Exploration of High-Dimensional Scalar Function
for Nuclear Reactor Safety Analysis and Visualization

Dan Maljovec, Bei Wang, Valerio Pascucci

Scientific Computing and Imaging Institute, University of Utah

{maljovec,beiwang,pascucci}@sci.utah.edu

Peer-Timo Bremer

Lawrence Livermore National Laboratory

bremer5@llnl.gov

Michael Pernice, Diego Mandelli, Robert Nourgaliev

Idaho National Laboratory

{michael.pernice,diego.mandelli,robert.nourgaliev}@inl.gov

ABSTRACT

The next generation of methodologies for nuclear reactor Probabilistic Risk Assessment (PRA) explicitly accounts for the time element in modeling the probabilistic system evolution and uses numerical simulation tools to account for possible dependencies between failure events. The Monte-Carlo (MC) and the Dynamic Event Tree (DET) approaches belong to this new class of dynamic PRA methodologies. A challenge of dynamic PRA algorithms is the large amount of data they produce which may be difficult to visualize and analyze in order to extract useful information. We present a software tool that is designed to address these goals. We model a large-scale nuclear simulation dataset as a high-dimensional scalar function defined over a discrete sample of the domain. First, we provide structural analysis of such a function at multiple scales and provide insight into the relationship between the input parameters and the output. Second, we enable exploratory analysis for users, where we help the users to differentiate features from noise through multi-scale analysis on an interactive platform, based on domain knowledge and data characterization. Our analysis is performed by exploiting the topological and geometric properties of the domain, building statistical models based on its topological segmentations and providing interactive visual interfaces to facilitate such explorations. We provide a user's guide to our software tool by highlighting its analysis and visualization capabilities, along with a use case involving data from a nuclear reactor safety simulation.

Key Words: high-dimensional data analysis, computational topology, nuclear reactor safety analysis, visualization

1 INTRODUCTION

Dynamic Probabilistic Risk Assessment (PRA) methodologies [1] couple numerical simulation tools and time-dependent stochastic models (i.e., probabilistic failure models or parameter uncertainties), to perform system safety analysis. Widely used dynamic PRA methodologies are based on Monte-Carlo [2] or Dynamic Event Tree algorithms [3]. The common underlying idea is to run a large number of simulations (by employing system simulators) where values of system stochastic parameters (e.g., timing of failure of a specific component or an uncertain parameter) are sampled from their own distribution at each run. This type of PRA analysis can be very time-consuming when a large number of stochastic parameters are considered and when large and complex system simulators are used. Moreover, a large volume of data is typically

generated. Such large amounts of information can be difficult to organize for extracting useful information. Furthermore, it is often not sufficient to merely calculate a quantitative value for the risk and its associated uncertainties. The development of risk insights that can increase system safety and improve system performance requires the interpretation of scenario evolutions and the principal characteristics of the events that contribute to the risk.

The need for software tools able to both analyze and visualize large amount of data generated by Dynamic PRA methodologies has been emerging only in recent years. A first step has been shown in [4, 5] using clustering-based algorithms which focus more on the analysis part than the visualization side.

In this paper, we present a software tool that provides scientists and domain experts with an interactive analysis and visualization environment for understanding the structures of high-dimensional nuclear simulation data. Our tool adapts and extends the innovative framework called HDViz first proposed by Gerber et. al. [6] in exploring high-dimensional scalar functions, and applies the underlying techniques to nuclear reactor safety analysis and visualization. Our tool includes a host of various analysis and visualization capabilities. We describe each of these capabilities on a modular basis, by explaining the underpinning theories and presenting usage cases. The software segments the domain of a high-dimension function into regions of uniform gradient flow by decomposing the data based on its approximate Morse-Smale complex. Points belonging to a particular segment have similar geometric and topological properties, and from these we can create compact statistical summaries of each segment. Such summaries are then presented to the user in an intuitive manner that highlights features of the dataset which are otherwise hidden in a global view of the data. In addition, the visual interfaces provided by the system are highly interactive and tightly integrated, providing users with the ability to explore various aspects of the datasets for both analysis and visualization purposes.

2 PRELIMINARIES

We provide relevant background in Morse-Smale complex, its approximation in high dimension and persistence simplification. We provide minimal technical definitions and illustrate the related concepts by examples whenever possible. We hope these intuitive examples are sufficient to convey a basic understanding of these concepts for non-specialists.

Morse-Smale Complex. The body of work presented in this paper relies heavily on the topological structure known as the Morse-Smale complex [7–9]. The Morse-Smale complex is based on Morse theory [10]. Let $\mathbb{M}$ be a smooth manifold embedded in $\mathbb{R}^n$ without boundary and $f : \mathbb{M} \to \mathbb{R}$ be a smooth function with gradient ∇f. A point $x \in \mathbb{M}$ is called *critical* if $\nabla f(x) = 0$, otherwise it is *regular*. At any regular point x the gradient is well-defined and integrating it in both (ascending/descending) directions traces out an integral line, which is a maximal path whose tangent vectors agree with the gradient [7]. Each integral line begins and ends at critical points of f. Therefore, all regular points can trace their ascending integral line to a local maximum. Similarly, tracing the descending integral line of a regular point will associate a point with a local minimum. The *unstable/stable manifolds* (or *ascending/descending manifolds*) of a critical point p are defined as all the points whose integral lines start/end at p. The set of intersections of unstable and stable manifolds creates the *Morse-Smale* complex of f. Each cell (*crystal*) of the Morse-Smale complex is a union of integral lines that all share the same origin and the same destination. In other words, all the points inside a single crystal have uniform gradient flow behavior. These crystals yield a decomposition into monotonic, non-overlapping regions of the domain. Figure 1 illustrates these concepts associated with a height function defined on a 2D domain. For such a 2D function, critical points include local maxima (red), local minima
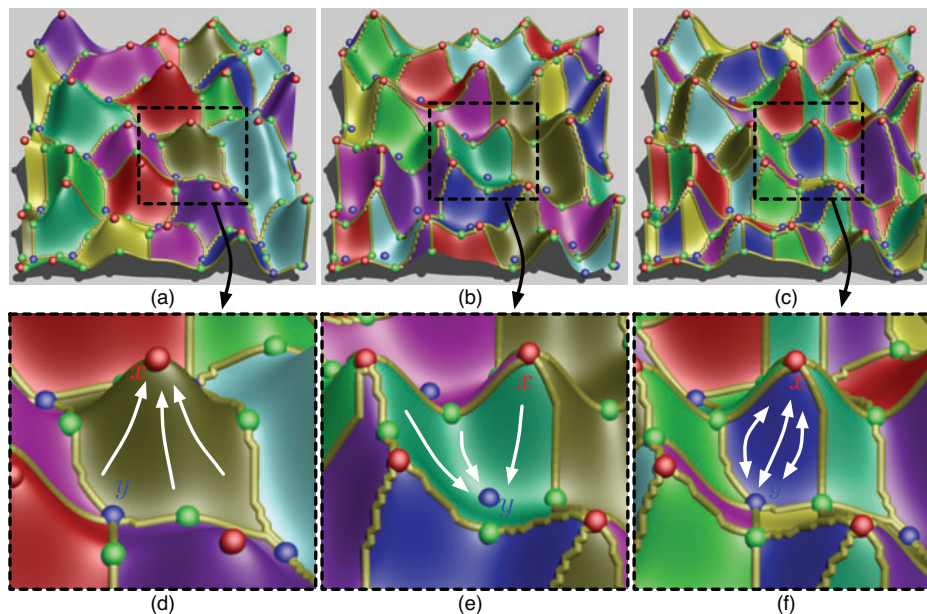
Figure 1. Top: (a) unstable manifolds, (b) stable manifolds and (c) Morse-Smale complex. Bottom: focused views of (d) an unstable manifold surrounding a local maximum x, (e) a stable manifold surrounding a local minimum y and (f) a Morse-Smale crystal associated with the maximum-minimum pair (x, y).

(blue) and saddles (green). (a) illustrates the unstable manifolds with a focused view surrounding a (local) maximum x in (d). (b) illustrates the stable manifolds with a focused view surrounding a (local) minimum y in (e). (c) illustrates the Morse-Smale complex created by the intersection of unstable and stable manifolds, and (f) focuses on a Morse-Smale crystal associated with the maximum-minimum pair (x, y). Each partition is colored by its corresponding gradient behavior, where points share the same color if: (d) their ascending gradient flow (white arrow) end at the same local maximum; (e) their descending gradient flow (white arrow) end at the same local minimum and (f) their ascending and descending gradient flow (white bidirectional arrow) correspond to the same maximum-minimum pair.

Approximating the Morse-Smale Complex in High Dimension. Suppose our input domain is a finite set of points $\mathbb{X}$ in $\mathbb{R}^n$ sampled from some unknown n-manifold. To approximate the Morse-Smale complex of a high-dimensional scalar function f defined on $\mathbb{X}$, our first task is to estimate the gradient at the input points, $\mathbb{X}$. First, we compute a neighborhood graph such as the k-nearest neighbor graph of $\mathbb{X}$ (that is, there is an undirected edge between points $x, y \in \mathbb{X}$ if x is among the k-nearest neighbors of y or y is among the k-nearest neighbors of x). At each point in $\mathbb{X}$, we choose the steepest ascending/descending edge to represent the gradient. With this gradient approximation, we can determine the local extrema by labeling all points with no neighbors of higher values as local maximum and all points with no neighbors of lower values as local minimum. We then label all points in $\mathbb{X}$ according to the local extrema at which its ascending and descending gradients terminate. Subsequently, we collect all vertices with the same pair of labels into crystals and add the extrema to all crystals that share the corresponding label. Therefore the set of sampled points $\mathbb{X}$ are then clustered into crystals $\mathbb{X}_i$ (where $\bigcup \mathbb{X}_i = \mathbb{X}$), with function values observed as $\mathbb{Y}_i = f(\mathbb{X}_i)$. These crystals then serve as an approximation of the Morse-Smale complex, as first introduced in [6]. Some research is underway [11,12] to understand how different neighborhood graphs may impact our approximations, and what sampling conditions we should impose on the data to guarantee the approximation quality. Figure 2(a)-(b) illustrate how our approximation works for a 2D height function where the function values are only given at a finite set of sampled points in the domain. The graph shown is part of a k-nearest neighbor graph where $k = 4$. Point x (y) is a local maximum (local minimum) since it has no neighbors in

the graph with higher (lower) function values. By following the approximated gradients, all the cyan points belong to the same crystal since they share the same pairs of maximum-minimum labels, e.g. (x, y).
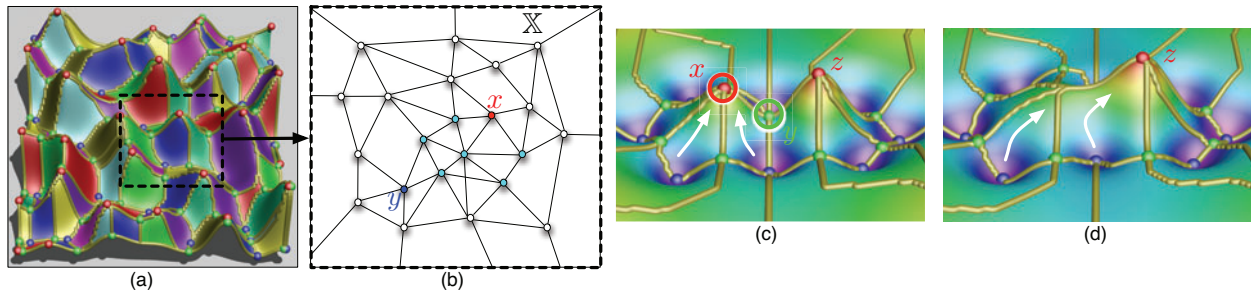

Figure 2. (a): The Morse-Smale complex of a 2D height function f. (b): The approximation of a Morse-Smale crystal of f, where the values of f are only given at a finite set of sampled points in the 2D domain. (c)-(d): A 2D Morse-Smale complex before (c) and after (d) persistence simplification.

Persistence Simplification. In real datasets, there is often noise which may manifest itself as small topological artifacts, either spurious extrema that may not truly exist in the data or small features the user does not deem relevant. To account for this and allow the user to select a *scale* (complexity level) appropriate for the specified dataset, we use the notion of *persistence simplification* whereby less salient features are merged with neighboring, more significant features. In the case of the Morse-Smale complex, we assign a persistence to each critical point in the complex which intuitively describes the scale at which a critical point would disappear through simplification. Here we omit the technical definition of persistence and refer the curious readers to [13, 14].

We illustrate the idea of persistence through the following example, as shown in Figure 2(c)-(d) for a 2D height function. As usual, red, blue and green points denote local maxima, local minima and saddles respectively. In (c), the left peak at the local maximum x is considered less important topologically than its nearby peak at local maximum z, since x is lower. Therefore, at a certain scale, we would like to represent this feature as a single peak instead of two separate peaks, as shown in (d), by redirecting ascending gradient flow (white arrow) that originally terminates at x to terminate at z. In this way, we simplify the function by removing (canceling) the local maximum x with its nearby saddle y. Such a simplification procedure is referred to as the persistence simplification.

A Complete Illustrative Example. Finally for a complete intuitive understanding of the above concepts, we give the following example of another 2D height function, as shown in Figure 3. The top of the figure illustrates its corresponding unstable manifolds, stable manifolds and Morse-Smale complex, while the middle of the figure showcases the progressive simplification of its critical points: at the finest level there are four maxima in the data, but as we increase the scale, neighboring topological features are merged where the ascending gradient flow is directed to its more salient neighboring maxima. Circled points correspond to the critical points that are removed at each simplification step. We thus build a filtration of segmentations where Morse-Smale crystals are merged based on the persistence value of their associated extrema.

3   ANALYSIS AND VISUALIZATION MODULES

In this section, we describe each analysis and visualization module within our integrated system that are either part of the original capabilities provided by HDViz [6] or part of our extension. We demonstrate our infrastructure with a 6D example dataset from nuclear plant safety analysis. The data is extracted from
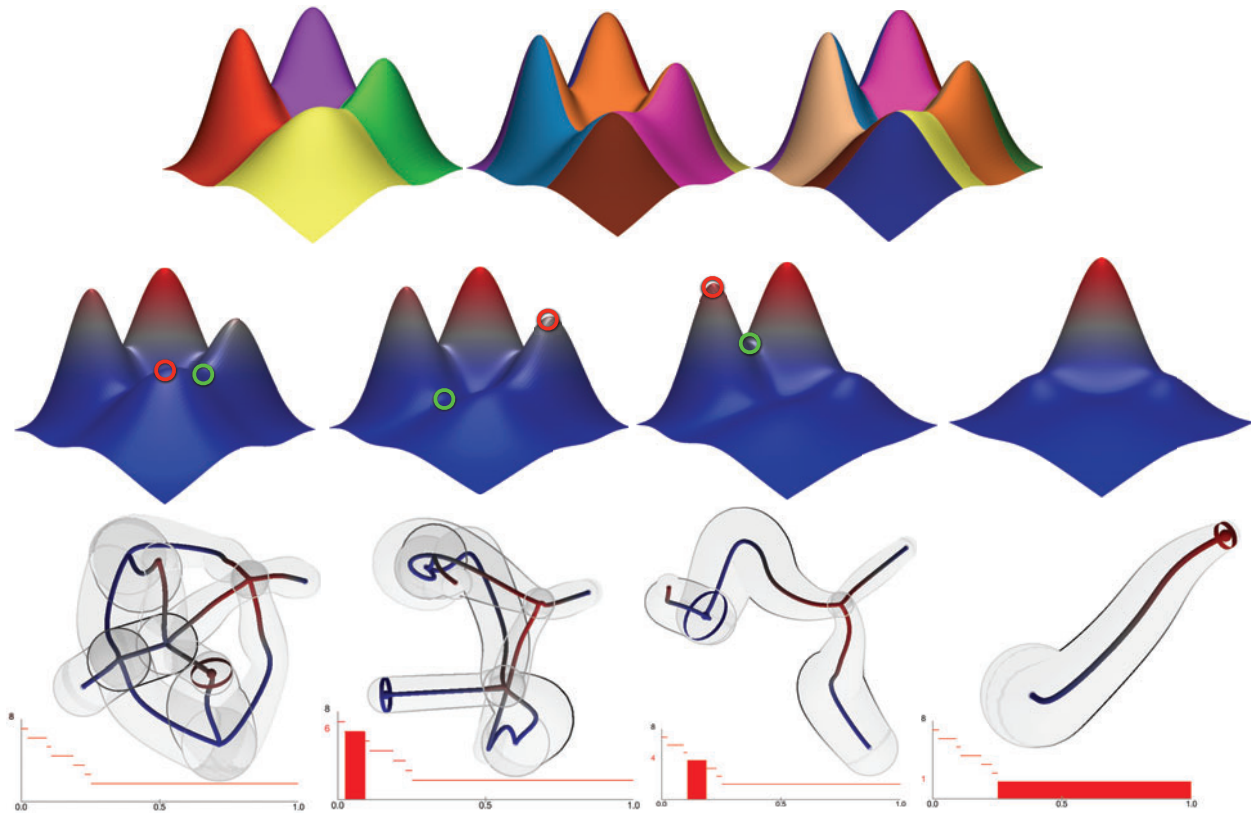
Figure 3. An illustrative example based on a 2D height function. Top: unstable manifolds, stable manifolds and Morse-Smale complex. Middle: progressive simplification of its critical points (maxima are red, saddles are green). Bottom: topological summaries shown at various scales (see Section 3).

a VR_2^+ nuclear reactor simulator and represents an ensemble of 10000 simulation trials where a SCRAM is simulated due to a failure in the system. A SCRAM event is when the control rods of the reactor are inserted into the core in order to prevent overheating of the reactor core. The output variable is the peak coolant temperature (PCT*), measured in Kelvin. The domain scientists are interested in what combination of conditions (in the form of input parameters) can cause potential reactor failure (i.e. nuclear meltdown witnessed by PCT exceeding a threshold value). The input space is defined by six parameters:

- **PumpTripPre** - the minimum pressure (MPa) in the heat exchange pump causing the SCRAM to trip
- **PumpStopTime** - the relaxation time (sec) of pump's phase-out
- **PumpPow** - end power of the pump
- **SCRAMtemp** - the maximum temperature in the system causing the SCRAM to trip
- **CRinject** - the control rod position at the end of SCRAM
- **CRtime** - the relaxation time (sec) of the control rod system.

3.1 Topological Summary

The visual interface designed for topological summary is inherited and extended from the capabilities provided by HDViz [6] with additional capacity for user interactivity. Considered as the main visual display of our software, this interface summarizes each Morse-Smale crystal into a 1D curve in high-dimensional

---

*Note that PCT used here does not stand for peak clad temperature.

space which is then projected onto a viewable 3D space. The interface encodes three steps (detailed in [6], with a high level description below), to arrive at a 3D representation for analysis and visualization of the d-dimensional scalar function f, defined on a set of sampled points $\mathbb{X}$. (1) *Morse-Smale approximation*: We approximate the Morse-Smale crystals, $\mathbb{X}_i$ and $\mathbb{Y}_i = f(\mathbb{X}_i)$, in high dimension using an approximate k-nearest neighbor graph [15]. (2) *Geometric summaries:* For each crystal of the Morse-Smale complex, since each point has similar approximated monotonic gradient behavior, a geometric summary is constructed by an inverse regression, yielding a 1D curve r_i in the d-dimensional domain of f. Intuitively, $r_i(y)$ (where $y \in \mathbb{Y}_i$) yields a representation of the crystal as the average of the function values of level sets within the crystal, see [6] for its detailed derivation. (3) *Dimension reduction:* The set of regression curves can be represented by a graph embedded in $\mathbb{R}^d$ with each edge corresponding to a curve and vertices corresponding to extremal points. For visualization, we embed this graph into 2D preserving the spatial relation among extrema and the geometry of the crystals that connect them. First, vertices are embedded using PCA or ISOMAP [16]. Second, edges are embedded individually through their first two principle components. And third, the resulting 2D curves are attached to the projected vertices through affine transformations. The third dimension is reserved for the output parameter.
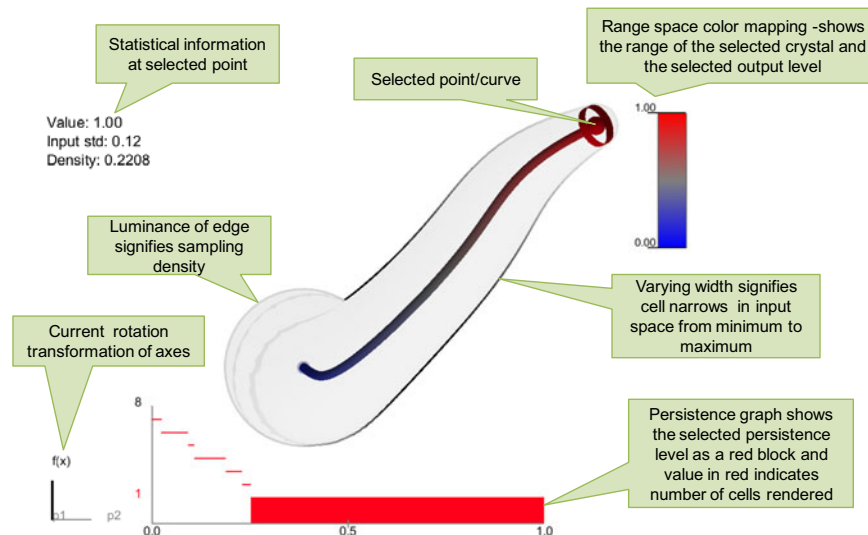

Figure 4. The topological summary visual interface of a simple 2D function.

Visual Components. The visual components of the above framework are shown in Figure 4 for a simple 2D function with one crystal (example revisited from Figure 3). Each curve is encased in a transparent tube where the width of the tube represents the "spread" of the data at a particular scale, and the luminance of the tube encodes the density of data points within each crystal. Users are given the flexibility to view the topological summary of a high-dimensional function by switching between PCA and ISOMAP projections, and using affine transformations to manipulate the projection directly on the screen. In order to preserve the "width" of a crystal at a given scale, we compute the standard deviation with respect to each input parameter and also a single *average* standard deviation across all input parameters. The latter is direction-independent and can therefore be used as a generalized width of the data at a particular output level. The radius of the outer transparent tubes are defined by this direction-independent standard deviation. The last visual cue is the darkness of the edge of the transparent tube. Where the sampling density is high, the outline of the tube is drawn black and as the sampling density decreases, the luminance of this edge increases. To enable multi-scale analysis, we use a modified version of the persistence diagram [13], referred to as the *persistence graph*. This is shown as a visual component at the bottom of Figure 4. It shows the number of Morse-Smale crystals (y-axis) as a function of scale (i.e. x-axis, persistence threshold normalized by the range of the

dataset). A selected scale is drawn with a red box and a corresponding number of Morse-Smale crystals is displayed along the y-axis in red. Stable features are considered as those that exist over a large range of scales (i.e. a sequence of persistence simplification with increasing scales), which correspond to long horizontal lines in the persistence graph.

We demonstrate how the topological summary visual interface allows the users to explore the data at multiple scales. We revisit our simple dataset under several levels of persistence simplification in Figure 3 bottom. Here the leftmost image shows the full resolution of data with four local maxima, while each subsequent image reduces the number of maxima by one until we are left with a single crystal describing the gradient flow from the global minimum to the global maximum. The numbers in red, shown in the persistence graphs, indicate the total number of crystals displayed, from left to right, as 8, 6, 4 and 1. Instead of giving the users a representation of the data at a fixed scale, we provide an interactive platform to help them differentiate features from noise through multi-scale analysis and to choose the appropriate scale based on domain knowledge and data characterization.

A Simple Example of Geometric Summary. Figure 5 shows several examples of simple 2D surfaces represented using the techniques described in this section. Each function has one local maximum and a single local minimum surrounded by a flat area. Using an evenly sampled grid, the density of points available at each level is mapped to the color on the edge of the transparent tubes. The widths of the tubes vary with respect to the spread of the data at each level set. Note how the sharp spike in (a) is represented in the topological display. The steep peak covers a small area and the small width of the transparent tube demonstrates this in the area near the maximum. (b) has a wider area surrounding the peak and while its representing curve remains a single straight line, the tube surrounding it changes behavior to account for the width of each level set. The plateau function in (c) even maps the discontinuity in the curve by making a sudden jump from a low value to a high value. Note how the edge of the tube has high luminance in the middle section denoting a lack of data used to compute this section, whereas the most densely sampled region is at the maximum value which has a black outline.
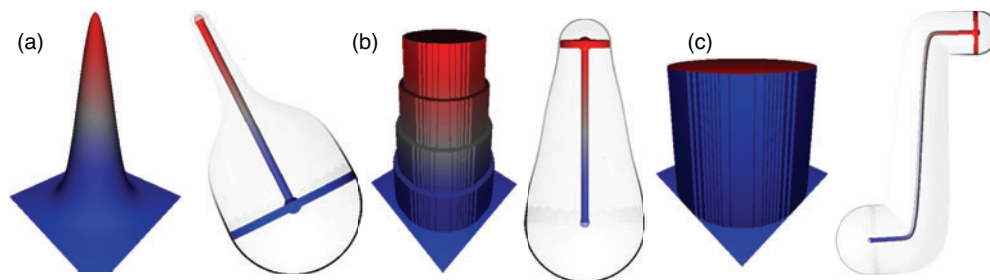

Figure 5. Three simple surfaces represented using geometric summary tubes.

6D Demo Dataset. As a final example, we illustrate our interface with our 6D demo example in Figure 6 under multiple scales. This dataset contains 753 individual crystals at the finest level, though due to low persistence and visual clutter caused by these crystals, we restrict our analysis to only a handful of the most salient crystals. During the persistence simplification steps, the topological summaries consist of 6, 3 and 1 Morse-Smale crystal(s), respectively, based on their long horizontal lines in the persistence graph (that likely correspond to stable features). At each of the shown persistence levels, the dataset is characterized by a single global minimum with high sampling density. The widths of all the crystals at the minimum is expansive compared to the widths at the maxima. We could infer from this analysis result that most of the data points (simulations) represent lower PCTs and the conditions to reach these lower PCTs varies widely in the domain space. On the other hand, the maximum PCTs are reached at more specific input ranges which is made clear through the narrow widths of the tubes (of one standard deviation) surrounding the red maxima.

For these chosen scales, further explorations could be done to understand the different combinations and correlations among input parameters that lead to 6, 3 and 1 local maxima of PCT, respectively.
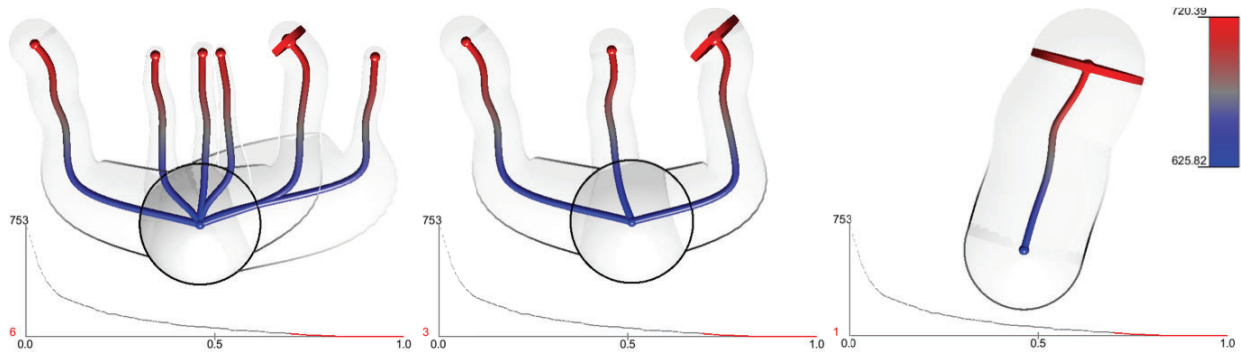

Figure 6. 6D demo dataset: Topological summaries shown at various scales.

3.2 Statistical Summary

The visual interface shown in Figure 7 demonstrates statistical and geometric (i.e. gradient) information associated with the selected point in the topological summary interface. Each input dimension, or coordinate, is viewed as an inverse function of the output parameter. In the left column of the statistical summary window, each horizontal axis describes the range of values of each input parameter and the coordinate mean and coordinate standard deviation associated with the selected point. The right column encodes the gradient information, that is, the change in the output with respect to the change in each input parameter. In Figure 7, we can see that three parameters have quite large standard deviations, PumpStopTime, SCRAMtemp, and CRtime, whereas the other parameters vary much less. This indicates a certain characteristic associated with the chosen crystal with respect to these three parameters.
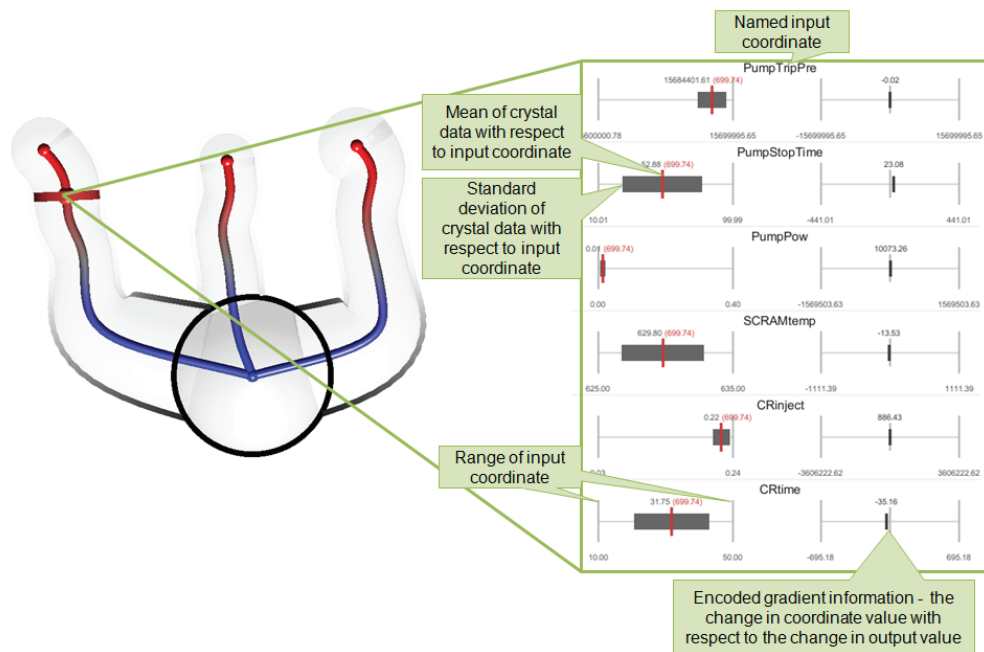

Figure 7. 6D demo dataset: A snapshot of the statistical summary with highlighted visual components.

3.3 Inverse Coordinate Plots

In the inverse coordinate plots, each input parameter is considered as a 1D function of the output variable. This visual interface is shown in Figure 8 for our 6D example. On the left, at the persistence level with three crystals, the interface displays the inverse coordinate plot for data points associated with the selected Morse-Smale crystal in the topological summary interface. On the right, at the persistence level with six crystals, the interface shows combined inverse coordinate plots associated with all the crystals that share the same local minimum. For the selected crystal(s), the regression curve is drawn in the inverse coordinate plots with a grey tube representing the parameter-specific standard deviation.

From the right image in Figure 8, we see the top set of axes where each of the six regression curves is readily distinguishable with respect to different ranges of values for PumpTripPre, whereas in the lower plots they vary little toward the left and only slightly at the right. On the other hand, parameters such as PumpPow and CRinject result in high temperatures only at specific levels. This conclusion is based upon the tight configuration of points resulting in high PCT and it is also supported by the consistent locations of the mean values across all crystals and the low standard deviations of these parameters.
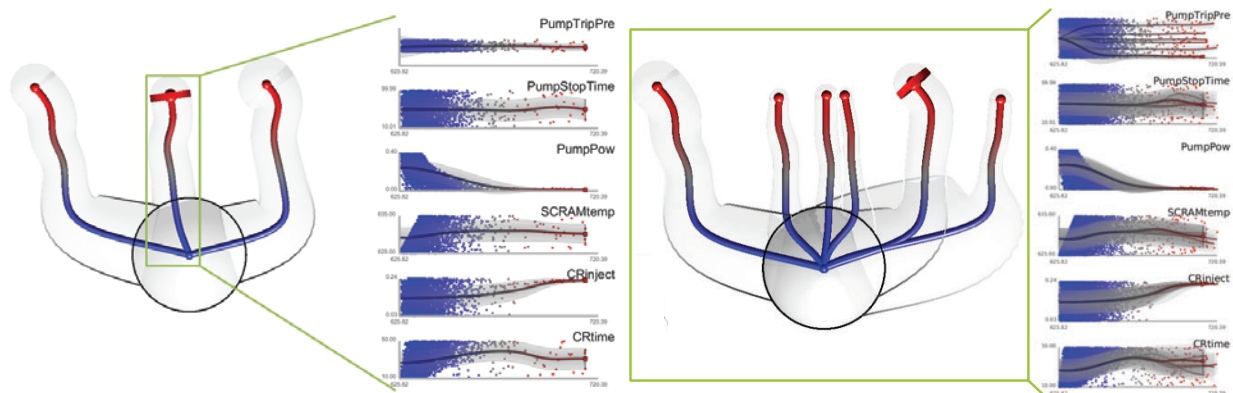

Figure 8. 6D demo dataset: Inverse coordinate plots shown for the highlighted crystal for the three crystal (left) and six crystal (right) case.

3.4 Interactive Projection, Parallel Coordinate Plots and Pairwise Scatter Plots

As the number of dimensions increases for a high-dimensional function defined on the sampled point set, selecting the interesting projection dimensions while interacting with the high-dimensional function could become counter-intuitive. We design an interactive visualization interface that provides simple and fully explanatory pictures that give comprehensive insights into the global structure of the high-dimensional function. Based on hypervolume visualization techniques developed in [17], our basic idea is a generalization of direct parallel projection methods. We first create an independent viewing system that scales with the number of dimensions where the user is allowed to manipulate how each axis is projected. We then apply these manipulations to project the geometric summaries from the functional space to the screen space [17]. Users are provided with a wheel of labeled axes which they can manipulate by stretching, contracting, and rotating. The geometry of the inverse regression curves therefore is meaningfully preserved. With a single manipulated projection displayed at a time, the interactive nature of such a tool provides the user with intuition that is otherwise lost in a PCA or ISOMAP projection. An example for the 6D demo dataset is shown in Figure 9. Currently, this system only supports a view of the geometric summaries, but a possible

extension is to create a full hypervolume visualization of the raw high-dimensional data points as proposed in [17].
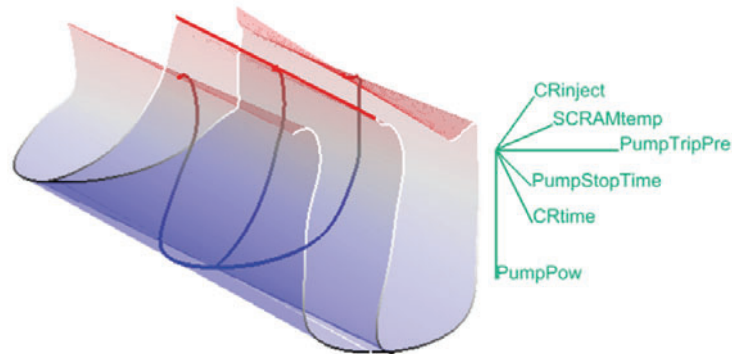

Figure 9. 6D demo dataset: Interactive Projection Visual Interface for the three crystal case. Left: projection of the high-dimensional inverse regression curves and their associated standard deviation tubes, based on the mapping shown in the right. Right: each input dimension is mapped to a single green line segment, which the user can manipulate by stretching and rotation, to emphasize or diminish the effect of a particular dimension on the projection.

Furthermore, we also use parallel coordinate [18] plots (Figure 10 left) where each input parameter is represented as a vertical axis, and a single line connecting parallel axes represents the adjacent dimensions of a given data point. In addition, we present a matrix of scatter plots (Figure 10 right) where each pair of input parameters are plotted as the x and y axes with the output variable mapped by color. We can see in both plots, the selected crystal of our 6D demo dataset has the defining characteristic of only extracting a portion of the domain values for the PumpTripPre parameter. Such features in the visualization can help us infer correlations among input parameters, and their influence on the output parameter.
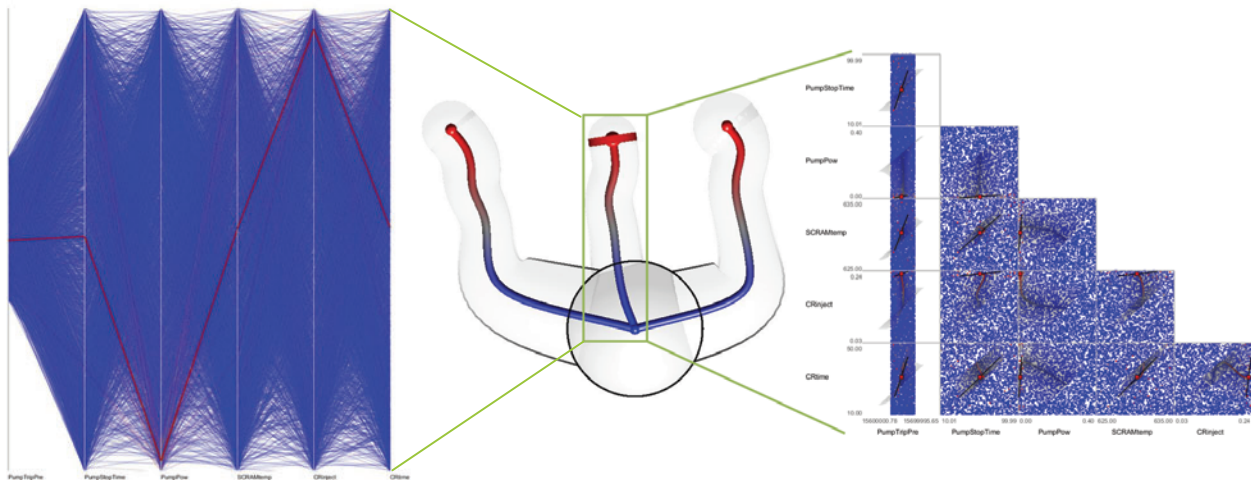

Figure 10. 6D demo dataset, with parallel coordinate plots (left) and pairwise scatter plots (right) for a selected crystal.

4   CONCLUSIONS

In this paper, we have presented a software tool suitable to analyze and visualize datasets generated by safety analysis codes. In particular, our software could be interfaced with Dynamic PRA algorithms. In such a configuration, a large number of runs of the safety analysis code are performed, where during each run, initial conditions and timing/sequencing of events are changed according to their statistical distribution. For large complex systems, this type of code generates a large amount of data, modeled as high-dimensional scalar functions. Our software is designed to assist the user in the process of extracting useful information from such functions.

This extraction is performed by exploring the correlations between uncertain input parameters and simulation outcome. These correlations are modeled as a Morse-Smale complex and then summarized in various visual display modules, where the reconstructed simulation outcome is directly linked to the input parameters. From an uncertainty quantification point of view, such an analysis tool allows the user to identify and consequently rank variables based on their correlations with simulation outcome (e.g., maximum fuel temperature). Such software becomes even more effective for the analysis of complex systems such as nuclear power plants where the number of variables is very large.

We have described each analysis and visualization module in the system, by explaining the theories and techniques based on a specific application of the software in a 6D demo example involving nuclear safety simulation. Our example is based on code performed on 10000 simulations of a simplified pressurized water reactor (PWR) system for a SCRAM scenario. Our software allows us to identify and visualize the correlations among six parameters and the maximum coolant temperature. We first perform topological analysis to identify the underlying topological structure of the dataset. Statistical information is then summarized and linked to the topological structures extracted from the data, allowing the users to identify correlations between timing/sequencing of events and simulation outcome. We have obtained some intuitive understanding of the structures of such a high-dimensional function, although a more comprehensive interpretation of our results depends on extensive testing and explorations from our collaborators. We expect to work closely with the domain scientists and obtain feedback from the end users regarding (a) the interpretation of the testing datasets using our software, (b) the limitations and potential improvements of current analysis techniques, and (c) the potential improvements of the visualization interfaces in terms of usability and interactivity.

REFERENCES

[1] J. Devooght, "Dynamic reliability," *Advances in Nuclear Science and Technology*, vol. 25, pp. 215–278, 1997.

[2] E. Z. M. Marseguerra, J. Devooght, and P. Labeau, "A concept paper on dynamic reliability via monte carlo simulation," *Mathematics and Computers in Simulation*, vol. 47, pp. 371–382, 1998.

[3] E. M. G. Cojazzi, J. M. Izquierdo and M. S. Perea, "The reliability and safety assessment of protection systems by the use of dynamic event trees: the dylam-treta package," *Proceedings XVIII Annual Meeting Spanish Nuclear Society*, 1992.

[4] D. Mandelli, A. Yilmaz, and T. Aldemir, "Data processing methodologies applied to dynamic PRA: an overview," in *Proceeding of Probabilistic Safety Analysis*, 2011.

[5] D. Mandelli, A. Yilmaz, and T. Aldemir, "Scenario analysis and pra: Overview and lessons learned," in *Proceedings of European Safety and Reliability Conference*, 2011.

[6] S. Gerber, P.-T. Bremer, V. Pascucci, and R. Whitaker, "Visual exploration of high dimensional scalar functions," *IEEE Transactions on Visualization and Computer Graphics*, vol. 16, pp. 1271–1280, 2010.

[7] H. Edelsbrunner, J. Harer, and A. J. Zomorodian, "Hierarchical Morse-Smale complexes for piecewise linear 2-manifolds," *Discrete and Computational Geometry*, vol. 30, pp. 87–107, 2003.

[8] H. Edelsbrunner, J. Harer, V. Natarajan, and V. Pascucci, "Morse-Smale complexes for piecewise linear 3-manifolds," *Proceedings 19th ACM Symposium on Computational Geometry*, pp. 361–370, 2003.

[9] A. Gyulassy, V. Natarajan, V. Pascucci, and B. Hamann, "Efficient computation of Morse-Smale complexes for three-dimensional scalar functions.," *IEEE Transactions on Visualization and Computer Graphics*, vol. 13, pp. 1440–1447, 2007.

[10] J. Milnor, *Morse Theory*. New Jersey, NY, USA: Princeton University Press, 1963.

[11] C. D. Correa and P. Lindstrom, "Towards robust topology of sparsely sampled data," *IEEE Transactions on Visualization and Computer Graphics*, vol. 17, pp. 1852–1861, 2011.

[12] D. Maljovec, A. Saha, P. Lindstrom, P.-T. Bremer, B. Wang, C. Correa, and V. Pascucci, "A comparative study of morse complex approximation using different neighborhood graphs." Workshop on Topological Methods in Data Analysis and Visualization (accepted), 2013.

[13] H. Edelsbrunner, D. Letscher, and A. J. Zomorodian, "Topological persistence and simplification," *Discrete and Computational Geometry*, vol. 28, pp. 511–533, 2002.

[14] H. Edelsbrunner and J. Harer, "Persistent homology - a survey," *Contemporary Mathematics*, vol. 453, pp. 257–282, 2008.

[15] S. Arya, D. M. Mount, N. S. Netanyahu, R. Silverman, and A. Y. Wu, "An optimal algorithm for approximate nearest neighbor searching fixed dimensions," *Journal of the ACM*, vol. 45, pp. 891–923, 1998.

[16] J. B. Tenenbaum, V. D. Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science*, vol. 290, pp. 2319–2323, 2000.

[17] C. L. Bajaj, V. Pascucci, G. Rabbiolo, and D. Schikorc, "Hypervolume visualization: a challenge in simplicity," *Proceedings IEEE Symposium on Volume Sisualization*, pp. 95–102, 1998.

[18] A. Inselberg, *Parallel Coordinates: Visual Multidimensional Geometry and its Applications.* Springer, 2009.