Linear Algebra and its Applications 602 (2020) 179–205

Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa

Manifold curvature learning from hypersurface integral invariants

Javier Álvarez-Vizoso [a,*], Michael Kirby [b], Chris Peterson [b]

[a] *Max-Planck-Institut für Sonnensystemforschung, Justus-von-Liebig-Weg 3, 37077 Göttingen, Germany*
[b] *Department of Mathematics, Colorado State University, 841 Oval Drive, Fort Collins, CO 80523, USA*

ARTICLE INFO

Article history:
Received 20 September 2019
Accepted 13 May 2020
Available online 21 May 2020
Submitted by P. Semrl

MSC:
53B99
55A07
62H25

Keywords:
Riemann curvature tensor
Principal component analysis
Local eigenvalue decomposition
Manifold learning

ABSTRACT

Integral invariants obtained from Principal Component Analysis on a small kernel domain of a submanifold encode important geometric information classically defined in differential-geometric terms. We generalize to hypersurfaces in any dimension major results known for surfaces in space, which in turn yield a method to estimate the extrinsic and intrinsic curvature tensor of an embedded Riemannian submanifold of general codimension. In particular, integral invariants are defined by the volume, barycenter, and the EVD of the covariance matrix of the domain. We obtain the asymptotic expansion of such invariants for a spherical volume component delimited by a hypersurface and for the hypersurface patch created by ball intersections, showing that the eigenvalues and eigenvectors can be used as multi-scale estimators of the principal curvatures and principal directions. This approach may be interpreted as performing statistical analysis on the underlying point-set of a submanifold in order to obtain geometric descriptors at scale with potential applications to Manifold Learning and Geometry Processing of point clouds.

* Corresponding author.
E-mail addresses: javizoso@alumni.colostate.edu (J. Álvarez-Vizoso), kirby@math.colostate.edu (M. Kirby), peterson@math.colostate.edu (C. Peterson).

https://doi.org/10.1016/j.laa.2020.05.020

1. Introduction

Manifold learning has as its prime goal the local characterization and reconstruction of manifold geometry from the study of the underlying point set, usually embedded as a submanifold in an ambient space, typically Euclidean. To obtain theoretical results that can serve as tools for this endeavor, it is assumed that the complete continuous point set is known so that local statistical invariants on given domains can be shown to be related to the relevant local geometry, whereas in practice only a finite cloud of points, probably with noise, is available. In geometry processing, the development of these methods provides us with descriptors that serve as geometry estimators, guide a possible reconstruction, or provide feature detectors. The integral invariant point of view attempts to overcome some of the difficulties of computational geometry when facing the task of extracting information that is classically defined as a differential invariant, like curvature, since its discrete version reduces to, e.g., sums instead of finite differences. The multi-scale behavior and averaging nature of these invariants is also of importance in applications and their possible stability and robustness with respect to noise.

Series expansion of the volume of small geodesic balls within a manifold [1], and volumes cut out by a hypersurface inside a ball of the ambient space [2], have been shown to be given in terms of the manifold curvature scalar invariants. In order to obtain local adaptive Galerkin bases for large-dimensional dynamical systems, the eigenvalue decomposition of covariance matrices of spherical intersection domains on the invariant manifold was introduced in [3], [4,5] to provide estimates of the dimension of the manifold and a suitable decomposition of phase space at every point. In the case of curves, the Frenet-Serret apparatus is recovered with explicit formulas at scale to obtain descriptors of the generalized curvatures in terms of the eigenvalues of the covariance matrix [6]. Integral invariants were already introduced and employed in geometry processing applications by [7], [8,9], [10,11], [12,13]. Local principal component analysis of this type has been studied primarily for the case of curves and surface in 2D and 3D in [10,11], [13], [14], [15], [16], [17], as a means to determine relevant local geometric information while maintaining stability with respect to noise [18], [19,20], e.g., for feature and shape detection using point clouds or meshes in computer graphics. Voronoi-based feature estimation [21,22] has also taken advantage of the PCA covariance matrix approach. Those methods study embedded manifolds whereas intrinsic probability and statistical analysis using geometric measurements inside a Riemannian manifold have also been developed [23,24] and could be used to do covariance analysis of submanifolds embedded in curved ambient spaces.

The complementary side of this framework is the study of finite point clouds and how their discrete PCA covariance matrices converge with the number of points to the exact analytical result of the smooth case, as studied in our work. Methods using geometric measure theory and harmonic analysis have been developed [25,26], [27,28] in order to study noisy samples from probability distributions supported on submanifolds of a high-dimensional Euclidean space [29]. In these works, ranges of scales are determined, taking into account curvature, for the covariance matrices to be most informative and close to the noisy empirical matrices. The approach of [29] is complemented by ours in the sense that we obtain explicitly the next to leading order terms of the eigenvalue expansion for the complete smooth data set providing the direct theoretical link between curvature and covariance. Since [29] develops an explicit algorithm for the estimation of the dimension of the manifold, a natural next step would be to expand these multiscale methods in order to apply them to our main theorems and thus to estimate curvature from noisy point clouds. Our descriptor algorithm to estimate the Riemann curvature provides the theoretical result to fulfill this task in practice.

In this paper we follow and generalize the major theoretical results of [20] for surfaces in space to hypersurfaces in any dimension, which in turn allows for the extension of their approach to obtain descriptors of the extrinsic and intrinsic curvature at a given scale for any Riemannian submanifold of general codimension in Euclidean space. Future work will show how the analysis for the ball intersection patch case further extends to general codimension, [30,31], establishing the connection between the generalized third fundamental form and the integral invariants, i.e. between local Riemannian geometry and local covariance integrals.

The structure of the paper is as follows: In section 2, PCA integral invariants and geometric descriptors are introduced to show how the study of hypersurfaces is sufficient to study the curvature of Riemannian submanifolds of any dimension by applying the analysis to k hypersurface projections (where k is the codimension of the submanifold). In section 3, an explicit toy example of the correspondence between the differential-geometric curvature and the integral invariant covariance is detailed. In section 4, these integral invariants are analytically computed for a volume region delimited by a hypersurface inside a ball; the asymptotic expansions of the invariants with respect to the scale of the ball are shown to be given in terms of the principal curvatures and the dimension, and the eigenvectors of the covariance matrix are shown to converge in the limit to the principal directions. In section 5, the analogous analysis is carried out for the integral invariants of the hypersurface patch cut out by the ball. In section 6, we see how these asymptotic formulas can be inverted to yield geometric descriptors at scale of the principal curvatures and principal directions for hypersurfaces, thus establishing concrete formulas to use in our final algorithm for curvature descriptors of Riemannian submanifolds. The notation and technical results needed for all computations are summarized in appendix A.

2. Integral invariants and descriptors

Our approach generalizes the theoretical part of the seminal work [20] with a focus on the analytical expansion of integral invariants to get descriptors of manifold curvature in any dimension. The local integral invariants considered are integrals over small kernel domains determined by balls and the hypersurface. In particular, we will focus on the Principal Component Analysis of a $(n+1)$-dimensional region delimited by the hypersurface inside a ball centered at a point on the hypersurface, and the n-dimensional patch on the submanifold cut out by such a ball. In general, one can define invariants for a measurable domain by computing the moments of the coordinates of the points inside, which leads us to

Definition 2.1. Let D be a measurable domain in $\mathbb{R}^n$, the *integral invariants* associated to the moments of order 0, 1 and 2 of the coordinate functions of the points of D are: the volume

$$V(D) = \mathbb{E}[1 \cdot \chi_D(\boldsymbol{X})] = \int_D 1 \text{ dVol}, \tag{1}$$

the barycenter

$$\boldsymbol{s}(D) = \mathbb{E}[\boldsymbol{X} \cdot \chi_D(\boldsymbol{X})] = \frac{1}{V(D)} \int_D \boldsymbol{X} \text{ dVol}, \tag{2}$$

and the eigenvalue decomposition of the covariance matrix

$$C(D) = \mathbb{E}[(\boldsymbol{X} - \boldsymbol{s}(D)) \otimes (\boldsymbol{X} - \boldsymbol{s}(D))^T \cdot \chi_D(\boldsymbol{X})] = \int_D (\boldsymbol{X} - \boldsymbol{s}(D)) \otimes (\boldsymbol{X} - \boldsymbol{s}(D))^T \text{ dVol}. \tag{3}$$

Here dVol is the measure on D induced by restriction of the Euclidean measure, and the tensor product is to be understood as the outer product of the components in a chosen basis. $\mathbb{E}$ represents taking the expectation value over all possible $\boldsymbol{X}$ in their domain, i.e. $\mathbb{R}^n$, and χ_D is the characteristic function of the set D (i.e., 1 if and only if $\boldsymbol{X} \in D$, zero otherwise).

An *integral invariant descriptor* $F(D)$ of some feature F of a measurable domain D is any expression for F completely given in terms of $V(D)$, $\boldsymbol{s}(D)$, the eigenvalue decomposition of $C(D)$ or other integral invariants. If the domain D is determined by a region of a hypersurface $\mathcal{S}$, the main geometric descriptors are any principal curvature estimators $\kappa_\mu(D)$ of $\kappa_\mu(p)$, and principal and normal direction estimators $\boldsymbol{e}_\mu(D)$, $\boldsymbol{N}(D)$ of $\boldsymbol{e}_\mu(p), \boldsymbol{N}(p)$, for some known point $p \in \mathcal{S}$. If the domain D is determined by a region of an embedded manifold $\mathcal{M}$, the main geometric descriptor is any second fundamental form estimator, $\mathbf{II}(D)$ of $\mathbf{II}_p$, for some known point $p \in \mathcal{M}$. Since our domain D of interest will possess a natural scale ε determined by the size of the ball that shall define it, we shall talk about *descriptors at scale*. Moreover, throughout all the paper we consider ε to be small enough so that we can approximate the hypersurface $\mathcal{S}$ by the local graph representation of its osculating quadric at p, which is sufficient to obtain the leading terms of the asymptotic expansions with scale of the integral invariants.

These descriptors become valuable tools to perform manifold learning, feature detection and shape estimation when only partial knowledge of the complete set of points is known or when noise is present. In this regard, [19,20,17] carried out experimental and theoretical analysis of the stability of these and other descriptors in the case of curves and surfaces in $\mathbb{R}^3$, reporting for example that the invariants of the spherical component domain are more robust with respect to noise than the patch region ones. It is to be expected that the same stability behavior holds in the hypersurface case due to the sensitivity to small changes of an n-dimensional patch compared to an $(n+1)$-dimensional volume of which the perturbed patch is only part of its boundary.

When the asymptotic expansions with respect to scale of hypersurface integral invariants are available to high enough order, curvature information can be extracted by truncating the series and inverting the relations in order to obtain a computable multi-scale estimator of the actual curvatures. In particular, the eigenvalues of the covariance matrix will provide such a descriptor for the principal curvatures of a smooth hypersurface, $\kappa_\mu(D)$, and its eigenvectors $\{\boldsymbol{e}_\mu(D)\}_{\mu=1}^n$, and $\boldsymbol{e}_{n+1}(D)$, will do the same for the normal direction. In order to produce analogous descriptors for an embedded Riemannian manifold of higher codimension, we just need to apply the procedure to the k hypersurfaces created by projecting the manifold down to $(n+1)$ linear subspaces determined by its tangent space and each of the normal directions.

Lemma 2.2. *Let $\mathcal{M} \subset \mathbb{R}^{n+k}$ be an n-dimensional embedded Riemannian manifold, and fix an orthonormal basis $\{\boldsymbol{e}_\mu\}_{\mu=1}^n$ of the tangent space $T_p\mathcal{M}$, and an orthonormal basis $\{\boldsymbol{N}_j\}_{j=1}^k$ of the normal space $N_p\mathcal{M}$ at $p \in \mathcal{M}$. Consider a ball $B_p^{(n+k)}(\varepsilon)$ for small enough $\varepsilon > 0$, such that the projections of $\mathcal{M} \cap B_p^{(n+k)}(\varepsilon)$ onto the linear subspaces $T_p\mathcal{M} \oplus \langle \boldsymbol{N}_i \rangle$, for all $i = 1, \dots, k$, are smooth hypersurfaces $\mathcal{S}_i$. Then, if $\kappa_\mu^{(i)}(D)$, $\{\boldsymbol{e}_\mu^{(i)}(D)\}_{\mu=1}^n$ are descriptors of the principal curvatures and principal directions at p for each of the hypersurfaces $\mathcal{S}_i$, then the second fundamental form of $\mathcal{M}$ at p has a descriptor:*

$$\mathbf{II}_p(D)(\boldsymbol{e}_\mu, \boldsymbol{e}_\nu) = \sum_{i=1}^{k} [V_i(D) K_i(D) V(D)_i^T]_{\mu\nu} \, \boldsymbol{N}_i \,, \qquad \mu, \nu = 1, \dots, n, \tag{4}$$

where $[V_i(D)]$ are the matrices whose columns are the components of $\{\boldsymbol{e}_\mu^{(i)}(D)\}_{\mu=1}^n$ in the chosen basis $\{\boldsymbol{e}_\mu\}_{\mu=1}^n$, and $[K_i(D)]$ is the diagonal matrix of principal curvature estimators. In turn, the Riemann curvature tensor of $\mathcal{M}$ at p acquires a descriptor:

$$\langle \boldsymbol{R}(D)(\boldsymbol{e}_\mu, \boldsymbol{e}_\nu)\boldsymbol{e}_\alpha, \, \boldsymbol{e}_\beta \rangle = \sum_{i=1}^{k} \left([V_i K_i V_i^T]_{\mu\beta} [V_i K_i V_i^T]_{\nu\alpha} - [V_i K_i V_i^T]_{\mu\alpha} [V_i K_i V_i^T]_{\nu\beta} \right). \tag{5}$$

In fact, the matrices $[V_i(D) K_i(D) V_i^T(D)]$ are a descriptor of the local Hessian of $\mathcal{S}_j$ at p.

Proof. By the implicit function theorem, there is a neighborhood of $U_p \subset T_p\mathcal{M}$ such that the manifold can be locally given by a graph $\boldsymbol{x} \mapsto (\boldsymbol{x}, f_1(\boldsymbol{x}), \dots, f_k(\boldsymbol{x}))$, where $\boldsymbol{x} \in U_p$, p corresponds to $\mathbf{0}$, and $\nabla f_i(\mathbf{0}) = \mathbf{0}$. From this, the projection hypersurfaces $\mathcal{S}_i$ are just $(\boldsymbol{x}, f_i(\boldsymbol{x}))$ within the linear subspace $T_p\mathcal{M} \oplus \langle N_i \rangle$, for $i = 1, \dots, k$. It can be shown

[32, vol. II ex. 3.3.] that the second fundamental form of $\mathcal{M}$ at p is precisely the linear combination of the second fundamental forms of each of the hypersurface projections weighed by the corresponding normal vector, i.e.,

$$\mathbf{II}_p(\boldsymbol{e}_\mu, \boldsymbol{e}_\nu) = \sum_{i=1}^{k} \left[\frac{\partial^2 f_i}{\partial x_\mu \partial x_\nu}(p) \right] \boldsymbol{N}_i$$

Analyzing each of those hypersurfaces in $T_p\mathcal{M} \oplus \langle \boldsymbol{N}_i \rangle \cong \mathbb{R}^{n+1}$, to obtain descriptors $\kappa_\mu^{(i)}(D)$, $\{\boldsymbol{e}_\mu^{(i)}(D)\}_{\mu=1}^n$ for every i, we obtain precisely a descriptor of the eigenvalue decomposition of each Hessian, i.e., Hess $f_i|_p(D) = [V_i(D)K_i(D)V(D)_i^T]$ is an estimator of the second fundamental form of $\mathcal{S}_i$ at p in the original basis. Applying Gauß equation

$$\langle \boldsymbol{R}(\boldsymbol{e}_\mu, \boldsymbol{e}_\nu)\boldsymbol{e}_\alpha,\, \boldsymbol{e}_\beta \rangle = \langle\, \mathbf{II}(\boldsymbol{e}_\mu, \boldsymbol{e}_\beta),\, \mathbf{II}(\boldsymbol{e}_\nu, \boldsymbol{e}_\alpha)\,\rangle - \langle\, \mathbf{II}(\boldsymbol{e}_\mu, \boldsymbol{e}_\alpha),\, \mathbf{II}(\boldsymbol{e}_\nu, \boldsymbol{e}_\beta)\,\rangle \tag{6}$$

yields a corresponding descriptor for the Riemann tensor. □

A concrete application of this result in a toy computation is presented in the next section. We summarize the steps for the applicability of these ideas for arbitrary embedded Riemannian manifolds in the following algorithm.

Algorithm 1 Curvature descriptors from ball intersections.

Input: Point-set $\mathcal{M} \subset \mathbb{R}^{n+k}$, point $p \in \mathcal{M}$, radius $\varepsilon > 0$
Output: 2nd fundamental form descriptor $\mathbf{II}_p(\varepsilon)$, Riemann tensor descriptor $R_{\mu\nu\alpha\beta}(\varepsilon)$ at p
if current basis is not known to split $T_p\mathcal{M} \oplus N_p\mathcal{M}$ into a tangent and normal basis **then**
- Find the eigenvectors $\{\boldsymbol{e}_\mu\}_{\mu=1}^n \cup \{\boldsymbol{N}_i\}_{i=1}^k$ of $\mathbb{E}[(\boldsymbol{X} - p) \otimes (\boldsymbol{X} - p)^T \cdot \chi_{B_p(\varepsilon)\cap\mathcal{M}}(\boldsymbol{X})]$
{Dimension n and split basis are determined by different scaling of eigenvalues, cf. [5]}
- Update the basis to the normalized eigenvector basis
end if
for $i = 1$ to n **do**
- Project $\mathcal{M}$ to a hypersurface $\mathcal{S}_i$ in the linear subspace $\langle \{\boldsymbol{e}_\mu\}_{i=1}^n, \boldsymbol{N}_i \rangle \cong \mathbb{R}^{n+1} \subset \mathbb{R}^{n+k}$
- Determine region $D_i := V_p^+(\varepsilon)$ (cf. Lemma 4.2) or $D_p(\varepsilon)$ (cf. section 5) of this $\mathcal{S}_i$
- Compute the induced volume $V(D_i) = \mathbb{E}[1 \cdot \chi_{D_i}(\boldsymbol{X})]$
- Compute the barycenter $\boldsymbol{s}(D_i) = \mathbb{E}[\boldsymbol{X} \cdot \chi_{D_i}(\boldsymbol{X})]/V(D_i)$
- Compute the eigenvectors $\{\boldsymbol{e}_\mu^{(i)}(\varepsilon)\}_{i=1}^n \cup \boldsymbol{N}_i(\varepsilon)$, and eigenvalues $\{\lambda_\mu^{(i)}(\varepsilon)\}_{\mu=1}^{n+1}$ of the covariance matrix $C(D_i) = \mathbb{E}[(\boldsymbol{X} - \boldsymbol{s}(D_i)) \otimes (\boldsymbol{X} - \boldsymbol{s}(D_i))^T \cdot \chi_{D_i}(\boldsymbol{X})]$
- Obtain the principal curvatures $\kappa_\mu^{(i)}(\varepsilon)$ from $\lambda_\mu^{(i)}(\varepsilon)$ using Corollary 6.1 or Corollary 6.2
- Set V_i to the matrix whose columns are $\{\boldsymbol{e}_\mu^{(i)}(\varepsilon)\}_{i=1}^n$
- Set K_i to the diagonal matrix of $\{\kappa_\mu^{(i)}(\varepsilon)\}_{\mu=1}^n$
- Determine the Hessian matrix descriptor at p of $\mathcal{S}_i$ by $II_p^{(i)}(\varepsilon) = [V_i \cdot K_i \cdot V_i^T]$
end for
- Obtain the 2nd fundamental form estimator: $\mathbf{II}_p(\varepsilon)(\boldsymbol{e}_\mu, \boldsymbol{e}_\nu) = \sum_{i=1}^k [II_p^{(i)}(\varepsilon)]_{\mu\nu}\, \boldsymbol{N}_i(\varepsilon)$
- Obtain the Riemann curvature tensor estimator:
$R_{\mu\nu\alpha\beta}(\varepsilon) = \langle \boldsymbol{R}(\varepsilon)(\boldsymbol{e}_\mu, \boldsymbol{e}_\nu)\boldsymbol{e}_\alpha,\, \boldsymbol{e}_\beta \rangle = \sum_{i=1}^k \Big([II_p^{(i)}(\varepsilon)]_{\mu\beta}[II_p^{(i)}(\varepsilon)]_{\nu\alpha} - [II_p^{(i)}(\varepsilon)]_{\mu\alpha}[II_p^{(i)}(\varepsilon)]_{\nu\beta} \Big)$

3. Example of the covariance-curvature correspondence

Let us study a simple analytic toy example to understand the integral invariant approach to differential geometry. The classical approach follows [33]. Let $\mathcal{M} \subset \mathbb{R}^4$ be an embedded smooth surface such that around a point $p \in \mathcal{M}$ it has a local graph expression in a neighborhood U_p given to second order by its osculating quadric $\mathcal{M}_p^{(2)}$ (indeed, note that curvature is fully determined by this second order truncation, but our general analysis will take into account the errors from the full series):

$$\mathcal{M}_p^{(2)} = \{(x,y,f_1,f_2) \in \mathbb{R}^4 \mid f_1(x,y) = -\frac{1}{2}x^2 - 3xy - \frac{1}{2}y^2, f_2(x,y) = \frac{9}{4}x^2 + \frac{\sqrt{3}}{2}xy + \frac{7}{4}y^2\}.$$

This can always be done for arbitrary dimension and we can think of $f_i(\boldsymbol{x})$ as the leading order truncation of the Taylor expansions of the local graph functions of $\mathcal{M}$ over $T_p\mathcal{M}$, with $\boldsymbol{x}$ as coordinates of this tangent space. Given an orthonormal tangent basis $\{\boldsymbol{e}_\mu\}_{\mu=1}^n$ and normal basis $\{\boldsymbol{N}_i\}_{i=1}^k$, the extrinsic curvature of $\mathcal{M}$ at p is encoded by the second fundamental form, determined by the Hessians $II_p^{(i)}$ of $f_i(\boldsymbol{x})$ as seen in the previous section. Thus, for this example it can be written as the following quadratic form with values in the normal space $N_p\mathcal{M}$, for any $\boldsymbol{a}, \boldsymbol{b} \in T_p\mathcal{M}$:

$$\mathbf{II}_p(\boldsymbol{a},\boldsymbol{b}) = \sum_{i=1}^{2} II_p^{(i)}(\boldsymbol{a},\boldsymbol{b})\boldsymbol{N}_i = \left(\boldsymbol{a}^T \cdot \begin{bmatrix} -1 & -3 \\ -3 & -1 \end{bmatrix} \cdot \boldsymbol{b}\right) \boldsymbol{N}_1 + \left(\boldsymbol{a}^T \cdot \begin{bmatrix} \frac{9}{2} & \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & \frac{7}{2} \end{bmatrix} \cdot \boldsymbol{b}\right) \boldsymbol{N}_2. \tag{7}$$

The mean curvature vector is then $\boldsymbol{H}_p = \operatorname{tr}\mathbf{II}_p = H_1\boldsymbol{N}_1 + H_2\boldsymbol{N}_2 = -2\boldsymbol{N}_1 + 8\boldsymbol{N}_2$, so we can consider $||\boldsymbol{H}_p|| = 2\sqrt{17} \approx 8.2462112512$ a scalar characterization of the extrinsic curvature at p. The (intrinsic) Riemann curvature tensor for a surface, even embedded in higher dimension, has only 1 independent component, the Gaußian curvature $R_{1221} = 7$. We can see this by computing the curvature operators at p in our orthonormal tangent basis using eq. (7) in eq. (6):

$$\boldsymbol{R}_p(\boldsymbol{e}_1,\boldsymbol{e}_2) = -\boldsymbol{R}_p(\boldsymbol{e}_2,\boldsymbol{e}_1) = \begin{bmatrix} 0 & -7 \\ 7 & 0 \end{bmatrix}, \qquad \boldsymbol{R}_p(\boldsymbol{e}_1,\boldsymbol{e}_1) = \boldsymbol{R}_p(\boldsymbol{e}_2,\boldsymbol{e}_2) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}. \tag{8}$$

The Ricci operator is $\boldsymbol{\mathcal{R}ic}_p = \begin{bmatrix} 7 & 0 \\ 0 & 7 \end{bmatrix}$, so the scalar curvature is $\mathcal{R}_p = \operatorname{tr}\boldsymbol{\mathcal{R}ic}_p = 14$, again the only independent intrinsic invariant (same as R_{1221} upon normalization by $\dim\mathcal{M}$, a convention which we do not follow here). These computations yield the classical geometric invariants determined by the differential structure of $\mathcal{M}$ around p.

From the integral invariant point of view, the aim is to obtain the same information by completely different means: using the eigenvalue decomposition of the covariance matrix of domains determined by the underlying point-set of $\mathcal{M}$ around p, i.e. instead of employing (covariant) derivatives on $\mathcal{M}$, the same information is hidden within the asymptotics of integrals over local domains. Conceptually this signifies that we need only know the measure function over a domain within $\mathcal{M}$, i.e. a means to evaluate expectation values $\mathbb{E}[\cdots\chi_D(\boldsymbol{X})]$, which is equivalent to knowing only one volume element function, e.g. $\sqrt{\det g(\boldsymbol{x})}d^n\boldsymbol{x}$ in local coordinates, instead of the $(n+k)$ embedding functions $\boldsymbol{X}(\boldsymbol{x})$, or the $n(n+1)/2$ components of the metric tensor $g_{\mu\nu} = \langle\frac{\partial\boldsymbol{X}}{\partial x^\mu}, \frac{\partial\boldsymbol{X}}{\partial x^\nu}\rangle$. Therefore the integral invariant approach furnishes in principle a dictionary between covariance and curvature that requires less actual knowledge about the parametrization focusing on the underlying point-set, thus providing a more adequate methodology when only a discrete point cloud sample is available.

In our toy example we can analytically show this correspondence to high precision since the embedding functions are known and the integral invariants of eq. (1), eq. (2) and eq. (3) can be expressed explicitly and computed numerically. Since there are two normal directions there are no canonical principal directions and principal curvatures for a surface in $\mathbb{R}^4$, but one such set for each normal direction. Choosing $\boldsymbol{N}_1$ and $\boldsymbol{N}_2$, the principal directions and curvatures are just the eigenvalue decomposition of the Hessian matrices in eq. (7). Therefore, we can consider the projected local (hyper)surfaces $\mathcal{S}_i$ given by the graphs $(\boldsymbol{x}, f_i(\boldsymbol{x}))$ in the linear subspaces $T_p\mathcal{M} \oplus \langle \boldsymbol{N}_i \rangle$, with corresponding principal directions $\{\boldsymbol{e}_\mu^{(i)}(p)\}_{\mu=1}^n$ and principal curvatures $\{\kappa_\mu^{(i)}(p)\}_{\mu=1}^n$, for $i = 1 \dots k$, here $k = 2$. This yields two local surfaces in two $\mathbb{R}^3$ subspaces of $\mathbb{R}^4$, with principal directions and principal curvatures given by the eigenvalue decomposition of each $\left[\frac{\partial^2 f_i}{\partial x^\mu \partial x^\nu}(0)\right]$:

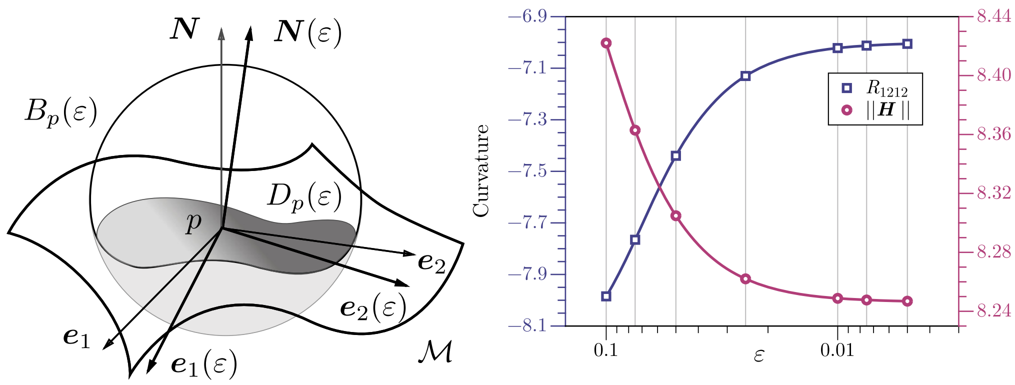
$$\kappa_1^{(1)}(p) = 2,\ \boldsymbol{e}_1^{(1)}(p) = \frac{1}{\sqrt{2}}(-1,\, 1, 0)^T \quad \text{and} \quad \kappa_2^{(1)}(p) = -4,\ \boldsymbol{e}_2^{(1)}(p) = \frac{1}{\sqrt{2}}(1,\, 1, 0)^T, \tag{9}$$

$$\kappa_1^{(2)}(p) = 3,\ \boldsymbol{e}_1^{(2)}(p) = (\frac{-1}{2}, \frac{\sqrt{3}}{2}, 0)^T \quad \text{and} \quad \kappa_2^{(2)}(p) = 5,\ \boldsymbol{e}_2^{(2)}(p) = (\frac{\sqrt{3}}{2}, \frac{1}{2}, 0)^T. \tag{10}$$

Determining these for each $\mathcal{S}_i$ from integral invariants makes it possible to recover the Hessian matrices and hence the full second fundamental form $\mathbf{II}_p$ of $\mathcal{M}$, and the Riemann tensor thereafter. As domain D_i for our descriptors we shall use the intersection regions $B_p^{(n+1)}(\varepsilon) \cap \mathcal{S}_i$, which cuts out a patch domain in $\mathcal{S}_i$ around p using the ambient space ball of radius $\varepsilon > 0$, see Fig. 1a and section 5. The moments of inertia of the shaded region in the figure yield approximations at scale ε to the principal and normal directions and the principal curvatures at p. In our example $D_i = \{(x, y, z) \in T_p\mathcal{M} \oplus \langle \boldsymbol{N}_i \rangle \mid z = f_i(x, y),\ x^2 + y^2 + f_i(x, y)^2 \leq \varepsilon^2\}$. These regions are well-defined for ε small enough which is sufficient to establish the asymptotic behavior of the integral invariants with the scale of the domain, as developed in the sections below, so that the covariance-curvature correspondence is well-defined in the limit to provide curvature estimators in the general case when the local graph expression of the manifold is of course unknown and the domain may be bigger. In that case the D_i are known only as point-sets and the expectation values can only be computed numerically. According to our results in section 5, the volumes and barycenters to leading order are:

$$Vol(D_1) = \pi\varepsilon^2(1 + \frac{9}{8}\varepsilon^2 + \dots), \qquad Vol(D_2) = \pi\varepsilon^2(1 + \frac{1}{8}\varepsilon^2 + \dots),$$

$$\boldsymbol{s}(D_1) = (0, 0, \frac{-\varepsilon^2}{4} + \dots)^T, \qquad \boldsymbol{s}(D_2) = (0, 0, \varepsilon^2 + \dots)^T.$$



(a) Principal direction descriptors at scale ε (b) Curvature descriptors at scale ε

Fig. 1. (a) The covariance EVD of domains determined by ball intersections at scale provide descriptors that approximate the principal and normal directions and principal curvatures of a hypersurface at a generic point. (b) In general codimension, this covariance analysis for projection hypersurfaces of the embedded manifold can be used to estimate the 2nd fundamental form and Riemann tensor. The simple example of the text shows the asymptotic convergence to the exact value of extrinsic and intrinsic curvature ($||\boldsymbol{H}||$ and R_{1212} resp.)

More importantly, the covariance matrices in this example have the following analytical form:

$$C(D_i) = \mathbb{E}[(\boldsymbol{X} - \boldsymbol{s}(D_i)) \otimes (\boldsymbol{X} - \boldsymbol{s}(D_i))^T \cdot \chi_D(\boldsymbol{X})]$$

$$= \int_{x^2+y^2+f_i^2 \leq \varepsilon^2} \begin{bmatrix} x^2 & xy & x(f_i - \frac{H_i\varepsilon^2}{8}) \\ yx & y^2 & y(f_i - \frac{H_i\varepsilon^2}{8}) \\ x(f_i - \frac{H_i\varepsilon^2}{8}) & y(f_i - \frac{H_i\varepsilon^2}{8}) & (f_i - \frac{H_i\varepsilon^2}{8})^2 \end{bmatrix} \sqrt{\det g_i(x,y)}\, dxdy,$$

where the induced volume elements on $\mathcal{S}_i$ are:

$$\mathrm{dVol}_1 = \sqrt{\det g_1(x,y)}\, dxdy = \sqrt{1 + 10x^2 + 12xy + 10y^2}\, dxdy,$$

$$\mathrm{dVol}_2 = \sqrt{\det g_2(x,y)}\, dxdy = \sqrt{1 + 21x^2 + 8\sqrt{3}xy + 13y^2}\, dxdy.$$

For instance, for $\varepsilon = 0.01$ numerical integration yields:

$$C(D_1) = \begin{bmatrix} 7.8544392039 \cdot 10^{-9} & -3.9289733304 \cdot 10^{-13} & -1.6500538629 \cdot 10^{-49} \\ -3.9289733304 \cdot 10^{-13} & 7.8544392039 \cdot 10^{-9} & -3.2345707458 \cdot 10^{-49} \\ -1.6500538629 \cdot 10^{-49} & -3.2345707458 \cdot 10^{-49} & 1.2431757559 \cdot 10^{-12} \end{bmatrix},$$

$$C(D_2) = \begin{bmatrix} 7.8516906374 \cdot 10^{-9} & -4.5347390337 \cdot 10^{-13} & 9.9391924526 \cdot 10^{-50} \\ -4.5347390337 \cdot 10^{-13} & 7.8522142640 \cdot 10^{-9} & 5.3725545096 \cdot 10^{-49} \\ 9.9391924526 \cdot 10^{-50} & 5.3725545096 \cdot 10^{-49} & 1.1769570544 \cdot 10^{-12} \end{bmatrix}.$$

From Theorem 5.4 below their respective eigenvectors approximate the principal and normal directions of $\mathcal{S}_i$ at this scale, and the corresponding eigenvalues serve as input for the formulas of Corollary 6.2 to estimate the principal curvatures as well:

$$\kappa_1^{(1)}(\varepsilon) = 1.9947, \qquad \boldsymbol{e}_1^{(1)}(\varepsilon) = (-0.7071067812,\ 0.7071067812,\ -3.238 \cdot 10^{-41})^T$$
$$\kappa_2^{(1)}(\varepsilon) = -3.99306, \quad \boldsymbol{e}_2^{(1)}(\varepsilon) = (0.7071067812,\ 0.7071067812,\ 2.750 \cdot 10^{-41})^T$$
$$\kappa_1^{(2)}(\varepsilon) = 2.99845, \qquad \boldsymbol{e}_1^{(2)}(\varepsilon) = (-0.5000000000,\ -0.8660254038,\ 5.097 \cdot 10^{-35})^T$$
$$\kappa_2^{(2)}(\varepsilon) = 4.9982, \qquad \boldsymbol{e}_2^{(2)}(\varepsilon) = (0.8660254038,\ 0.5000000000,\ 3.696 \cdot 10^{-35})^T.$$

Compare these descriptive values at scale $\varepsilon = 0.01$ with the exact analytical values of eq. (9) and eq. (10). Now, since these are supposed to be as well the eigenvalue decomposition of the Hessian matrices of f_i, we have essentially achieved an integral reconstruction of the second fundamental form of $\mathcal{M}$ at p, at scale $\varepsilon = 0.01$:

$$\mathbf{II}_p(\varepsilon) = \begin{bmatrix} -0.9991798599 & -2.9938785021 \\ -2.9938785021 & -0.9991798599 \end{bmatrix} \boldsymbol{N}_1 + \begin{bmatrix} 4.4982644881 & 0.8659210756 \\ 0.8659210756 & 3.4983849558 \end{bmatrix} \boldsymbol{N}_2,$$

so $||\boldsymbol{H}_p(\varepsilon)|| = 8.2425629448$. Therefore the procedure yields an estimation of the Riemann curvature tensor:

$$\boldsymbol{R}_p(\varepsilon)(\boldsymbol{e}_1, \boldsymbol{e}_2) = -\boldsymbol{R}_p(\varepsilon)(\boldsymbol{e}_2, \boldsymbol{e}_1) = \begin{bmatrix} 0 & -7.02189341 \\ 7.02189341 & 0 \end{bmatrix},$$

to be compared with the exact tensors of eq. (7) and eq. (8).

By repeating this procedure at different scales ε one can clearly see the convergence of the covariance-curvature correspondence in this example, cf. Fig. 1b. The steps followed here generalize to any dimension and are summarized in Algorithm 1. In the rest of the paper we develop the technical machinery needed to establish the validity, formulae and algorithm of this asymptotic correspondence, providing the theoretical foundation of this methodology for practical applications in manifold learning.

4. Hypersurface spherical component integral invariants

The following domain is introduced in [2] to study the relation between the mean curvature of hypersurfaces and the volume of sections of balls (we reserve their notation $B_p^+(\varepsilon)$ for the half-ball).

Definition 4.1. Let $\mathcal{S}$ be a smooth hypersurface in $\mathbb{R}^{n+1}$ with a locally chosen normal vector field $\boldsymbol{N} : \mathcal{S} \to \mathbb{R}^{n+1}$. Let $B_p^{(n+1)}(\varepsilon)$ be a ball of radius $\varepsilon > 0$ centered at a point $p \in \mathcal{S}$, for small enough ε the hypersurface always separates this ball into two connected components. Define the region $V_p^+(\varepsilon)$ to be that component such that $\boldsymbol{N}(p)$ is oriented towards its interior.

All the methods and results of [20] for surfaces using this domain generalize because to approximate integrals of functions over this type of region in $\mathbb{R}^3$, the formula developed in their work makes use of the hypersurface approximations of [2], valid in any dimension.

Note we start to use the notation from the appendix. Also, in the rest of the paper we shall say that a smooth function $f(x)$ has order $\mathcal{O}(x^n)$ if its local power series around $x = 0$ begins at order n, i.e., the nth-order derivative at $x = 0$ is nonzero whereas all the lower order derivatives are zero at that point; this generalizes to functions of several variables by referring to its first nonzero higher-order partial derivative at the center of the expansion.

Lemma 4.2. *Let $f : \mathbb{R}^{n+1} \to \mathbb{R}$ be a function of order $\mathcal{O}(\rho^k z^l)$ in cylindrical coordinates $\boldsymbol{X} = (\boldsymbol{x}, z) = (\rho\overline{\boldsymbol{x}}, z)$, $\overline{\boldsymbol{x}} \in \mathbb{S}^{n-1}$, let $\mathcal{S}$ be a graph hypersurface given by the function $z(\boldsymbol{x})$ whose normal at the origin points in the positive z-axis, and $V_p^+(\varepsilon)$ the spherical component delimited by this $\mathcal{S}$, then*

$$\int\limits_{V_p^+(\varepsilon)} f(\boldsymbol{X})\mathrm{dVol} = \int\limits_{B_p^+(\varepsilon)} f(\boldsymbol{X})\mathrm{dVol} - \int\limits_{B_p^n(\varepsilon)} \left[\int\limits_{z=0}^{z=\frac{1}{2}\sum_{\mu=1}^n \kappa_\mu x_\mu^2} f(\boldsymbol{x}, z)\, dz \right] d^n\boldsymbol{x} + \mathcal{O}(\varepsilon^{k+2l+n+3}) \tag{11}$$

where the half-ball $B_p^+(\varepsilon)$ consists of the points of $B_p^{n+1}(\varepsilon)$ such that $z \geq 0$.

Proof. We approximate $z(\boldsymbol{x})$ by its osculating quadric at the origin, $\frac{1}{2}\sum_{\mu=1}^n \kappa_\mu x_\mu^2$, and remove from the complete half-ball integral of $f(\boldsymbol{X})$ its contribution from below the paraboloid. The exact integration domain is determined by the sphere intersection with the hypersurface, $\{\|\boldsymbol{x}\|^2 + z(\boldsymbol{x})^2 \leq \varepsilon^2\}$, and what can be computed exactly is the integral over the cylinder $\{\rho \leq \varepsilon\}$, so that for every $\boldsymbol{x} \in B_p^{(n)}(\varepsilon) \subset T_p\mathcal{S}$, we can remove the contribution of $\int_0^z f(\boldsymbol{x}, z)dz$. Then:

$$\int\limits_{V_p^+(\varepsilon)} f(\boldsymbol{X})\mathrm{dVol} \approx \int\limits_{B_p^+(\varepsilon)} f(\boldsymbol{X})\mathrm{dVol} \; - \int\limits_{B_p^n(\varepsilon)} \left[\int\limits_{z=0}^{z(\boldsymbol{x})} f(\boldsymbol{x}, z(\boldsymbol{x}))\, dz \right] d^n\boldsymbol{x}.$$

We need to find the order of the error in this approximation. The volume in the second integral extends outside the ball that defines $V_p^+(\varepsilon)$, which is inscribed in the cylinder, and thus the integral below the hypersurface is subtracting an extra contribution from the region Ω, that lies outside the sphere but inside the cylinder and is bounded by the hypersurface. Thus

$$\int\limits_{\Omega} f(\boldsymbol{X})\mathrm{dVol} \;\; \leq \;\; \max_{\boldsymbol{X}\in\Omega} |f(\boldsymbol{X})| \cdot \mathrm{Vol}(\Omega).$$

Since $z(\rho\overline{\boldsymbol{x}}) \sim \mathcal{O}(\rho^2)$, then $\max_{\boldsymbol{X}\in\Omega} |f(\boldsymbol{X})| \sim \mathcal{O}(\rho^k(\rho^2)^l)$. To bound the volume of Ω, notice ρ is bounded by ε from the cylinder and by approximately $\varepsilon - C\varepsilon^3$ from the intersection of the sphere with the hypersurface, for some constant C (cf. Lemma 5.1 below or the estimation in [2]). This maximum thickness $\mathcal{O}(\varepsilon^3)$ is added up for every point of the base sphere, whose area is $\sim \mathcal{O}(\varepsilon^{n-1})$. Now, the maximum height in the z direction of Ω is of order $\mathcal{O}(\varepsilon^2)$ because it is given by the intersection of the cylinder with the hypersurface. Therefore, $\mathrm{Vol}(\Omega) \sim \mathcal{O}(\varepsilon^2\varepsilon^{n-1}\varepsilon^3) \sim \mathcal{O}(\varepsilon^{n+4})$. The total error of this approximation is then $\mathcal{O}(\varepsilon^{k+2l+n+4})$. Finally, the graph function $z(\boldsymbol{x})$ is to be approximated as a quadric, truncating the terms $\mathcal{O}(\rho^3)$ from its Taylor series. This makes a new error in the second integral of our formula, given by the integral over the region in between the quadric and the actual hypersurface, which has height given by the $\mathcal{O}(\rho^3)$ Therefore, the integral we are neglecting by this truncation makes an error

$$\int\limits_{\mathbb{S}^{n-1}} \int\limits_{\rho=0}^{\rho=\varepsilon} \mathcal{O}(\rho^k(\rho^2)^l)\mathcal{O}(\rho^3)\rho^{n-1} d\rho\, d\,\mathbb{S} \sim \mathcal{O}(\varepsilon^{k+2l+n+3})$$

which is the leading order of the two errors studied for the original integral. □

This type of approximations were used by [2] to obtain the first integral invariant.

Proposition 4.3 *(Hulin and Troyanov). The volume of the spherical component cut by a hypersurface has the asymptotic expansion, with the mean curvature H_p appearing to second order:*

$$V(V_p^+(\varepsilon)) = \frac{V_{n+1}(\varepsilon)}{2} - \frac{\varepsilon^2\, V_n(\varepsilon)}{2(n+2)} H_p + \mathcal{O}(\varepsilon^{n+3}). \tag{12}$$

Proposition 4.4. *The barycenter of the spherical component is of the form:*

$$\boldsymbol{s}(V_p^+(\varepsilon)) = [\,0, \ldots, 0,\ 2\frac{V_n(\varepsilon)}{V_{n+1}(\varepsilon)}\frac{\varepsilon^2}{n+2}\left(1 + \frac{V_n(\varepsilon)}{V_{n+1}(\varepsilon)}\frac{\varepsilon^2}{n+2}H_p\right)\,]^T\ + \mathcal{O}(\varepsilon^3). \tag{13}$$

Proof. Notice that $\int_{V_p^+(\varepsilon)} \boldsymbol{x}\, \mathrm{dVol} = \mathcal{O}(\varepsilon^{n+4})$ because applying Lemma 4.2, $\int_{B_p^+(\varepsilon)} \boldsymbol{x}\, d^n\boldsymbol{x}\, dz$ is zero, and the second integral is of monomials of odd degree. Then the normal component

$$\begin{aligned}[V(V_p^+(\varepsilon))\boldsymbol{s}(V_p^+(\varepsilon))]_z &= \int\limits_{B_p^+(\varepsilon)} z\, d^n\boldsymbol{x}\, dz - \int\limits_{B_p^n(\varepsilon)} \frac{1}{2}\left[\sum_{\mu=1}^{n} \kappa_\mu x_\mu^2\right]^2 d^n\boldsymbol{x} + \mathcal{O}(\varepsilon^{n+4}) \\ &= D_1^{(n+1)} + \mathcal{O}(\varepsilon^{n+4})\end{aligned}$$

where we have discarded the second integral since its order is $\mathcal{O}(D_4^{(n)}) = \mathcal{O}(D_{22}^{(n)}) \sim \mathcal{O}(\varepsilon^{n+4})$, which leaves the same order $\mathcal{O}(\varepsilon^3)$ as the error after dividing by the volume. The final expression follows from inverting the volume formula and using $D_1^{(n+1)}$ from the appendix. □

Theorem 4.5. *The covariance matrix $C(V_p^+(\varepsilon))$ has eigenvalues with the following series expansion, for all $\mu = 1, \dots, n$:*

$$\lambda_\mu(V_p^+(\varepsilon)) = V_{n+1}(\varepsilon)\frac{\varepsilon^2}{2(n+3)} - V_n(\varepsilon)\frac{\varepsilon^4}{2(n+2)(n+4)}(2\kappa_\mu(p) + H_p) + \mathcal{O}(\varepsilon^{n+5}), \quad (14)$$

$$\lambda_{n+1}(V_p^+(\varepsilon)) = V_{n+1}(\varepsilon)\frac{\varepsilon^2}{2(n+3)} - 2\frac{V_n(\varepsilon)^2}{V_{n+1}(\varepsilon)}\frac{\varepsilon^4}{(n+2)^2}\left(1 + \frac{V_n(\varepsilon)}{V_{n+1}(\varepsilon)}\frac{\varepsilon^2}{n+2}H_p\right) + \mathcal{O}(\varepsilon^{n+5}). \quad (15)$$

Moreover, in the limit $\varepsilon \to 0^+$, when the principal curvatures are different, the corresponding eigenvectors $\boldsymbol{e}_\mu(V_p^+(\varepsilon))$ converge linearly to the principal directions of $\mathcal{S}$ at p, and $\boldsymbol{e}_{n+1}(V_p^+(\varepsilon))$ converges quadratically to the hypersurface normal vector $\boldsymbol{N}$ at p.

Proof. Working in the basis formed by the principal directions and the normal vector of the hypersurface at the fixed point p, we shall compute the entries of the covariance matrix and see that it is diagonal to all orders smaller than $\mathcal{O}(\varepsilon^{n+5})$, precisely the error we get in the diagonal elements, therefore the eigenvalues coincide with those diagonal terms up to that error since differences between eigenvalues of symmetric matrices are bounded by the matrix norm metric. The covariance matrix splits into the first two terms of

$$C(V_p^+(\varepsilon)) = \int\limits_{V_p^+(\varepsilon)} \boldsymbol{X}\otimes\boldsymbol{X}^T \,\mathrm{dVol} - \int\limits_{V_p^+(\varepsilon)} \boldsymbol{X}\otimes\boldsymbol{s}^T \,\mathrm{dVol} - \int\limits_{V_p^+(\varepsilon)} \boldsymbol{s}\otimes\boldsymbol{X}^T \,\mathrm{dVol} + \int\limits_{V_p^+(\varepsilon)} \boldsymbol{s}\otimes\boldsymbol{s}^T \,\mathrm{dVol},$$

because the last three terms become the same upon integration. To compute the term left we can use the expression for $V\boldsymbol{s}$ from the proof of the barycenter formula to get:

$$\int\limits_{V_p^+(\varepsilon)} \boldsymbol{s}\otimes\boldsymbol{s}^T \,\mathrm{dVol} = V(V_p^+(\varepsilon))\boldsymbol{s}\otimes\boldsymbol{s}^T = \left[\begin{array}{c|c} \mathcal{O}(\varepsilon^{n+7})_{n\times n} & \mathcal{O}(\varepsilon^{n+5})_{n\times 1} \\ \hline \mathcal{O}(\varepsilon^{n+5})_{1\times n} & V(V_p^+(\varepsilon))s_z^2 \end{array}\right]$$

where $V(V_p^+(\varepsilon))s_z^2 = \dfrac{[D_1^{(n+1)}]^2}{V(V_p^+(\varepsilon))} + \mathcal{O}(\varepsilon^{n+5})$. The other contribution to the last matrix entry is

$$\begin{aligned}\int\limits_{V_p^+(\varepsilon)} z^2 \,\mathrm{dVol} &= \int\limits_{B_p^+(\varepsilon)} z^2\, d^n\boldsymbol{x}\, dz - \frac{1}{24}\int\limits_{B_p^n(\varepsilon)} \left[\sum_{\mu=1}^{n}\kappa_\mu x_\mu^2\right]^3 d^n\boldsymbol{x} + \mathcal{O}(\varepsilon^{n+7}) \\ &= \frac{D_2^{(n+1)}}{2} + \mathcal{O}(\varepsilon^{n+6}),\end{aligned}$$

in which we have neglected the second integral for being of higher order than the barycenter matrix error, whose subtraction yields the stated result for the normal eigenvalue. Notice that the other elements in the last column and row of the complete covariance matrix are $\mathcal{O}(\varepsilon^{n+5})$ since the remaining contributions come from $\int_{V_p^+(\varepsilon)} x_\mu z \text{ dVol} \sim \mathcal{O}(\varepsilon^{n+6})$, and its approximation formula has all monomials with odd powers in x.

Now, we compute the tangent coordinates block. This can be done at once for any $\mu, \nu = 1, \ldots, n$, noticing that when $\mu \neq \nu$, the integrals of Lemma 4.2 are of monomials of odd degree in tangent coordinates so the off-diagonal elements are $\mathcal{O}(\varepsilon^{n+5})$, (we use that $D_4 = 3D_{22}$):

$$\begin{aligned}\int\limits_{V_p^+(\varepsilon)} x_\mu^2 \text{ dVol} &= \int\limits_{B_p^+(\varepsilon)} x_\mu^2 \, d^n\boldsymbol{x}\, dz - \int\limits_{B_p^n(\varepsilon)} x_\mu^2 \left(\frac{1}{2}\sum_{\alpha=1}^{n} \kappa_\alpha x_\alpha^2\right) d^n\boldsymbol{x} + \mathcal{O}(\varepsilon^{n+5}) \\ &= \frac{D_2^{(n+1)}}{2} - \frac{D_4^{(n)}}{2}\kappa_\mu - \frac{D_{22}^{(n)}}{2}\sum_{\alpha\neq\mu}\kappa_\alpha + \mathcal{O}(\varepsilon^{n+5}) \\ &= \frac{D_2^{(n+1)}}{2} - \frac{D_{22}^{(n)}}{2}(2\kappa_\mu + H_p) + \mathcal{O}(\varepsilon^{n+5}).\end{aligned}$$

The perturbation theory of Hermitian matrices [34], [35] shows the convergence of the eigenvectors to the principal directions in the case of no multiplicity: truncating $C(V_p^+(\varepsilon))$ to order lower than $\mathcal{O}(\varepsilon^{n+5})$, that is precisely the order of the perturbation with respect to the exact diagonalized matrix. Fixing an eigenvalue $\lambda_\mu(V_p^+(\varepsilon))$ with $\mu \neq n+1$, the minimum difference to the other eigenvalues is of order $\sim \varepsilon^{n+4}(\kappa_\mu - \kappa_\nu)$, whereas for the last eigenvalue its distance to all the others is already at leading order $\sim \varepsilon^{n+3}$. Therefore, from the $\sin\theta$ theorem [34], the perturbation $\mathcal{O}(\varepsilon^{n+5})$ changes the eigenvectors $\{\boldsymbol{e}_\mu(V_p^+(\varepsilon))\}_{\mu=1}^n$ with respect to the principal directions as $\mathcal{O}(\varepsilon^{n+5})/\mathcal{O}(\varepsilon^{n+4}(\kappa_\mu - \kappa_\nu)) \sim \frac{\varepsilon}{\kappa_\mu - \kappa_\nu}$, and changes the eigenvector $\boldsymbol{e}_{n+1}(V_p^+(\varepsilon))$ with respect to the normal as $\mathcal{O}(\varepsilon^{n+5})/\mathcal{O}(\varepsilon^{n+3}) \sim \varepsilon^2$, i.e., in the limit $\varepsilon \to 0^+$ the eigenvectors of $C(V_p^+(\varepsilon))$ get a vanishing correction with respect to the principal directions. □

Therefore, we may write the covariance matrix as:

$$C(V_p^+(\varepsilon)) = \frac{V_{n+1}(\varepsilon)\,\varepsilon^2}{2(n+3)}\,\mathrm{Id}_{n+1} - \frac{V_n(\varepsilon)\,\varepsilon^4}{(n+2)(n+4)}\left(\begin{array}{c|c} \widehat{\boldsymbol{S}} + \frac{H}{2}\mathrm{Id}_n & 0_{n\times 1} \\ \hline 0_{1\times n} & 2\frac{V_n(\varepsilon)(n+4)}{V_{n+1}(\varepsilon)(n+2)}\end{array}\right) + \mathcal{O}(\varepsilon^{n+5}),$$

where the Weingarten operator $\widehat{\boldsymbol{S}}$ at p is $\mathrm{diag}(\kappa_1(p), \ldots, \kappa_n(p))$ in our basis.

5. Hypersurface patch integral invariants

Now, we shall compute the asymptotic expansions of the integral invariants of the hypersurface patch cut out by a ball centered at p and radius $\varepsilon > 0$, i.e. over the domain $D_p(\varepsilon) = \mathcal{S} \cap B_p^{n+1}(\varepsilon)$. Since a parametrization of the region is needed to perform the integrals locally, we need to find local parametric equations of the boundary $\partial(\mathcal{S} \cap$

$B_p^{n+1}(\varepsilon)$) to high enough order in ε so that we can expand asymptotically the integral invariants in terms of the geometric information of the hypersurface at the point. The strategy of [20], hinted in [2], obtaining a cylindrical coordinate approximation for the boundary radius of the patch, works in general dimension as follows.

Lemma 5.1. *In cylindrical coordinates $(\rho, \phi_1, \ldots, \phi_{n-1}, z)$ over the tangent space $T_p\mathcal{S}$, fixing the basis to the principal directions and the normal vector of $\mathcal{S}$ at p, the parametric equations of a point $\boldsymbol{X} = (\rho\overline{x}_1, \ldots, \rho\overline{x}_n, z)^T$ in $\partial D_p(\varepsilon) = \mathcal{S} \cap \mathbb{S}_p^n(\varepsilon)$, are*

$$r(\overline{\boldsymbol{x}}) := \rho(\overline{x}_1, \ldots, \overline{x}_n) = \varepsilon - \frac{1}{8}\kappa^2(\overline{\boldsymbol{x}})\varepsilon^3 + \mathcal{O}(\varepsilon^4), \qquad z(\overline{x}_1, \ldots, \overline{x}_n) = \frac{1}{2}\kappa^2(\overline{\boldsymbol{x}})\varepsilon^2 + \mathcal{O}(\varepsilon^3), \tag{16}$$

where $\overline{x}_1, \ldots, \overline{x}_n$ are the coordinates of points on $\mathbb{S}^{n-1} \subset T_p\mathcal{S}$, and $\kappa(\overline{\boldsymbol{x}}) = \kappa(\overline{x}_1, \ldots, \overline{x}_n) = \sum_{\mu=1}^n \kappa_\mu \overline{x}_\mu^2$ is the normal curvature of $\mathcal{S}$ at p cut by a normal plane in the direction of $\overline{\boldsymbol{x}}$.

Proof. In this coordinate system the expansion of the function that locally defines $\mathcal{S}$ is $z(\boldsymbol{x}) = \frac{1}{2}\sum_{\mu=1}^n \kappa_\mu x_\mu^2 + \mathcal{O}(x^3) = \frac{1}{2}\kappa(\overline{\boldsymbol{x}})\rho^2 + \mathcal{O}(\rho^3)$ since $x_\mu = \rho\overline{x}_\mu$, and because $\mathbf{II}_p$ is diagonal in our basis with $\overline{\boldsymbol{x}}$ a unit vector, the curve curvature cut by a normal plane is by Euler's formula $\kappa(\overline{\boldsymbol{x}}) = \mathbf{II}_p(\overline{\boldsymbol{x}}, \overline{\boldsymbol{x}}) = \sum_{\mu=1}^n \kappa_\mu \overline{x}_\mu^2$. Now, a point $\boldsymbol{X} = (\rho\overline{x}_1, \ldots, \rho\overline{x}_n, z)^T$ in $\mathcal{S} \cap \mathbb{S}_p^n(\varepsilon)$ satisfies the equation of the sphere $\rho^2 + z^2 = \varepsilon^2$. Substituting the expansion of $z(\boldsymbol{x})$ above, we obtain $\frac{1}{4}\kappa(\overline{\boldsymbol{x}})^2\rho^4 + \rho^2 - \varepsilon^2 + \mathcal{O}(\rho^5) = 0$, which up to order 4 is a biquadratic equation in ρ whose positive solution is the following and leads to the mentioned approximation:

$$\rho^2 = \frac{2}{\kappa(\overline{\boldsymbol{x}})^2}\left(-1 + \sqrt{1 + \kappa(\overline{\boldsymbol{x}})^2\varepsilon^2}\right) = \varepsilon^2 - \frac{1}{4}\kappa(\overline{\boldsymbol{x}})^2\varepsilon^4 + \mathcal{O}(\varepsilon^6),$$

then by extracting a common factor ε^2 and taking the square root, the approximation expression follows. □

The volume or mass of the domain can be expressed as a correction to the volume of the n-ball in terms of the extrinsic mean curvature H_p and intrinsic curvature $\mathcal{R}_p$ of $\mathcal{S}$ at the point, as it depends on the embedding. This compares to the case of the volume of a geodesic ball domain inside a manifold [1], which exhibits a correction only dependent on the intrinsic scalar curvature $\mathcal{R}_p$.

Proposition 5.2. *The n-dimensional area of the hypersurface patch expands as*

$$V(D_p(\varepsilon)) = V_n(\varepsilon)\left[1 + \frac{\varepsilon^2}{8(n+2)}(H_p^2 - 2\mathcal{R}_p) + \mathcal{O}(\varepsilon^3)\right]. \tag{17}$$

Proof. Computing the induced metric tensor using Lemma 5.1, the volume around p becomes

$$\mathrm{dVol}|_{D_p(\varepsilon)} = \sqrt{\det g(\boldsymbol{x})}\, dx_1 \cdots dx_n = \left[1 + \frac{1}{2}\sum_{\mu=1}^{n} \kappa_\mu^2 x_\mu^2 + \mathcal{O}(x^3)\right] dx_1 \cdots dx_n,$$

since $\|\nabla z(\boldsymbol{x})\|^2$ can be considered small for small enough $\varepsilon > 0$, because in our coordinates $\nabla z(\mathbf{0}) = 0$. With this and the cylindrical measure, eq. (A.1), the integration becomes

$$V(D_p(\varepsilon)) = \int\limits_{\mathcal{S}\cap B_p^{n+1}(\varepsilon)} \mathrm{dVol} = \int\limits_{\mathbb{S}^{n-1}} d\,\mathbb{S} \int\limits_0^{r(\overline{\boldsymbol{x}})} \left[1 + \frac{1}{2}\sum_{\mu=1}^{n} \kappa_\mu^2\, \rho^2 \overline{x}_\mu^2 + \mathcal{O}(\rho^3)\right] \rho^{n-1}\, d\rho$$

$$= \int\limits_{\mathbb{S}^{n-1}} d\,\mathbb{S} \left[\frac{1}{n}(\varepsilon - \frac{\kappa(\overline{\boldsymbol{x}})^2 \varepsilon^3}{8} + \mathcal{O}(\varepsilon^4))^n \right.$$

$$\left. + \frac{1}{2}\sum_{\mu=1}^{n} \frac{\kappa_\mu^2 \overline{x}_\mu^2}{n+2}(\varepsilon - \frac{\kappa(\overline{\boldsymbol{x}})^2 \varepsilon^3}{8} + \mathcal{O}(\varepsilon^4))^{n+2} + \mathcal{O}(\varepsilon^{n+4})\right]$$

after integrating over ρ up to the boundary radius. Expanding the binomial series and the square of the normal curvature, all the remaining integrals are in Theorem A.4, leading to

$$V(D_p(\varepsilon)) = V_n(\varepsilon) - \frac{\varepsilon^{n+2}}{8} \int\limits_{\mathbb{S}^{n-1}} d\,\mathbb{S} \left[\sum_{\mu=1}^{n} \kappa_\mu^2 \overline{x}_\mu^4 + 2\sum_{\mu<\nu}^{n} \kappa_\mu \kappa_\nu \overline{x}_\mu^2 \overline{x}_\nu^2\right] + \frac{C_2\, \varepsilon^{n+2}}{2(n+2)} \sum_{\mu=1}^{n} \kappa_\mu^2$$

$$+ \mathcal{O}(\varepsilon^{n+3})$$

$$= V_n(\varepsilon) + \frac{\varepsilon^{n+2}}{n+2}\left[\left(\frac{C_2}{2} - \frac{n+2}{8} C_4\right) \sum_{\mu=1}^{n} \kappa_\mu^2 - C_{22}\frac{n+2}{8} 2 \sum_{\mu<\nu}^{n} \kappa_\mu \kappa_\nu\right]$$

$$+ \mathcal{O}(\varepsilon^{n+3}),$$

where the final expression is obtained upon recognizing the mean and scalar curvature in terms of the principal curvatures, and using the relations among the coefficients from the appendix. □

The center of mass in this case turns out to deviate, to leading order in ε, only in the normal direction with respect to the center of the ball.

Proposition 5.3. *The barycenter of the patch region has coordinates in the principal basis with respect to p given by*

$$\boldsymbol{s}(D_p(\varepsilon)) = [\,\mathcal{O}(\varepsilon^4), \ldots, \mathcal{O}(\varepsilon^4),\ \frac{\varepsilon^2}{2(n+2)} H_p + \mathcal{O}(\varepsilon^3)\,]^T. \tag{18}$$

Proof. When integrating any tangent component x_α of $\boldsymbol{X}$, only factors with an odd power in some components are produced because the computable terms (see previous proof) now contain products $\overline{x}_\alpha \overline{x}_\mu^2$, $\overline{x}_\alpha \overline{x}_\mu^4$ and $\overline{x}_\alpha \overline{x}_\mu^2 \overline{x}_\nu^2$, which always have an odd power factor regardless of the subindices combination. Therefore the first n components of $V(D_p(\varepsilon))\boldsymbol{s}(D_p(\varepsilon))$ are of order $\mathcal{O}(\varepsilon^{n+4})$, coming from the error inside $r(\overline{\boldsymbol{x}})^{n+1}$ after integrating radially the first term $x_\alpha \rho^{n-1} d\rho$. The normal component of $\boldsymbol{X}$ integrates as

$$\int\limits_{\mathcal{S}\cap B_p^{n+1}(\varepsilon)} z \,\mathrm{dVol} = \int\limits_{\mathbb{S}^{n-1}} d\,\mathbb{S} \int\limits_0^{r(\overline{\boldsymbol{x}})} \left[\frac{1}{2}\kappa(\overline{\boldsymbol{x}})\rho^2 + \mathcal{O}(\rho^3)\right]\left[1 + \frac{1}{2}\sum_{\mu=1}^n \kappa_\mu^2 \rho^2 \overline{x}_\mu^2 + \mathcal{O}(\rho^3)\right] \rho^{n-1}\, d\rho$$

$$= \int\limits_{\mathbb{S}^{n-1}} d\,\mathbb{S}\left[\frac{\kappa(\overline{\boldsymbol{x}})}{2(n+2)}(\varepsilon - \frac{\kappa(\overline{\boldsymbol{x}})^2}{8}\varepsilon^3 + \mathcal{O}(\varepsilon^4))^{n+2} + \mathcal{O}(\varepsilon^{n+3})\right] = C_2 \frac{\varepsilon^{n+2}}{2(n+2)} H_p + \mathcal{O}(\varepsilon^{n+3}).$$

Then normalizing by the volume to lowest order cancels the coefficient $C_2\varepsilon^n$. □

Finally, the study of the covariance matrix of the patch domain shows a behavior similar to the spherical component, but where the next-to-leading order contribution to the eigenvalues includes only products of principal curvatures and no linear terms on them.

Theorem 5.4. *The covariance matrix $C(D_p(\varepsilon))$ has n eigenvalues that scale like ε^{n+2} as*

$$\lambda_\mu(D_p(\varepsilon)) = V_n(\varepsilon)\left[\frac{\varepsilon^2}{n+2} + \frac{\varepsilon^4}{8(n+2)(n+4)}(H_p^2 - 2\mathcal{R}_p - 4H_p\kappa_\mu(p))\right] + \mathcal{O}(\varepsilon^{n+5}), \quad (19)$$

for all $\mu = 1, \ldots, n$, and one eigenvalue scaling as ε^{n+4} with leading term

$$\lambda_{n+1}(D_p(\varepsilon)) = V_n(\varepsilon)\frac{\varepsilon^4}{2(n+2)(n+4)}\left(\frac{n+1}{n+2}H_p^2 - \mathcal{R}_p\right) + \mathcal{O}(\varepsilon^{n+5}). \quad (20)$$

Moreover, in the limit $\varepsilon \to 0^+$, if the principal curvatures at p are all different, the eigenvectors $\boldsymbol{e}_\mu(D_p(\varepsilon))$ corresponding to the first n eigenvalues converge to the principal directions of $\mathcal{S}$ at p, and the last eigenvector $\boldsymbol{e}_{n+1}(D_p(\varepsilon))$ converges to the hypersurface normal vector $\boldsymbol{N}(p)$.

Proof. We need to evaluate $\int_{D_p(\varepsilon)} \boldsymbol{X}(\boldsymbol{x}) \otimes \boldsymbol{X}(\boldsymbol{x})^T \sqrt{\det g}\, d^n\boldsymbol{x}$ and $V(D_p(\varepsilon))\boldsymbol{s}(D_p(\varepsilon)) \otimes \boldsymbol{s}(D_p(\varepsilon))^T$. The latter can be obtained from the previous proof:

$$[\,\mathcal{O}(\varepsilon^{n+4}), \ldots, \mathcal{O}(\varepsilon^{n+4}), \frac{C_2\varepsilon^{n+2}}{2(n+2)}H_p + \mathcal{O}(\varepsilon^{n+3})\,]^T$$

$$\otimes[\,\mathcal{O}(\varepsilon^4), \ldots, \mathcal{O}(\varepsilon^4), \frac{\varepsilon^2}{2(n+2)}H_p + \mathcal{O}(\varepsilon^3)],$$

resulting in all entries of the $n \times n$ block being $\mathcal{O}(\varepsilon^{n+8})$, the first n elements of the last column and last row being $\mathcal{O}(\varepsilon^{n+6})$, and the last element of the matrix becoming

$$[V(D_p(\varepsilon))\boldsymbol{s}(D_p(\varepsilon)) \otimes \boldsymbol{s}(D_p(\varepsilon))^T]_{(n+1),(n+1)} = \frac{V_n(\varepsilon)\,\varepsilon^4}{4(n+2)^2} H_p^2 + \mathcal{O}(\varepsilon^{n+5}),$$

(we already disregarded the term of $\mathcal{O}(\varepsilon^{n+6})$ that can be computed for this matrix entry because, as shown below, the other contributing term in that position has error $\mathcal{O}(\varepsilon^{n+5})$).

Now, the rest of the covariance matrix requires the longest computations so far. The entries of $\boldsymbol{X}(\boldsymbol{x}) \otimes \boldsymbol{X}(\boldsymbol{x})^T$ are of three types: $x_\mu x_\nu$, $x_\mu z(\boldsymbol{x})$ and $z(\boldsymbol{x})^2$. The first n entries of the last column and last row, $x_\mu z(\boldsymbol{x})$, contribute at order $\mathcal{O}(\varepsilon^{n+4})$. This implies that the matrix may not decompose at order $\mathcal{O}(\varepsilon^{n+4})$ as direct sum of a "tangent" $n \times n$ block, the integrals of $[x_\mu x_\nu]$, and a "normal" 1×1 block, the integral of $z(\boldsymbol{x})^2$. Hence, the argument in the proof of Theorem 4.5 to equate the diagonal elements of this expansion with that of the actual eigenvalues cannot be made here, since there are off-diagonal error elements at the same order as the diagonal approximation. Nevertheless, one can show, cf. [30], how these do not affect the eigenvalues at the order we are interested in by writing the eigenvalue-eigenvector equation as a series expansion order-by-order, which is always possible and converges for Hermitian matrices of converging power series elements [36], such as our $C(D_p(\varepsilon))$:

$$\begin{aligned}&[\, a\,\varepsilon^2 \left(\begin{array}{c|c} \mathrm{Id}_n & 0_{n\times 1} \\ \hline 0_{1\times n} & 0 \end{array}\right) + b\,\varepsilon^4 \left(\begin{array}{c|c} A_{n\times n} & B_{n\times 1} \\ \hline B_{1\times n} & C \end{array}\right) + \mathcal{O}(\varepsilon^5)\,][\,\boldsymbol{V}^{(0)} + \boldsymbol{V}^{(1)}\varepsilon + \boldsymbol{V}^{(2)}\varepsilon^2 + \dots] = \\ &= (\lambda^{(1)}\varepsilon^1 + \lambda^{(2)}\varepsilon^2 + \lambda^{(3)}\varepsilon^3 + \lambda^{(4)}\varepsilon^4 + \dots)[\,\boldsymbol{V}^{(0)} + \boldsymbol{V}^{(1)}\varepsilon + \boldsymbol{V}^{(2)}\varepsilon^2 + \dots].\end{aligned}$$

Therefore, expanding the components of $C(D_p(\varepsilon))$ shall yield exactly the actual eigenvalues to order $\mathcal{O}(\varepsilon^{n+4})$. The last matrix element expands the integrals into the following terms

$$\begin{aligned}\int\limits_{\mathcal{S}\cap B_p^{n+1}(\varepsilon)} z^2\,\mathrm{dVol} &= \frac{1}{4}\left[\sum_{\alpha=1}^{n}\kappa_\alpha^2 \int\limits_{\mathbb{S}^{n-1}} \overline{x}_\alpha^4\,d\,\mathbb{S} + 2\sum_{\alpha<\beta}^{n}\kappa_\alpha\kappa_\beta \int\limits_{\mathbb{S}^{n-1}} \overline{x}_\alpha^2\overline{x}_\beta^2\,d\,\mathbb{S}\right]\frac{\varepsilon^{n+4}}{n+4} + \mathcal{O}(\varepsilon^{n+5}) \\ &= \frac{\varepsilon^{n+4}}{4(n+4)}\left[C_4(H_p^2 - \mathcal{R}_p) + C_{22}\mathcal{R}_p\right] + \mathcal{O}(\varepsilon^{n+5}),\end{aligned}$$

whereof subtracting the barycenter matrix contribution, the last eigenvalue becomes

$$\lambda_{n+1}(p,\varepsilon) = \frac{C_2\,\varepsilon^{n+4}}{4(n+2)(n+4)}\left[3H_p^2 - 2\mathcal{R}_p\right] - \frac{C_2\,\varepsilon^{n+4}}{4(n+2)^2} H_p^2 + \mathcal{O}(\varepsilon^{n+5}).$$

The "tangent" block entries can be computed simultaneously for any $\mu, \nu = 1, \dots, n$:

$$\int_{\mathcal{S}\cap B_p^{n+1}(\varepsilon)} x_\mu x_\nu \,\mathrm{dVol} = \int_{\mathbb{S}^{n-1}} d\,\mathbb{S} \int_0^{r(\overline{\boldsymbol{x}})} \rho^2 \overline{x}_\mu \overline{x}_\nu \rho^{n-1} \left[1 + \frac{1}{2}\sum_{\alpha=1}^n \kappa_\alpha^2 \rho^2 \overline{x}_\alpha^2 + \mathcal{O}(\rho^3)\right] d\rho$$

$$= \frac{\varepsilon^{n+2}}{n+2}\left[\delta_{\mu\nu} C_2 - \frac{\varepsilon^2(n+2)}{8}\int_{\mathbb{S}^{n-1}} d\,\mathbb{S}\,\overline{x}_\mu \overline{x}_\nu \left(\sum_{\alpha=1}^n \kappa_\alpha^2 \overline{x}_\alpha^4 + 2\sum_{\alpha<\beta}^n \kappa_\alpha \kappa_\beta \overline{x}_\alpha^2 \overline{x}_\beta^2\right)\right] +$$

$$+ \frac{\varepsilon^{n+4}}{n+4}\frac{\delta_{\mu\nu}}{2}\left(\kappa_\mu^2 \int_{\mathbb{S}^{n-1}} \overline{x}_\mu^4 \, d\,\mathbb{S} + \sum_{\alpha\neq\mu}^n \kappa_\alpha^2 \int_{\mathbb{S}^{n-1}} \overline{x}_\alpha^2 \overline{x}_\mu^2 \, d\,\mathbb{S}\right) + \mathcal{O}(\varepsilon^{n+5}),$$

where the $\delta_{\mu\nu}$ appears because the monomials get an odd power if $\mu \neq \nu$. Now, the different integrals inside the indexed sums result in different constants depending on the different monomials that the terms $\overline{x}_\mu^2 \overline{x}_\alpha^4$ and $\overline{x}_\mu^2 \overline{x}_\alpha^2 \overline{x}_\beta^2$ can combine into, and after some algebraic manipulations the above integral is equal to

$$C_2 \frac{\varepsilon^{n+2}}{n+2}\delta_{\mu\nu} + \frac{\varepsilon^{n+4}}{n+4}\delta_{\mu\nu}\left[\left(\frac{C_4}{2} - \frac{n+4}{8}C_6\right)\kappa_\mu^2 + \left(\frac{C_{22}}{2} - \frac{n+4}{8}C_{24}\right)\sum_{\alpha\neq\mu}^n \kappa_\alpha^2 \right.$$
$$\left. - \frac{n+4}{8}\left(2C_{24}\sum_{\alpha\neq\mu}^n \kappa_\mu\kappa_\alpha + C_{222}\sum_{\substack{\alpha\neq\beta \\ \alpha,\beta\neq\mu}}^n \kappa_\alpha\kappa_\beta\right)\right] + \mathcal{O}(\varepsilon^{n+5}).$$

Notice that the summations in the last equation are all over indices that must be different from μ, so we can add and subtract the corresponding missing terms to those sums as long as we subtract them in the correct place. Doing this, and using the crucial relationships between the constants from the appendix, each of the different terms under the big braces simplify to:

$$\left(\frac{C_4}{2} - \frac{n+4}{8}C_6 - \frac{C_{22}}{2} + \frac{n+4}{8}C_{24}\right)\kappa_\mu^2 = -\frac{C_2}{2(n+2)}\kappa_\mu^2(p),$$

$$\left(\frac{C_{22}}{2} - \frac{n+4}{8}C_{24}\right)\sum_{\alpha=1}^n \kappa_\alpha^2 = \frac{C_2}{8(n+2)}(H_p^2 - \mathcal{R}_p),$$

$$-\frac{n+4}{8}\left((2C_{24} - 2C_{222})\sum_{\alpha\neq\mu}^n \kappa_\mu\kappa_\alpha + 2C_{222}\sum_{\alpha<\beta}^n \kappa_\alpha\kappa_\beta\right)$$
$$= -\frac{C_2}{2(n+2)}R_{\mu\mu}(p) + \frac{C_2}{8(n+2)}\mathcal{R}_p,$$

recalling that the diagonal components of the Ricci tensor for hypersurfaces are $R_{\mu\mu}(p) = \sum_{\alpha\neq\mu}^n \kappa_\alpha(p)\kappa_\mu(p)$, and the scalar curvature is $\mathcal{R}_p = 2\sum_{\mu<\nu}^n \kappa_\mu(p)\kappa_\nu(p)$. Finally, these add up into the expression

$$\int_{D_p(\varepsilon)} x_\mu x_\nu \mathrm{dVol} = \delta_{\mu\nu} V_n(\varepsilon)\left[\frac{\varepsilon^2}{n+2} + \frac{\varepsilon^4}{8(n+2)(n+4)}(H_p^2 - 2\mathcal{R}_p - 4\kappa_\mu^2 - 4R_{\mu\mu})\right] + \mathcal{O}(\varepsilon^{n+5}),$$

and since $\kappa_\mu^2(p) + R_{\mu\mu}(p) = \kappa_\mu(p) H_p$ the stated formula for the tangent eigenvalues follows from the diagonal of this block. Therefore, we can write $C(D_p^+(\varepsilon)) =$

$$\frac{V_n(\varepsilon)\,\varepsilon^2}{n+2}\left(\begin{array}{c|c} \mathrm{Id}_n & 0_{n\times 1} \\ 0_{1\times n} & 0 \end{array}\right) + \frac{V_n(\varepsilon)\,\varepsilon^4}{2(n+2)(n+4)}\left(\begin{array}{c|c} \frac{H^2-2\mathcal{R}}{4}\mathrm{Id}_n - H_p\,\widehat{\boldsymbol{S}} & A_{n\times 1} \\ \hline A_{1\times n} & \frac{n+1}{n+2}H_p^2 - \mathcal{R}_p \end{array}\right) + \mathcal{O}(\varepsilon^{n+5}),$$

so the Weingarten operator appears inside the covariance matrix in this case as well. □

6. Multi-scale curvature descriptors

By solving the second term in the expansion from our integral invariants, we can extract the curvature information they encode and write it in terms of the volume and eigenvalues at a fixed scale. This means that these local statistical measurements of the underlying point set can be employed to reconstruct or estimate its differential geometry, e.g., from a discrete sample cloud of points. These estimators can be used in geometry processing to ignore details below a given scale and act as feature detectors.

Employing the asymptotic expressions of section 4, we invert the relations and solve for the principal curvatures.

Corollary 6.1. *Abbreviating the integral invariants of the spherical component as* $\lambda_\mu(p,\varepsilon) \equiv \lambda_\mu(V_p^+(\varepsilon)), V_p(\varepsilon) \equiv V(V_p^+(\varepsilon))$*, then the corresponding descriptors of the principal curvatures, at scale* $\varepsilon > 0$ *and point* $p \in \mathcal{S}$*, are given by*

$$\kappa_\mu(V_p^+(\varepsilon)) = \frac{n+4}{\varepsilon^4 V_n(\varepsilon)}\left[\frac{\varepsilon^2 V_{n+1}(\varepsilon)}{n+3} - (n+1)\lambda_\mu(p,\varepsilon) + \sum_{\alpha\neq\mu}^{n}\lambda_\alpha(p,\varepsilon)\right], \tag{21}$$

or equivalently, using the mean curvature H*, by*

$$H(V_p^+(\varepsilon)) = \frac{(n+2)V_{n+1}(\varepsilon)}{\varepsilon^2 V_n(\varepsilon)}\left(1 - 2\frac{V_p(\varepsilon)}{V_{n+1}(\varepsilon)}\right), \tag{22}$$

$$\kappa_\mu(V_p^+(\varepsilon)) = \frac{(n+2)(n+4)}{\varepsilon^4 V_n(\varepsilon)}\left(\frac{\varepsilon^2 V_{n+1}(\varepsilon)}{2(n+3)} - \lambda_\mu(p,\varepsilon)\right) + \frac{1}{2}H(V_p^+(\varepsilon)), \tag{23}$$

with corresponding errors $|H_p - H(V_p^+(\varepsilon))| \leq \mathcal{O}(\varepsilon)$*, and* $|\kappa_\mu(p) - \kappa_\mu(V_p^+(\varepsilon))| \leq \mathcal{O}(\varepsilon)$*, for any* $\mu = 1,\ldots,n$*. The eigenvectors* $\boldsymbol{e}_\mu(V_p^+(\varepsilon))$ *and* $\boldsymbol{e}_{n+1}(V_p^+(\varepsilon))$ *are descriptors of the principal and normal directions respectively.*

Proof. Let us define the coefficients $a = \dfrac{\varepsilon^2 V_{n+1}(\varepsilon)}{2(n+3)}$, $b = -\dfrac{\varepsilon^4 V_n(\varepsilon)}{2(n+2)(n+4)}$, then the tangent eigenvalues from eq. (14) solve the principal curvatures

$$\kappa_\mu = \frac{\lambda_\mu - a}{2b} - \frac{1}{2}H_p + \mathcal{O}(\varepsilon).$$

Fixing one $\mu = 1, \dots, n$, and subtracting any two such equations with $\mu \neq \alpha$ results in

$$\kappa_\alpha = \frac{\lambda_\alpha - \lambda_\mu}{2b} + \kappa_\mu + \mathcal{O}(\varepsilon),$$

and inserting this into the definition of H yields

$$\kappa_\mu(V_p^+(\varepsilon)) = \frac{\lambda_\mu - a}{b(n+2)} - \sum_{\alpha\neq\mu}^{n} \frac{\lambda_\alpha - \lambda_\mu}{2b(n+2)} = \frac{1}{2b(n+2)}\left(-2a + (n+1)\lambda_\mu - \sum_{\alpha\neq\mu}^{n}\lambda_\alpha\right).$$

The truncation error is given by the order of $\mathcal{O}(\varepsilon^{n+5})/b \sim \mathcal{O}(\varepsilon)$. Alternatively, one can solve the Hulin-Troyanov relation, eq. (12), to obtain a descriptor of H_p, and then use this in the expression of κ_μ in terms of λ_μ and H above. □

An analogous inversion process can be carried out with the series expansions of section 5.

Corollary 6.2. *Denoting by $\lambda(p,\varepsilon) \equiv \lambda(D_p(\varepsilon)), V_p(\varepsilon) \equiv V(D_p(\varepsilon))$ the integral invariants of the hypersurface patch domain, then the corresponding curvature descriptors at scale $\varepsilon > 0$ and point $p \in \mathcal{S}$, for any $\mu = 1, \dots, n$, are*

$$\mathcal{R}(D_p^+(\varepsilon)) = 2(n+2)^2(n+4)\frac{\lambda_{n+1}(p,\varepsilon)}{n\,\varepsilon^4\,V_n(\varepsilon)} - \frac{8(n+1)(n+2)}{n\,\varepsilon^2}\left(\frac{V_p(\varepsilon)}{V_n(\varepsilon)} - 1\right) \tag{24}$$

$$H(D_p^+(\varepsilon)) = (\pm)\sqrt{4(n+2)^2(n+4)\frac{\lambda_{n+1}(p,\varepsilon)}{n\,\varepsilon^4 V_n(\varepsilon)} + \frac{8(n+2)^2}{n\,\varepsilon^2}\left(1 - \frac{V_p(\varepsilon)}{V_n(\varepsilon)}\right)}, \tag{25}$$

$$\kappa_\mu(D_p^+(\varepsilon)) = \frac{2(n+2)}{\varepsilon^2 H(D_p^+(\varepsilon))}\left[\frac{V_p(\varepsilon)}{V_n(\varepsilon)} + \frac{n+4}{\varepsilon^2}\left(\frac{\varepsilon^2}{n+2} - \frac{\lambda_\mu(p,\varepsilon)}{V_n(\varepsilon)}\right) - 1\right], \tag{26}$$

where the overall sign can be chosen by fixing a normal orientation from

$$(\pm) = \operatorname{sgn}\langle\, \boldsymbol{e}_{n+1}(D_p(\varepsilon)),\, \boldsymbol{s}(D_p(\varepsilon))\,\rangle.$$

The eigenvectors $\boldsymbol{e}_\mu(D_p(\varepsilon))$ and $\boldsymbol{e}_{n+1}(D_p(\varepsilon))$ are descriptors of the principal and normal directions respectively. The corresponding errors are $|H_p^2 - H(D_p(\varepsilon))^2| \leq \mathcal{O}(\varepsilon)$, $|\mathcal{R}_p - \mathcal{R}(D_p(\varepsilon))| \leq \mathcal{O}(\varepsilon)$, and $|\kappa_\mu^2(p) - \kappa_\mu(D_p(\varepsilon))^2| \leq \mathcal{O}(\varepsilon)$.

Proof. By solving the second term in eq. (17) and eq. (20), let us define coefficients

$$A = \frac{8(n+2)}{\varepsilon^2}\left(\frac{V_p(\varepsilon)}{V_n(\varepsilon)} - 1\right) + \mathcal{O}(\varepsilon), \quad B = 2(n+2)(n+4)\frac{\lambda_{n+1}(p,\varepsilon)}{\varepsilon^4 V_n(\varepsilon)} + \mathcal{O}(\varepsilon),$$

so that we have the system of equations $A = H_p^2 - 2\mathcal{R}_p$, $B = \frac{n+1}{n+2}H_p^2 - \mathcal{R}_p$, whose solution is

$$\mathcal{R}_p = \frac{1}{n}((n+2)B - (n+1)A), \quad H_p^2 = \frac{(n+2)}{n}(2B - A).$$

We can approximate the normal direction and orientation by using $\boldsymbol{e}_{n+1}(p,\varepsilon)$, and since the barycenter eq. (18) has normal component with leading order in terms of H_p, their mutual projection can serve to fix the orientation and overall relative sign of all the principal curvatures. The principal curvatures themselves are then solved from eq. (19) substituting the value of H_p above, resulting in $\kappa_\mu = \dfrac{1}{4H_p}(A - \Gamma_\mu)$, where

$$\Gamma_\mu = \frac{8(n+2)(n+4)}{\varepsilon^4}\left(\frac{\lambda_\mu(p,\varepsilon)}{V_n(\varepsilon)} - \frac{\varepsilon^2}{n+2}\right) + \mathcal{O}(\varepsilon).$$

The errors follow straightforwardly by the truncation of A, B, Γ_μ. □

In the spirit of the limit formula obtained in [6] for regular curves in $\mathbb{R}^n$, relating ratios of the covariance eigenvalues to the Frenet-Serret curvatures, we also state here analogous expressions for hypersurfaces using the ratios of the covariance eigenvalues, whose proofs are straightforward.

Corollary 6.3. *Let $p \in \mathcal{S}$ and consider the spherical component invariants. Then for any $\mu, \nu = 1, \ldots, n$, the first n eigenvalues, $\lambda_\mu(p,\varepsilon) \equiv \lambda_\mu(V_p^+(\varepsilon))$, of the covariance matrix $C(V_p^+(\varepsilon))$ satisfy the following limit ratio:*

$$\lim_{\varepsilon\to 0^+} \frac{V_{n+1}^2(\varepsilon)}{V_n(\varepsilon)} \frac{\lambda_\mu(p,\varepsilon) - \lambda_\nu(p,\varepsilon)}{\lambda_\mu(p,\varepsilon)\lambda_\nu(p,\varepsilon)} = \frac{4(n+3)^2}{(n+2)(n+4)}[\kappa_\nu(p) - \kappa_\mu(p)]. \tag{27}$$

Corollary 6.4. *Let $p \in \mathcal{S}$ and consider the hypersurface patch invariants. Then for any $\mu, \nu = 1, \ldots, n$, the first n eigenvalues, $\lambda_\mu(p,\varepsilon) \equiv \lambda_\mu(D_p(\varepsilon))$, of the covariance matrix $C(D_p(\varepsilon))$ satisfy the following limit ratio:*

$$\lim_{\varepsilon\to 0^+} V_n(\varepsilon)\frac{\lambda_\mu(p,\varepsilon) - \lambda_\nu(p,\varepsilon)}{\lambda_\mu(p,\varepsilon)\lambda_\nu(p,\varepsilon)} = \frac{n+2}{2(n+4)}[\kappa_\nu(p) - \kappa_\mu(p)]H_p, \tag{28}$$

and the last eigenvalue satisfies:

$$\lim_{\varepsilon\to 0^+} V_n(\varepsilon)\frac{\lambda_{n+1}(p,\varepsilon)}{\lambda_\mu(p,\varepsilon)\lambda_\nu(p,\varepsilon)} = \frac{n+2}{2(n+4)}\left[\frac{n+1}{n+2}H_p^2 - \mathcal{R}_p\right]. \tag{29}$$

These ratios can be used as well to define descriptors solving for the curvature variables aided by the volume descriptor, like in the preceding corollaries.

7. Conclusions

In this paper we have generalized major PCA methods and results known for surfaces in space to establish the asymptotic relationship between integral invariants and the principal curvatures and principal directions of hypersurfaces of any dimension, which furnishes a method to obtain geometric descriptors at any given scale using the eigenvalue decomposition of the covariance matrix. We have seen that these methods are sufficient to provide also estimators of the Riemann curvature tensor of embedded submanifolds of higher-codimension, using its hypersurface projections onto the linear subspaces in ambient space spanned by the tangent space and each of the normal vectors from an orthonormal basis. These results establish a theoretical foundation for the implementation of the computational integral invariant approach to study the geometry of point clouds of high dimensionality, which should be a helpful tool for manifold learning and geometry processing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank Louis Scharf for very helpful discussions during the writing of this paper. This paper is based on research partially supported by the National Science Foundation under Grants No. DMS-1513633, and DMS-1322508.

Appendix A. Integration of monomials over spheres

Let $\boldsymbol{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$, and denote the sphere and ball of radius ε in $\mathbb{R}^n$ by:

$$\mathbb{S}^{n-1}(\varepsilon) = \{\boldsymbol{x} \in \mathbb{R}^n : \|\boldsymbol{x}\| = \varepsilon\}, \quad B^n(\varepsilon) = \{\boldsymbol{x} \in \mathbb{R}^n : \|\boldsymbol{x}\| \leq \varepsilon\},$$

where we set $\mathbb{S}^{n-1} = \mathbb{S}^{n-1}(1)$. Using generalized spherical coordinates $(r, \phi_1, \ldots, \phi_{n-1})$, where $r = \|\boldsymbol{x}\|$, $\overline{x}_\mu = x_\mu/r \in \mathbb{S}^{n-1}$, i.e.,

$$\overline{x}_1 = \cos\phi_1, \ldots, \quad \overline{x}_{n-1} = \sin\phi_1 \cdots \sin\phi_{n-2}\cos\phi_{n-1}, \quad \overline{x}_n = \sin\phi_1 \cdots \sin\phi_{n-2}\sin\phi_{n-1},$$

the Euclidean measure over the unit sphere and ball of any radius can be written as

$$d\,\mathbb{S}^{n-1} = d\phi_{n-1} \prod_{\mu=1}^{n-2} \sin^{n-1-\mu}(\phi_\mu) d\phi_\mu, \qquad d^n B = dx_1 \cdots dx_n = r^{n-1} dr\, d\,\mathbb{S}^{n-1}. \quad \text{(A.1)}$$

Definition A.1. For any integers $\alpha_1, \dots, \alpha_n \in \{0, 1, 2, \dots\}$, the integrals of the monomials $x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ over the unit sphere and the ball of radius ε are denoted by:

$$C^{(n)}_{\alpha_1 \dots \alpha_n} = \int_{\mathbb{S}^{n-1}} x_1^{\alpha_1} \cdots x_n^{\alpha_n}\, d\,\mathbb{S}^{n-1}, \qquad D^{(n)}_{\alpha_1 \dots \alpha_n} = \int_{B^n(\varepsilon)} x_1^{\alpha_1} \cdots x_n^{\alpha_n}\, d^n B. \quad \text{(A.2)}$$

These can be computed directly in spherical coordinates by collecting factors and separating the integrals into a product of integrals of powers of sines and cosines which can be given in terms of the Beta function, that then telescopes and simplifies; other shorter proof uses the usual exponential trick, see for example [37], resulting in the following formula.

Theorem A.2. *Denoting $\beta_\mu = \frac{1}{2}(\alpha_\mu + 1)$, the values of the integrals eq. (A.2) over spheres are*

$$C^{(n)}_{\alpha_1 \dots \alpha_n} = \begin{cases} 0, & \text{if some } \alpha_\mu \text{ is odd}, \\ 2\dfrac{\Gamma(\beta_1)\Gamma(\beta_2)\cdots\Gamma(\beta_n)}{\Gamma(\beta_1 + \beta_2 + \cdots + \beta_n)}, & \text{if all } \alpha_\mu \text{ are even}, \end{cases} \quad \text{(A.3)}$$

and the integrals over balls become

$$D^{(n)}_{\alpha_1 \dots \alpha_n} = \frac{\varepsilon^{n + (\alpha_1 + \cdots + \alpha_n)}}{n + (\alpha_1 + \cdots + \alpha_n)}\, C^{(n)}_{\alpha_1 \dots \alpha_n}. \quad \text{(A.4)}$$

Notice that the values of the integrals of these monomials only depend on the combination of powers, not on which particular coordinates have those powers. Using these formulas we compute the relevant integrals that are needed for our work.

Remark A.3. Unless integrals over spheres of different dimension appear in the same expression, we shall abbreviate and omit the superscript $^{(n)}$ to be understood from the context.

Example A.4. Using the factorial property of the gamma function, $\Gamma(z + 1) = z\Gamma(z)$, and the value $\Gamma(\frac{1}{2}) = \sqrt{\pi}$, the integrals of monomials of even powers of order 2, 4 and 6, have the following relations (shortening $d\,\mathbb{S}^{n-1}$ as $d\,\mathbb{S}$):

$$C_2 = \int_{\mathbb{S}^{n-1}} x_1^2\, d\,\mathbb{S} = 2\frac{\Gamma(\frac{3}{2})\Gamma(\frac{1}{2})^{n-1}}{\Gamma(\frac{3}{2} + \frac{n-1}{2})} = \frac{\pi^{n/2}}{\Gamma(\frac{n}{2} + 1)}, \qquad C_{22} = \int_{\mathbb{S}^{n-1}} x_1^2 x_2^2\, d\,\mathbb{S} = \frac{C_2}{n+2},$$

$$C_4 = \int\limits_{\mathbb{S}^{n-1}} x_1^4 \, d\mathbb{S} = \frac{3\,C_2}{n+2} = 3\,C_{22}, \qquad C_{222} = \int\limits_{\mathbb{S}^{n-1}} x_1^2 x_2^2 x_3^2 \, d\mathbb{S} = \frac{C_2}{(n+2)(n+4)},$$

$$C_{24} = \int\limits_{\mathbb{S}^{n-1}} x_1^2 x_2^4 \, d\mathbb{S} = \frac{3\,C_2}{(n+2)(n+4)} = 3\,C_{222}, \quad C_6 = \int\limits_{\mathbb{S}^{n-1}} x_1^6 \, d\mathbb{S} = \frac{15\,C_2}{(n+2)(n+4)}.$$

The value of C_2 is related to the n-dimensional volume of the ball of radius ε, and the $(n-1)$-dimensional area of the unit sphere by $V_n(\varepsilon) = \mathrm{Vol}(B^n(\varepsilon)) = \varepsilon^n\, C_2$, and $S_{n-1} = \mathrm{Area}(\mathbb{S}^{n-1}) = n\, C_2$. The integrals over balls needed in our work are:

$$D_2 = \int\limits_{B^n(\varepsilon)} x_1^2 \;\, dx_1 \cdots dx_n = \frac{\varepsilon^{n+2}}{n+2} C_2 = \frac{\varepsilon^2}{n+2} V_n(\varepsilon),$$

$$D_{22} = \int\limits_{B^n(\varepsilon)} x_1^2 x_2^2 \;\, dx_1 \cdots dx_n = \frac{\varepsilon^{n+4}}{(n+2)(n+4)} C_2 = \frac{\varepsilon^4}{(n+2)(n+4)} V_n(\varepsilon),$$

$$D_4 = \int\limits_{B^n(\varepsilon)} x_1^4 \;\, dx_1 \cdots dx_n = \frac{3\,\varepsilon^{n+4}}{(n+2)(n+4)} C_2 = \frac{3\,\varepsilon^4}{(n+2)(n+4)} V_n(\varepsilon).$$

We also need the integral of monomials over half-balls $B^+(\varepsilon)$ (without loss of generality we can consider the half-ball is defined by $x_1 \geq 0$). If all the α_i are even then nothing changes in the proof of Theorem A.2 except that now we integrate over half the domain and an extra factor of $\frac{1}{2}$ is needed. If any α_i is odd for $i \neq 1$, the integration over those variables is still carried out over the same domain so the overall integral is still 0. However, if α_1 is odd the corresponding integral of that coordinate does not cancel out, and the main formula still holds with $\beta_1 = 1$ but without the factor of 2.

Example A.5. Using the formula in the mentioned adjusted form, we define and compute

$$D_1^{(n)} = \int\limits_{B^+(\varepsilon)} x_1 \;\, dx_1 \cdots dx_n = \frac{\varepsilon^{n+1}\, \pi^{\frac{n-1}{2}}}{2\Gamma(\frac{n+3}{2})},$$

which gives the constant needed in our main text $D_1^{(n+1)} = \int\limits_{B^+(\varepsilon)} x_1 \;\, dx_1 \cdots dx_{n+1} = \frac{\varepsilon^2}{n+2} V_n(\varepsilon)$. When integrating $\int_{B^+(\varepsilon)} x_1^2 \mathrm{dVol}$, we shall just write $\frac{D_2}{2}$ to be consistent with our notation.

References

[1] A. Gray, L. Vanhecke, Riemannian geometry as determined by the volumes of small geodesic balls, Acta Math. 142 (1979) 157–198.

[2] D. Hulin, M. Troyanov, Mean curvature and asymptotic volume of small balls, Am. Math. Mon. 110 (2003).
[3] D. Broomhead, R. Indik, A. Newell, D. Rand, Local adaptive Galerkin bases for large-dimensional dynamical systems, Nonlinearity 4 (2) (1991) 159–197.
[4] F. Solis, Geometric aspects of local adaptive Galerkin bases, PhD. Thesis, University of Arizona, 1993.
[5] F. Solis, Geometry of local adaptive Galerkin bases, Appl. Math. Optim. 41 (2000) 331–342.
[6] J. Álvarez-Vizoso, R. Arn, M. Kirby, C. Peterson, B. Draper, Geometry of curves in $\mathbb{R}^n$ from the local singular value decomposition, Linear Algebra Appl. 571 (2019) 180–202.
[7] M. Connolly, Measurement of protein surface shape by solid angles, J. Mol. Graph. 4 (1) (1986) 3–6.
[8] F. Cazals, M. Pouget, Estimating differential quantities using polynomial fitting of osculating jets, Symp. Geometry Processing, Eurographics (2003) 177–178.
[9] F. Cazals, M. Pouget, Molecular shape analysis based upon the Morse-Smale complex and the Connolly function, Proc. Symp. Comp. Geometry (2003) 351–360.
[10] U. Clarenz, M. Griebel, M. Rumpf, M. Schweitzer, A. Telea, Feature sensitive multiscale editing on surfaces, Vis. Comput. 20 (5) (2004) 329–343.
[11] U. Clarenz, M. Rumpf, A. Telea, Robust feature detection and local classification for surfaces based on moment analysis, IEEE Trans. Vis. Comput. Graph. 10 (2003) 516–524.
[12] S. Manay, D. Cremers, B.-W. Hong, A.J. Yezzi, Integral invariants for shape matching, IEEE Trans. Pattern Anal. Mach. Intell. 28 (10) (2006) 1602–1618.
[13] S. Manay, B.-W. Hong, A.J. Yezzi, S. Soatto, Integral invariant signatures, in: T. Pajdla, J. Matas (Eds.), Computer Vision - ECCV 2004, Springer Berlin Heidelberg, 2004, pp. 87–99.
[14] P. Alliez, D. Cohen-Steiner, Y. Tong, M. Desbrun, Voronoi-based variational reconstruction of unoriented point sets, in: Proceedings of the Fifth Eurographics Symposium on Geometry Processing, SGP'07, Eurographics Association, Aire-la-Ville, Switzerland, 2007, pp. 39–48.
[15] J. Berkmann, T. Caelli, Computation of surface geometry and segmentation using covariance techniques, IEEE Trans. Pattern Anal. Mach. Intell. 16 (1994) 1114–1116.
[16] H. Hoppe, T. DeRose, T. Duchamp, J. McDonald, W. Stuetzle, Surface reconstruction from unorganized points, SIGGRAPH Comput. Graph. 26 (2) (1992) 71–78.
[17] Y.-L. Yang, Y.-K.-K. Lai, S.-M. Hu, H. Pottmann, Robust principal curvatures on multiple scales, in: Proc. Eurographics Symposium on Geometry Processing, 2006, pp. 223–226.
[18] Y.-K. Lai, S.-M. Hu, T. Fang, Robust principal curvatures using feature adapted integral invariants, in: Proceedings - SPM 2009: SIAM/ACM Joint Conference on Geometric and Physical Modeling, 2009, pp. 325–330.
[19] H. Pottmann, J. Wallner, Q.-X. Huang, Y.-L. Yang, Integral invariants for robust geometry processing, Comput. Aided Geom. Des. 26 (1) (2008).
[20] H. Pottmann, J. Wallner, Y.-L. Yang, Y.-K. Lai, S.-M. Hu, Principal curvatures from the integral invariant viewpoint, Comput. Aided Geom. Des. 24 (2007) 428–442.
[21] Q. Mérigot, Geometric structure detection in point clouds, PhD. Thesis, Université Nice Sophia Antipolis, 2009.
[22] Q. Mérigot, M. Ovsjanikov, L. Guibas, Voronoi-based curvature and feature estimation from point clouds, IEEE Trans. Vis. Comput. Graph. 17 (2011) 743–756.
[23] X. Pennec, Probabilities and statistics on Riemannian manifolds: basic tools for geometric measurements, in: Int. Workshop on Nonlinear Signal and Image Processing (NSIP'99), 1999.
[24] X. Pennec, Intrinsic statistics on Riemannian manifolds: basic tools for geometric measurements, J. Math. Imaging Vis. 25 (2006) 127–154.
[25] A.V. Little, Estimating the Intrinsic Dimension of High-Dimensional Data Sets: A Multiscale, Geometric Approach, Ph.D. thesis, Duke University, 2011.
[26] A.V. Little, J. Lee, Y.-M. Jung, M. Maggioni, Estimation of intrinsic dimensionality of samples from noisy low-dimensional manifolds in high dimensions with multiscale SVD, in: IEEE/SP 15th Workshop on Statistical Signal Processing, 2009, SSP'09, IEEE, 2009, pp. 85–88.
[27] G. Chen, A.V. Little, M. Maggioni, L. Rosasco, Wavelets and Multiscale Analysis, Springer-Verlag, 2011, Ch. 10.
[28] A.V. Little, Y.-M. Jung, M. Maggioni, Multiscale estimation of intrinsic dimensionality of data sets, in: Proc. A.A.A.I., 2009.
[29] A.V. Little, M. Maggioni, L. Rosasco, Multiscale Geometric Methods for Data Sets I: Multiscale SVD, Noise and Curvature, Tech. Rep., MIT, 2012.

[30] J. Álvarez-Vizoso, M. Kirby, C. Peterson, Local Eigenvalue Decomposition for Embedded Riemannian Manifolds, submitted for publication, preprint, arXiv:1804.10425, 2019.
[31] J. Álvarez-Vizoso, M. Kirby, C. Peterson, Manifold curvature from covariance analysis, in: 2018 IEEE Statistical Signal Processing Workshop (SSP), 2018, pp. 353–357.
[32] S. Kobayashi, K. Nomizu, Foundations of Differential Geometry, John Wiley & Sons, Inc., 1969.
[33] M. Spivak, A Comprehensive Introduction to Differential Geometry, Publish or Perish, 1999.
[34] C. Davis, W. Kahan, The rotation of eigenvectors by a perturbation. III, SIAM J. Numer. Anal. 7 (1) (1970) 1–46.
[35] T. Kato, A Short Introduction to Perturbation Theory for Linear Operators, Springer-Verlag, 1982.
[36] F. Rellich, Perturbation Theory of Eigenvalue Problems, Gordon and Breach, 1969.
[37] G. Folland, How to integrate a polynomial over a sphere, Am. Math. Mon. 108 (5) (2001) 446–448.