta p|_\infty. \tag{2.4}$$

Proof: See Appendix A.1.

The following lemma is an extension of [29, Lemma 2].

LEMMA 2.3. *Let $T_1, T_2 : \ell^\infty \to \ell^\infty$. Suppose that there exist $C, A, \delta > 0$ and $\alpha > 1$ such that for all $p \in \ell^\infty$ with $|\Delta p|_\infty < \delta$, we have*

$$|\Delta T_1 p|_\infty \leqslant C|\Delta p|_\infty, \tag{2.5}$$

$$|T_1 p - T_2 p|_\infty \leqslant A|\Delta p|_\infty^\alpha. \tag{2.6}$$

If T_1 is bounded and linear, then for any $j \in \mathbb{N}$, there exist $\delta_j, C_j > 0$ such that when $|\Delta p|_\infty < \delta_j$,

$$|T_1^j p - T_2^j p|_\infty \leqslant C_j|\Delta p|_\infty^\alpha. \tag{2.7}$$

Proof: See Appendix A.2.

THEOREM 2.4. *Let $T_1, T_2 : \ell^\infty \to \ell^\infty$. Suppose that there exist $C, A, \delta > 0$, $\mu \in (0, 1)$ and $\alpha > 1$ such that: for all $p \in \ell^\infty$ with $|\Delta p|_\infty < \delta$, we have*

$$|\Delta T_1^j p|_\infty \leqslant C\mu^j|\Delta p|_\infty, \quad \forall j \in \mathbb{N}, \tag{2.8}$$

$$|T_1 p - T_2 p|_\infty \leqslant A|\Delta p|_\infty^\alpha. \tag{2.9}$$

If at least one of T_1 and T_2 is bounded and linear, then for any $\epsilon > 0$ there exist $\delta', C' > 0$ such that when $|\Delta p|_\infty < \delta'$,

$$|\Delta T_2^j p|_\infty \leqslant C'(\mu + \epsilon)^j|\Delta p|_\infty, \quad \forall j \in \mathbb{N}.$$

Proof: See Appendix A.3.

REMARK 2.5. In Theorem 2.4, we make the assumption that one of T_1 and T_2 is bounded and linear to accommodate the assumption (2.8) on T_1, which is weaker than (2.1) in Theorem 2.1.

3. Closest Point Projection Scheme for $\mathbb{S}^{m-1}$-valued Data. Let $\overline{S}$ be a linear interpolatory subdivision operator defined by

$$(\overline{S}y)_{2k} = y_k, \quad (\overline{S}y)_{2k+1} = \sum_{i=1}^{n} w_i\, y_{k+i+\ell}, \tag{3.1}$$

where $\ell \in \mathbb{Z}$ is fixed. We also assume that $\overline{S}$ reproduces Π_{p-1}. By linearity and the well-posedness of Lagrange interpolation, it is necessary that $p \leqslant n$, and there is a unique mask with $p = n$. Since we are in the business of proving smoothness equivalence, we assume that $\overline{S}$ has at least some smoothness: $s_\infty(\overline{S}) > 0$. This, in turn, implies that (i) $p \geqslant 1$, (ii):

$$\sum_{i=1}^{n} w_i = 1. \tag{3.2}$$

and (iii) there exists a subdivision operator $\overline{S}^{[1]}$ such that $\overline{S}^{[1]} \circ \Delta = \Delta \circ \overline{S}$. The special cases of Deslauriers-Dubuc schemes [12, 9] correspond to $n = 2L$ (i.e. n is even), $\ell = -L$ and $p = n$.

In the z-transform domain, the polynomial reproduction property is equivalent to the existence of a polynomial $b(z)$ such that

$$\sum_{i=1}^{n} w_i' z^{2i} + z^{2n'+1} = (1+z)^p\, b(z), \tag{3.3}$$

where $w'_i = w_{n+1-i}$ and $n' = n + \ell$. By taking derivatives of both sides of (3.3) and evaluating at $z = -1$, we get the sum rules:

$$\sum_{i=1}^{n} \binom{2i}{\gamma} w'_i = \binom{2n'+1}{\gamma}, \quad \gamma = 0, ..., p-1, \tag{3.4}$$

where $\binom{u}{v}$ is the standard binomial coefficient, i.e.

$$\binom{u}{v} = \begin{cases} \frac{u!}{v!(u-v)!} & \text{if } u \geqslant v \geqslant 0, \\ 0 & \text{otherwise.} \end{cases}$$

Let $\mathbb{S}^{m-1}$ be the unit sphere in $\mathbb{R}^m$, i.e. $\mathbb{S}^{m-1} = \{x \in \mathbb{R}^m : \|x\| = 1\}$. Let $P : \mathbb{R}^m \backslash \{0\} \to \mathbb{S}^{m-1}$ be the closest point projection operator onto the sphere, i.e. $P(x) = x/\|x\|$.

Consider the nonlinear subdivision operator S for $\mathbb{S}^{m-1}$-valued data defined by

$$(Sy)_{2k} = y_k, \quad (Sy)_{2k+1} = P(\sum_{i=1}^{n} w_i y_{k+i+\ell}). \tag{3.5}$$

Remark 3.1. The subdivision operation (3.5) is well-defined as long as $y : \mathbb{Z} \to \mathbb{S}^{m-1}$ is such that $\sum_{i=1}^{n} w_i \, y_{k+i+\ell} \neq 0$ for all $k \in \mathbb{Z}$. We show in Lemma 3.3 that as long as $|\Delta y|_\infty$ is small enough, $S^j y$ is well-defined for all j.

Let $x_1, ..., x_n \in \mathbb{S}^{m-1}$ with $\sum_{i=1}^{n} w_i \, x_i \neq 0$. For any $y \in \mathbb{R}^m \backslash \{0\}$,

$$\Big\|y - P(y)\Big\| = \Big\|y - y/\|y\|\Big\| = \Big|1 - \|y\|\Big| \leqslant \Big|1 - \|y\|^2\Big|. \tag{3.6}$$

Combined with the facts that $x_i \in \mathbb{S}^{m-1}$ and $\sum_i w_i = 1$, we have:

$$\Big\|\sum_{i=1}^{n} w_i x_i - P\Big(\sum_{i=1}^{n} w_i x_i\Big)\Big\| = \Big|1 - \Big\|\sum_{i=1}^{n} w_i x_i\Big\|\Big| \leqslant \Big|1 - \Big\|\sum_{i=1}^{n} w_i x_i\Big\|^2\Big| \tag{3.7}$$

$$= \Big|\sum_{i=1}^{n} w_i \langle x_i, x_i \rangle - \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \langle x_i, x_j \rangle\Big| = \Big|\sum_{i=1}^{n}\sum_{j=1}^{n} c^0_{i,j} \langle x_i, x_j \rangle\Big|, \tag{3.8}$$

where the $n \times n$ matrix $C_0 = (c^0_{i,j})$ is given by

$$c^0_{i,j} := \begin{cases} w_i - w_i^2, & i = j, \\ -w_i w_j, & i \neq j. \end{cases} \tag{3.9}$$

We shall come back to this matrix after a few remarks.

Remark 3.2. Using the identity $\langle x_i, x_j \rangle = 1 - \|x_i - x_j\|^2/2$, we can also rewrite the upper bound (3.7) as

$$\Big\|\sum_{i=1}^{n} w_i x_i - P(\sum_{i=1}^{n} w_i x_i)\Big\| \leqslant \Big|1 - \Big\|\sum_{i=1}^{n} w_i x_i\Big\|^2\Big| = \Big|\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \|x_i - x_j\|^2\Big|. \tag{3.10}$$

Lemma 3.3 (**Well-definedness of $S^j y$**). There exist $\delta^* > 0$ and $K > 0$ such that for any $y : \mathbb{Z} \to \mathbb{S}^{m-1}$ with $\|\Delta y\|_\infty \leqslant \delta^*$, $|\Delta S^j y|_\infty \leqslant K\delta^*$ and $S^j y$ is well-defined for all $j = 1, 2, \ldots$. **Proof:** The proof is easy and we just give the main idea: (3.10) says that as long as $|\Delta y|_\infty$ is small enough, $(Sy)_k$ stays away from the origin for every k. Since we assume $s_\infty(\overline{S}) > 0$, $|\Delta \overline{S}^j y|_\infty = O(2^{-j\nu})$ for any $0 < \nu < \min(1, s_\infty(\overline{S}))$, with (hidden constant) $\propto |\Delta y|_\infty$. Again by (3.10), S and $\overline{S}$ satisfy the proximity condition (2.9) in Theorem 2.4 (with $\alpha = 2$); hence by the theorem, $|\Delta S^j y|_\infty = O(2^{-j\nu^-})$ for any $\nu^- < \nu$, again with (hidden constant) $\propto |\Delta y|_\infty$. This means if $|\Delta y|_\infty$ is small enough to begin with, then all $|\Delta S^j y|_\infty$ are small (in fact decays exponentially fast as j increases) and all $S^j y$ are well-defined. ∎

Remark 3.4. We define convergence of $S^j y$ and the limit function $S^\infty y$ analogous to the linear case. Similar to (1.2), we define:

$$s_\infty(S) := \inf_y \sup\{\alpha : S^\infty y \in \mathrm{Lip}\,\alpha\},$$

where the infimum is taken over all sequences y for which $S^\infty y$ is well-defined. Since, for any such y, $|\Delta S^j y|_\infty$ decays and the smoothness of $S^\infty y$ does not depend on the behavior of $S^j y$ at coarse scales, there is no difference if we take the infimum over all y such that $|\Delta y|_\infty \leqslant \delta$, for any $0 < \delta \leqslant \delta^*$ where δ^* is given by Lemma 3.3.

3.1. Main proximity theorem. In this section, we show that the difference between Sy and $\overline{S}y$ can be bounded by the magnitudes of high order differences of y.

Lemma 3.5. *Let $y_1, ..., y_n, z_1, ..., z_n \in \mathbb{R}^m$. Suppose*

$$y_i = \sum_{j=1}^n f_{i,j} z_j, \quad i = 1, ..., n.$$

Let $F = (f_{i,j})$ and $A_0 = (a^0_{i,j})$ be $n \times n$ real matrices. Then

$$\sum_{i=1}^n \sum_{k=1}^n a^0_{i,k} \langle y_i, y_k \rangle = \sum_{i=1}^n \sum_{k=1}^n a^1_{i,k} \langle z_i, z_k \rangle,$$

where matrix $A_1 = (a^1_{i,k})$ is given by

$$A_1 = F^T A_0 F.$$

The proof is straightforward and we omit it.

Define $d^0_i = x_i$, $i = 1, \ldots, n$, and for $k = 1, ..., n$,

$$d^k_i = \begin{cases} d^{k-1}_i, & i = 1, \ldots, k \\ d^{k-1}_i - d^{k-1}_{i-1}, & i = k+1, \ldots, n. \end{cases} \tag{3.11}$$

It follows from Lemma 3.5 and (3.11) that we can further rewrite (3.8) as follow:

$$\sum_{i=1}^n \sum_{j=1}^n c^0_{i,j} \langle x_i, x_j \rangle = \sum_{i=1}^n \sum_{j=1}^n c^k_{i,j} \langle d^k_i, d^k_j \rangle, \quad k = 1, ..., n-1, \tag{3.12}$$

where $C_k = (c^k_{i,j})$ is given by

$$C_k = F^T_{k-1} C_{k-1} F_{k-1}, \quad k = 1, ..., n-1,$$

and

$$F_{k-1} = \begin{pmatrix} I_{k-1} & & & \\ & 1 & & \\ & \vdots & \ddots & \\ & 1 & \cdots & 1 \end{pmatrix}. \quad (I_{k-1} \text{ is the } (k-1)\times(k-1) \text{ identity matrix.}) \tag{3.13}$$

The next result says that the matrix C_k has many vanishing entries when the interpolatory subdivision mask reproduces polynomials; in particular, if $n = p$, i.e. the interpolatory mask has the highest possible order of polynomial reproducibility, then C_{n-1} has the following form:

$$C_{n-1} = \begin{bmatrix} 0 & 0 & \cdots & \cdots & 0 \\ 0 & \cdots & \cdots & 0 & * \\ \vdots & \vdots & \cdot & * & \vdots \\ \vdots & 0 & * & \cdots & \vdots \\ 0 & * & \cdots & \cdots & * \end{bmatrix}.$$

Lemma 3.6 (**Vanishing Property of C_k**). *If the interpolatory subdivision mask that defines the matrices C_k reproduces Π_{p-1}, then $c^k_{i,j} = 0$ if $i = 1$ or $j = 1$ or $i, j \leqslant k$, $i + j \leqslant p + 1$.* **Proof:** See Appendix A.4.

Lemma 3.6 leads to our main proximity result:

Theorem 3.7. *Let $\overline{S}$ be a linear interpolatory subdivision scheme which reproduces Π_{p-1}. There are constants $B_2, \dots, B_p$ such that for any $y : \mathbb{Z} \to \mathbb{S}^{m-1}$ such that (3.5) is well-defined, we have:*

$$\text{if } p \geqslant 1, \quad |\overline{S}y - Sy|_\infty \leqslant B_2 |\Delta y|^2_\infty; \tag{3.14}$$

$$\text{if } p \geqslant 3, \quad |\overline{S}y - Sy|_\infty \leqslant B_k \sum_{i=1}^{k-1} |\Delta^i y|_\infty |\Delta^{k-i} y|_\infty, \quad k = 3, ..., p. \tag{3.15}$$

Proof: Let $1 \leqslant k \leqslant p$. By Lemma 3.6, $c^k_{i,j} = 0$ when $i = 1$ or $j = 1$ or $i + j \leqslant k + 1$.

When $k = 1$, we have

$$\left| \sum_{i=1}^n \sum_{j=1}^n c^1_{i,j} \langle d^1_i, d^1_j \rangle \right| = \left| \sum_{i=2}^n \sum_{j=2}^n c^1_{i,j} \langle d^1_i, d^1_j \rangle \right| \leqslant \sum_{i=2}^n \sum_{j=2}^n |c^1_{i,j}| \left\| d^1_i \right\| \left\| d^1_j \right\| \leqslant B_2 |\Delta x|^2_\infty,$$

where B_2 is a constant that depends on $w_1, ..., w_n$.

If $p \geqslant 3$, then for $3 \leqslant k \leqslant p$,

$$\left| \sum_{i=1}^n \sum_{j=1}^n c^k_{i,j} \langle d^k_i, d^k_j \rangle \right| = \left| \sum_{\substack{i+j>k+1 \\ i>1, j>1}} c^k_{i,j} \langle d^k_i, d^k_j \rangle \right| \leqslant \sum_{\substack{i+j>k+1 \\ i>1, j>1}} |c^k_{i,j}| \left\| d^k_i \right\| \left\| d^k_j \right\| \leqslant B_k \sum_{i=1}^{k-1} |\Delta^i x|_\infty |\Delta^{k-i} x|_\infty,$$

where B_k is a constant that depends on $w_1, ..., w_n$.

Combined with (3.8) and (3.12), the above two estimates yield:

$$\text{if } p \geqslant 1, \quad \Big\| \sum_{i=1}^n w_i x_i - P(\sum_{i=1}^n w_i x_i) \Big\|_\infty \leqslant B_2 |\Delta x|^2_\infty;$$

$$\text{if } p \geqslant 3, \quad \Big\| \sum_{i=1}^n w_i x_i - P(\sum_{i=1}^n w_i x_i) \Big\|_\infty \leqslant B_k \sum_{i=1}^{k-1} |\Delta^i x|_\infty |\Delta^{k-i} x|_\infty, \quad k = 3, ..., p\,.$$

The above can be applied to any n consecutive entries of an infinite sequence y, so we have proved the theorem. ■

3.2. Smoothness equivalence. We now prove the main smoothness equivalence result:

Theorem 3.8. *If $\overline{S}$ is a linear interpolatory subdivision scheme with $s_\infty(\overline{S}) > 0$, then*

$$s_\infty(S) \geqslant s_\infty(\overline{S}).$$

To prove this theorem, we need to first recall two results.

The first one is well-known, see e.g. [5, 10], which basically says that one can characterize the Hölder smoothness of a continuous function based on its samples at dyadic points. If $f : \mathbb{R} \to \mathbb{R}$ is bounded and continuous, then for any $\alpha > 0$,

$$\begin{aligned} f \in \text{Lip}\,\alpha \quad &\Longleftrightarrow \quad \exists\, r \in \mathbb{Z}_+,\, r > \alpha, \text{ s.t. } |(\Delta^r f_j)_k|_\infty = O(2^{-j\alpha}) \\ &\Longleftrightarrow \quad \forall\, r \in \mathbb{Z}_+,\, r > \alpha,\, |(\Delta^r f_j)_k|_\infty = O(2^{-j\alpha}), \end{aligned} \tag{3.16}$$

where $f_j := f|_{2^{-j}\mathbb{Z}}$, i.e. $(f_j)_k := f(2^{-j}k)$. These equivalences also imply that the critical Hölder regularity exponent of f can be determined from the exact asymptotic decay rate of $|(\Delta^r f_j)_k|_\infty$ for a large enough differencing order r, i.e.

$$\sup\{\alpha : f \in \text{Lip}\,\alpha\} = \sup\Big\{\alpha : |(\Delta^r f_j)_k|_\infty = O(2^{-j\alpha})\Big\}. \tag{3.17}$$

The second result is the perturbation theorem [8, Theorem 3.3]. Originally derived to meet the needs of the analysis of a specific nonlinear subdivision algorithm, the theorem has been proved to be useful in the analysis of other nonlinear subdivision algorithms as well, see [33, 32, 23, 22, 36]. We restate this result in a form convenient to us:

THEOREM 3.9 (**[8, Theorem 3.3]**). *Let $\overline{S}$ be a linear subdivision operator with $s_\infty(\overline{S}) > 0$. Let S be a (linear or nonlinear) subdivision operator that maps $\mathcal{D}(S) \subseteq \ell^\infty$ into itself. Let $y \in \mathcal{D}(S)$. If there exists $\nu > 0$ such that*

$$\big|(S - \overline{S}\,)S^j y\big|_\infty = O(2^{-j\nu}),$$

then S is convergent and $s_\infty(S, y) := \sup\{\alpha : S^\infty y \in \operatorname{Lip} \alpha\} \geqslant \min(\nu, s_\infty(\overline{S}))$.

Proof of Theorem 3.8. By Remark 3.4, it suffices to prove $s_\infty(S, y) := \sup\{\alpha : S^\infty y \in \operatorname{Lip} \alpha\} \geqslant s_\infty(\overline{S})$ for those y such that $|\Delta y|_\infty$ is small.

1° For any $0 < \gamma < \min(1, s_\infty(\overline{S}))$, there exists a constant $C > 0$ such that

$$|\Delta\overline{S}^j y|_\infty \leqslant C\, 2^{-\gamma j}|\Delta y|_\infty, \quad \forall\, y \in \ell^\infty.$$

For any $\gamma' \in (0, \gamma)$, let $\epsilon := 2^{-\gamma'} - 2^{-\gamma} > 0$. Then it follows from (3.14) and Theorem 2.4 that there exist $\delta_{\gamma'} > 0$ and $C' > 0$ such that

$$|\Delta S^j y|_\infty \leqslant C'\,(2^{-\gamma} + \epsilon)^j|\Delta y|_\infty = C'\, 2^{-\gamma' j}|\Delta y|_\infty, \quad \forall\, y : \mathbb{Z} \to \mathbb{S}^{m-1} \text{ s.t. } |\Delta y|_\infty < \delta_{\gamma'}. \tag{3.18}$$

2° It suffices to consider a fixed y with $|\Delta y|_\infty$ small enough so that (3.18) can be applied and all $S^j y$ are well-defined. Recall Lemma 3.3 and Remark 3.4.

By (3.14) in Theorem 3.7, we have

$$|S(S^j y) - \overline{S}(S^j y)|_\infty \leqslant B_2|\Delta S^j y|_\infty^2 \overset{(3.18)}{=} O(2^{-2\gamma' j}).$$

Then, by Theorem 3.9, we have $s_\infty(S, y) \geqslant \min(2\gamma', s_\infty(\overline{S}))$. Since γ' can be arbitrarily close to $\min(1, s_\infty(\overline{S}))$, we get

$$s_\infty(S, y) \geqslant \min(2, s_\infty(\overline{S})). \tag{3.19}$$

So the theorem is proved if $s_\infty(\overline{S}) \leqslant 2$. So from now on we assume $s_\infty(\overline{S}) > 2$.

3° Let q be the unique integer such that

$$p \geqslant q + 1 \geqslant s_\infty(\overline{S}) > q \geqslant 2. \tag{3.20}$$

(Recall that $\overline{S}$ reproduces Π_{p-1} with $p \geqslant s_\infty(\overline{S})$ according to the theory of linear subdivision.)

We use induction to prove

$$s_\infty(S, y) \geqslant q. \tag{3.21}$$

Let $2 \leqslant k < q$. Assume that we have proved $s_\infty(S, y) \geqslant k$. By (3.16),

$$\begin{aligned} |\Delta^\ell S^j y|_\infty &= O(2^{-\ell j}), \quad \ell = 1, \ldots, k-1 \\ |\Delta^k S^j y|_\infty &= O(2^{-(k-\epsilon)j}). \end{aligned} \tag{3.22}$$

Since $p > k + 1$, Theorem 3.7 applies and we have

$$|\overline{S}S^j y - S^{j+1} y|_\infty \leqslant B_{k+1} \sum_{\ell=1}^{k} |\Delta^\ell S^j y|_\infty |\Delta^{k+1-\ell} S^j y|_\infty \overset{(3.22)}{=} O(2^{-(k+1-\epsilon)j}). \tag{3.23}$$

So by Theorem 3.9,

$$s_\infty(S, y) \geqslant \min(k + 1 - \epsilon, s_\infty(\overline{S})) \overset{(3.20)}{=} k + 1 - \epsilon$$

for any $\epsilon > 0$, which implies $s_\infty(S, y) \geqslant k+1$. This completes the induction.

Since we have now proved (3.21), (3.22) and (3.23) hold for $k = q$. Using Theorem 3.9 one more time gives:

$$s_\infty(S, y) \geqslant \min(q+1, s_\infty(\overline{S})) \overset{(3.20)}{=} s_\infty(\overline{S}).$$

This completes the proof. ∎

3.3. Near-closest projections onto the sphere. In this section, we show that for projections that are close to the closest point projection onto the sphere, the associated nonlinear schemes are also at least as smooth as the underlying interpolatory linear scheme. Such a result is to be expected: recall that the starting point of our proximity result is the simple inequality in (3.6), it is clear that we can relax P a little to achieve essentially the same upper bound on the right-hand side of (3.6). It is also clear that for any such projection operator, all the arguments for our smoothness equivalence result pertaining to the nonlinear subdivision operator $S = P \circ \overline{S}$ [2] go through verbatim.

In order for $S = P \circ \overline{S}$ to have a chance of being convergent, for any sequence y such that all $S^j y$ are well-defined, $|\Delta S^j y|_\infty$ must converge to zero as $j \to \infty$. Also, by (3.10), as long as consecutive points in a sequence $z : \mathbb{Z} \to \mathbb{S}^{m-1}$ are sufficiently close to each other, the points in $\overline{S}z$ can be made as close to the sphere as we want. Therefore, we only need to study the property of P in a neighborhood of the sphere as far as the smoothness analysis of S is concerned.

Lemma 3.10. *Let $P : \mathbb{R}^m \to \mathbb{S}^{m-1}$. If there exist $\delta, C > 0$ such that when $\big|\|x\| - 1\big| < \delta$,*

$$\cos(\angle(x, P(x))) \geqslant \frac{1 + \|x\|^2 - C(1 - \|x\|^2)^2}{2\|x\|}, \tag{3.24}$$

then $\|P(x) - x\| \leqslant \sqrt{C}\big|1 - \|x\|^2\big|$ when $\big|\|x\| - 1\big| < \delta$. **Proof:** Since $\langle x, P(x)\rangle = \|x\|\|P(x)\| \cos(\angle(x, P(x))) = \|x\| \cos(\angle(x, P(x)))$, it follows from (3.24) that

$$\langle x, P(x)\rangle \geqslant \frac{1 + \|x\|^2 - C(1 - \|x\|^2)^2}{2}.$$

Hence

$$\begin{aligned}
\|P(x) - x\|^2 &= \langle P(x) - x, P(x) - x\rangle \\
&= 1 - 2\langle x, P(x)\rangle + \|x\|^2 \\
&\leqslant 1 - (1 + \|x\|^2 - C(1 - \|x\|^2)^2) + \|x\|^2 \\
&= C(1 - \|x\|^2)^2.
\end{aligned}$$

Thus $\|P(x) - x\| \leqslant \sqrt{C}\,\big|1 - \|x\|^2\big|$. ∎

When P is the closest point projection, we always have $\angle(x, P(x)) = 0$. The above lemma shows that as long as P satisfies (3.24), the same bound (3.8) applies with an adjustment of hidden constant. Consequently, we have:

Theorem 3.11. *For any projection operator P satisfying (3.24) and any interpolatory linear subdivision scheme $\overline{S}$, the corresponding nonlinear subdivision S_P satisfies $s_\infty(S_P) \geqslant s_\infty(\overline{S})$.*

4. Projection Scheme for $SO(m)$-valued data and Extensions. In this section, we first extend our smoothness equivalence result to the Lie group of special orthogonal matrices:

$$SO(m) = \{Y \in \mathbb{R}^{m\times m} : YY^T = I,\ \det(Y) = 1\}.$$

In order to use the projection approach for data taking values in $SO(m)$, we need to (i) embed $SO(m)$ into an Euclidean space and (ii) define a projection operator P from the Euclidean space to the embedded surface. From a practical of view, we would also like such a P to be efficiently computable.

[2] We abuse notation and extend the definition of P to a map that maps sequences of m-vectors to sequences of points on the sphere; i.e. if y is a sequence of non-zero m-vectors, $P(y)$ is the sequence $P(y)_i = y_i/\|y_i\|$. We will abuse notation in a similar manner later without mention.

There is a natural way to embed $SO(m)$ in $\mathbb{R}^{m^2}$: simply treat every matrix in $SO(m)$ as a point in $\mathbb{R}^{m^2}$. It is not hard to prove that such a procedure indeed defines a smooth embedding; so $SO(m)$ now 'looks like' a $m(m-1)/2$ dimensional curved surface in $\mathbb{R}^{m^2}$, and we shall decide how to project a given point outside of this surface onto the surface.

For $X_1, X_2 \in \mathbb{R}^{m\times m}$, define $\langle , \rangle$ by

$$\langle X_1, X_2\rangle := \mathrm{trace}(X_1 X_2^T), \tag{4.1}$$

where $\mathrm{trace}(X) = \sum_{i=1}^m x_{i,i}$ is the trace of a matrix $X = (x_{i,j})$. This inner product induces the so-called Frobenius norm:

$$\|X\|_F := (\sum_{i=1}^m \sum_{j=1}^m x_{i,j}^2)^{1/2}. \tag{4.2}$$

Recall that for any orthogonal matrices U, V,

$$\|UXV\|_F^2 = \|X\|_F^2. \tag{4.3}$$

Note: If we identify $\mathbb{R}^{m\times m}$ with $\mathbb{R}^{m^2}$, then (4.1) and (4.2) are just the most standard inner product and Euclidean norm (respectively) in $\mathbb{R}^{m^2}$.

We can also extend the definitions of Δ and $|\cdot|_\infty$ to sequences with entries in $\mathbb{R}^{m\times m}$. For example, $|Y|_\infty := \sup_i \|Y_i\|_F$.

The closest point projection onto $SO(m)$ of a matrix with positive determinant can be computed efficiently using its singular value decomposition; a now-classical reference for (especially the computational aspect of) SVD is [15].

Proposition 4.1. *Let* $A \in \mathbb{R}^{m\times m}$ *with* $\det(A) > 0$. *Then*

$$P(A) := \operatorname*{argmin}_{X\in SO(m)} \|A - X\|_F = UV^T = (AA^T)^{-1/2}A \tag{4.4}$$

where $A = U\Sigma V^T$ *is a singular value decomposition of* A. **Proof:** Let $A = U\Sigma V^T$ be a singular value decomposition of A, where U, V are orthogonal matrices and D is a diagonal matrix with positive diagonal entries $\sigma_1 \geqslant \cdots \geqslant \sigma_m > 0$. Then

$$\begin{aligned} \operatorname*{argmin}_{X\in SO(m)} \|A - X\|_F &= \operatorname*{argmin}_{X\in SO(m)} \|U\Sigma V^T - X\|_F \\ &= U(\operatorname*{argmin}_{X\in SO(m)} \|\Sigma - U^T XV\|_F)V^T \\ &= U(\operatorname*{argmin}_{R\in SO(m)} \|\Sigma - R\|_F)V^T. \end{aligned}$$

It is easy to show that $\mathrm{argmin}_{R\in SO(m)} \|\Sigma - R\|_F = I$:

$$\|\Sigma - R\|_F^2 = \sum_{i=1}^m \Big[(\sigma_i - R_{ii})^2 + \sum_{j\neq i} R_{ij}^2\Big] = \sum_{i=1}^m \big[(\sigma_i - R_{ii})^2 + (1 - R_{ii}^2)\big] = \sum_{i=1}^m (\sigma_i^2 + 1 - 2\sigma_i R_{ii}).$$

Since $R_{ii} \leqslant 1$, the right-hand side is minimized when $R_{ii} = 1$ which, since $R \in SO(m)$, also implies $R = I$. Notice also that $(AA^T)^{-1/2}A = (U\Sigma^2U^T)^{-1/2}U\Sigma V^T = (U\Sigma^{-1}U^T)U\Sigma V^T = UV^T$. ■

Remark 4.2. Proposition 4.1 is essentially published in [2, 28] and seems to be known to others as well. We present our proof anyway not only because it is short and elementary but also because we want to address a subtle point. First of all, we note that the projection operator defined by (4.4) has the invariance property: $P(R_1AR_2) = R_1P(A)R_2$ for any $R_1, R_2 \in SO(m)$. It implies that our resulted subdivision algorithm has the desirable property that it does not depend on the artificial choice of orthogonal coordinate system for representing m-dimensional rotations. Our presentation above is more elementary than that in [2] because we are unconcerned with invariance at the beginning and simply think of $SO(m)$ as a regular surface in $\mathbb{R}^{m^2}$

and use the plain Euclidean metric in $\mathbb{R}^{m^2}$. The approach in [2], instead, considers $SO(m)$ as a subgroup $GL(m)$ and uses a (left-)invariant Riemannian metric of $GL(m)$ in defining 'closest'. While the two different points of view yield the same projector $P(A) = (AA^T)^{-1/2}A$, the coincidence is due to (4.3) and is specific to $SO(m)$. We will revisit this issue in Section 4.2.

Let $X = (X_1, ..., X_n)$ with $X_i \in SO(m)$ and $w_1, ..., w_n$ be as in (3.1). Then it follows from $\sum_{i=1}^n w_i = 1$ that

$$\|\sum_{i=1}^n w_i X_i - X_1\|_F = \|\sum_{i=1}^n w_i(X_i - X_1)\|_F \leqslant (\sum_{j=1}^n |w_j|) \max_i \|X_i - X_1\|_F. \tag{4.5}$$

Since $\det(X)$ is a continuous function of $X \in \mathbb{R}^{m\times m}$ and $\det(X_1) = 1 > 0$, it follows from (4.5) that there exists $\delta > 0$ such that when $\max_i \|X_i - X_1\|_F < \delta$,

$$\det(\sum_{i=1}^n w_i X_i) > 0,$$

and consequently Proposition 4.1 can be applied to define $P(\sum_{i=1}^n w_i X_i)$.

Let $\sum_{i=1}^n w_i X_i = U\Sigma V^T$ be the singular value decomposition of $\sum_{i=1}^n w_i X_i$, where U, V are orthogonal matrices and Σ is a diagonal matrix with positive diagonal entries $\sigma_1 \geqslant \sigma_2 \geqslant ... \geqslant \sigma_m > 0$. Then it follows from Proposition 4.1 that

$$\begin{aligned}\Big\|\sum_{i=1}^n w_i X_i - P(\sum_{i=1}^n w_i X_i)\Big\|_F &= \|U\Sigma V^T - UV^T\|_F = \|\Sigma - I\|_F = \Big(\sum_{\ell=1}^m (\sigma_\ell - 1)^2\Big)^{1/2} \\ &\leqslant \sqrt{m}\max_\ell |\sigma_\ell - 1| \leqslant \sqrt{m}\max_\ell |\sigma_\ell^2 - 1|.\end{aligned} \tag{4.6}$$

Since $(\sum_{i=1}^n w_i X_i)V = U\Sigma$, it follows that $(\sum_{i=1}^n w_i X_i)v_\ell = \sigma_\ell\, u_\ell$, where $u_\ell, v_\ell \in \mathbb{S}^{m-1}$ are the columns of U, V respectively. So

$$\sigma_\ell^2 = \|\sigma_\ell u_\ell\|^2 = \|(\sum_{i=1}^n w_i X_i)v_\ell\|^2 = \Big\langle(\sum_{i=1}^n w_i X_i)v_\ell, (\sum_{i=1}^n w_i X_i)v_\ell\Big\rangle = \sum_{i=1}^n\sum_{j=1}^n w_i w_j v_\ell^T X_j^T X_i v_\ell.$$

Note that $\|(X_i - X_j)v_\ell\|^2 = \langle X_i v_\ell - X_j v_\ell, X_i v_\ell - X_j v_\ell\rangle = 2 - 2v_\ell^T X_j^T X_i v_\ell$, hence

$$\begin{aligned}\sigma_\ell^2 &= \sum_{i=1}^n\sum_{j=1}^n w_i w_j (1 - \frac{1}{2}\|(X_i - X_j)v_\ell\|^2) \\ &= 1 - \frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n w_i w_j \|(X_i - X_j)v_\ell\|^2.\end{aligned}$$

Thus

$$|\sigma_\ell^2 - 1| = \Big|\frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n w_i w_j \|(X_i - X_j)v_\ell\|^2\Big|.$$

Now fix ℓ and let $x_i := X_i v_\ell$, $i = 1, ..., n$. Then

$$|\sigma_\ell^2 - 1| = \Big|\frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n w_i w_j \|x_i - x_j\|^2\Big|. \tag{4.7}$$

Note that the right-hand side of (4.7) looks exactly the same as the right-hand side of (3.10). Following exactly the same arguments there, we can show that:
If $p \geqslant 1$,

$$\Big|\frac{1}{2}\sum_{i=1}^n\sum_{j=1}^n w_i w_j \|x_i - x_j\|^2\Big| \leqslant \sum_{i=2}^n\sum_{j=2}^n |c_{i,j}^1|\, \|d_i^1\|\, \|d_j^1\|\,. \tag{4.8}$$

If $p \geqslant 3$, then for $3 \leqslant k \leqslant p$,

$$\Big|\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \|x_i - x_j\|^2\Big| \leqslant \sum_{\substack{i+j>k+1\\ i>1,j>1}} |c_{i,j}^k| \left\|d_i^k\right\| \left\|d_j^k\right\|, \tag{4.9}$$

where $c_{i,j}^k$ and d_i^k are as defined in Section 3.

Parallel to the definition of d_i^k, we define $D_i^0 = D_i$, $i = 1, \dots, n$, and for $k = 1, ..., n$,

$$D_i^k = \begin{cases} D_i^{k-1}, & i = 1, \dots, k \\ D_i^{k-1} - D_{i-1}^{k-1}, & i = k+1, \dots, n. \end{cases}$$

Then for any k and i,

$$d_i^k = D_i^k v_\ell.$$

So

$$\|d_i^k\| = \|D_i^k v_\ell\| \leqslant \|D_i^k\|_2 \|v_\ell\| = \|D_i^k\|_2 \leqslant \|D_i^k\|_F, \tag{4.10}$$

where $\|.\|_2$ denotes the 2-norm of a matrix and we used the fact that $\|Y\|_2 \leqslant \|Y\|_F$ for any matrix Y.

Combining (4.10) with (4.7), (4.8) and (4.9), we get

$$(p \geqslant 1) \qquad |\sigma_\ell^2 - 1| \leqslant \sum_{i=2}^{n}\sum_{j=2}^{n} |c_{i,j}^1| \left\|D_i^1\right\|_F \left\|D_j^1\right\|_F \leqslant B_2 |\Delta X|_\infty^2$$

and for $3 \leqslant k \leqslant p$,

$$(p \geqslant 3) \qquad |\sigma_\ell^2 - 1| \leqslant \sum_{\substack{i+j>k+1\\ i>1,j>1}} |c_{i,j}^k| \|D_i^k\|_F \|D_j^k\|_F \leqslant B_k \sum_{i=1}^{k-1} |\Delta^i X|_\infty |\Delta^{k-i} X|_\infty,$$

where $B_2, B_3, \dots, B_p$ are constants that only depend on $w_1, ..., w_n$. Combining with (4.6), we get

$$\Big\|\sum_{i=1}^{n} w_i X_i - P(\sum_{i=1}^{n} w_i X_i)\Big\|_F \tag{4.11}$$

$$\leqslant \begin{cases} \sqrt{m}\, B_2 |\Delta X|_\infty^2, & \text{if } p \geqslant 1 \\ \sqrt{m}\, B_k \sum_{i=1}^{k-1} |\Delta^i X|_\infty |\Delta^{k-i} X|_\infty, & k = 3, ..., p, \quad \text{if } p \geqslant 3. \end{cases} \tag{4.12}$$

This is essentially the same as (3.14) and (3.15) in Theorem 3.7 of which the proof of Theorem 3.7 is based on; this also means that we have proved:

THEOREM 4.3. *For any interpolating linear subdivision $\overline{S}$, the corresponding closest point projection scheme S for $SO(m)$-valued data satisfies $s_\infty(S) \geqslant s_\infty(\overline{S})$.*

4.1. Extensions to related Lie groups. We consider rigid body displacements:

$$SE(m) = \{T_{A,b} : \mathbb{R}^m \to \mathbb{R}^m \mid T_{A,b}(x) = Ax + b,\ A \in SO(m),\ b \in \mathbb{R}^m\}.$$

There is a standard way to smoothly embed this matrix Lie group into $\mathbb{R}^{(m+1)\times(m+1)}$:

$$T_{A,b} \mapsto \begin{bmatrix} A & b \\ 0 & 1 \end{bmatrix}.$$

This embedding is also a group homomorphism from $SE(m)$ to the general linear group $GL(m+1)$, as $T_{A',b'} \circ T_{A,b} = T_{A'A,A'b+b}$ and

$$\begin{bmatrix} A' & b' \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A & b \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} A'A & A'b + b' \\ 0 & 1 \end{bmatrix}.$$

So, once again, we are in the situation as discussed in Remark 4.2: for the purpose of constructing a subdivision scheme for $SE(m)$-valued data based on a *linear* subdivision scheme $\overline{S}$, we take the point of view that $SE(m)$ is a regular surface in the *linear space* $\mathbb{R}^{(m+1)\times(m+1)}$; however, for the purpose of constructing a projection operator P with a sensible invariance property, we should take the point of view that $SE(m)$ is embedded in the Lie group $GL(m)$ and define a projection operator based on a (left-)invariant metric with respect to the group operation. In this case, the projection operator is given by:

$$P\left(\begin{bmatrix} A & b \\ 0 & 1 \end{bmatrix}\right) = \begin{bmatrix} UV^T & b \\ 0 & 1 \end{bmatrix}, \tag{4.13}$$

where $A = U\Sigma V^T$ is a SVD of A. However, again as in the case of $SO(m)$, it does not quite matter whether we think of "closest point projection" in terms of the standard Euclidean metric in $\mathbb{R}^{m^2}$ or an invariant metric in $GL(m+1)$.

Afterall, the most important fact is that the nonlinear subdivision operator $S = P \circ \overline{S}$ with P given by (4.13) and $\overline{S}$ a linear interpolatory subdivision scheme acting componentwise is well-defined when applied to any sequence $Y : \mathbb{Z} \to SE(m)$ with a small enough $|\Delta Y|_\infty$; moreover S has the desirable property that it is invariant under any change of orthogonal reference frame for representing rigid motions in an m-dimensional space.

Motivated by motion design, we are also interested in direct products of $SO(m)$, e.g. one can model the combined motion of 17 major human joints as an element in

$$\begin{aligned} &\underbrace{SO(3)}_{\text{neck}} \times \underbrace{SO(3) \times SO(3)}_{\text{shoulders}} \times \underbrace{SO(2) \times SO(2)}_{\text{elbows}} \times \underbrace{SO(3) \times SO(3)}_{\text{wrists}} \\ &\quad \times \underbrace{SO(3) \times SO(3)}_{\text{hips}} \times \underbrace{SO(2) \times SO(2)}_{\text{knees}} \times \underbrace{SO(3) \times SO(3)}_{\text{ankles}} \times \underbrace{SO(3) \times SO(3) \times SO(3) \times SO(3)}_{\text{spine}}. \end{aligned}$$

See [19, Figure 3.2] for a graphical illustration.

It is obvious that we can extend the projection approach to subdivide data taking values on such a direct product. It is also obvious that we can extend Theorem 4.3 to $SE(m)$ and such direct products. For the record, let us state it formally:

THEOREM 4.4. *Let $\mathcal{M} = SE(m)$ or $\prod_{i=1}^{k} SO(m_i) \times \prod_{j=1}^{k'} SE(n_j)$. For any interpolating linear subdivision $\overline{S}$, the corresponding closest point projection scheme $S = P \circ \overline{S}$ for $\mathcal{M}$-valued data satisfies $s_\infty(S) \geqslant s_\infty(\overline{S})$.*

4.2. $SL(m)$. We recall once again the dual view first discussed in Remark 4.2. This time we consider the matrix Lie group of all measure and orientation preserving linear transformations:

$$SL(m) := \{Y \in \mathbb{R}^{m\times m} : \det(Y) = 1\}.$$

$SL(m)$ has a natural embedding as a regular hypersurface in $\mathbb{R}^{m^2}$, but is also a subgroup of $GL(m)$. Unlike the cases of $SO(m)$ and $SE(m)$, the two points of views give different projectors. Using the latter setup which offers invariance, the resulting projector P is given by [13]:

$$A \mapsto A/\det A^{1/m}. \tag{4.14}$$

Note that $P(UAV) = UP(A)V$ for all $U, V \in SL(m)$. On the other hand, solving $\min_{X\in SL(m)} \|A - X\|_F$ is in principle a straightforward application of the method of Lagrange multipliers but gives a very complicated projector – which also lacks invariance – even in dimension $m = 2$.

Our computational experiment (akin to those "smoothness equivalence experiments" found in [33, 23, 24]) clearly indicates that the nonlinear subdivision scheme $S = P\circ\overline{S}$ enjoys the same smoothness equivalence property as in the case of Theorems 3.8, 4.3, 4.4. A proof is yet to be found.

5. Conclusions and Discussions. Interpolation of manifold-valued data is a fundamental problem that has applications in many fields. Linear subdivision method is an efficient and very well-studied method for interpolating or approximating real-valued data in a multiresolution fashion. We described in Section 1 three sets of approaches for adapting a linear subdivision scheme to subdivide manifold-valued data. The mathematical analysis of such nonlinear subdivision schemes is at its infancy. We mentioned a number of articles which offer some low degree smoothness equivalence results for certain nonlinear subdivision schemes. To the best of our knowledge, this is the first article that attacks the arbitrary degree smoothness equivalence conjectures.

We suspect that Theorems 3.8, 4.3, 4.4. can be extended to any C^∞ k-dimensional regular surface in $\mathbb{R}^n$ with any near-closest projection operator.

We mention a recent smoothness **non**-equivalence result in the nonlinear subdivision literature: In [37], a seemingly non-adaptive nonlinear subdivision scheme is shown to have a fairly strong data-dependent property, *unlike* any linear subdivision scheme or the weakly nonlinear subdivision schemes such as the ones studied in this article. Specifically, it is shown in [37] that a nonlinear convexity preserving subdivision scheme produces limit curves with critical Hölder regularity depending on the initial data and the regularity can be anywhere between 1 and 2.

We discuss potential applications of our results in two seemingly unrelated problems:

- *Conics-reproducing subdivision scheme.* A standard complaint of standard B-splines and standard linear subdivision scheme is that they can only reproduce polynomials but not conic sections. The industrial standard NURBS uses rational B-splines because rational polynomials can reproduce conics while polynomials cannot. But it is widely argued that NURBS methods lack some of the key advantages of subdivision methods.[3] A linear but *non-stationary* 4-point interpolatory scheme is derived in [14, 26] which reproduces $\text{span}(1, x, \cos(x), \sin(x))$ instead of the usual $\Pi_3 = \text{span}(1, x, x^2, x^3)$, and hence can reproduces circles when the initial control polygon is sampled uniformly from a circle. Given the result in this paper, it seems like that a better way to solve this problem is to use the projection approach. When one demands a subdivision scheme to exactly reproduces a circle, or any conic section, or any other pre-specified shape $\mathcal{C}$, our proposed method is to use a nonlinear but stationary scheme of the form $S = P_{\mathcal{C}} \circ \overline{S}$, as opposed to the linear but non-stationary scheme proposed in [14, 26]. The projection approach seems more general and flexible: it does not require uniform sampling, it can be used in conjunction with any interpolatory subdivision scheme (not just 4-point) and, for the circle at least, Theorem 3.8 says that the nonlinear scheme is as smooth as the underlying linear scheme. (The exact Hölder regularity of the specific non-stationary 4-point scheme in [14, 26] is not known, but the scheme is shown to be at least C^1.)
- *Normal multiresolution of curves.* Underlying the method of normal multiresolution of curves in [8] is a nonlinear subdivision scheme of almost exactly the same form as those studied in this paper, i.e. $S = P \circ \overline{S}$. The key difference is that the P's in this article are such that $P(y)_i$ is only dependent on y_i, whereas the P in [8] is more data-adaptive: $P(y)_{2i} = y_{2i}$ and $P(y)_{2i+1} =$ an intersection point of $\mathcal{C}$ with the line passing through y_{2i+1} and normal to the line $\overline{y_{2i}\, y_{2i+2}}$. See [8, Figure 2] for a graphical illustration. (Here $\mathcal{C}$ is a planar curve subject to a normal multiresolution analysis.) It is conjectured that the parametrization induced by a normal multisolution has exactly the same smoothness as that of the underlying interpolatory subdivision scheme. Similar to the other low degree smoothness equivalence results mentioned earlier, Daubechies *et al* only prove the smoothness equivalence when (in the notation of this paper) $s_\infty(\overline{S}) \leqslant 2$. It seems possible to adapt the ideas and results in this paper to solve the full smoothness equivalence conjecture pertaining to normal multiresolution.

Yet another extension is to consider Hermite subdivision schemes on manifolds. See, e.g. [20], for the interests of Hermite interpolation in Lie groups arising from geometric integration of ODEs. Hermite subdivision schemes in the linear setting is quite well-studied, see [21, 34, 16] and the references therein. It is not hard to construct Hermite subdivision schemes on Lie groups, but the analysis of such schemes is likely to be difficult.

Finally, it is needless to say that similar smoothness equivalence results for the more intrinsic linearization methods described in Section 1 are waiting to be developed.

REFERENCES

[3]The debate, however, is mostly on surface modeling, not curve modeling.

[1] D. Asimov. The grand tour: a tool for viewing multidimensional data. *SIAM J. Scient. and Stat. Comp.*, 6(1):128–143, 1985.

[2] C. Belta and V. Kumar. An SVD-based projection method for interpolation on $SE(3)$. *IEEE Transaction on Robotics and Automation*, 18(3):334–345, 2002.

[3] W. M. Boothby. *An Introduction to Differentiable Manifolds and Riemannian Geometry.* Academic Press, 2nd edition, 2002.

[4] A.S. Cavaretta, W. Dahmen, and C.A. Micchelli. Stationary subdivision. *Mem. Amer. Math. Soc.*, 453, 1991. American Math. Soc, Providence.

[5] Z. Ciesielski. Approximation of splines and its application to Lipschitz classes and to stochastic processes. In *Approximation Theory of Functions, Proc. of the conference in Kaluga, 1975*, pages 397–400. Nauka, Moscow, 1977.

[6] I. Daubechies and J. Lagarias. Two-scale difference equations I. existence and global regularity of solutions. *SIAM J. Math. Anal.*, 22(5):1388–1410, 1991.

[7] I. Daubechies and J. Lagarias. Two-scale difference equations II. local regularity, infinite products of matrices and fractals. *SIAM J. Math. Anal.*, 23(4):1031–1079, 1992.

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

[9] G. Deslauriers and S. Dubuc. Symmetric iterative interpolation processes. *Constructive Approximation*, 5:49–68, 1989.

[10] Z. Ditzian. Moduli of smoothness using discrete data. *J. Approx. Theory*, 49:115–129, 1987.

[11] M. P. do Carmo: translated by F. Flaherty. *Riemannian Geometry.* Birkhäuser Boston, 1992.

[12] S. Dubuc. Interpolation through an iterative scheme. *Journal of Mathematical Analysis and Applications*, 114:185–204, 1986.

[13] T. Duchamp. Personal communication. December 2005.

[14] N. Dyn and D. Levin. Subdivision schemes in geometric modelling. *Acta Numerica*, 11:73–144, 2002.

[15] G. Golub and C. van Loan. *Matrix computations.* The Johns Hopkins University Press, London, 3rd edition, 1996.

[16] B. Han, T. P.-Y. Yu, and Y. Xue. Non-interpolatory Hermite subdivision schemes. *Mathematics of Computation*, 74(251):1345–1367, 2005.

[17] A. Iserles, H. Z. Munthe-Kaas, S. P. Nrsett, and A. Zanna. Lie-group methods. *Acta Numerica*, 9:215–365, 2000.

[18] A. Iserles and S.P. Nrsett. On the solution of linear differential equations on Lie groups. *Phil. Trans Royal Soc. A*, 357:983–1019, 1999.

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

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

[21] J. L. Merrien. A family of Hermite interpolants by bisection algorithms. *Numerical Algorithms*, 2:187–200, 1992.

[22] P. Oswald. Smoothness of a nonlinear subdivision scheme. In A. Cohen, J.-L. Merrien, and L. L. Schumaker, editors, *Curves and Surface Fitting: Saint-Malo 2002*, pages 323–332. Nashboro Press, 2003.

[23] P. Oswald. Smoothness of nonlinear median-interpolation subdivision. *Advances in Computational Mathematics*, 20(4):401–423, 2004.

[24] I. Ur Rahman, I. Drori, V. C. Stodden, D. L. Donoho, and P. Schröder. Multiscale representations for manifold-valued data. *Multiscale Modeling and Simulation*, 4(4):1201–1232, 2005.

[25] O. Rioul. Simple regularity criteria for subdivision schemes. *SIAM J. Math. Anal.*, 23(6):1544–1576, November 1992.

[26] S. Schaefer. A factored, interpolatory subdivision scheme for surfaces of revolution. Master's thesis, Department of Computer Science, Rice University, 2003.

[27] C. H. Sequin and J. A. Yen. Fair and robust curve interpolation on the sphere. *Sketches and Application, SIGGRAPH'01*, page 182, 2001.

[28] J. Wallner. Gliding spline motions and applications. *Computer Aided Geometric Design*, 21:3–21, 2004.

[29] J. Wallner. Smoothness analysis of subdivision schemes by proximity, 2005. Preprint.

[30] J. Wallner and N. Dyn. Convergence and C^1 analysis of subdivision schemes on manifolds by proximity. *Computer Aided Geometric Design*, 22(7):593–622, 2005.

[31] J. Wallner and H. Pottmann. Intrinsic subdivision with smooth limits for graphics and animation. *ACM Trans. Graphics*, x(x):x, 2005.

[32] G. Xie and T. P.-Y. Yu. On a linearization principle for nonlinear p-mean subdivision schemes. In M. Neamtu and E. B. Saff, editors, *Advances in Constructive Approximation*, pages 519–533. Nasboro Press, 2004.

[33] G. Xie and T. P.-Y. Yu. Smoothness analysis of nonlinear subdivision schemes of homogeneous and affine invariant type. *Constructive Approximation*, 22(2):219–254, 2005.

[34] T. P.-Y. Yu. On the regularity analysis of interpolatory Hermite subdivision schemes. *Journal of Mathematical Analysis and Applications*, 302(1):201–216, 2005.

[35] T. P.-Y. Yu. Approximation order equivalence properties of manifold-valued data subdivision schemes. In preparation, 2006.

[36] T. P.-Y. Yu. Cutting corners on the sphere. In Guanrong Chen and Ming-Jun Lai, editors, *Wavelets and Splines: Athens 2005*, pages 496–506. Nasboro Press, 2006.

[37] T. P.-Y. Yu. How data dependent is a nonlinear subdivision scheme? – a case study based on convexity preserving subdivision. *SIAM Journal on Numerical Analysis*, 44(3):936–948, 2006.

Appendix A. Appendix.

A.1. Proof of Lemma 2.2. Since $|\Delta x - \Delta y|_\infty \leqslant 2|x-y|_\infty$ for any $x, y \in \ell^\infty$, it follows that

$$|\Delta T_2 p - \Delta T_1 p|_\infty \leqslant 2|T_2 p - T_1 p|_\infty.$$

Combining with (2.3), we get

$$|\Delta T_2 p - \Delta T_1 p|_\infty \leqslant 2A|\Delta p|_\infty^\alpha, \quad \text{for all } p \in \ell^\infty, |\Delta p|_\infty < \delta.$$

So

$$|\Delta T_2 p|_\infty \leqslant |\Delta T_1 p|_\infty + 2A|\Delta p|_\infty^\alpha, \quad \text{for all } p \in \ell^\infty, |\Delta p|_\infty < \delta.$$

Since we have (2.2), it follows that for all $p \in \ell^\infty$ with $|\Delta p|_\infty < \min(\delta, 1)$, we have

$$|\Delta T_2 p|_\infty \leqslant C|\Delta p|_\infty + 2A|\Delta p|_\infty^\alpha = (C + 2A|\Delta p|_\infty^{\alpha-1})|\Delta p|_\infty < (C+2A)|\Delta p|_\infty. \tag{A.1}$$

Therefore (2.4) holds for $C' = C + 2A$ and $\delta' = \min(\delta, 1)$. ∎

A.2. Proof of Lemma 2.3. We use induction. For $j = 1$, (2.7) follows immediately from (2.6) by choosing $\delta_1 = \delta$ and $C_1 = A$. Now suppose (2.7) holds for some $j \geqslant 1$.

It follows from Lemma 2.2 that there exists $C' > 1$ and $\delta' > 0$ such that when $|\Delta p|_\infty < \delta'$,

$$|\Delta T_2 p|_\infty \leqslant C'|\Delta p|_\infty.$$

So for any $j \in \mathbb{N}$,

$$|\Delta T_2^j p|_\infty \leqslant C'^j |\Delta p|_\infty < \delta, \quad \text{if } |\Delta p|_\infty < \min(\delta, \delta')C'^{-j}.$$

Since T_1 is bounded and linear, it follows from (2.6) that for $q \in \ell^\infty$ satisfying $|\Delta q|_\infty < \delta$, we have

$$|T_1 p - T_2 q|_\infty \leqslant |T_1 p - T_1 q|_\infty + |T_1 q - T_2 q|_\infty \leqslant |T_1|_\infty |p - q|_\infty + A|\Delta q|^\alpha.$$

Hence if $|\Delta p|_\infty < \min(\min(\delta, \delta')C'^{-j}, \delta_j)$, then

$$\begin{aligned}
|T_1^{j+1} p - T_2^{j+1} p|_\infty &\leqslant |T_1|_\infty |T_1^j p - T_2^j p|_\infty + A|\Delta T_2^j p|_\infty^\alpha \\
&\leqslant |T_1|_\infty |T_1^j p - T_2^j p|_\infty + AC'^{j\alpha}|\Delta p|_\infty^\alpha \\
&\leqslant (|T_1|_\infty C_j + AC'^{j\alpha})|\Delta p|_\infty^\alpha.
\end{aligned}$$

This means (2.7) holds for $j+1$. By induction, (2.7) holds for any $j \in \mathbb{N}$. ∎

A.3. Proof of Theorem 2.4. Theorem 2.1 already covers the case of $C \leqslant 1$, so we can assume $C > 1$.

For any $\epsilon \in (0, 1-\mu)$, we can find $N \in \mathbb{N}$ such that $C^{1/N} < 1 + \epsilon/\mu$. So $C\mu^N < (\mu + \epsilon)^N < 1$. Hence when $|\Delta p|_\infty < \delta$,

$$|\Delta T_1^{jN} p|_\infty \leqslant C\mu^{jN}|\Delta p|_\infty \leqslant (C\mu^N)^j |\Delta p|_\infty, \quad \forall j \in \mathbb{N}.$$

It follows from (2.8), (2.9) and Lemma 2.2 that there exist $\tilde{C}, \tilde{\delta} > 0$ such that when $|\Delta p|_\infty < \tilde{\delta}$,

$$|\Delta T_2 p|_\infty \leqslant \tilde{C}|\Delta p|_\infty. \tag{A.2}$$

Since one of T_1 and T_2 is bounded linear, it follows from (2.8), (A.2), (2.9) and Lemma 2.3 that there exist $C_N > 0$ and $\delta_N > 0$ such that when $|\Delta p|_\infty < \delta_N$,

$$|T_1^N p - T_2^N p|_\infty \leqslant C_N |\Delta p|_\infty^\alpha.$$

Now we can apply Theorem 2.1 to operators T_1^N and T_2^N. We have for any $0 < \epsilon_0 < (\mu+\epsilon)^N - C\mu^N$, there exists $0 < \delta_0 < \delta$ such that

$$|\Delta T_2^N p|_\infty \leqslant (C\mu^N + \epsilon_0)|\Delta p|_\infty \leqslant (\mu + \epsilon)^N |\Delta p|_\infty, \quad \text{if } |\Delta p|_\infty < \delta_0.$$

Therefore for $k = 0, 1, ...,$ and $r = 0, 1, ..., N-1$

$$|\Delta T_2^{kN+r} p|_\infty \leqslant (\mu+\epsilon)^{kN} |\Delta T_2^r p|_\infty, \quad \text{if } |\Delta T_2^r p|_\infty < \delta_0. \tag{A.3}$$

It follows from (A.2) that

$$|\Delta T_2^r p|_\infty \leqslant \tilde{C}^r |\Delta p|_\infty < \delta_0 \quad \text{if } |\Delta p|_\infty < \min(\delta_0, \tilde{\delta})\tilde{C}^{-r}. \tag{A.4}$$

Combining (A.3) and (A.4), we have

$$|\Delta T_2^{kN+r} p|_\infty \leqslant (\mu+\epsilon)^{kN} \tilde{C}^r |\Delta p|_\infty \leqslant (\mu+\epsilon)^{kN+r} (\mu+\epsilon)^{-N} \tilde{C}^N |\Delta p|_\infty$$

if $|\Delta p| < \delta' := \min(\delta_0, \tilde{\delta})\tilde{C}^{-r}$. Let $C' = (\mu+\epsilon)^{-N}\tilde{C}^N$. Then when $|\Delta p|_\infty < \delta'$,

$$|\Delta T_2^j p|_\infty \leqslant C'(\mu+\epsilon)^j |\Delta p|_\infty$$

for any $j \in \mathbb{N}$ and $|\Delta p|_\infty < \delta'$. ■

A.4. Proof of Lemma 3.6. By (3.9), $c^0_{i,j} = c^0_{j,i}$ and for $j = 1, ..., n$,

$$\sum_{i=1}^{n} c^0_{i,j} = c^0_{j,j} + \sum_{\substack{i=1\\ i\neq j}}^{n} c^0_{i,j} = w_j - w_j^2 - \sum_{\substack{i=1\\ i\neq j}}^{n} w_i w_j = w_j - w_j^2 - w_j(1-w_j) = 0. \tag{A.5}$$

So the first row and first column of C_0 both sum to zero.

LEMMA A.1. *Let $A_0 = (a^0_{i,j})$ be a $n\times n$ real matrix. For $k = 0, ..., n-2$, define*

$$A_{k+1} = F_k^T A_k F_k,$$

where F_k is defined by (3.13). Then for $k = 0, ..., n-2$, $A_{k+1} = (a^{k+1}_{\alpha,\beta})$ is given by

$$a^{k+1}_{\alpha,\beta} = \begin{cases} \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} \binom{i-1}{\alpha-1}\binom{j-1}{\beta-1} a^0_{i,j}, & \alpha,\beta \leqslant k+1; \\ \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} \binom{i-\alpha+k}{k}\binom{j-1}{\beta-1} a^0_{i,j}, & \alpha > k+1, \beta \leqslant k+1; \\ \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} \binom{i-1}{\alpha-1}\binom{j-\beta+k}{k} a^0_{i,j}, & \alpha \leqslant k+1, \beta > k+1; \\ \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} \binom{i-\alpha+k}{k}\binom{j-\beta+k}{k} a^0_{i,j}, & \alpha,\beta > k+1. \end{cases} \tag{A.6}$$

Proof: We prove (A.6) by induction. It follows from $A_1 = F_0^T A_0 F_0$, that

$$a^1_{\alpha,\beta} = \sum_{i=\alpha}^{n}\sum_{j=\beta}^{n} a^0_{i,j} = \sum_{i=1}^{n}\sum_{j=1}^{n} \binom{i-\alpha}{0}\binom{j-\beta}{0} a^0_{i,j}.$$

This shows that (A.6) is true for $k = 0$.

Suppose (A.6) is true for $k = q-1$, i.e.

$$a^{q}_{\alpha,\beta} = \begin{cases} \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} \binom{i-1}{\alpha-1}\binom{j-1}{\beta-1} a^0_{i,j}, & \alpha,\beta \leqslant q; \\ \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} \binom{i-\alpha+q-1}{q-1}\binom{j-1}{\beta-1} a^0_{i,j}, & \alpha > q, \beta \leqslant q; \\ \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} \binom{i-1}{\alpha-1}\binom{j-\beta+q-1}{q-1} a^0_{i,j}, & \alpha \leqslant q, \beta > q; \\ \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} \binom{i-\alpha+q-1}{q-1}\binom{j-\beta+q-1}{q-1} a^0_{i,j}, & \alpha,\beta > q. \end{cases} \tag{A.7}$$

It follows from $A_{q+1} = F_q^T A_q F_q$ that

$$a^{q+1}_{\alpha,\beta} = \begin{cases} a^q_{\alpha,\beta}, & \alpha,\beta \leqslant q; \\ \sum\limits_{s=\alpha}^{n} a^q_{s,\beta}, & \alpha > q, \beta \leqslant q; \\ \sum\limits_{t=\beta}^{n} a^q_{\alpha,t}, & \alpha \leqslant q, \beta > q; \\ \sum\limits_{s=\alpha}^{n}\sum\limits_{t=\beta}^{n} a^q_{s,t}, & \alpha,\beta > q. \end{cases} \tag{A.8}$$

Substituting (A.7) into (A.8), we get

$$a^{q+1}_{\alpha,\beta} = \begin{cases} \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} \binom{i-1}{\alpha-1}\binom{j-1}{\beta-1} a^0_{i,j}, & \alpha,\beta \leqslant q; \\ \sum\limits_{s=\alpha}^{n}\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} \binom{i-s+q-1}{q-1}\binom{j-1}{\beta-1} a^0_{i,j}, & \alpha > q, \beta \leqslant q; \\ \sum\limits_{t=\beta}^{n}\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} \binom{i-1}{\alpha-1}\binom{j-t+q-1}{q-1} a^0_{i,j}, & \alpha \leqslant q, \beta > q; \\ \sum\limits_{s=\alpha}^{n}\sum\limits_{t=\beta}^{n}\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} \binom{i-s+q-1}{q-1}\binom{j-t+q-1}{q-1} a^0_{i,j}, & \alpha,\beta > q. \end{cases}$$

Using the following identities on combinatorial numbers:

$$\sum_{s=\alpha}^{n} \binom{i-s+q-1}{q-1} = \binom{i-\alpha+q}{q}, \quad \sum_{t=\beta}^{n} \binom{j-t+q-1}{q-1} = \binom{j-\beta+q}{q},$$

we get

$$a^{q+1}_{\alpha,\beta} = \begin{cases} \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} \binom{i-1}{\alpha-1}\binom{j-1}{\beta-1} a^0_{i,j}, & \alpha,\beta \leqslant q; \\ \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} \binom{i-\alpha+q}{q}\binom{j-1}{\beta-1} a^0_{i,j}, & \alpha > q, \beta \leqslant q; \\ \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} \binom{i-1}{\alpha-1}\binom{j-\beta+q}{q} a^0_{i,j}, & \alpha \leqslant q, \beta > q; \\ \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} \binom{i-\alpha+q}{q}\binom{j-\beta+q}{q} a^0_{i,j}, & \alpha,\beta > q. \end{cases} \tag{A.9}$$

It can be easily verified that (A.9) agrees with (A.6) when $k = q$. This concludes the proof. ■

It follows from Lemma A.1 that for $k = 1, ..., n-1$,

$$c^k_{\alpha,\beta} = \begin{cases} \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} \binom{i-1}{\alpha-1}\binom{j-1}{\beta-1} c^0_{i,j}, & \alpha,\beta \leqslant k; \\ \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} \binom{i-\alpha+k-1}{k-1}\binom{j-1}{\beta-1} c^0_{i,j}, & \alpha > k, \beta \leqslant k; \\ \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} \binom{i-1}{\alpha-1}\binom{j-\beta+k-1}{k-1} c^0_{i,j}, & \alpha \leqslant k, \beta > k; \\ \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} \binom{i-\alpha+k-1}{k-1}\binom{j-\beta+k-1}{k-1} c^0_{i,j}, & \alpha,\beta > k. \end{cases} \tag{A.10}$$

Combining with (A.5), we get for $k = 1, ..., n-1$ and $\ell = 1, ..., n$

$$c^k_{\ell,1} = c^k_{1,\ell} = \left\{ \begin{array}{ll} \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} \binom{j-1}{\ell-1} c^0_{i,j}, & \ell \leqslant k \\ \sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n} \binom{j-\ell+k-1}{k-1} c^0_{i,j}, & \ell > k \end{array} \right\} = 0. \tag{A.11}$$

For any $z \in \mathbb{R}$ and $\ell \in \mathbb{N} \cup \{0\}$, the generalized binomial coefficient $\binom{z}{\ell}$ is defined as

$$\binom{z}{\ell} = \frac{z(z-1)\cdots(z-\ell+1)}{\ell!}.$$

For each fixed ℓ, let $q_\ell(z) = \binom{z}{\ell}$. Then $q_\ell(z)$ is the unique polynomial in z of degree ℓ satisfying

$$q_\ell(0) = q_\ell(1) = \cdots = q_\ell(\ell-1) = 0, \ q_\ell(\ell) = 1.$$

Furthermore, $q_0(z), q_1(z), ..., q_\ell(z)$ form a basis of the polynomial space Π_ℓ. So for each $\gamma \in \mathbb{N} \cup \{0\}$, there exist constants $\tau_0^\gamma, ..., \tau_\gamma^\gamma$ satisfying

$$\binom{n-\frac{z}{2}}{\gamma} = \sum_{j=0}^{\gamma} \tau_j^\gamma \binom{z}{j}.$$

Combining with (3.4), we get for $\gamma = 0, ..., p-1$

$$\sum_{i=1}^{n} \binom{n-i}{\gamma} w_i' = \sum_{i=1}^{n}\sum_{j=0}^{\gamma} \tau_j^\gamma \binom{2i}{j} w_i' = \sum_{j=0}^{\gamma} \tau_j^\gamma \binom{2n'+1}{j} = \binom{n-n'-\frac{1}{2}}{\gamma}. \tag{A.12}$$

More generally, for each $\gamma_1, \gamma_2 \in \mathbb{N} \cup \{0\}$, there exist constants $\tau_0^{\gamma_1,\gamma_2}, ..., \tau_{\gamma_1+\gamma_2}^{\gamma_1,\gamma_2}$ satisfying

$$\binom{n-\frac{z}{2}}{\gamma_1}\binom{n-\frac{z}{2}}{\gamma_2} = \sum_{j=0}^{\gamma_1+\gamma_2} \tau_j^{\gamma_1,\gamma_2} \binom{z}{j}, \ \forall z \in \mathbb{R}.$$

Together with (3.4), we get for $\gamma_1 + \gamma_2 \leqslant p-1$

$$\begin{aligned}\sum_{i=1}^{n} \binom{n-i}{\gamma_1}\binom{n-i}{\gamma_2} w_i' &= \sum_{i=1}^{n}\sum_{j=0}^{\gamma_1+\gamma_2} \tau_j^{\gamma_1,\gamma_2} \binom{2i}{j} w_i' = \sum_{j=0}^{\gamma_1+\gamma_2} \tau_j^{\gamma_1,\gamma_2} \binom{2n'+1}{j} \\ &= \binom{n-n'-\frac{1}{2}}{\gamma_1}\binom{n-n'-\frac{1}{2}}{\gamma_2}.\end{aligned} \tag{A.13}$$

Therefore it follows from (A.10),(3.9),(A.12) and (A.13) that for $\alpha, \beta \leqslant k$ and $\alpha + \beta \leqslant p+1$, we have

$$\begin{aligned} c_{\alpha,\beta}^k &= \sum_{i=1}^{n}\sum_{j=1}^{n} \binom{i-1}{\alpha-1}\binom{j-1}{\beta-1} c_{i,j}^0 \\ &= \sum_{i=1}^{n} \binom{i-1}{\alpha-1}\binom{i-1}{\beta-1} w_i - \sum_{i=1}^{n}\sum_{j=1}^{n} \binom{i-1}{\alpha-1}\binom{j-1}{\beta-1} w_i w_j \\ &= \sum_{i=1}^{n} \binom{n-i}{\alpha-1}\binom{n-i}{\beta-1} w_i' - (\sum_{i=1}^{n} \binom{n-i}{\alpha-1} w_i')(\sum_{j=1}^{n} \binom{n-j}{\beta-1} w_j') \\ &= \binom{n-n'-\frac{1}{2}}{\alpha-1}\binom{n-n'-\frac{1}{2}}{\beta-1} - \binom{n-n'-\frac{1}{2}}{\alpha-1}\binom{n-n'-\frac{1}{2}}{\beta-1} \\ &= 0.\end{aligned}$$

We have completed the proof of Lemma 3.6. ∎