

# A non-linear dimension reduction methodology for generating data-driven stochastic input models<sup>1</sup>

Nicholas Zabaras and Baskar Ganapathysubramanian  
Materials Process Design and Control Laboratory  
Sibley School of Mechanical and Aerospace Engineering  
101 Frank H.T. Rhodes Hall  
Cornell University  
Ithaca, NY 14853-3801

Stochastic analysis of random heterogeneous media provides information of significance only if realistic input models of the topology and material property variations are used. This work introduces a framework to construct such input stochastic models for the topology, thermal diffusivity and permeability variations in heterogeneous media using a data-driven strategy. Given a set of microstructure realizations (input samples) generated from given statistical information about the medium topology, the framework constructs a reduced-order stochastic representation of the topology and material properties. This problem of constructing a low-dimensional stochastic representation of property variations is analogous to the problem of manifold learning and parametric fitting of hyper-surfaces encountered in image processing and psychology.

Denote by  $\mathbf{M}$  the set of microstructures that satisfy the given experimental statistics. A non-linear dimension reduction strategy is utilized to map  $\mathbf{M}$  to a low-dimensional region,  $\mathbf{A}$ . We first show that  $\mathbf{M}$  is a compact manifold embedded in a high-dimensional input space  $\mathbb{R}^n$ . An isometric mapping  $F$  from  $\mathbf{M}$  to a low-dimensional, compact, connected set  $A \subset \mathbb{R}^d, d \ll n$  is constructed. Given only a finite set of samples of the data, the methodology uses arguments from graph theory and differential geometry to construct the isometric transformation  $F: \mathbf{M} \rightarrow A$ . Asymptotic convergence of the representation of  $\mathbf{M}$  by  $A$  is shown. This mapping  $F$  serves as an accurate, low-dimensional, data-driven representation of the property variations.

The reduced-order model of the material topology and property variations is subsequently used as an input in the solution of stochastic PDEs with multiscale features that describe the evolution of the dependent variables. A sparse grid collocation strategy (Smolyak algorithm) with piecewise multi-linear hierarchical interpolation functions is utilized to solve these stochastic equations efficiently. We showcase the methodology by constructing low-dimensional input stochastic models to represent thermal diffusivity and permeability in porous media. These models are then used in analyzing the effects of topological variations on the evolution of temperature and flow in random heterogeneous media.

---

<sup>1</sup> Presented at the 'Algorithms and Analysis in Uncertainty Quantification' symposium in the 2008 SIAM Annual Meeting, (Xiaoliang Wan and G. Karniadakis, organizers), San Diego CA, July 7-11, 2008.