

RESEARCH ARTICLE

WILEY

Grassmannian diffusion maps based surrogate modeling via geometric harmonics

Ketson R. M. dos Santos¹ | Dimitris G. Giovanis² | Katiana Kontolati² |
Dimitrios Loukrezis^{3,4} | Michael D. Shields²

¹Earthquake Engineering and Structural Dynamics Laboratory, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

²Department of Civil & Systems Engineering, Johns Hopkins University, Baltimore, Maryland, USA

³Institute for Accelerator Science and Electromagnetic Fields (TEMF), Technische Universität Darmstadt, Darmstadt, Germany

⁴Centre for Computational Engineering, Technische Universität Darmstadt, Darmstadt, Germany

Correspondence

Michael D. Shields, Department of Civil & Systems Engineering, Johns Hopkins University, Latrobe Hall, 3400 N Charles St. 205, Baltimore, MD 21218, USA.
Email: michael.shields@jhu.edu

Funding information

U.S. Department of Energy, Grant/Award Number: DE-SC0020428

Abstract

A novel surrogate model based on the Grassmannian diffusion maps (GDMaps) and utilizing geometric harmonics (GH) is developed for predicting the response of complex physical phenomena. The method utilizes GDMaps to obtain a low-dimensional representation of the underlying behavior of physical/mathematical systems with respect to uncertain input parameters. Using this representation, GH, an out-of-sample extension technique, is employed to create a global map from the input parameter space to a Grassmannian diffusion manifold. GH is further employed to locally map points on the diffusion manifold onto the tangent space of a Grassmann manifold. The exponential map is then used to project the points in the tangent space onto the Grassmann manifold, where reconstruction of the full solution is performed. The performance of the proposed surrogate model is verified with three examples. The first problem is a toy example used to illustrate the technique. In the second example, errors associated with the various mappings are assessed by studying response predictions of the electric potential of a dielectric cylinder in a homogeneous electric field. The last example applies the method for uncertainty prediction in the strain field evolution in a model amorphous material using the shear transformation zone theory of plasticity.

KEYWORDS

diffusion maps, geometric harmonics, Grassmann manifold, manifold learning, surrogate modeling

1 | INTRODUCTION

Surrogate models (aka emulators or metamodels) have become an important tool for uncertainty quantification because they afford a computationally efficient means of approximating (often complex) input–output relations generated from high-fidelity computational models. Surrogate models are constructed by training (or learning) a mathematical model from a finite set of input–output observations referred to as the training set. Given a new set of input parameters, where the solution of the model is unknown, the surrogate can then be used to predict the solution at minimal cost. Surrogate models are typically classified as either intrusive or non-intrusive.^{1,2} Intrusive methods, such as those based on Galerkin schemes, typically provide good convergence,^{3,4} but their complexity limits their flexibility⁵ and requires the

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *International Journal for Numerical Methods in Engineering* published by John Wiley & Sons Ltd.

development of new numerical schemes and, hence, entirely new solvers. In contrast, non-intrusive methods are trained from realizations of a deterministic solver (i.e., they can leverage pre-existing numerical methods and solvers) at selected sample points generated from the uncertain parameters. Moreover, surrogate models often provide an improvement over statistical approaches based on Monte Carlo simulation (MCS). Although MCS is a versatile and non-intrusive method, it typically offers slow convergence with the number of samples.⁶ To circumvent this limitation, quasi Monte Carlo methods,⁷ adaptive sampling,^{8,9} and other intelligent sampling methods leveraging variance reduction techniques^{10–13} are employed. Increasingly, these enhanced sampling methods are being leveraged to improve the training efficiency of surrogate models; enabling surrogates to be developed from far fewer training data.

Among the most widely used surrogate models in UQ are polynomial chaos expansions (PCE)³ and Gaussian process (GP) regression (Kriging).^{14,15} PCE was originally proposed by Wiener¹⁶ based on the projection of random solutions onto a basis of Hermite polynomials, which are orthogonal with respect to the Gaussian measure. Ghanem et al.³ introduced the Stochastic Galerkin projection for PCE, an intrusive method that requires the formulation of a system of algebraic equations for distinct classes of problems. The generalized PCE (gPCE)¹⁷ provides improved convergence and enhanced flexibility by utilizing polynomial bases from the Wiener–Askey scheme. Moreover, gPCE can be used in the construction of non-intrusive surrogate based on collocation schemes aiming at enhanced versatility and convergence properties. On the other hand, a GP approximates the input–output relation as a Gaussian stochastic process completely described by its mean and correlation function. Learning a GP thus consists of determining the mean and correlation functions. However, the computational complexity of the GP training process with n observations is on the order $\mathcal{O}(n^3)$, whereas the memory requirements scale with $\mathcal{O}(n^2)$.

It is well known that the relationship between data dimensionality and model fidelity has a strong influence on the computational performance of UQ for large-scale models. In this regard, dimension reduction techniques have become attractive due to their ability to represent high-dimensional data in a low-dimensional and more informative space (manifold).^{18–21} Dimension reduction techniques can be classified into linear and nonlinear methods. Linear methods include principal component analysis (PCA),²² locality preserving projections (LPP),²³ linear regression,²⁴ and singular value decomposition (SVD).²⁵ On the other hand, nonlinear methods are useful for constructing nonlinear maps between the high- and low-dimensional spaces. Nonlinear methods include Isomaps,²⁶ locally linear embedding (LLE),^{27,28} Kernel PCA,²⁹ and diffusion maps (DMaps).^{30,31} It is worth noting that several methods based on DMaps have been proposed in the literature such as the method proposed by Soize and Ghanem³² for sampling a random vector whose probability distribution is constrained to an Euclidean manifold. They used DMaps for discovering the underlying structure of the dataset. Moreover, DMaps has been used in the development of surrogate models either based on local polynomial interpolations and the Nyström out-of-sample extension,³³ or based on neural networks and the Laplacian pyramids.³⁴ Further, a subspace extension of DMaps, the Grassmannian diffusion maps (GDMaps), was introduced by dos Santos et al.³⁵ In this technique, a model dimension hyper-reduction is achieved by combining a pointwise linear dimensionality reduction technique, projecting the high-dimensional data onto a low-dimensional Grassmann manifold, and a multi-point nonlinear dimension reduction using DMaps which reveals the intrinsic structure of the data on the manifold. Recently, Kontolati et al.³⁶ employed GDMaps and PCE in the development of a manifold learning based method for UQ in systems describing complex processes.

In this work, we leverage the GDMaps to construct surrogate models for very high-dimensional systems. More specifically, a set of low-dimensional coordinates embedding the high-dimensional model data on the Grassmann manifold is obtained via GDMaps. This low-dimensional diffusion space serves as a connecting space between the parameter space and the Grassmann manifold, where full solution reconstruction can be performed. To connect these spaces, maps are constructed using the idea of out-sampling extension.^{37–40} Inspired by the Nyström extension,^{41,42} Coifman and Lafon⁴³ introduced a scheme, referred to as geometric harmonics (GH) for extending empirical functions only available at few locations. They demonstrated that this process relates the function complexity and its extension, with important implications to the construction of the lifting and restriction operators. In this article, GH is employed in the construction of an out-of-sample extension to create maps between spaces of interest. A global GH surrogate is constructed between the parameter and diffusion spaces, while local GH surrogates are constructed between the diffusion space and the Grassmann manifold via the tangent space, a flat inner-product space allowing local exponential mapping onto the Grassmannian, where the predicted model response can be constructed. Moreover, GH is also used to create a map between the parameter space and the space of singular values used in the construction of the model prediction.

This article is organized as follows. Section 2 discusses important background on the Grassmann manifold. Section 3 provides an overview of the GDMaps technique. Section 4 introduces GH as an out-of-sample function extension technique. Section 5 contains a detailed description of the surrogate modeling approach, referred to as Grassmannian-geometric harmonics maps (Grassmannian-GHMaps), developed herein. In Section 6, three examples

are provided. The first is a simple example used to explain the proposed method in a manner that is conceptually understandable and easy to visualize. The second example uses the GDMaps surrogates to predict the electric potential for a dielectric cylinder in homogeneous electric field, and is used to perform error analysis on the proposed method. The third example develops a surrogate model for the evolution of the strain field of amorphous solids under simple shear using the shear transformation zone (STZ) theory of plasticity. Finally, concluding remarks are provided in Section 7. Furthermore, the algorithms presented in this article have been implemented in UQpy (uncertainty quantification with python),⁴⁴ a general purpose Python toolbox for modeling uncertainty in physical and mathematical systems.

2 | GRASSMANN MANIFOLD

The concepts presented in this section are essential for the development of a surrogate model based on the GDMaps. Let us begin by defining a p -plane as a p -dimensional subspace, and a p -frame as a coordinate system that spans that subspace. Based on these two concepts one can define two important manifolds, the Stiefel and the Grassmann manifold. The Stiefel manifold $\mathcal{V}(p, n)$ is the set of all p -frames in \mathbb{R}^n such that $\mathcal{V}(p, n) = \{\mathbf{X} \in \mathbb{R}^{n \times p} : \mathbf{X}^\top \mathbf{X} = \mathbf{I}_p\}$, where $\mathbf{I}_p \in \mathbb{R}^{p \times p}$ is the identity matrix and $\mathbf{X} \in \mathbb{R}^{n \times p}$ is an orthonormal matrix.⁴⁵ The Grassmann manifold (or Grassmannian) $\mathcal{G}(p, n)$ is the set of p -planes in \mathbb{R}^n , where a point is given by $\mathcal{X} = \text{span}(\Psi)$, with $\Psi \in \mathcal{V}(p, n)$.⁴⁶ It is worth noting that \mathcal{X} is identified as an equivalence class of $n \times p$ matrices under orthogonal transformation of the Stiefel manifold.⁴⁶⁻⁴⁸ Therefore, a point on the Grassmann manifold is represented by an orthonormal matrix $\Psi \in \mathbb{R}^{n \times p}$ (the Stiefel representation).

2.1 | Tangent space: Exponential and logarithmic maps

The Grassmann manifold is a smooth and continuously differentiable manifold, which enables numerous mathematical operations such as differentiation and optimization.^{49,50} Given that the Grassmann manifold is smooth and continuously differentiable, one can define a trajectory $\gamma(z)$, $z \in [0, 1]$, known as geodesic, defining the shortest path between two points, $\gamma(0) = \mathcal{X}_0 = \text{span}(\Psi_0)$ and $\gamma(1) = \mathcal{X}_1 = \text{span}(\Psi_1)$, on the manifold $\mathcal{G}(p, n)$.⁴⁹ The derivative of this trajectory at any point \mathcal{X} (represented by Ψ) defines the tangent space ($\mathcal{T}_{\mathcal{X}}\mathcal{G}(p, n)$), which is given by the set of all tangent vectors Γ such that $\mathcal{T}_{\mathcal{X}}\mathcal{G}(p, n) = \{\Gamma \in \mathbb{R}^{n \times p} : \Gamma^\top \Psi = 0\}$. Therefore, given a tangent space $\mathcal{T}_{\mathcal{X}_0}\mathcal{G}(p, n)$ at \mathcal{X}_0 , one can map Γ_1 onto the Grassmannian point $\gamma(1) = \mathcal{X}_1$ represented by Ψ_1 via the exponential map

$$\Psi_1 = \exp_{\mathcal{X}_0}(\Gamma_1) = \exp_{\mathcal{X}_0}(\mathbf{U}\mathbf{S}\mathbf{V}^T) = \Psi_0 \mathbf{V} \cos(\mathbf{S}) \mathbf{Q}^T + \mathbf{U} \sin(\mathbf{S}) \mathbf{Q}^T, \quad (1)$$

where $\mathbf{Q} \in \mathbb{R}^{n \times n}$ is an orthogonal matrix satisfying the following expressions.

$$\mathbf{V} \cos(\mathbf{S}) \mathbf{Q}^T = \Psi_0^T \Psi_1, \quad (2)$$

and

$$\mathbf{U} \sin(\mathbf{S}) \mathbf{Q}^T = \Psi_1 - \Psi_0 \Psi_0^T \Psi_1. \quad (3)$$

After appropriate manipulation, one can obtain the following expression:

$$\mathbf{U} \tan(\mathbf{S}) \mathbf{V}^T = (\Psi_1 - \Psi_0 \Psi_0^T \Psi_1) (\Psi_0^T \Psi_1)^{-1}. \quad (4)$$

Consequently, one can write the logarithmic map from the Grassmannian to the tangent space, $\log_{\mathcal{X}} : \mathcal{G}(p, n) \rightarrow \mathcal{T}_{\mathcal{X}}\mathcal{G}(p, n)$ as

$$\log_{\mathcal{X}}(\Psi_1) = \mathbf{U} \tan^{-1}(\mathbf{S}) \mathbf{V}^T. \quad (5)$$

2.2 | Grassmannian distance

The properties of the Grassmann manifold further afford a notion of distance between points on it. Many definitions of distance exist and can be expressed in terms of the principal angles between subspaces.⁵¹ One can easily see that the

cosine of the principal angles $\theta_i \in [0, \pi/2]$ between two subspaces $\mathcal{X} = \text{span}(\Psi_x)$ and $\mathcal{Y} = \text{span}(\Psi_y)$ can be computed from the singular values of $\Psi_x^T \Psi_y = \bar{\mathbf{U}} \bar{\mathbf{S}} \bar{\mathbf{V}}^T$, where $\bar{\mathbf{U}} \in O(k)$, $\bar{\mathbf{V}} \in O(l)$, and $\bar{\mathbf{S}} = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_p)$, with $p = \min(k, l)$. Thus, the principal angles are computed as $\theta_i = \cos^{-1}(\sigma_i)$.⁵² A well-known and commonly used distance is the geodesic distance, $d_{\mathcal{G}(p,n)}(\mathcal{X}, \mathcal{Y})$, corresponding to the distance along the geodesic curve $\gamma(z)$ parameterized by $z \in [0, 1]$, and given by $d_{\mathcal{G}(p,n)}(\mathcal{X}, \mathcal{Y}) = \|\mathbf{s}\|_2$,^{46,53,54} where $\mathbf{s} = (\theta_1, \theta_2, \dots, \theta_p)$ is the vector of principal angles. See Reference 35 for more definitions of distances on the Grassmann manifold.

2.3 | Karcher mean

Consider a set of points (subspaces) on $\mathcal{G}(p, n)$. The Riemannian center of mass of these points is known as Karcher mean, $\mu_{\mathcal{G}(p,n)}$, and corresponds to the point $\mathcal{Y} \in \mathcal{G}(p, n)$ that minimizes the cost function $\sigma_{\mathcal{G}(p,n)}^2 : \mathcal{G}(p, n) \rightarrow \mathbb{R}^{+55,56}$ given by

$$\sigma_{\mathcal{G}(p,n)}^2(\mathcal{Y}) = \int_{\mathcal{G}(p,n)} d_{\mathcal{G}(p,n)}^2(\mathcal{Y}, \mathcal{X}) dP(\mathcal{X}), \quad (6)$$

where $dP(\mathcal{X}) = \rho(\mathcal{X}) d\mathcal{G}(p, n)$ is a probability measure on the Grassmann manifold with probability density function $\rho(\mathcal{X})$. Thus, the Karcher mean $\mu_{\mathcal{G}(p,n)} = \text{span}(\mathbf{M})$ can be computed by solving the optimization problem

$$\mu_{\mathcal{G}(p,n)} \doteq \underset{\mathcal{Y} \in \mathcal{G}(p,n)}{\text{argmin}} \int_{\mathcal{G}(p,n)} d_{\mathcal{G}(p,n)}^2(\mathcal{Y}, \mathcal{X}) dP(\mathcal{X}). \quad (7)$$

One can easily notice a similarity between Equation (6) and the variance of a continuous random variable, which gives a notion of dispersion around the mean. The Karcher variance $\sigma_{\mathcal{G}(p,n)}^2$ and the Karcher mean $\mu_{\mathcal{G}(p,n)}$ can therefore be interpreted as the mean and variance of a continuous random variable on $\mathcal{G}(p, n)$. With this interpretation, the Karcher mean can be estimated from a discrete set of points on the Grassmann manifold $S = \{\mathcal{X}_1, \dots, \mathcal{X}_N\} \subset \mathcal{G}(p, n)$, by solving the following minimization

$$\hat{\mu}_{\mathcal{G}(p,n)} \approx \underset{\mathcal{Y} \in \mathcal{G}(p,n)}{\text{argmin}} \frac{1}{N} \sum_{i=1}^N d_{\mathcal{G}(p,n)}^2(\mathcal{Y}, \mathcal{X}_i). \quad (8)$$

2.4 | Grassmannian kernels

Kernel-based dimensionality reduction techniques such as the conventional DMaps depend on the appropriate definition of a real-valued positive semi-definite kernel $k(x_i, x_j)$ with $\sum_{i,j} c_i c_j k(x_i, x_j) \leq 0$, where $c_i, c_j \in \mathbb{R}$. In this regard, the Gaussian kernel (Equation 9) is perhaps the most widely used

$$k(\mathbf{X}_i, \mathbf{X}_j) = \exp \left(-\frac{\|\mathbf{X}_i - \mathbf{X}_j\|_2^2}{4\epsilon} \right), \quad (9)$$

where \mathbf{X}_i and \mathbf{X}_j are the high-dimensional data and ϵ is a length-scale parameter. However, the Gaussian kernel is not suitable to represent the underlying subspace structure of datasets. On the other hand, Grassmannian kernels are endowed with this feature, which is advantageous in the analysis of high-dimensional data. A Grassmannian kernel is defined as a real symmetric map $k : \mathcal{G}(p, n) \times \mathcal{G}(p, n) \rightarrow \mathbb{R}$ embedding the Grassmann manifold into a reproducing kernel Hilbert space. Moreover, a Grassmannian kernel is invariant to the choice of basis and is positive semi-definite. Several families of Grassmannian kernels, with different characteristics, are proposed in the literature (see References 51,57,58). However, the most popular kernels are the Binet–Cauchy and projection kernels. The Binet–Cauchy kernel is constructed by embedding the Grassmann manifold $\mathcal{G}(p, n)$ into a projective space $\mathbb{P}(\wedge^p \mathbb{R}^n)$. Considering two subspaces $\mathcal{X} = \text{span}(\Psi_x)$ and $\mathcal{Y} = \text{span}(\Psi_y)$, the Binet–Cauchy kernel is given by

$$k_{bc}(\mathcal{X}, \mathcal{Y}) = \det(\Psi_x^T \Psi_y)^2, \quad (10)$$

or equivalently in terms of principal angles^{51,58}

$$k_{bc}(\mathcal{X}, \mathcal{Y}) = \prod_{i=1}^p \cos^2(\theta_i). \quad (11)$$

Another kernel frequently used in kernel-based methods on the Grassmann manifold is the projection kernel. It is constructed based on the projection embedding $\Pi : \mathcal{G}(p, n) \rightarrow \mathbb{R}^{n \times n}$ such that $\Pi(\Psi) = \Psi\Psi^T$. This kernel is defined as

$$k_{pr}(\mathcal{X}, \mathcal{Y}) = \|\Psi_x^T \Psi_y\|_F^2, \quad (12)$$

or equivalently in terms of principal angles^{51,58}

$$k_{pr}(\mathcal{X}, \mathcal{Y}) = \sum_{i=1}^p \cos^2(\theta_i). \quad (13)$$

Further discussion of Grassmannian kernels and their specific use for GDMaps can be found in dos Santos et al.⁵⁹

3 | GRASSMANNIAN DIFFUSION MAPS

The GDMaps³⁵ is a nonlinear dimension reduction technique that uses DMaps to learn the low-dimensional structure of a dataset on the Grassmann manifold. GDMaps is a two-stage dimension reduction. The first dimensional reduction is a pointwise projection of the elements of a dataset onto the Grassmann manifold. Next, a connected graph is created across the data on the Grassmann manifold where a random walk is performed to embed the data into a low-dimensional Euclidean space. This dimension reduction is illustrated in Figure 1 and the procedure is detailed herein.

Given a set of high-dimensional data $S_X = \{\mathbf{X}_1, \dots, \mathbf{X}_N\}$, with $\mathbf{X}_i \in \mathbb{R}^{n \times m}$ possessing a low-rank structure, one can project \mathbf{X}_i onto a Grassmann manifold via SVD such that $\mathbf{X}_i = \mathbf{U}_i \mathbf{S}_i \mathbf{V}_i^T$, with $\mathcal{U}_i = \text{span}(\mathbf{U}_i) \in \mathcal{G}(p, n)$, and $\mathcal{V}_i = \text{span}(\mathbf{V}_i) \in \mathcal{G}(p, m)$. Next, considering a positive semi-definite Grassmannian kernel $k : \mathcal{G}(p, \cdot) \times \mathcal{G}(p, \cdot) \rightarrow \mathbb{R}$ (i.e., the projection or Binet–Cauchy kernel), we construct the kernel matrices $\mathbf{K}_{\mathcal{V}} = [k_{\mathcal{V},ij}] = [k(\mathcal{U}_i, \mathcal{U}_j)] \in \mathbb{R}^{N \times N}$ and

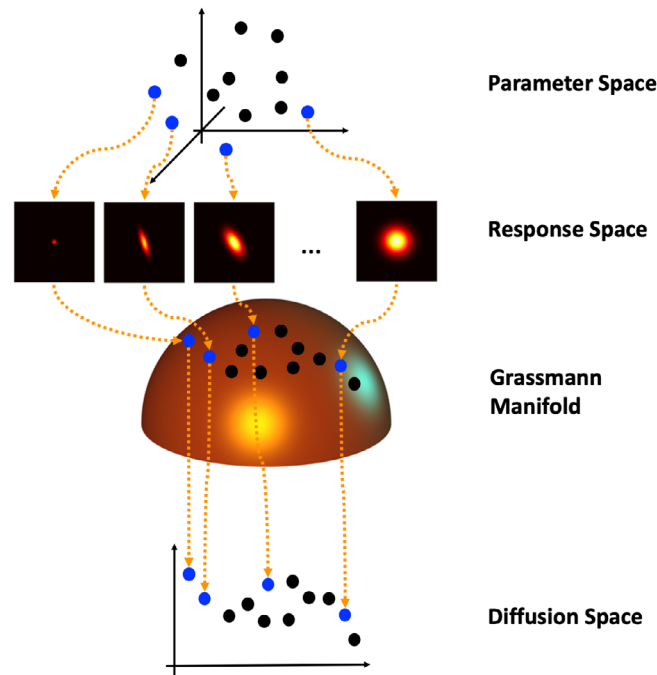


FIGURE 1 Conceptual illustration of the Grassmannian diffusion maps

$\mathbf{K}_\mathcal{V} = [k_{\mathcal{V},ij}] = [k(\mathcal{V}_i, \mathcal{V}_j)] \in \mathbb{R}^{N \times N}$. Next, we compile the composite kernel matrix as either $\mathbf{K} = \mathbf{K}_\mathcal{U} + \mathbf{K}_\mathcal{V}$ or $\mathbf{K} = \mathbf{K}_\mathcal{U} \circ \mathbf{K}_\mathcal{V}$ where \circ denotes the Hadamard product. We use this composite kernel to define a random walk over the data on the manifold, denoted $W = (\{S_\mathcal{U}, S_\mathcal{V}\}, F_\Theta, \mathbf{P})$, where $S_\mathcal{U} = \{\mathcal{U}_1, \dots, \mathcal{U}_N\}$ and $S_\mathcal{V} = \{\mathcal{V}_1, \dots, \mathcal{V}_N\}$, F_Θ is the joint probability distribution of the input parameters, and \mathbf{P} is the transition probability matrix. The matrix $\mathbf{P} = [P_{ij}]$ is constructed by first building the following diagonal matrix $\mathbf{D} = [D_{ii}] \in \mathbb{R}^{N \times N}$ as

$$D_{ii} = \sum_{j=1}^N k_{ij}, \quad (14)$$

such that the stationary distribution of the random walk is given by

$$\pi_i = \frac{D_{ii}}{\sum_{k=1}^N D_{kk}}. \quad (15)$$

Next, by normalizing the kernel k_{ij} as follows

$$\kappa_{ij} = \frac{k_{ij}}{\sqrt{D_{ii}D_{jj}}}, \quad (16)$$

the transition matrix \mathbf{P} can be constructed as

$$P_{ij} = \frac{\kappa_{ij}}{\sum_{k=1}^N \kappa_{ik}}. \quad (17)$$

The eigendecomposition of \mathbf{P} yields a set of eigenvectors $\Phi = [\phi_0, \dots, \phi_N]$ and their respective eigenvalues $\Lambda = \{\lambda_0, \dots, \lambda_N\}$. Thus, every element \mathbf{X}_i of $S_\mathbf{X}$ has a representation on a low-dimensional Euclidean space defined by the points $\xi_i = [\lambda_0 \Phi_{i0}, \dots, \lambda_r \Phi_{ir}]^T$, where $r < N$ due to the decaying spectrum $\{\lambda_0, \dots, \lambda_N\}$.

4 | GEOMETRIC HARMONICS

The construction of a mapping between the reduced (i.e., diffusion space) and ambient (i.e., parameter space, response space) space has been discussed in the literature by several authors.^{18-20,60,61} In this regard, the construction of the lifting (from the reduced space to the ambient space) and restriction (from the ambient space to the reduced space) operators relies on the extension of empirical functions only known at specific locations of the domain, also known as out-of-sample extension.

For example, the Nyström extension, which is very closely related to GP regression,⁶² is commonly used to construct the restriction operator within the conventional DMaps framework.^{18,60} In particular, given a new sample \mathbf{X}_k in the dataset $S_\mathbf{X}$, the diffusion coordinates can be obtained as follows

$$\xi_i(\mathbf{X}_k) = \lambda_k^{-1} \sum_{\mathbf{X} \in S_\mathbf{X}} P(\mathbf{X}_k, \mathbf{X}) \xi_i(\mathbf{X}), \quad (18)$$

where $P(\cdot, \cdot)$ is the transition matrix. The Nyström extension has some disadvantages such as the computational complexity of the required diagonalization. Moreover, it can become ill-conditioned,⁶² so instead we leverage a variation on the Nyström extension, known as GH,⁴³ for constructing both the lifting and restriction operators.

The GH aim to extend a function f defined on a set Ω to a set $\bar{\Omega}$ such that $\Omega \subset \bar{\Omega}$. This extension depends on the selection of an appropriate positive semi-definite kernel $k(\cdot, \cdot)$ (such as the Gaussian kernel in Equation 9), which defines a unique reproducing kernel Hilbert space \mathcal{H} of functions in $\bar{\Omega}$. Therefore, restricting $k(\cdot, \cdot)$ to Ω one can define an operator $\mathbf{K} : L^2(\Omega, d\mu) \rightarrow \mathcal{H}$, such that

$$\mathbf{K}f(\bar{\omega}) = \int_{\Omega} k(\bar{\omega}, \omega) f(\omega) d\mu(\omega), \quad (19)$$

where $d\mu$ is a measure with $d\mu < +\infty$, $\omega \in \Omega$, and $\bar{\omega} \in \bar{\Omega}$. A lemma presented by Coifman and Lafon⁴³ shows that the adjoint operator $\mathbf{K}^* : \mathcal{H} \rightarrow L^2(\Omega, d\mu)$ is in fact the restriction operator. Moreover, as this operator is self-adjoint and

compact, its eigendecomposition exists and one can write the GH as

$$\psi_i(\bar{\omega}) = \lambda_i^{-1} \int_{\Omega} k(\bar{\omega}, \omega) \psi(\omega) d\mu(\omega), \quad (20)$$

which can be summarized by the following expressions

$$\mathbf{K}\psi_i = \lambda_i \bar{\psi}_i, \quad (21)$$

and

$$\mathbf{K}^* \bar{\psi}_i = \psi_i, \quad (22)$$

where $L_{\delta}^2 = \text{span}\{\psi_i, i \in D_{\delta}\}$ and $\mathcal{H}_{\delta} = \text{span}\{\bar{\psi}_i, i \in D_{\delta}\}$, with $D_{\delta} = \{i, \lambda_i \geq \delta \lambda_0\}$ and $\delta > 0$. Therefore, the mechanization of the extension algorithm is given by two main steps.⁴³ First, f is projected onto $L_{\delta}^2 = \text{span}\{\psi_i, i \in D_{\delta}\}$. Such that,

$$f \mapsto \mathbf{P}_{\delta} f = \sum_{j \in D_{\delta}} \langle f, \psi_j \rangle_{\Omega} \psi_j. \quad (23)$$

Second, the extension $\mathbf{E}f$ is given by

$$\mathbf{E}f(\bar{\omega}) = \sum_{j \in D_{\delta}} \langle f, \psi_j \rangle_{\Omega} \bar{\psi}_j(\bar{\omega}). \quad (24)$$

Details on implementation of GH can be found in Algorithm 1. In this algorithm, one can observe that the sets S_x and S_y as well as S_x^* and S_y^* are composed of column vectors. Therefore, if the Euclidean space of interest is the space of matrices, one can transform an element of this space into vectors by stacking its columns. Further, one can easily observe that the restriction operator between a point on the Grassmann manifold $\mathcal{G}(p, n)$ and the Grassmannian diffusion space exists since appropriate kernels can be defined on the Grassmann manifold, see Section 3. On the other hand, there is no guarantee that the inverse map (lifting operator) exists due to the orthogonality constraints of the Grassmann manifold. However, assuming that the Grassmann manifold is locally approximated by a flat inner-product space (i.e., the tangent space $\mathcal{T}_{\mathcal{X}_0} \mathcal{G}(p, n)$) constructed at \mathcal{X}_0 , one can define a local lifting operator from the Grassmannian diffusion space defined on a set $\Omega \in \mathbb{R}^k$ to $\mathcal{T}_{\mathcal{X}_0} \mathcal{G}(p, n)$. It is then straightforward to apply the exponential mapping to project the extended sampling points in $\mathcal{T}_{\mathcal{X}_0} \mathcal{G}(p, n)$ onto $\mathcal{G}(p, n)$ (see Section 2.1).

Algorithm 1. Geometric harmonics

Input: The set $S_x = \{\mathbf{X}_1, \dots, \mathbf{X}_N\} \subset \mathbb{R}^n$ represented by the matrix $\mathbf{S}_x = [\mathbf{X}_1, \dots, \mathbf{X}_N]^T \subset \mathbb{R}^{N \times n}$, the target set $S_y = \{\mathbf{Y}_1, \dots, \mathbf{Y}_N\} \subset \mathbb{R}^m$ represented by the matrix $\mathbf{S}_y = [\mathbf{Y}_1, \dots, \mathbf{Y}_N]^T \subset \mathbb{R}^{N \times m}$, a positive semi-definite kernel $k(\cdot, \cdot)$, and a new subset $S_{x^*} = \{\mathbf{X}_1^*, \dots, \mathbf{X}_N^*\} \subset \mathbb{R}^n$ of S_x represented by the matrix $\mathbf{S}_{x^*} = [\mathbf{X}_1^*, \dots, \mathbf{X}_N^*]^T \subset \mathbb{R}^{M \times n}$.

- 1: Compute the kernel matrix $\mathbf{K} = [k(\mathbf{X}_i, \mathbf{X}_j)]$.
- 2: Kernel eigendecomposition: $\mathbf{K} \bar{\psi}_i = \bar{\lambda}_i \bar{\psi}_i$, with $i = 1, \dots, r$, with $r \leq N$. Eigenvectors and eigenvalues can be written as the matrices $\bar{\Psi} \in \mathbb{R}^{N \times r}$ and $\bar{\Lambda} = \text{diag}(\bar{\lambda}_1, \dots, \bar{\lambda}_r) \in \mathbb{R}^{r \times r}$, respectively.
- 3: Compute $\mathbf{B} = \bar{\Psi} \bar{\Lambda}^{-1} \bar{\Psi}^T \mathbf{S}_y \in \mathbb{R}^{N \times m}$.
- 4: Compute the kernel values for the new element: $\hat{\mathbf{K}} = [k(\mathbf{X}_i^*, \mathbf{X}_j)] \in \mathbb{R}^{M \times N}$.
- 5: Extension: $\mathbf{S}_{y^*} = \hat{\mathbf{K}} \mathbf{B} \in \mathbb{R}^{M \times m}$.

Output: a new set \mathbf{S}_{y^*} .

5 | GRASSMANNIAN-GEOMETRIC HARMONICS MAPS

Consider the random vector $\Theta \in \mathbb{R}^Q$ having joint probability distribution $F_{\Theta}(\Theta_1, \dots, \Theta_Q)$ as the input parameters to a model $\mathcal{M}(\cdot)$. One can obtain samples Θ_i as elements of a set $S_{\Theta} = \{\Theta_1, \dots, \Theta_N\} \subset \Pi$ from $F_{\Theta}(\Theta_1, \dots, \Theta_Q)$, where Π is the parameter space. For each element of S_{Θ} , the model $\mathcal{M}(\cdot)$ (e.g., finite element model) produces a high-dimensional response $\mathbf{X}_i \in \mathbb{R}^{n \times m}$ such that $\mathbf{X}_i = \mathcal{M}(\Theta_i)$. Therefore, a set $S_X = \{\mathcal{M}(\Theta_1), \dots, \mathcal{M}(\Theta_N)\} = \{\mathbf{X}_1, \dots, \mathbf{X}_N\} \subset \Xi$ is obtained,

where Ξ is the response space. With the set S_X and assuming that \mathbf{X}_i has a low-rank structure, we begin by projecting \mathbf{X}_i onto a Grassmann manifold. This operation is performed via SVD, as presented in Section 2. Thus, one can decompose \mathbf{X}_i as $\mathbf{X}_i = \mathbf{U}_i \mathbf{S}_i \mathbf{V}_i^T$, with $\mathcal{U}_i = \text{span}(\mathbf{U}_i) \in \mathcal{G}(p, n)$, and $\mathcal{V} = \text{span}(\mathbf{V}) \in \mathcal{G}(p, m)$. Moreover, a set $S_S = \{\mathbf{S}_1, \dots, \mathbf{S}_N\} \subset \Sigma$ of singular values is obtained, where Σ is the space of singular values and $\mathbf{S}_i \in \mathbb{R}^{p \times p}$.

Selecting an appropriate Grassmannian kernel,³⁵ we next construct a connected graph on the sets $S_U = \{\mathcal{U}_1, \dots, \mathcal{U}_N\} \subset \mathcal{G}(p, n)$ and $S_V = \{\mathcal{V}_1, \dots, \mathcal{V}_N\} \subset \mathcal{G}(p, m)$ and apply the procedure of Section 3 to determine the new coordinates (Grassmannian diffusion coordinates) embedding the data on the Grassmann manifolds $\mathcal{G}(p, n)$ and $\mathcal{G}(p, m)$ into a low-dimensional Euclidean space (Grassmannian diffusion space). Once the Grassmannian diffusion coordinates $\xi = \{\xi_1, \dots, \xi_N\} \in \Delta$ are obtained, we construct a global map (surrogate) using GH (see Section 4) between Π and Δ ($GH_0 : \Pi \rightarrow \Delta$), although local maps can also be constructed if necessary. For this mapping, the Gaussian kernel $k(\boldsymbol{\Theta}_i, \boldsymbol{\Theta}_j)$ (Equation 9) is used because we want to construct a map between Euclidean spaces. A second map between Π and Σ ($GH_1 : \Pi \rightarrow \Sigma$) is constructed using GH as well. With both maps constructed, we can predict the dimension reduced response of $\mathcal{M}(\cdot)$ for any new set of input parameters. In other words, considering that a new set of input parameters $\boldsymbol{\Theta}^*$ is sampled from $F_{\boldsymbol{\Theta}}$, we estimate the corresponding Grassmannian diffusion coordinates in Δ by $\xi^* = GH_0(\boldsymbol{\Theta}^*)$. Simultaneously, we estimate the singular values by $\mathbf{S}^* = GH_1(\boldsymbol{\Theta}^*)$.

Once the estimated coordinate ξ^* in Δ is obtained, it is necessary to expand this reduced dimension solution from the Grassmannian diffusion manifold back to the full-dimensional solution in the ambient space. The first step of this decoding is to define a mapping between Δ and the Grassmann manifolds $\mathcal{G}(p, n)$ and $\mathcal{G}(p, m)$. To achieve this, we define a series of local GH lifting operators $\Lambda_0 : \Delta \rightarrow \mathcal{T}_{\hat{\mu}_u} \mathcal{G}(p, n)$ and $\Lambda_1 : \Delta \rightarrow \mathcal{T}_{\hat{\mu}_v} \mathcal{G}(p, m)$ where $\hat{\mu}_u$ and $\hat{\mu}_v$ are the reference points on the Grassmann manifold where the tangent spaces are constructed. The local operators are determined by identifying the k nearest neighbors to the point ξ^* . These points define the vicinity of ξ^* on the Grassmann manifold, and the tangent space can be constructed either around their Karcher mean (see Section 2.3); or around the nearest neighbor of ξ^* on the Grassmann manifold, which is a computationally efficient method since no optimization is performed. These local data are then used to construct the GH lifting operator. Using these lifting operators, we obtain the points $\Gamma_u^* = \Lambda_0(\xi^*)$ and $\Gamma_v^* = \Lambda_1(\xi^*)$, where $\Gamma_u^* \in \mathbb{R}^{m \times p}$ and $\Gamma_v^* \in \mathbb{R}^{m \times p}$ represent points on their respective tangent spaces. Next, we apply the exponential map to obtain the corresponding points on $\mathcal{G}(p, n)$ and $\mathcal{G}(p, m)$ (see Section 2.1) as $\mathbf{U}^* = \exp_{\hat{\mu}_u}(\Gamma_u^*)$ and $\mathbf{V}^* = \exp_{\hat{\mu}_v}(\Gamma_v^*)$, where $\mathcal{U}^* = \text{span}(\mathbf{U}^*) \in \mathcal{G}(p, n)$ and $\mathcal{V}^* = \text{span}(\mathbf{V}^*) \in \mathcal{G}(p, m)$. Finally, the solution \mathbf{X}^* for the new set of input parameters $\boldsymbol{\Theta}^*$ can be predicted by the following matrix product

$$\mathbf{X}^* = \mathbf{U}^* \mathbf{S}^* \mathbf{V}^{*T}. \quad (25)$$

Next, two algorithms are presented summarizing this method. Algorithm 2 describes the construction of the maps between the spaces of interest (training), and Algorithm 3 shows how to predict the response using the constructed maps. Moreover, the proposed surrogate modeling approach is illustrated conceptually in Figure 2.

Algorithm 2. Grassmannian-GHMaps: training

Input: a model $\mathcal{M}(\boldsymbol{\Theta})$; a set of N vectors of input parameters $S_{\boldsymbol{\Theta}} = \{\boldsymbol{\Theta}_1, \dots, \boldsymbol{\Theta}_N\} \subset \Pi$, with $\boldsymbol{\Theta} \subset \mathbb{R}^Q$; and a response set $S_X = \{\mathcal{M}(\boldsymbol{\Theta}_1), \dots, \mathcal{M}(\boldsymbol{\Theta}_N)\} = \{\mathbf{X}_1, \dots, \mathbf{X}_N\} \subset \Xi$, with $\mathbf{X}_i \in \mathbb{R}^{n \times m}$.

- 1: **for** $i \in 1, \dots, N$ **do**
- 2: Compute the thin singular value decomposition: $\mathbf{X}_i = \mathbf{U}_i \mathbf{S}_i \mathbf{V}_i^T$, where $\mathcal{U}_i = \text{span}(\mathbf{U}_i) \in \mathcal{G}(p, n)$ and $\mathcal{V}_i = \text{span}(\mathbf{V}_i) \in \mathcal{G}(p, m)$.
- 3: **end for**
- 4: Construct the sets $S_U = \{\mathcal{U}_1, \dots, \mathcal{U}_N\} \subset \mathcal{G}(p, n)$, $S_V = \{\mathcal{V}_1, \dots, \mathcal{V}_N\} \subset \mathcal{G}(p, m)$, and $S_S = \{\mathbf{S}_1, \dots, \mathbf{S}_N\} \subset \Sigma$.
- 5: For every pair $[\mathcal{U}_i, \mathcal{U}_j]$ and $[\mathcal{V}_i, \mathcal{V}_j]$ compute the entries of k_{ij} of the kernel matrices $k_{ij}(\mathcal{U})$ and $k_{ij}(\mathcal{V})$, either using Equation (10) or (12) (or Equation (11) or (13), equivalently).
- 6: If necessary, compute the composed kernel matrix $k(\mathcal{U}, \mathcal{V}).k(\mathcal{U}, \mathcal{V}) = k_{ij}(\mathcal{U}) + k_{ij}(\mathcal{V})$ or $k(\mathcal{U}, \mathcal{V}) = k_{ij}(\mathcal{U}) \circ k_{ij}(\mathcal{V})$, where \circ is the Hadamard product.

- 7: Apply the approach of Section 3 on $k(\mathcal{U}, \mathcal{V})$ to get the Grassmannian diffusion coordinates $\xi = \{\xi_1, \dots, \xi_N\} \in \Delta$.
- 8: Construct the GH map from Π to Δ , $GH_0 : \Pi \rightarrow \Delta$. (Algorithm 1 with $(\boldsymbol{\Theta}, \xi)$ as training data.)
- 9: Construct the GH map from Π to Σ , $GH_1 : \Pi \rightarrow \Sigma$. (Algorithm 1 with $(\boldsymbol{\Theta}, \mathbf{S})$ as training data.)

Output: GH maps: GH_0 and GH_1 .

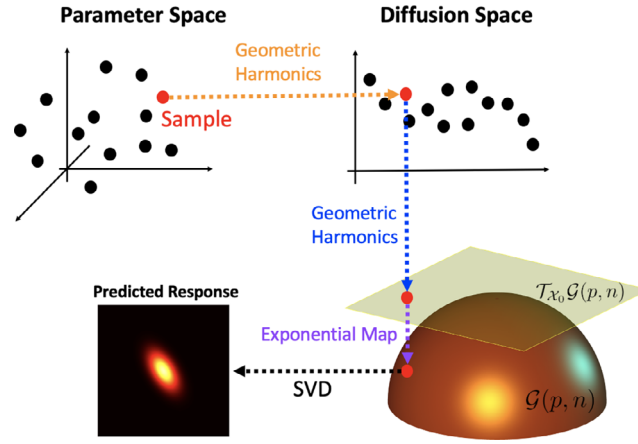


FIGURE 2 Conceptual illustration of the Grassmannian diffusion maps based surrogate modeling: sampling the parameter space and mapping to response prediction

Algorithm 3. Grassmannian-GHMaps: prediction

Input: the GH maps GH_0 and GH_1 (Algorithm 2), and a new vector of input parameters Θ^* to predict $\mathbf{X}^* = \mathcal{M}(\Theta^*)$.

- 1: Estimate the diffusion coordinates ξ^* corresponding to Θ^* : $\xi^* = GH_0(\Theta^*)$.
- 2: Estimate the singular values \mathbf{S}^* corresponding to Θ^* : $\mathbf{S}^* = GH_1(\Theta^*)$.
- 3: Find the k -neighbors (indices I_k) of ξ^* in ξ to create the subset $K_\xi = \{\xi_k | k \in I_k\}$.
- 4: From the points with indices in I_k , find the reference points (e.g., nearest neighbor, Karcher mean) $\hat{\mu}_u$ and $\hat{\mu}_v$ on $\mathcal{G}(p, n)$ and $\mathcal{G}(p, m)$, respectively.
- 5: **for** $i \in I_k$ **do**
- 6: Map the corresponding points on $\mathcal{G}(p, n)$ and $\mathcal{G}(p, m)$ to the tangent spaces $\mathcal{T}_{\hat{\mu}_u}$ and $\mathcal{T}_{\hat{\mu}_v}$, respectively: $\Gamma_i^{(U)} = \log_{\hat{\mu}_u}(\mathcal{U}_i)$ and $\Gamma_i^{(V)} = \log_{\hat{\mu}_v}(\mathcal{V}_i)$.
- 7: **end for**
- 8: Use Algorithm 1 and the set K_ξ to create the local maps $\Lambda_0(\cdot)$ and $\Lambda_1(\cdot)$ between Δ and $\mathcal{T}_{\hat{\mu}_u}$ and $\mathcal{T}_{\hat{\mu}_v}$ in the vicinity of $\hat{\mu}_u$ and $\hat{\mu}_v$, respectively.
- 9: Compute the matrices corresponding to the points on $\mathcal{T}_{\hat{\mu}_u}$ and $\mathcal{T}_{\hat{\mu}_v}$: $\Gamma_u^* = \Lambda_0(\xi^*)$ and $\Gamma_v^* = \Lambda_1(\xi^*)$.
- 10: Use the exponential map to project the points in the tangent spaces onto their respective Grassmann manifolds: $\mathbf{U}^* = \exp_{\hat{\mu}_u}(\Gamma_u^*)$ and $\mathbf{V}^* = \exp_{\hat{\mu}_v}(\Gamma_v^*)$, where $\mathcal{U}^* = \text{span}(\mathbf{U}^*) \in \mathcal{G}(p, n)$ and $\mathcal{V}^* = \text{span}(\mathbf{V}^*) \in \mathcal{G}(p, m)$.

Output: predicted solution $\mathbf{X}^* = \mathbf{U}^* \mathbf{S}^* \mathbf{V}^{*T}$.

6 | EXAMPLES

In this section, three examples are considered to demonstrate the versatility of the proposed surrogate modeling approach. We begin with a toy example in which structured points on the Grassmann manifold can be easily visualized as points on the unit sphere. In the second example, the electrical potential field of an infinitely long dielectric cylinder suspended in a homogeneous electric field is predicted considering that the cylinder's radius r_0 and the strength of the electric field E_∞ are random variables. The third example considers the evolution of the strain field in an amorphous solid under simple shear using the STZ theory of plasticity. In each case, we consider examples with low-dimensional input random variables because the emphasis of the proposed method is on quantifying uncertainty when output dimension is high. When the input dimension is large, additional methods for dimension reduction are often required, which we study elsewhere.⁶³

The projection kernel in Equation (12) is adopted in all examples presented in this section, and the kernel composition by the Hadamard product is considered. Moreover, the accuracy of the predicted solutions is evaluated by using the entry-wise relative error for matrices (Equation 26) and the relative error in a Euclidean norm sense (Equation 27).

$$\text{error}_{rel} = \left| \frac{\mathbf{X}^* - \mathbf{X}_{exact}}{\mathbf{X}_{exact}} \right|, \quad (26)$$

$$\text{error}_{\|\cdot\|} = \frac{\|\mathbf{X}^* - \mathbf{X}_{exact}\|}{\|\mathbf{X}_{exact}\|}, \quad (27)$$

where the norm $\|\cdot\|$ in Equation (27) denotes either the Frobenius norm $\|\cdot\|_F$, if its argument is a matrix, or the Euclidean L_2 -norm $\|\cdot\|_2$, if its argument is a vector.

Further, the dimensions p and n (or m) of $\mathcal{G}(p, n)$ (or $\mathcal{G}(p, m)$) must be determined. In this regard, n (or m) is fixed based on the shape of the element $\mathbf{X}_i \in \mathbb{R}^{n \times m}$ of the set $S_X = \{\mathbf{X}_1, \dots, \mathbf{X}_N\} \subset \Xi$. Moreover, all the points \mathbf{X}_i are projected onto the same Grassmann manifold $\mathcal{G}(p, n)$ (or $\mathcal{G}(p, m)$), which means that p is fixed based either on the user selection or determined by

$$p = \max[\text{rank}(\mathbf{X}_1), \dots, \text{rank}(\mathbf{X}_N)]. \quad (28)$$

The computational performance of the surrogate modeling technique proposed in this article is assessed in Example 3. To this aim, a MacBook Pro with a 2.3 GHz 8-core Intel core i9 with 32 GB (2400 MHz DDR4) of memory.

6.1 | Structured data on the unit sphere in \mathbb{R}^3

Consider the following set of equations,

$$\begin{aligned} x &= |r| \sin(t) \cos(s), \\ y &= r \sin(t) \sin(s), \\ z &= |r| \cos(t); \end{aligned} \quad (29)$$

such that r is uniformly distributed in the interval $[-2, 2]$, t is uniformly distributed in the interval $[-\pi/2, \pi/2]$, and $s = \sin^{-1}(\cos(t)^2)$. We draw $N = 3000$ sample pairs $(r, t) \in \Pi$ to obtain a collection of N points constrained on two cone-like structures in \mathbb{R}^3 as presented in Figure 3A, with the colors representing the magnitude $\sqrt{x^2 + y^2 + z^2}$. In effect, we have a model (Equation 29) that maps two random variables onto a surface in \mathbb{R}^3 .

Each point is represented by a column vector $\mathbf{X}_i = [x_i, y_i, z_i]^T$, which together compose the set $S_X = \{\mathbf{X}_1, \dots, \mathbf{X}_N\}$. One can readily see that these points can be projected onto the Grassmann manifold $\mathcal{G}(1, 3)$, which is the unit sphere \mathbb{S}^3 . A point \mathcal{X}_i on $\mathcal{G}(1, 3)$ is given by the unit vector obtained from the normalization of \mathbf{X}_i such that $\mathcal{X}_i = \mathbf{X}_i / \|\mathbf{X}_i\|_2$, which reveals two inverted teardrop shaped structures on the sphere as illustrated in Figure 3B. Applying DMaps to these points on the Grassmann manifold, we see that a well-defined parametrization is obtained, as revealed by the Grassmannian diffusion coordinates in Figure 3C.

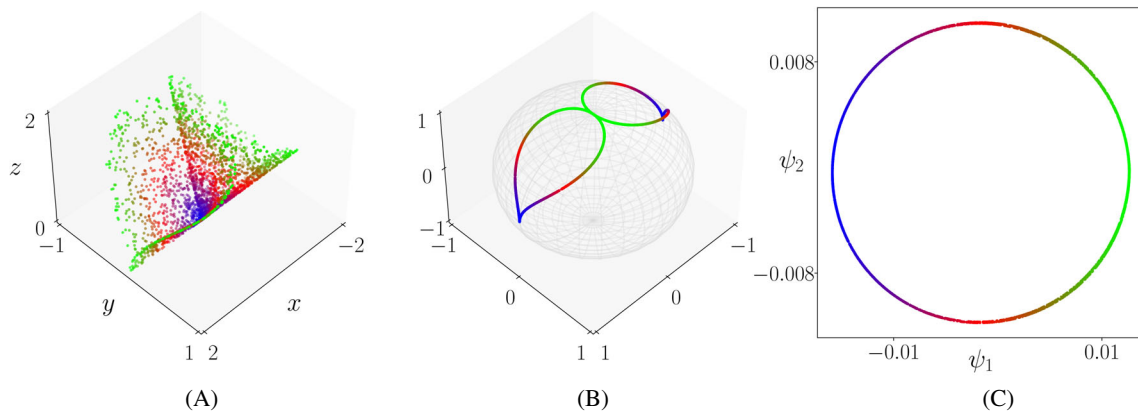


FIGURE 3 Example 1: Collection of $N = 3000$ random points constrained by Equation (29) (A) in the ambient space \mathbb{R}^3 , (B) in $\mathcal{G}(1, 3)$ or \mathbb{S}^3 , (C) in Grassmannian diffusion coordinates. In (A) points are colored by Euclidean distance from the origin. In (B, C) points are colored by the first Grassmannian diffusion coordinate

GH is used to create a map $GH_0 : \Pi \rightarrow \Delta$ from the parameter space Π to the Grassmannian diffusion manifold Δ . Therefore, it can be considered as a manifold learning technique, where the position on the Grassmannian diffusion manifold (Figure 3C) can be predicted for any point Θ in the parameter space Π . To verify the accuracy of this learning process, we draw 3000 additional samples Θ (Figure 4). One can easily see in Figure 4 that the trained GH can reliably predict the shape of the Grassmannian diffusion manifold.

We obtain a new parameter vector $\Theta^* = (r, t) = (1, 1.3)$ by sampling Π , and the GDMaps-based surrogate model is used to predict the vector \mathbf{X}^* . Using the map $GH_1 : \Pi \rightarrow \Sigma$ one can predict the first two nontrivial diffusion coordinates as $\xi^* = [-1.3657 \times 10^{-2}, -6.5931 \times 10^{-3}]$, represented by the red star in Figure 5A. Moreover, we observe that $S^* = r = 1$ determines the magnitude of the predicted point in \mathbb{R}^3 . Selecting the $k = 3$ nearest neighbors of ξ^* in Algorithm 3, we predict the point in the ambient space (i.e., on the cone structure) corresponding to Θ^* by mapping ξ^* onto the tangent space \mathcal{T} , constructed in the closest neighbor of ξ^* (one can also use the Karcher mean alternatively), for posterior projection onto the Grassmann manifold $\mathcal{G}(1, 3)$, as illustrated by the red star in Figure 5B. This point coincides very closely with true point denoted by the black dot. From the point projected onto the Grassmann manifold and considering the magnitude of the point given by $|r|$, we predict the point $\mathbf{X}^* = [0.96117, 0.06895, 0.2675]^T$ represented by the red star in the ambient space in Figure 5C, where again the black dot is the true value $\mathbf{X}_{exact} = [0.9611, 0.0689, 0.2675]$. In this case, we obtain $error_{rel} = [8.5947 \times 10^{-5}, 8.3515 \times 10^{-3}, 5.6326 \times 10^{-4}]$ and $error_{||\cdot||} = 6 \times 10^{-4}$.

Next, we draw $N = 10,000$ pairs $(r, t) \in \Pi$ to assess the overall performance of the proposed surrogate modeling technique. Using this new set of input parameters we predict N points on the cone-like structure and compare them with

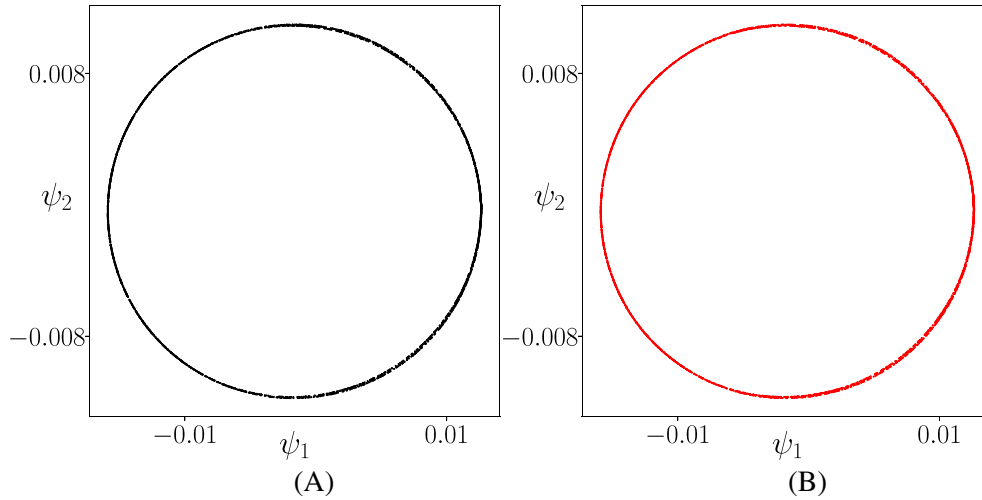


FIGURE 4 Example 1: Grassmannian diffusion manifold: (A) Training set for GH, and (B) predicted Grassmannian diffusion manifold for 3000 additional samples

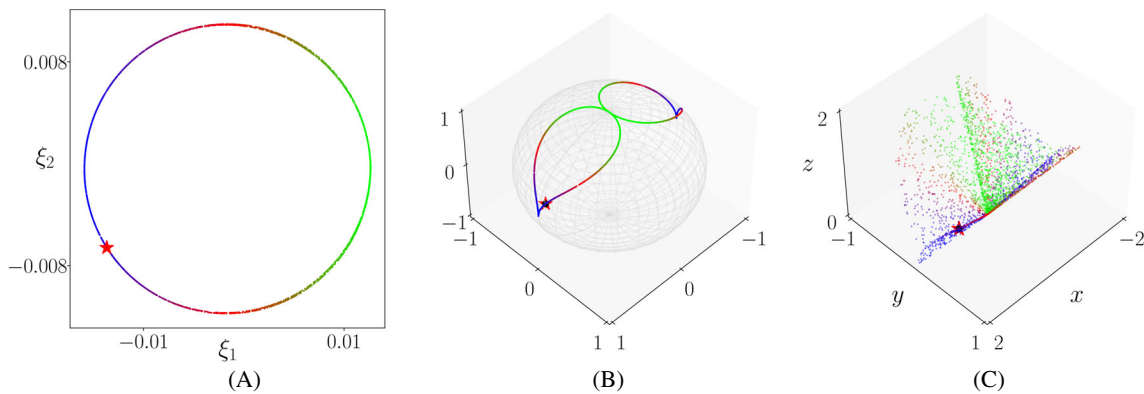


FIGURE 5 Example 1: (A) Point (red star) predicted in the Grassmannian diffusion space, (B) projected real point (black dot) and predicted point (red star) on the Grassmann manifold, and (C) the predicted point (red star) and the real point (black dot) in the Euclidean space

the exact points corresponding to the set of input parameters. Figure 6 shows the predicted cone-like structures from these 10,000 surrogate model evaluations. Comparing with Figure 3, we can see that the points closely match the true structure.

To assess the overall quality of the predictions, the marginal probability density functions (PDF) for each dimension (x, y, z) are estimated using the kernel density estimation (KDE) and shown in Figure 7 for both the true samples and the surrogate predictions. The PDFs for the surrogate predictions match those of the true samples very closely.

6.2 | Dielectric cylinder in homogeneous electric field

In this example, we study variations in the electrical potential of an infinitely long dielectric cylinder suspended in a homogeneous electric field, resulting from uncertainty in the input parameters. The problem is defined over a rectangular domain $\Omega = [-1, 1] \times [-1, 1]$ with the embedded cylinder domain $D_c = \{\mathbf{x} = (x, y) \mid \sqrt{x^2 + y^2} \leq r_0\}$, where r_0 is the cylinder's radius. We assume Dirichlet boundary conditions, Γ_D , on the left and right boundaries and Neumann boundary conditions Γ_N on the top and bottom boundaries.

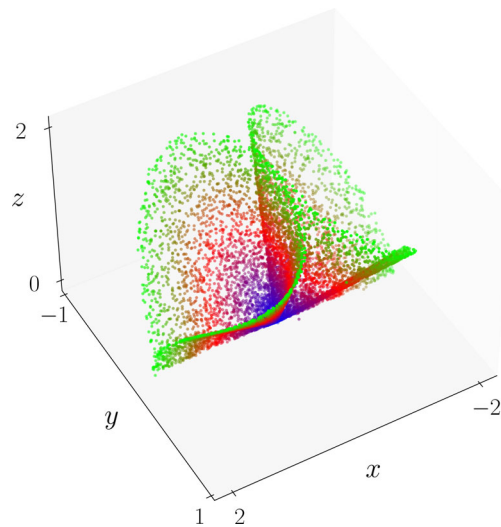


FIGURE 6 Example 1: Predicted points on the cone-like structure in the ambient space using the surrogate model. Color scale indicates Euclidean distance from the origin

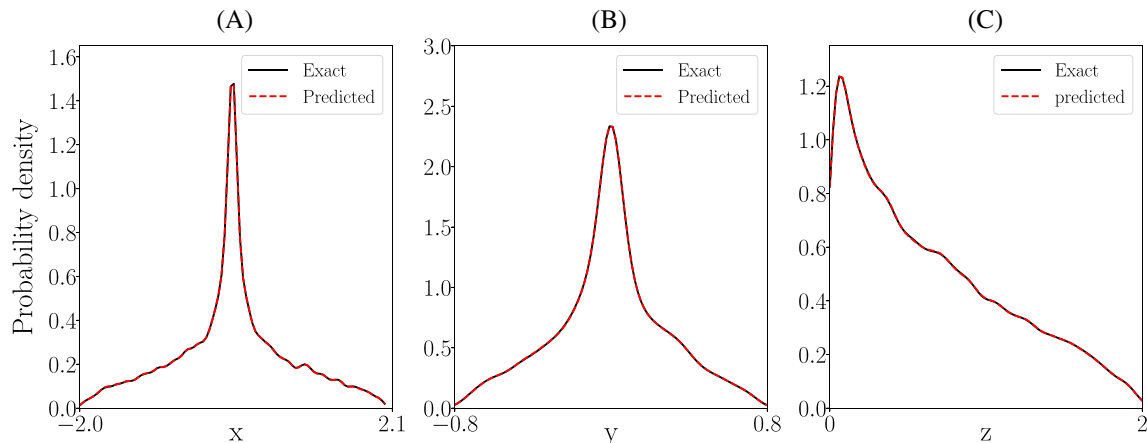


FIGURE 7 Example 1: Probability density functions for dimensions (A) x , (B) y , and (C) z

The electric potential $u(\mathbf{x})$ in Ω can be computed by solving the Laplace equation

$$-\nabla \cdot (\varepsilon(\mathbf{x}) \nabla u(\mathbf{x})) = 0, \quad \mathbf{x} \in \Omega, \quad (30a)$$

$$u(\mathbf{x}) = u^*(\mathbf{x}), \quad \mathbf{x} \in \Gamma_D, \quad (30b)$$

$$(\nabla u(\mathbf{x})) \cdot \mathbf{n} = (\nabla u^*(\mathbf{x})) \cdot \mathbf{n}, \quad \mathbf{x} \in \Gamma_N, \quad (30c)$$

where \mathbf{n} denotes the outer normal unit vector. The permittivity $\varepsilon(\mathbf{x})$ is given by

$$\varepsilon(\mathbf{x}) = \begin{cases} \varepsilon_c, & \mathbf{x} \in D_c, \\ \varepsilon_o, & \mathbf{x} \in \Omega \setminus D_c, \end{cases} \quad (31)$$

and u^* , which is also the analytical solution to the problem, is given by

$$u^*(\mathbf{x}) = -E_\infty x \begin{cases} 1 - \frac{\varepsilon_c/\varepsilon_o - 1}{\varepsilon_c/\varepsilon_o + 1} \frac{r_0^2}{x^2 + y^2}, & \mathbf{x} \in \Omega \setminus D_c, \\ \frac{2}{\varepsilon_c/\varepsilon_o + 1}, & \mathbf{x} \in D_c. \end{cases} \quad (32)$$

where E_∞ is the strength of the homogeneous electric field. We consider as random variables the cylinder's radius r_0 , and the strength of the electric field E_∞ . These variables are uniformly distributed as defined in Table 1.

The GDMaps based surrogate model presented in this article is composed of three distinct maps based on GH: (1) A global map between the parameter space and the Grassmannian diffusion manifold ($GH_0 : \Pi \rightarrow \Delta$); (2) a global map between the parameter space and the singular values space ($GH_1 : \Pi \rightarrow \Sigma$); and (3) a local map between the Grassmannian diffusion manifold and the tangent space of a region of the Grassmann manifold $\Lambda : \Delta \rightarrow \mathcal{T}$. Moreover, an additional map is built for the projection of the points in the tangent space onto the Grassmann manifold. Considering this sequence of mappings, one can expect that errors can propagate within the framework pipeline, negatively affecting the response prediction. Therefore, the identification of the source of errors is relevant for increasing the accuracy of the predicted outcomes. Therefore, the model presented in this example is used to demonstrate how the errors can be reduced by an appropriate selection of parameters for the construction of maps based on GH.

For the error analysis presented in this section, let's assume a training set of input parameters $\Theta = (r_0, E_\infty)$ of $N = 300$ samples and the corresponding electrical potential fields $\mathbf{X} = \mathcal{M}(\Theta) \in \mathbb{R}^{m \times n}$ over the domain Ω discretized into $(n \times m) = (300 \times 300) = 90,000$ mesh points. In this example, the Grassmann manifolds $\mathcal{G}(p, n) = \mathcal{G}(74, 300)$ and $\mathcal{G}(p, m) = \mathcal{G}(74, 300)$ are sufficient to encode the geometric structure of the electrical potential fields \mathbf{X} according to Equation (28). Using Algorithm 2, we obtain points on the Grassmannian diffusion manifold; and according to the parsimonious representation of the DMaps,²¹ only the first non-trivial diffusion coordinate is relevant to encode the geometrical information of the underlying physical phenomenon. Therefore, a dimension reduction from $n \times m = 90,000$ to $r = 1$ is achieved using GDMaps. This dimension reduction occurs because the electric potential $u^*(\mathbf{x})$ is linearly proportional to the strength of the electric field E_∞ . Therefore, E_∞ only influences the magnitude of $u^*(\mathbf{x})$ and not its shape. On the other hand, r_0 changes the geometry of the solution \mathbf{X} . Therefore, the intrinsic structure of the model's solution is in fact one-dimensional, and it is reflected in the discovered low-dimensional space.

To assess the ability of the global map $GH_0 : \Pi \rightarrow \Delta$ to learn the relationship between points in the parameter space and the structure of the data on the Grassmannian diffusion manifold (Δ), we start by sampling 1000 additional points

TABLE 1 Details of the state variables and input parameters of the dielectric cylinder suspended in homogeneous electric field

Description of variables/parameters		Uncertainty/value
Cylinder radius	r_0	$\sim \mathcal{U}(0.20, 0.70)$
Strength of electric field	E_∞	$\sim \mathcal{U}(8, 18)$
Relative permittivity of cylinder's material	ε_c	3
Relative permittivity of surrounding space	ε_o	1

Note: All sizes are expressed in SI units.

from Π to be mapped in Δ using GH_0 . In Figure 8, which shows both training data and predicted points in the Grassmannian diffusion coordinates, we see that GH_0 has adequately learned the shape of the data on the Grassmannian diffusion manifold using Gaussian kernel with length-scale equal to $\epsilon_G = 1$, and retaining $q = 50$ eigenvalues in the GH framework. However, these parameter, together with the number of samples in the training set, can influence the accuracy of the GH maps. Next, the influence of (ϵ_G) and q on the accuracy of both $GH_0 : \Pi \rightarrow \Delta$ and $GH_1 : \Pi \rightarrow \Sigma$ is analyzed.

We start by plotting the decay of the eigenvalues for GH_0 for different values of ϵ_G . It is clear from Figure 9 that as ϵ_G increases, the eigenvalues tends to decay quicker. This behavior will have a strong influence on the prediction error of GH_0 and GH_1 because the construction of the matrix \mathbf{B} in Algorithm 1 depends on the reciprocal of the corresponding eigenvalues. Therefore, if a larger number of very small eigenvalues are retained, their reciprocal could lead to large errors and numerical instabilities. Thus, the selection of ϵ_G is inherently connected with the number of eigenvalues one should retain for the construction of \mathbf{B} . This analysis is presented in Figure 10A for GH_0 , and in Figure 10B for GH_1 where we show the average error in the GH predictions for different combinations of ϵ_G and q . In both cases, we see that for large values of ϵ_G , a smaller number of eigenvectors and their respective eigenvalues should be retained in the construction of the matrix \mathbf{B} , which has a direct influence in its rank. Selecting a larger ϵ_G along with a high q will introduce large errors.

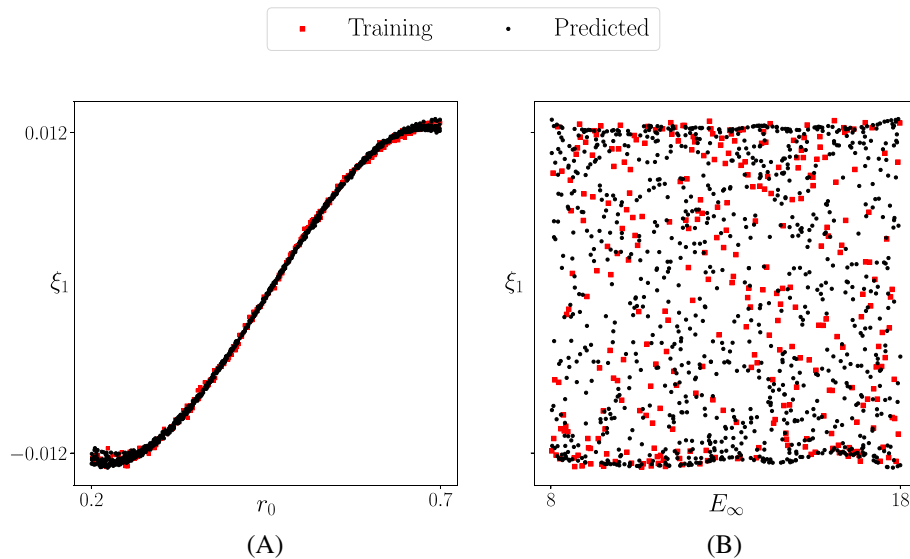


FIGURE 8 Example 2: The first Grassmannian diffusion coordinate ξ_1 , and the predicted coordinates of 1000 additional samples, as a function of (A) r_0 and (B) E_∞

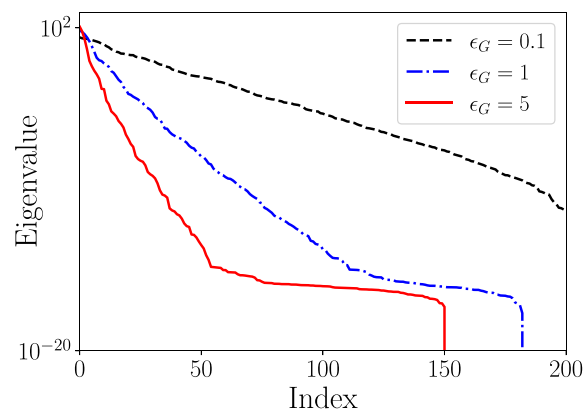


FIGURE 9 Example 2: Decay of the eigenvalues for the geometric harmonics surrogate GH_0 considering different length-scale parameters ϵ_G in the Gaussian kernel

Next, the influence of the size of the training set N is investigated. In this regard, the initial 300 samples and their respective diffusion coordinates are split into a training and a testing set to which the predicted Grassmannian diffusion coordinates can be compared. Considering N_{train} varying from 20 to 200, and keeping $\epsilon_G = 1$ and $q = 50$ constants, the average error ($\text{error}_{\|\cdot\|}$) as a function of N_{train} is presented in Figure 11A, for GH_0 , and in Figure 11B, for GH_1 . In both cases the error reduces, although at a limited rate, after a certain value of N due to the fact that a residual error remains due to the selected values for ϵ_G and q .

Next, we investigate the accuracy of the local maps, $\Lambda_0(\cdot)$ and $\Lambda_1(\cdot)$, between Δ (Grassmannian diffusion space) and the tangent spaces $\mathcal{T}_{\hat{\mu}_u}$ and $\mathcal{T}_{\hat{\mu}_v}$, respectively; together with the mapping from the tangent spaces to the Grassmann manifold, and from the Grassmann manifold to the response space via the SVD-like reconstruction. The exact diffusion coordinates for each of the 300 training points are used, and the response prediction of the corresponding points are obtained using the local maps with different numbers of neighbors (k) used on the construction of the tangent space. The probability density functions (PDFs) for $\text{error}_{\|\cdot\|}$ are estimated using kernel density estimation for $k = 3, 5$, and 10 (closest neighbors points) in Figure 12A. We clearly see that the errors induced by these local maps are minimal and that they are not strongly influenced by k in this specific problem, because a point is predicted in a region close to the reference point where the tangent space is constructed on. Note also that the length-scale parameter for the local maps is taken to be 0.25 times the square value of the median of the pairwise distances of the k neighbors.

The cumulative error associated with the full process is analyzed by drawing 1000 sample points from the parameter space and computing posterior error estimation ($\text{error}_{\|\cdot\|}$) of the predicted response. The estimated PDF of the error is shown in Figure 12B, where the mean is equal to 1.8489×10^{-3} and the standard deviation is equal to 2.7598×10^{-3} . This reveals that the overall errors in the prediction solutions are very small compared to their true solutions, even considering a training set of only 300 points.

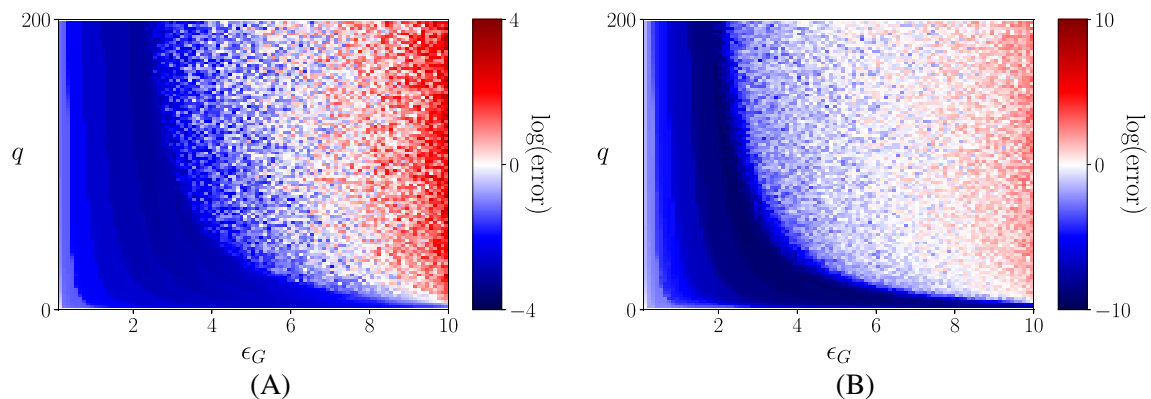


FIGURE 10 Example 2: Average error from 100 test samples for different combinations of the number of eigenvalues retained in the construction of the geometric harmonics (q) and the value to the length-scale parameter (ϵ_G) for (A) surrogate GH_0 , and (B) surrogate (GH_1)

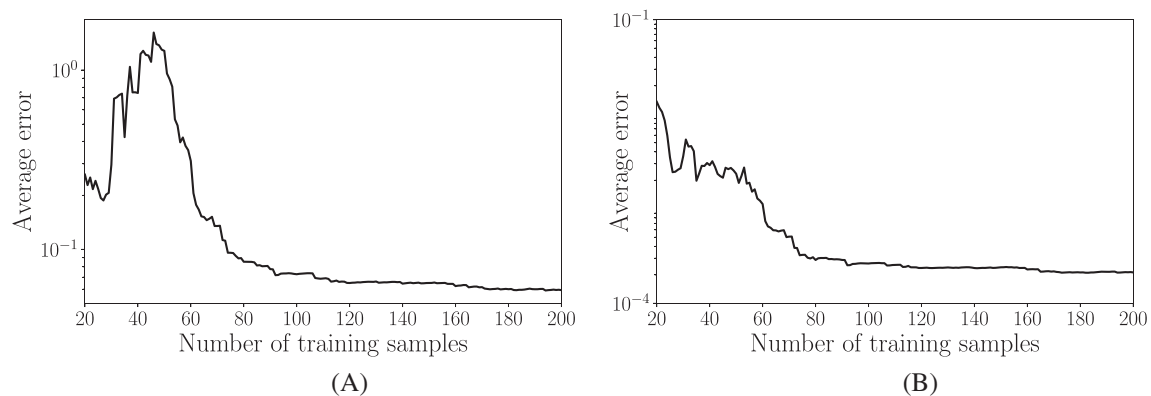


FIGURE 11 Example 2: Average error from surrogates GH_0 (A) and GH_1 (B) for increasing size of the training data set

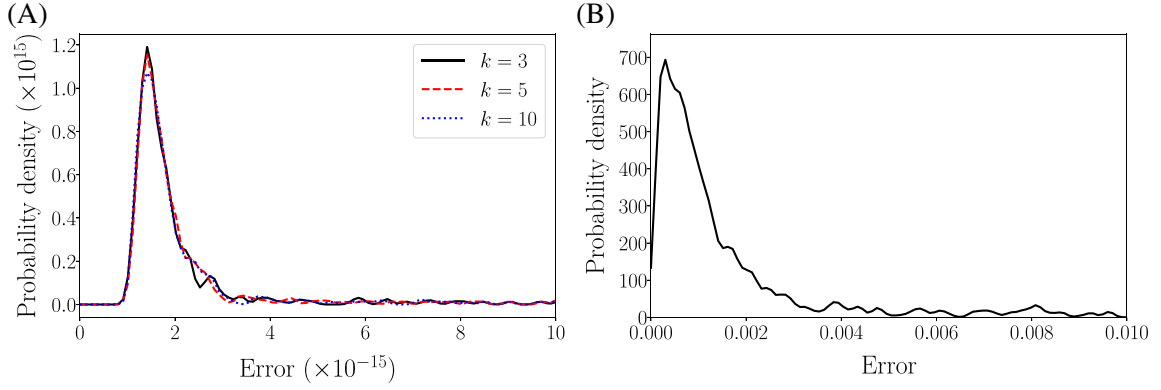


FIGURE 12 Example 2: PDFs of the errors for: (A) Local maps for different k (300 samples), (B) overall response prediction using all maps from 1000 additional samples

Finally, for illustration we consider three test points $\Theta_a^* = (r_0, E_\infty) = (0.25, 17)$, $\Theta_b^* = (r_0, E_\infty) = (0.40, 10)$ and $\Theta_c^* = (r_0, E_\infty) = (0.65, 15)$. The exact electrical potential fields are obtained by numerically solving the model in Equation (30) (Figure 13A,D,G) on a 300×300 meshed domain. The predict electric fields (Figure 13B,E,H) are obtained using the surrogate model developed herein. The exact and predicted solutions are compared by error_{rel} in Figure 13C,F,I, where the corresponding error ($\text{error}_{||\cdot||}$) are 5.4862×10^{-4} , 4.3704×10^{-4} , and 1.8306×10^{-3} for Θ_a^* , Θ_b^* , and Θ_c^* , respectively. We can see that our surrogate model, which reduces the dimension of the solution from 90,000 spatial points to a single Grassmannian diffusion coordinate is very accurate.

6.3 | Continuum modeling of plasticity in an amorphous solid

An important theoretical hypothesis about the behavior of amorphous solids (e.g., metallic glasses) says that irreversible plastic deformation is mediated through atomic rearrangements in small clusters of atoms known as STZs.⁶⁴ Consequently, amorphous materials subject to large shear stresses are often prone to the formation of shear bands due to the rearrangements of STZs in localized regions. It has been proposed that one can connect these large-scale plastic deformations to an effective temperature⁶⁵

$$T_{eff} = \frac{\partial U_c}{\partial S_c}, \quad (33)$$

where U_c and S_c are the potential energy and entropy of the configurational degrees of freedom under the assumption that both the total energy U and total entropy S are separated into kinetic and configurational components, that is, $U = U_c + U_k$ and $S = S_c + S_k$, respectively. This effective temperature provides a measure of the degree of structural disorder (characterizing the density of STZs) and can be dedimensionalized as

$$\chi = \frac{k_B T_{eff}}{E_z}, \quad (34)$$

where $k_B = 1.38 \times 10^{-23}$ is the Boltzmann factor, E_z is the STZ formation energy.

Given a spatially varying initial effective temperature field χ on a material domain, the STZ theory defines two coupled equations to model the evolution of plastic strain in the material. The first describes a plastic flow rule that relates the plastic rate of deformation tensor \mathbf{D}^{pl} to the effective temperature as:

$$\mathbf{D}^{pl} = \frac{1}{\tau_0} \exp \left\{ - \left(\frac{E_z}{k_B \chi} + \frac{\Delta_\star}{k_B T} \right) \right\} \cosh \left(\frac{\Omega \epsilon_0 \bar{\sigma}}{k_B T} \right) \left(1 - \frac{\sigma_y}{\bar{\sigma}} \right), \quad (35)$$

Note that this flow rule is monotonic with respect to $\bar{\sigma}/\sigma_y$, with $\bar{\sigma} = |\sigma_\theta|$ given as the magnitude of the deviatoric shear stress $\sigma_\theta = \sigma = \frac{1}{3} \mathbf{1} \text{Tr}(\sigma)$. Therefore, plastic deformation does not occur when $\bar{\sigma}/\sigma_y < 1$. The second equation describes the evolution of χ as

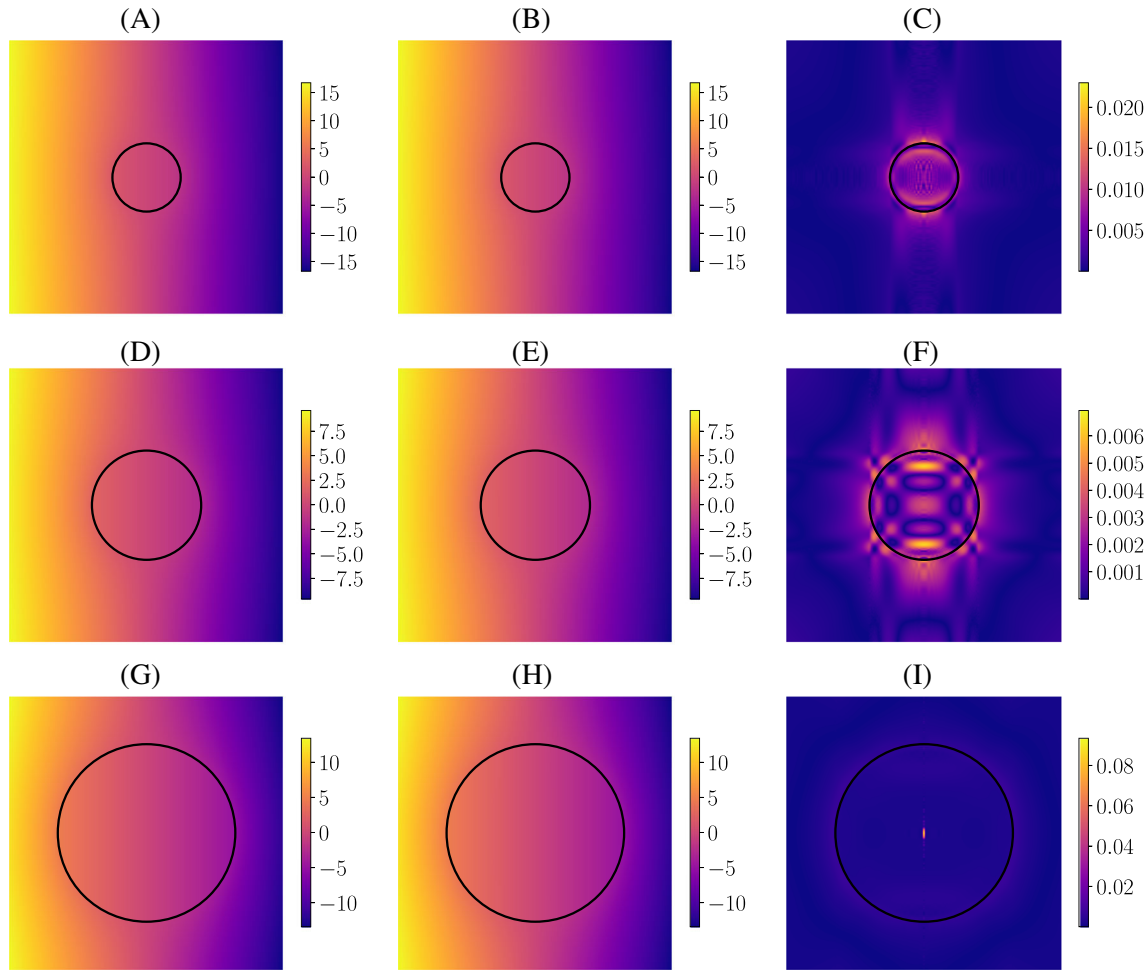


FIGURE 13 Example 2: Left—Exact electric potential for (A) Θ_a^* , (D) Θ_b^* , and (G) Θ_c^* . Center—Predicted electric potential for (B) Θ_a^* , (E) Θ_b^* , and (H) Θ_c^* . Right—Relative errors for (C) Θ_a^* , (F) Θ_b^* , and (I) Θ_c^*

$$c_0 \dot{\chi} = \frac{1}{\sigma_y} (\mathbf{D}^{pl} : \boldsymbol{\sigma}_0) (\chi_\infty - \chi) + \nabla \cdot D_\chi \nabla \chi, \quad (36)$$

where $D_\chi = l^2 \sqrt{\mathbf{D}^{pl} : \mathbf{D}^{pl}}$. Other parameters are defined in Table 2.

In this example, a numerical scheme developed by Rycroft et al.^{66,67} is used to solve the system of Equations (35) and (36). This method utilizes an Eulerian finite-difference method under quasi-static conditions. As mentioned previously, the STZ theory assumes that the effective temperature has a spatial distribution that influences the material response. Therefore, the evolution of Equations (35) and (36) depends on the initial χ field. This initial χ field is defined from an atomic system simulated using molecular dynamics (MD) and coarse-grained onto an equally spaced grid, where the weighted average potential energy of each atom within a radius r_c around each node is determined using a Gaussian windowing with length scale l_χ . The value of χ is determined through an affine transformation of the weighted average atomic potential energies. See References 68,69 for a detailed explanation and methods to obtain the parameters of this affine transformation. Despite these efforts, there remains significant uncertainty as to the appropriate range of χ for these fields. We therefore, consider that a realization of the χ can be obtained from MD simulations with uncertain mean and standard deviation.

Observations further suggest that this field can be approximated as Gaussian.^{56,68,69} To approximate the initial χ field using the Gaussian assumption, a sample of a zero-mean delta-correlated Gaussian field is obtained using the Box-Muller transform. This sample is then smoothed with a bi-variate Gaussian of length scale ($l_\chi = 10\text{\AA}$) centered at the nodes of the grid, where in this example $r_c = 5$. Next, the smoothed field is multiplied by σ_χ , normalized by T times a Gaussian factor, and added to μ_χ , which is also normalized by T . Herein, it is assumed that μ_χ and $c_\chi = \sigma_\chi / \mu_\chi$ are both uncertain

with uniform distributions as described in Table 3. Further, only the uncertainties in μ_χ and c_χ , and not the uncertainties due to realizations of the random field, are considered because we want to illustrate the influence of the relative scale of fluctuations on the material behavior in the evolution of the STZ model. In the simulations, a simple shear up to 50% strain is imposed to a simulation box of size $400\text{\AA} \times 400\text{\AA}$. A grid of size 32×32 is considered in the discretization, where each element has a size of $12.5\text{\AA} \times 12.5\text{\AA}$. Therefore, each snapshot of this simulation is given by a matrix $\mathbf{X}_i \in \mathbb{R}^{32 \times 32}$.

We obtain $N = 196$ samples of the pair $\Theta = (\mu_\chi, c_\chi)$ via stratified sampling to train a surrogate model for full evolution of the plastic strain field. The evolution of the plastic strain field for a given pair $\Theta = (\mu_\chi, c_\chi)$ is presented in Figure 14 as a sequence of 101 snapshots of size L_x and L_y at discrete values of the imposed shear strain $\bar{\epsilon}$. A matrix $\mathbf{Y} \in \mathbb{R}^{1024 \times 101}$ is then constructed for a given pair $\Theta = (\mu_\chi, c_\chi)$, where each column of \mathbf{Y} correspond to the vectorized snapshot of the plastic strain field.

In this problem, two Grassmann manifolds given by $\mathcal{G}(p, n)$ (left manifold) and $\mathcal{G}(p, m)$ (right manifold) are associated with the left and right singular vectors of the matrices \mathbf{Y} ; where $n = 1024$; $m = 101$; and $p = 10$, which is the maximum rank among the matrices \mathbf{Y} , suffices to encode the geometric structure of each data point. Using GDMaps, we obtain a

TABLE 2 Parameters for the STZ plasticity model for a bulk metallic glass material

Parameter	Unit	Value	Description
σ_y	GPa	0.7	Yield stress
τ_0	s	10^{-13}	Molecular vibration timescale
ϵ_0	-	0.333	Typical local strain at STZ transition
Δ_*/k_B	K	7948	Typical activation temperature
Ω/k_B	\AA^3	349	Typical activation volume
T	K	97	Bath temperature
χ_∞	K	1050.6	Steady-state effective temperature
E_z/k_B	K	21,000	STZ formation energy
c_0	-	0.414	Plastic work fraction
l	\AA	10	Diffusion length scale

TABLE 3 Probability distributions of the STZ random field parameters

Description of variables/parameters	Uncertainty/value
Mean	μ_χ $\sim \mathcal{U}(500, 700)$
Coefficient of variation	c_χ $\sim \mathcal{U}(0, 0.1)$

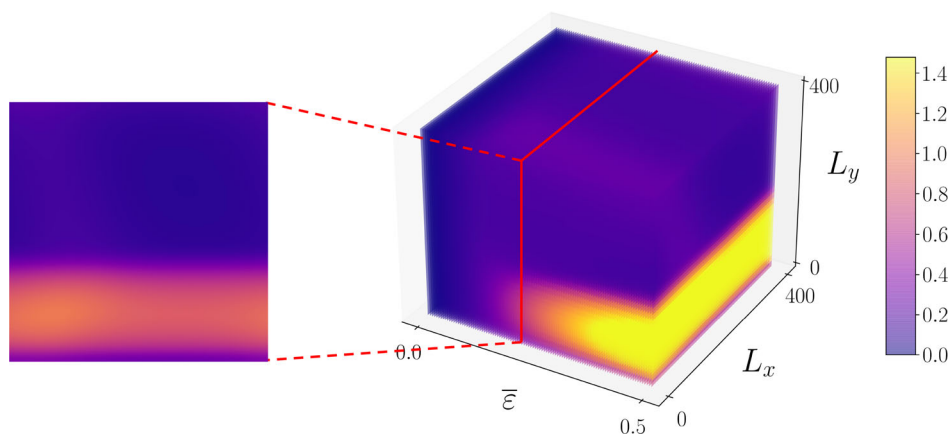


FIGURE 14 Evolution of plastic strain with a snapshot of the strain field at a given strain level extracted

set of 196 Grassmannian diffusion coordinates embedding the high-dimensional data into a low-dimensional Euclidean space. Figure 15 shows the first three non-trivial diffusion coordinates. In total, we use a total of $r = 8$ diffusion coordinates estimated by the number of eigenvalues whose values are larger than 10% of the value of the first non-trivial eigenvalue. The hyper-parameters q of the GH kernel for GH_0 and GH_1 are set equal to 50. Further, their length scale parameters are equal to the median squared of the pairwise distances of the points in S_Θ and the number k of neighbors points in the diffusion space used to support the construction of the tangent space on the corresponding Grassmann manifold is equal to 5. We recommend that k is not large because we want to reduce the distortions due to the exponential mapping from the tangent space to the Grassmann manifold. Moreover, standard searching techniques such as direct grid search or a randomized search could be employed to tune the hyper-parameters.

Once the surrogate model is trained using the 196 samples obtained previously, we can predict the full evolution of the plastic strain field for any pair $\Theta = (\mu_\chi, c_\chi)$. Considering a representative case with $\Theta = (530.1748, 0.0792)$, the simulated and predicted evolution of the plastic strain field, as well as their relative error, are presented in Figure 16 for five different levels of imposed strain (0%, 12.5%, 25%, 37.5%, 50%). The error ($\text{error}_{\|\cdot\|}$) for this plastic strain field is equal to 2.0473×10^{-3} . Next, considering 100 additional samples we compute the mean and standard deviation of the plastic strain fields at the different levels of imposed strain (25%, 30%, 35%, 40%, 45%, 50%) as presented in Figures 17, and 18. These figures include the statistical characterization obtained by using the numerical model and the surrogate model, along with the relative errors. From these results, we see that the surrogate model developed herein can predict the uncertain response of a complex model with high-accuracy, by taking advantage of low-dimensional subspace structure of the problem to reduce the computational burden associated with running high-fidelity models for UQ.

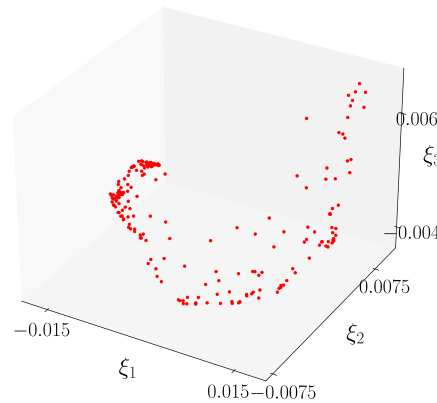


FIGURE 15 Grassmannian diffusion coordinates

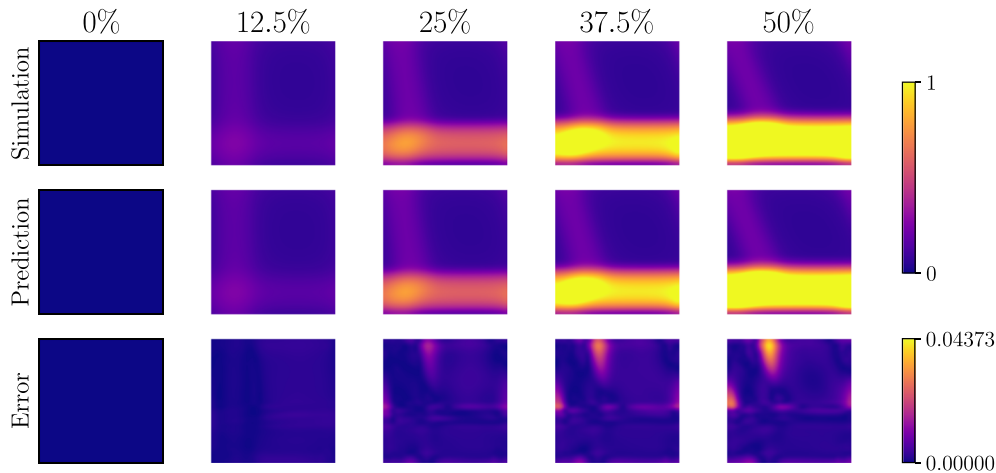


FIGURE 16 Exact and simulated strain field evolution for $(\mu_\chi, c_\chi) = (530.1748, 0.079)$ and the corresponding errors

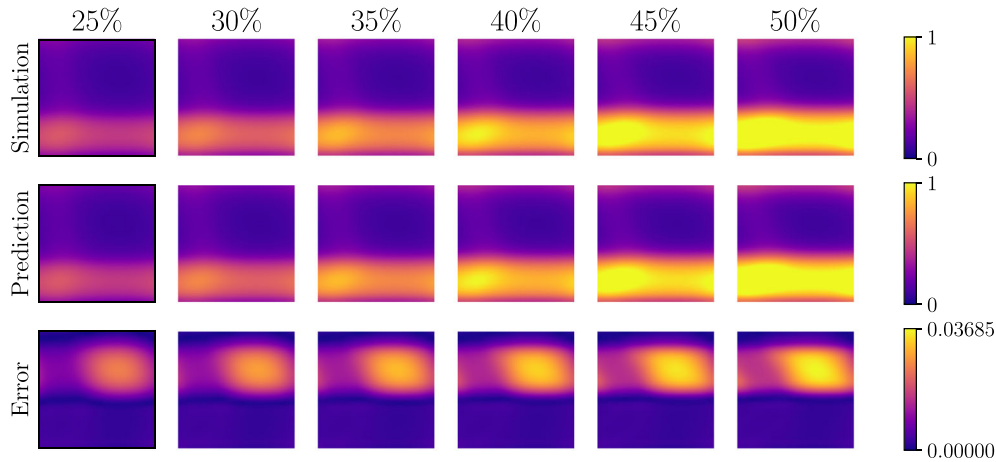


FIGURE 17 Simulated and predicted evolution of the mean of the strain field for 100 additional samples and the corresponding errors

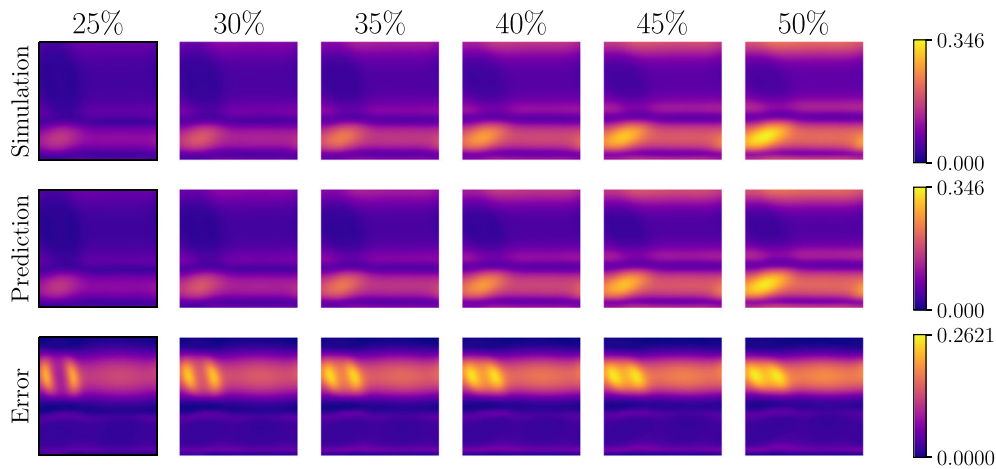


FIGURE 18 Simulated and predicted evolution of the standard deviation of the strain field for 100 additional samples and the corresponding errors

This example is also used to demonstrate the computational performance of the proposed surrogate modeling technique. Here, the computational performance of the training and prediction phases are statistically compared with the processing time required to obtain the evolution of the strain field by solving Equations (35) and (36) using a finite differences scheme. To obtain a statistically consistent measure of the surrogate model training time from 196 data points, the training is repeated 100 times. The average training time is equal to 9.7335 s with standard deviation equal to 0.7638 s. Next, 100 additional samples of input parameters are selected at random, and the average prediction time for the surrogate model is 0.0594 s, with standard deviation equal to 0.0173 s. However, for the same 100 additional samples of input parameters, the numerical scheme based on finite-differences takes 4.2863 s, on average, to estimate the full strain field, with standard deviation of 0.4806 s. This means that the prediction using the surrogate model is more than 72 times faster than the high-fidelity model. Furthermore, this finite difference model is not a particularly expensive computational model. The very rapid surrogate model predictions would be even more advantageous computationally for more expensive models.

7 | CONCLUSIONS

This article introduces a fully data-driven surrogate model for uncertainty quantification of high-dimensional models of complex physical/engineering systems. It takes advantage of the low-dimensional representation of high-dimensional

input/output data obtained via GDMaps to create a set of GH based maps. A global map is constructed to predict the Grassmannian diffusion coordinates corresponding to any new element in the set of input parameters with good accuracy. Once the Grassmannian diffusion coordinates corresponding to a new set of input parameters are predicted, the k -nearest neighbors points in the Grassmannian diffusion space and their associated points on the Grassmann manifold are utilized to estimate, via GH, a local map from the Grassmannian diffusion space to a tangent space. Next, the exponential map project the point onto the Grassmann manifold, a result used to predict the high-fidelity solution of the problem.

The method developed herein used the descriptive power of the GDMaps and the computational performance of GH to provide an efficient and accurate prediction of the solution of complex systems described by algebraic equations and partial/ordinary differential equations. Three examples were considered to evaluate the performance of this technique. The first one consisted of a toy example to demonstrate the ability of the technique to predict data with complex geometry using spectral methods in a way that is easy to understand and visualize. In the second example, the performance of the surrogate modeling developed herein was verified in a physical model (i.e., electric potential of a cylinder in homogeneous electric field) with high-dimensional response, also considering discontinuities in the system response. It was demonstrated that some parameters such as the length-scale parameter and the number of retained eigenvalues, both for the Gaussian kernel used in the GH framework; as well as the amount of data in the training set, are important quantities affecting the accuracy of the presented technique. The third problem analyzed in this article evaluated the plastic deformation of amorphous solids using the STZ theory of plasticity. The uncertainty was imposed in the mean and coefficient of variation of the initial nondimensionalized effective temperature field (χ). In this case, the evolution of the strain field with the strain level is also taken into consideration, and the uncertainties of the plastic strain field are predicted accurately.

In all the cases considered herein, a good accuracy was identified in the predicted solutions in comparison with the exact ones. The method proves advantageous due to its computational performance and ability to make reliable predictions for high-dimensional responses considering a highly sparse set of points.

ACKNOWLEDGMENT

This material is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research under Award Number DE-SC0020428.

CONFLICT OF INTEREST

The authors declare no potential conflict of interests.

ORCID

Ketson R. M. dos Santos  <https://orcid.org/0000-0002-4049-678X>

Dimitris G. Giovanis  <https://orcid.org/0000-0003-2272-2584>

Katiana Kontolati  <https://orcid.org/0000-0003-2027-9638>

Dimitrios Loukrezis  <https://orcid.org/0000-0003-1264-1182>

Michael D. Shields  <https://orcid.org/0000-0003-1370-6785>

REFERENCES

1. Le Maître OP, Knio OM. *Non-Intrusive Methods*. Springer; 2010:45-72.
2. Smith R. *Uncertainty Quantification: Theory, Implementation, and Applications*. Computational Science and Engineering. SIAM; 2013.
3. Ghanem R, Spanos P. *Stochastic Finite Elements: A Spectral Approach*. Civil, Mechanical and Other Engineering Series. Dover Publications; 2003.
4. Babuška I, Nobile F, Tempone R. A stochastic collocation method for elliptic partial differential equations with random input data. *SIAM Rev.* 2010;52(2):317-355. doi:10.1137/100786356
5. Chen P, Quarteroni A, Rozza G. Reduced order methods for uncertainty quantification problems. Technical report, Zürich, Switzerland; 2015.
6. Fishman G. *Monte Carlo*. Springer Series in Operations Research and Financial Engineering. Springer; 1996.
7. Caflisch RE. Monte Carlo and quasi-Monte Carlo methods. *Acta Numerica*. 1998;7:1-49. doi:10.1017/S0962492900002804
8. Shields MD, Teferra K, Hapij A, Daddazio RP. Refined stratified sampling for efficient Monte Carlo based uncertainty quantification. *Reliab Eng Syst Saf.* 2015;142:310-325. doi:10.1016/j.res.2015.05.023
9. Shields MD. Adaptive Monte Carlo analysis for strongly nonlinear stochastic systems. *Reliab Eng Syst Saf.* 2018;175:207-224. doi:10.1016/j.res.2018.03.018

10. McKay MD, Beckman RJ, Conover WJ. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics*. 1979;21(2):239-245.
11. Echard B, Gayton N, Lemaire M. AK-MCS: an active learning reliability method combining Kriging and Monte Carlo simulation. *Struct Saf*. 2011;33(2):145-154. doi:10.1016/j.strusafe.2011.01.002
12. dos Santos KRM, Beck AT. A benchmark study on intelligent sampling techniques in Monte Carlo simulation. *Latin Am J Solids Struct*. 2015;12:624-648.
13. Botev Z, Ridder A. *Variance Reduction*. American Cancer Society; 2017:1-6.
14. Krige D. A statistical approach to some basic mine valuation problems on the Witwatersrand. *J South Afr Inst Min Metall*. 1951;52(6):119-139. doi:10.10520/AJA0038223X_4792
15. Rasmussen CE. *Gaussian Processes in Machine Learning*. Springer; 2004:63-71.
16. Wiener N. The homogeneous chaos. *Am J Math*. 1938;60(4):897-936.
17. Xiu D, Karniadakis GE. Modeling uncertainty in flow simulations via generalized polynomial chaos. *J Comput Phys*. 2003;187(1):137-167. doi:10.1016/S0021-9991(03)00092-5
18. Nadler B, Lafon S, Coifman RR, Kevrekidis IG. Diffusion maps, spectral clustering and reaction coordinates of dynamical systems. *Appl Comput Harmonic Anal*. 2006;21(1):113-127. Special Issue: Diffusion Maps and Wavelets. doi:10.1016/j.acha.2005.07.004
19. Erban R, Frewen TA, Wang X, et al. Variable-free exploration of stochastic models: a gene regulatory network example. *J Chem Phys*. 2007;126(15):155103. doi:10.1063/1.2718529
20. Coifman RR, Kevrekidis IG, Lafon S, Maggioni M, Nadler B. Diffusion maps, reduction coordinates, and low dimensional representation of stochastic systems. *Multisc Model Simul*. 2008;7(2):842-864. doi:10.1137/070696325
21. Dsilva CJ, Talmon R, Coifman RR, Kevrekidis IG. Parsimonious representation of nonlinear dynamical systems through manifold learning: a chemotaxis case study. *Appl Computat Harmonic Anal*. 2015;44(3):759-773. doi:10.1016/j.acha.2015.06.008
22. Pearson K. LIII. On lines and planes of closest fit to systems of points in space. *Lond Edinburgh Dublin Philos Mag J Sci*. 1901;2(11):559-572.
23. He X, Yan S, Hu Y, Niyogi P, Zhang HJ. Face recognition using Laplacian faces. *IEEE Trans Pattern Anal Mach Intell*. 2005;27(3):328-340.
24. Rangarajan L, Nagabhushan P. Linear regression for dimensionality reduction and classification of multi dimensional data. In: Pal SK, Bandyopadhyay S, Biswas S, eds. *Pattern Recognition and Machine Intelligence*. Springer; 2005:193-199.
25. Strang G. *Introduction to Linear Algebra*. Wellesley-Cambridge Press; 2016.
26. Tenenbaum JB, Silva V, Langford JC. A global geometric framework for nonlinear dimensionality reduction. *Science*. 2000;290(5500):2319-2323. doi:10.1126/science.290.5500.2319
27. Roweis ST, Saul LK. Nonlinear dimensionality reduction by locally linear embedding. *Science*. 2000;290(5500):2323-2326. doi:10.1126/science.290.5500.2323
28. Donoho DL, Grimes C. Hessian eigenmaps: locally linear embedding techniques for high-dimensional data. *Proc Natl Acad Sci*. 2003;100(10):5591-5596. doi:10.1073/pnas.1031596100
29. Lataniotis C, Marelli S, Sudret B. Extending classical surrogate modeling to high dimensions through supervised dimensionality reduction: a data-driven approach. *Int J Uncertain Quantif*. 2020;10(1):55-82.
30. Coifman RR, Lafon S, Lee AB, et al. Geometric diffusions as a tool for harmonic analysis and structure definition of data: diffusion maps. *Proc Nat Acad Sci*. 2005;102(21):7426-7431. doi:10.1073/pnas.0500334102
31. Coifman RR, Lafon S. Diffusion maps. *Appl Comput Harmonic Anal*. 2006;21(1):5-30. Special Issue: Diffusion Maps and Wavelets. doi:10.1016/j.acha.2006.04.006
32. Soize C, Ghanem R. Data-driven probability concentration and sampling on manifold. *J Comput Phys*. 2016;321:242-258. doi:10.1016/j.jcp.2016.05.044
33. Kalogeris I, Papadopoulos V. Diffusion maps-based surrogate modeling: an alternative machine learning approach. *Int J Numer Methods Eng*. 2020;121(4):602-620. doi:10.1002/nme.6236
34. Kalogeris I, Papadopoulos V. Diffusion maps-aided neural networks for the solution of parametrized PDEs. *Comput Methods Appl Mech Eng*. 2021;376:113568. doi:10.1016/j.cma.2020.113568
35. dos Santos KRM, Giovanis DG, Shields MD. Grassmannian diffusion maps based dimension reduction and classification for high-dimensional data; 2021. arXiv:2009.07547.
36. Kontolati K, Loukrezis D, dos Santos KRM, Giovanis DG, Shields MD. Manifold learning-based polynomial chaos expansions for high-dimensional surrogate models; 2021. arXiv:2107.09814.
37. Burt P, Adelson E. The Laplacian PYRAMID AS A COMPACT IMAGE CODE. *IEEE Trans Commun*. 1983;31(4):532-540. doi:10.1109/TCOM.1983.1095851
38. Broomhead D, Lowe D. Multivariable functional interpolation and adaptive networks. *Complex Syst*. 1988;2:321-355.
39. Majdisova Z, Skala V. Radial basis function approximations: comparison and applications. *Appl Math Model*. 2017;51:728-743. doi:10.1016/j.apm.2017.07.033
40. Leeb W. Properties of Laplacian pyramids for extension and denoising; 2019. arXiv:1909.07974.
41. Williams C, Seeger M. Using the Nyström method to speed up kernel machines. *Neural Inf Process Syst*. 2001;13:682-688.
42. Heimowitz A, Eldar YC. The nystrom extension for signals defined on a graph; 2018:4199-4203.
43. Coifman RR, Lafon S. Geometric harmonics: a novel tool for multiscale out-of-sample extension of empirical functions. *Appl Comput Harmonic Anal*. 2006;21(1):31-52. Special Issue: Diffusion Maps and Wavelets. doi:10.1016/j.acha.2005.07.005
44. Olivier A, Giovanis D, Aakash B, Chauhan M, Vandanapu L, Shields M. UQpy: a general purpose Python package and development environment for uncertainty quantification. *J Comput Sci*. 2020;101204.

45. Auslander L, MacKenzie R. *Introduction to Differentiable Manifolds*. Dover Books on Mathematics. Dover Publications; 2012.
46. Ye K, Lim LH. Schubert varieties and distances between subspaces of different dimensions; 2014. arXiv:1407.0900.
47. Ye K, Wong KSW, Lim LH. Optimization on flag manifolds; 2019. arXiv:1907.00949.
48. Lim LH, Sze-Wai Wong K, Ye K. Numerical algorithms on the affine Grassmannian. *SIAM J Matrix Anal Appl.* 2019;40(2):371-393. doi:10.1137/18M1169321
49. Edelman A, Arias TA, Smith ST. The geometry of algorithms with orthogonality constraints. *SIAM J Matrix Anal Appl.* 1998;20(2):303-353.
50. Amsallem D, Farhat C. Interpolation method for adapting reduced-order models and application to aeroelasticity. *AIAA J.* 2008;46(7):1803-1813.
51. Hamm J, Lee DD. Grassmann discriminant analysis: a unifying view on subspace-based learning. Proceedings of the 25th International Conference on Machine Learning; 2008:376-383; Association for Computing Machinery.
52. Miao J, Ben-Israel A. On principal angles between subspaces in R^n . *Linear Algebra Appl.* 1992;171:81-98.
53. Wong YC. Differential geometry of Grassmann manifolds; 1967:589-594.
54. Giovanis D, Shields M. Uncertainty quantification for complex systems with very high dimensional response using Grassmann manifold variations. *J Comput Phys.* 2018;364:393-415. doi:10.1016/j.jcp.2018.03.009
55. Karcher H. Riemannian center of mass and mollifier smoothing. *Commun Pure Appl Math.* 1977;30(5):509-541. doi:10.1002/cpa.3160300502
56. Giovanis D, Shields M. Data-driven surrogates for high dimensional models using Gaussian process regression on the Grassmann manifold. *Comput Methods Appl Mech Eng.* 2020;370:113269. doi:10.1016/j.cma.2020.113269
57. Hamm J, Lee DD. Extended Grassmann kernels for subspace-based learning. In: Koller D, Schuurmans D, Bengio Y, Bottou L, eds. *Advances in Neural Information Processing Systems 21*. Curran Associates, Inc; 2009:601-608.
58. Harandi MT, Salzmann M, Jayasumana S, Hartley R, Li H. Expanding the family of Grassmannian kernels: an embedding perspective; 2014. arXiv:1407.1123.
59. dos Santos KR, Brudastova O, Kougiumtzoglou IA. Spectral identification of nonlinear multi-degree-of-freedom structural systems with fractional derivative terms based on incomplete non-stationary data. *Struct Saf.* 2020;86:101975.
60. Chiavazzo E, Gear CW, Dsilva CJ, Rabin N, Kevrekidis IG. Reduced models in chemical kinetics via nonlinear data-mining. *Processes.* 2014;2(1):112-140.
61. Erichson NB, Mathelin L, Brunton SL, Kutz JN. Diffusion Maps meet Nyström; 2018. arXiv:1802.08762.
62. Bermanis A, Averbuch A, Coifman RR. Multiscale data sampling and function extension. *Appl Comput Harmonic Anal.* 2013;34(1):15-29. doi:10.1016/j.acha.2012.03.002
63. Kontolati K, Loukrezis D, Giovanis DG, Vandanapu L, Shields MD. A survey of unsupervised learning methods for high-dimensional uncertainty quantification in black-box-type problems; 2021. arxiv.
64. Falk ML, Langer JS. Dynamics of viscoplastic deformation in amorphous solids. *Phys Rev E.* 1998;57:7192-7205.
65. Bouchbinder E, Langer J. Nonequilibrium thermodynamics of driven amorphous materials. III. Shear-transformation-zone plasticity. *Phys Rev E.* 2009;80(3):031133.
66. Rycroft CH, Sui Y, Bouchbinder E. An Eulerian projection method for quasi-static elasto-plasticity. *Comput Phys.* 2008;30:1-14.
67. Boffi NM, Rycroft CH. Parallel three-dimensional simulations of quasi-static elastoplastic solids. *Comput Phys Commun.* 2020;70:107254.
68. Hinkle AR, Rycroft CH, Shields MD, Falk ML. Coarse graining atomistic simulations of plastically deforming amorphous solids. *Phys Rev E.* 2017;95:053001. doi:10.1103/PhysRevE.95.053001
69. Kontolati K, Alix-Williams D, Boffi NM, Falk ML, Rycroft CH, Shields MD. Manifold learning for coarse-graining atomistic simulations: application to amorphous solids. *Acta Mater.* 2021;117008.

How to cite this article: dos Santos KRM, Giovanis DG, Kontolati K, Loukrezis D, Shields MD. Grassmannian diffusion maps based surrogate modeling via geometric harmonics. *Int J Numer Methods Eng.* 2022;1-23. doi: 10.1002/nme.6977