Analysis and Exploration of High Dimensional Landscapes of Simulation Ensembles for Uncertainty Quantification

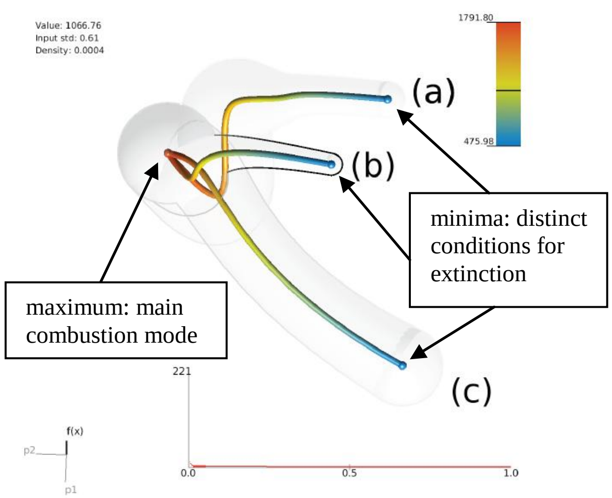
Topology for Statistical Modeling of Petascale Data

S. Gerber, P.-T. Bremer, R. Whitaker, and V. Pascucci
University of Utah
72 So. Central Campus Drive Room 3750
Salt Lake City, Utah 84112

Abstract

A paramount goal of scientific data analysis techniques for petascale data is to understand the behavior of complex systems or processed such as scientific simulations based on a discrete sampling sample of such system. In many instances it is possible to observe both input parameters and system outputs, and characterize important properties of the system as a high-dimensional merit function. Applications in which this setting is appropriate include UQ of large numerical simulations, energy landscapes in optimization problems, or the analysis of image data relating to biological or medical parameters. The example in the figure below shows the analysis of the heat release in a combustion simulation represented as function of the ten chemical species in its fuel composition. This can be intuitively understood as a ten-dimensional "landscape," which valleys (regions a,b,c around minima) indicate the condition for extinction while the global maximum represents the main combustion mode where fuel and oxidizer are present in stoichiometric proportions.

Our proposed approach allows analyzing and visualizing such data sets to provide novel quantitative and qualitative insight that complements classical UQ techniques. The method combines topological and geometric techniques to provide interactive visualizations of discretely sampled high-dimensional scalar fields. The method relies on a segmentation of the parameter space using an approximate Morse-Smale complex on the cloud of point samples, where each sample can be the result of running an entire simulation. For each crystal of the Morse-Smale complex, a regression of the system parameters with respect to the output yields a curve in the parameter space. The result is a simplified geometric representation of the Morse-Smale complex in the high dimensional input domain. Finally, the geometric representation is projected in 2D, using dimension reduction, to provide an interactive visualization platform. The geometric properties of the regression curves enable the visualization of additional information about each crystal such as local and global shape, width, length, and sampling densities. The method is validated on several synthetic examples and demonstrated on real datasets including: (a) two use cases, using data sets from the UCI machine learning repository, (b) the analysis of parameters of climate simulations and their relationship to predicted global energy flux, and (c) the concentrations of chemical species in a combustion simulation and their integration with temperature.



Ten-dimensional landscape of heat release as a function of the ten chemical species composing the fuel in a combustion simulation.