FINAL REPORT

DE-SC0011615, University of Utah

Project Title: Advanced Dynamically Adaptive Algorithms for Stochastic Simulations on Extreme Scales

PI: Dongbin Xiu

1 Project Background

The focus of the project is the development of mathematical methods and high-performance computational tools for stochastic simulations, with a particular emphasis on computations on extreme scales. The core of the project revolves around the design of highly efficient and scalable numerical algorithms that can adaptively and accurately, in high dimensional spaces, resolve stochastic problems with limited smoothness, even containing discontinuities.

During the course of the project, we have made tremendous progresses. The key results include

- Efficient high order edge detection method that can accurately detect discontinuities in high dimensional random space. More importantly, the adaptive version of this method scales linearly with dimensionality. This unique feature allows the detection of discontinuities in random space for complex systems.

- Least orthogonal interpolation method that allows stochastic collocation interpolation on arbitrary sample set. This is a major breakthrough in term of approximation theory, as rigorous framework for polynomial interpolation on arbitrary grids did not exist before. Our method also has an easy implementation procedure using numerical linear algebra. This method allows one to conduct accurate stochastic collocation approximation using any number of samples.

- Minimal element generalized polynomial chaos (mE-gPC) method. This method combines the aforementioned two key results. The mE-gPC allows one to adaptively determine if the stochastic simulation contains discontinuities. If it does, then the method automatically decompose the random space into smooth subdomains that are determined by the problem itself. Consequently, the number of smooth subdomains is minimal, and significantly less than any other existing domain decomposition methods. The least orthogonal interpolation is then used in each subdomain to construct accuracy stochastic models. The mE-gPC method is therefore the optimal method — one can not construct anything better.

2 Key Result: Edge Detection

This work was largely carried out by the PI, Dr. John Jakeman, who served as a post-doc during the course the project, and Dr. Richard Archibald, the PI of the counterpart project at Oak Ridge National Laboratory. The collaboration has led to the development of an efficient algorithm for detection of discontinuities in high dimensional random space.

The method is based on polynomial annihilation discontinuity detection method ([2]), which seeks to detect and approximate jumps, denoted as $[f](x)$, in a given function $f(x)$. The basic idea is: for any point x, we surround it with a local stencil S_x, consisting of $m+1$ local points, and then construct an approximation of $[f](x)$ via

$$L_m f(x) = \frac{1}{q_m(x)} \sum_{x_j \in S_x} c_j(x) f(x_j), \tag{1}$$

where the coefficients $c_j(x)$ are chosen to annihilate polynomials of degree up to $m-1$ and are determined by solving the linear system

$$\sum_{x_j \in S_x} c_j(x) p_i(x_j) = p_i^{(m)}(x), \qquad \forall\, i = 0, \cdots, m. \tag{2}$$

Here $\{p_i\}_{i=0}^m$ is a set of basis of polynomials of degree up to m, $p_i^{(m)}(x)$ denotes the m^{th} derivative of $p_i(x)$, and $q_m(x)$ is the normalization constant. As a high-order approximation to the jump $[f](x)$, $L_m f(x)$ converges to zero rapidly away from the jump. And this allows us to successfully detect any jump discontinuities, if they exist in $f(x)$.

The polynomial annihilation method was originally applied to low-dimensional problems, mostly as an edge detection method for image analysis in one or two dimensions. An initial attempt was made in [1] to extend the method to high dimensions. However, the procedure in [1] relies on examinations of the function values on (local) tensor grids and thus restricts the dimensionality that can be handled.

During the course of the project, we developed a much improved algorithm for high dimensional problems. The new algorithm targets the high-dimensional stochastic simulations typically encountered in uncertainty quantification. In particular, it is closely related to one of the most popular numerical implementations of polynomial chaos methods—sparse grid stochastic collocation [11]. The new algorithm uses an adaptive sparse grid approximation, where a local adaptive approach is used to determine and resolve the location of any discontinuities and then a dimensional adaptive approach is used to refine only in the dimensions that the discontinuities reside. By doing so, the new method invests the majority of function evaluations only in regions surrounding the discontinuities and neglects the grid points in the smooth regions and irrelevant dimensions.

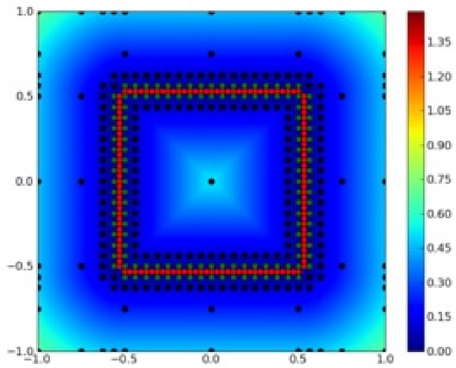
Illustrations of the method can be found in Fig. 1, where the results of a set of two-dimensional problems are shown. It is clear that the algorithm automatically detects the locations and structures of the discontinuities by using the points of sparse grids only close to the discontinuities.

A much more important issue to understand is how the algorithm performs in high dimensions. And this can be seen from the results of detecting a hyper spherical discontinuity in various dimensions. With a fixed level of resolution control, a set of tests were conducted for dimensions up to $d = 100$. The total number of points required to resolve the discontinuity is tabulated in Table 1. We observe that the growth of the number of points (N) is linear with respect to the dimension d, as shown even more clearly in Fig. 2.

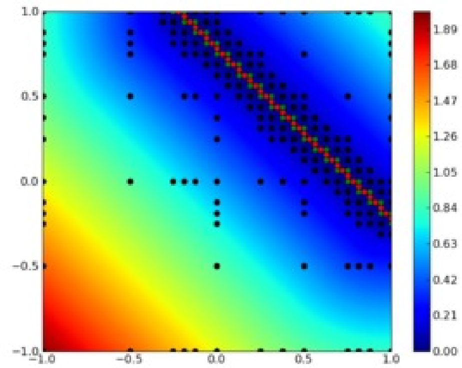
The significance of the linear growth of the number of grid points can not be overlooked, for it allows us to apply the method to practical problems, where the number of uncertain inputs is typically high and the simulation is expensive. In these cases, it is critical to keep the simulation effort as small as possible in high dimensions. And "linear growth" of simulation effort represents almost the best scenario. To this end, the current algorithm can be considered "optimal".

To summarize, the features of the algorithm include:

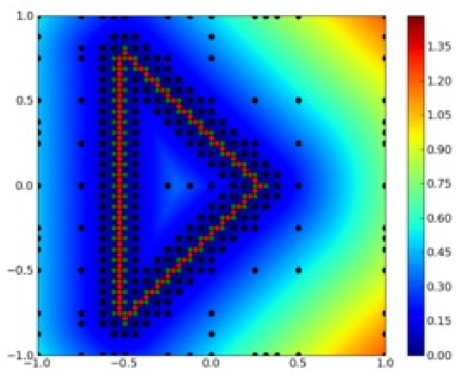
- It works in arbitrary number of dimensions and does not require any assumptions on the properties of the discontinuities.
- It is a high order method, in the sense that it can resolve the locations and structures of the discontinuities with high resolution.



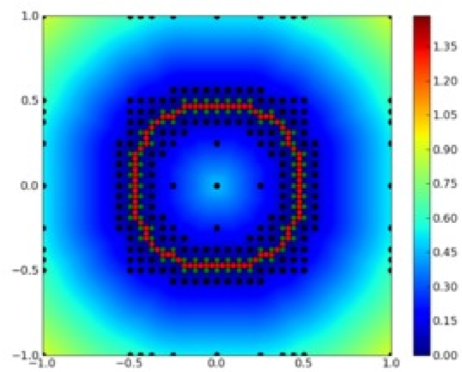
(a) Squared-shape discontinuity.



(b) linear-cut discontinuity.



(c) Triangle discontinuity.



(d) Circle discontinuity.

Figure 1: Illustrations of adaptive sparse grid discontinuity detection in two-dimensions with resolution level 2^{-4}. Black points represent points in the adaptive sparse grid, among which red points define the structure of the discontinuities.

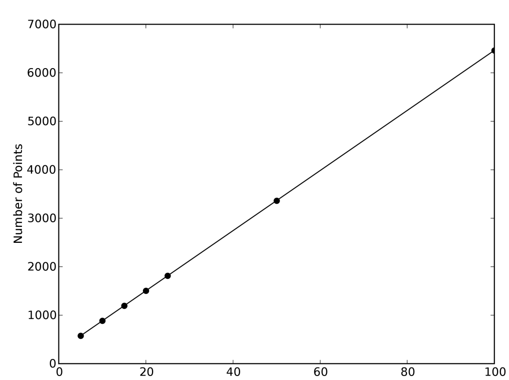


Figure 2: Graphical illustration of the growth of the number of sparse grids N with respect to the dimensionality d, based on the results in Table 1.

d	N	Resolution level
5	573	0.03206
10	883	0.03160
15	1193	0.03156
20	1503	0.03075
25	1813	0.03055
50	3363	0.03137
100	6463	0.03128

Table 1: The total number of sparse grids N required to resolve a spherical discontinuity in various dimensions d, with resolution level fixed around ~ 0.03.

- It is non-intrusive and would not require rewriting of existing simulation code.
- It is adaptive, not only in term of resolving the locations of discontinuities but also in term of guiding itself through the most relevant random dimension—dimensional adaptivity.
- It is highly efficient and in fact "optimal", in the sense that the **simulation effort of the algorithm grows linearly with the dimensionality of the random space**. This implies that the algorithm is free of the well known curse-of-dimensionality.

The features of the algorithm, especially the non-intrusiveness and dimensional adaptivity, make the method highly useful for practical problems. The fact that its simulation efforts depends only linearly on the dimensionality of the stochastic space makes the method extremely useful for large-scale simulations.

The algorithm is presented in a paper, which is published earlier this year [6].

- J. Jakeman, R. Archibald and D. Xiu, Characterization of Discontinuities in High-dimensional Stochastic Problmes on Adaptive Sparse Grids, Journal of Computational Physics, Vol. 230, 3977-3997, 2011.

3 Key Result: Least orthogonal interpolation

While our edge detection method can accurately and efficiently detect the discontinuities, accurate stochastic prediction still remains challenging. The difficulty arises from the fact that the discontinuities usually present themselves in the random space in a unpredictable manner. That is, their locations, geometrical structures and properties are determined by the underlying physical problems. Hence the smooth subdomains defined by the discontinuities possess rather complex and irregular structure. In the context of stochastic collocation methods, this implies that, no matter how structured the underlying collocation points are, their structure in the subdomains will be destroyed by the presence of the irregularly shaped discontinuities, which now serve as the boundaries of the subdomains. This can be clearly seen from Figure 1. Even though adaptive sparse grids (the black dots) are structured and identify the discontinuities. The collocation points in each subdomain become unstructured because of the irregular shape of the discontinuities. Also the number of collocation points in each subdomain becomes rather arbitrary. The straightforward use of any standard polynomial chaos approximation techniques is not possible. We therefore face the following problem:

> Given function values at the set of nodal points in a multi-dimensional space, where the number and locations of the points are arbitrary, how to construct an accurate polynomial approximation to the underlying function.

More specifically, the problem can be posed as following. Let $x_1, \ldots, x_N$, be a set of distinct point in $\mathbb{R}^d$, $d > 1$, and let $f_j = f(x_j)$, $j = 1, \ldots, N$, be given function values, then find a polynomial $p(x)$ such that $p(x_j) = f_j$, $j = 1, \ldots, N$. We remark that this now becomes a fundamental mathematical problem of multivariate interpolation on arbitrary grids. And to this day, there exists no sound practical solution. (When $d = 1$, the problem is the well understood polynomial interpolation on a real line.)

During the course of the project, the PI worked closely with Dr. Narayan, who worked as a post-doc at Purdue University, and made a fundamental breakthrough and developed a methodology of "least orthogonal interpolation". The work is based on an earlier work of "least interpolation", developed by de Boor Ron [3, 4] in 1990's. The work of de Boor and Ron uses monomials as basis function. It remains largely theoretical and is numerically unstable. Our present work of least orthogonal interpolation (LOI) is a much broader framework. It uses the classical orthogonal polynomials as basis functions and much more stable numerically. It also incorporates the work of de Boor and Ron as a special case.

Assume $f \in L^2_\omega$ is mean square integrable with respect to a probability measure ω, then its best L^2_ω approximation exists — its orthogonal projection, $f = P_\infty f$, where

$$P_n f = \sum_{|\mathbf{i}|=0}^{n} \langle f, \Phi_{\mathbf{i}} \rangle_\omega \Phi_{\mathbf{i}}, \qquad n \geq 0.$$

Here $\langle \cdot, \cdot \rangle_\omega$ denotes the inner product with respect to the probability measure ω, and $\mathbf{i} = (i_1, \ldots, i_d)$ is multi-index with $|\mathbf{i}| = i_1 + \cdots + i_d$. With this standard definition, we define $f_{\downarrow,\omega}$, called $f - least$, as the first non-zero order term in the series, i.e.,

$$f_{\downarrow,\omega} = P_m f, \qquad m = \min\{n : P_n f \neq 0\}.$$

For each collocation point (defined at arbitrary locations) x_j, we define

$$h_j(x) = \sum_{|\mathbf{i}|=0}^{\infty} \Phi_{\mathbf{i}}(x_j) \Phi_{\mathbf{i}}(x), \qquad j = 1, \ldots, N,$$

and subsequently define

$$H_{\downarrow,\omega} = \{h_{\downarrow,\omega} : h \in \text{span}(h_1, \ldots, h_N)\}.$$

Then, we can prove

> Theorem: The space $H_{\downarrow,\omega}$ in minimally total for interpolation on the nodal set $x_1, \ldots, x_N$.

This implies that one can interpolate any function values on the nodal set and all nodal values are used in the interpolation. More importantly, the following result holds.

> Corollary: There exists a set of orthogonal basis for $H_{\downarrow,\omega}$.

This implies that one can construct efficient interpolation using the orthogonal basis. In fact, we have devoted significant effort and developed a straightforward algorithm to construct the basis function for the space. The algorithm requires only elementary row operations on the interpolation matrix. Another fundamental result is stated in the following theorem.

Theorem: If the probability measure ω is i.i.d. standard Gaussian measure with zero mean and unit variance, the space $H_{\downarrow,\omega}$ constructed using Hermite orthogonal polynomials coincides with the least interpolation space by de Boor and Ron [3, 4].

This implies that the work of de Boor and Ron is a special case of the current least orthogonal interpolation, using Hermite polynomials.

Several numerical tests were conducted to examine the performance of the least orthogonal interpolation. Here we illustrate its properties using a simple and yet challenging example. Consider a targe unknown function in $d = 2$, $f(x) = \cos(\pi x_1)\cos(\pi y)$. We choose 20 equidistance points on a straight line skewed 30° clockwise from the $y-$axis. If one adopts the traditional interpolation method, this problem becomes unsolvable, because the interpolation matrix (aka, the Vandermonde matrix) is singular. However, this poses no problem for the current least orthogonal interpolation, because it is guaranteed to work on arbitrary grids. The results are illustrated in Fig. 3, where the solution using Hermite basis is shown on the top left and Legendre basis on the top right. The exact target function is shown at the bottom. We remark that even though both interpolations show noticeable difference from the target function, one should consider both interpolation as "correct". This is because both interpolation faithfully interpolate the function values at the 20 collocation nodes. Since the nodes lie on the straight line, no information of the target function is available away from the line. Hence any numerical methods can produce anything away from the line. This example illustrates an important feature of the least orthogonal interpolation: it can produce reliable interpolation even in the singular (by the traditional method) cases.

In Figure 4, we examine the interpolation accuracy in the whole random space and compare it against a more traditional method, the cubic least-square. The target function is a Gaussian function and we perform tests in various dimensions for up to $d = 15$. The results demonstrate the superior accuracy obtained by the least orthogonal interpolation.

To summarize, the features of the least orthogonal interpolation include:

- It is a fundamental work in approximation theory, particularly in multi-dimensional polynomial interpolation, and goes well beyond the field of stochastic collocation. The only available method is by de Boor and Ron [3, 4]. And the current work is a broader framework that incorporates it as a special case.

- It works for any set of nodes, whose location may be singular in the traditional sense. It also works for arbitrary number of nodes.

- It is a high order and nested method, in the sense that with additional nodes the interpolation polynomial becomes progressively higher order.

- The combination of arbitrary locations of the nodes, arbitrary number of the nodes, and nestedness of the nodes makes the method ideal for the most flexible adaptive implementation.

The method is presented in a paper [8].

- A. Narayan and D. Xiu, Stochastic Collocation Methods on Unstructured Grids in High Dimensions via Interpolation, SIAM Journal on Scientific Computing, Volume 34, A1729-A1752, 2012.

4 Key Result: Minimal Element GPC

One of the more widely adopted methodologies is generalized polynomial chaos (gPC) [12], an extension of the standard polynomial chaos (PC) method [5]. It is well known that the performance

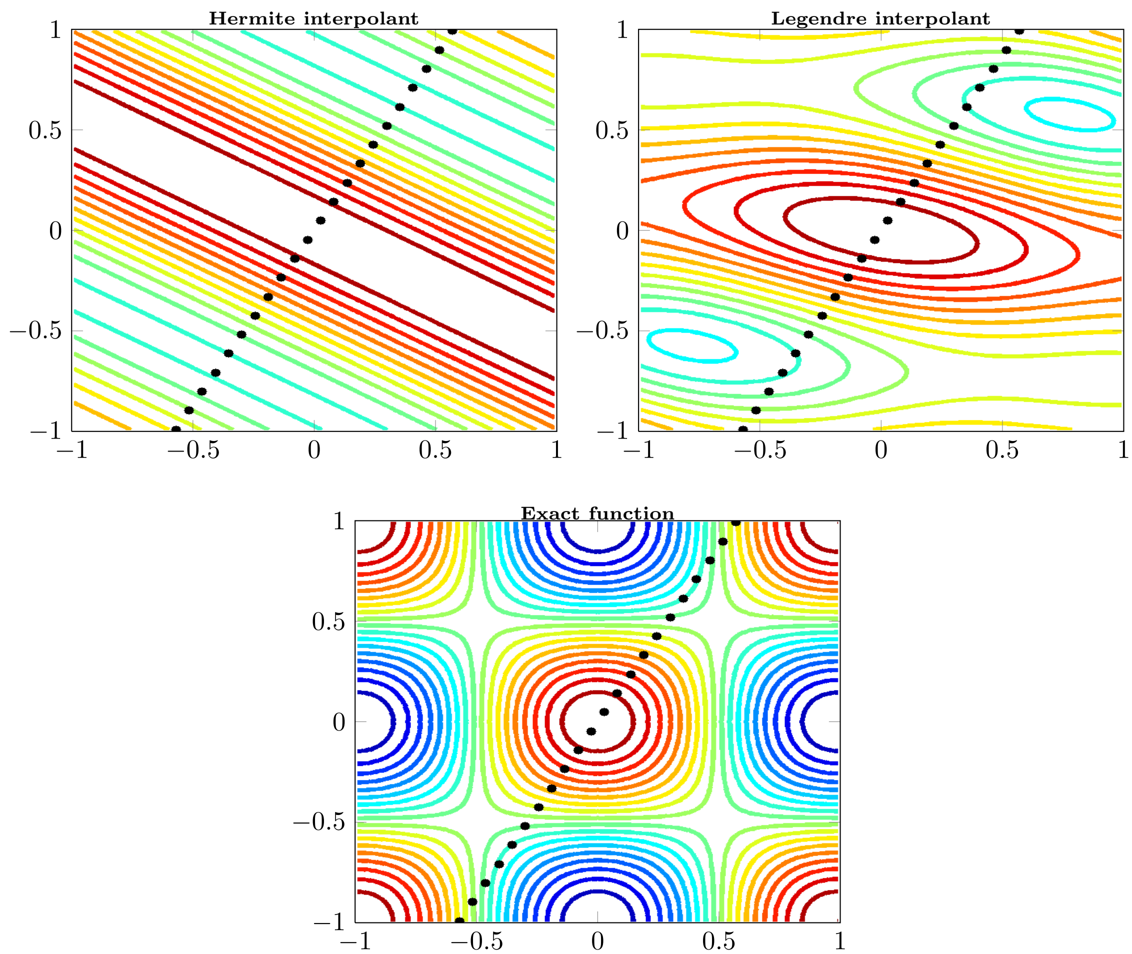


Figure 3: Contour plots for the least orthogonal interpolant for the Hermite (top left) and Legendre (top right) functions. The points where interpolation is enforced are marked on the left plots by 20 collinear black dots. Dark lines indicate lower values. The target function is at the bottom.

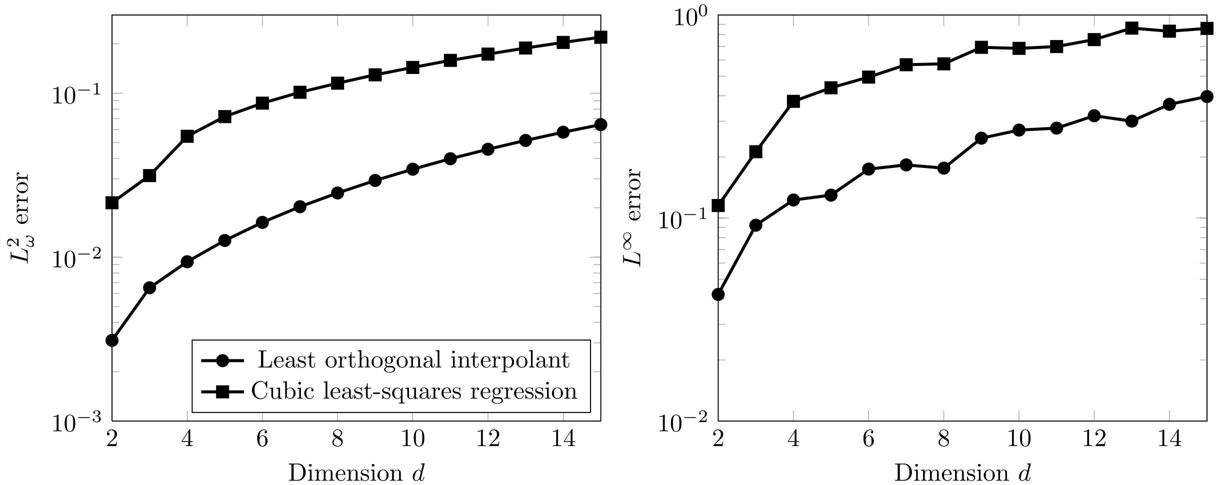


Figure 4: Interpolation accuracy for least orthogonal interpolation vs. cubic least-squares regression. The L^2_ω (left) and L^∞ (right) errors as measured on 10^6 Monte-Carlo nodes are shown.

of gPC based methods depend critically on the smoothness of the target function. Whenever the function is smooth in the random parameter space, these methods converge quickly and are highly efficient. On the other hand, their performance deteriorates when the target function lacks regularity, and in particular, possesses discontinuities. In these situations, the gPC methods using global polynomials suffer from Gibbs-type phenomenon and have very slow convergence. To circumvent the difficulty, multi-element gPC (ME-gPC) methods were developed [9, 10]. The idea is to decompose the random space into sub-domains, in each of which the target function is smooth and amenable to local gPC approximations. The undesirable impact of the discontinuities is thus confined in a limited number of elements surrounding the discontinuities, and the global solution can regain high accuracy away from them.

The challenge of the standard multi-element (ME) approaches is its simulation cost. One now needs to first resolve the stochastic problem in each of the elements. In the existing ME methods, the elements are constructed by splitting each axis, and then defining the corresponding hypercubes. A prominent drawback of this approach is that the construction inevitably utilizes a tensor structure, which results in a fast growth of the total number of elements in high dimensional random space. For example, if each axis is split into two parts (the minimal amount of splitting), then the total number of elements in d-dimensional random space is 2^d, where each element requires a solution of the original d-dimensional stochastic system. In high dimensions $d \gg 1$, the total simulation cost can be prohibitive. (In many cases, the axes are required to be split into more than two parts, though adaptive algorithms can reduce the number of splitting.)

During the course of this project, we developed a novel multi-element method that abandons the tensor structure in local element constructions. A distinct feature of the current method is that the local elements are defined by the underlying stochastic problem directly. More precisely, we seek to decompose the random space by splitting it into elements along where the discontinuities lie. By doing so, the total number of elements equals the number of smooth sub-domains defined by the underlying target function. In this sense the current method can be considered optimal. Hence the term minimal-element (mE) method.

The minimal-element method consists of the two important methods developed in this project. First, we need an algorithm to efficiently detect the existence of discontinuities; and if there is one, to locate its geometry and classify. To this end, we employ the high-order polynomial edge detection method ([6]), described in detail in Section 2. Once the random space has been decomposed into disjoint elements, defined by the smooth sub-domains of the target function, the next task is to construct accurate polynomial approximations in each elements. The challenge here is that the elements are of irregular shape, because of the arbitrary geometry of the underlying discontinuities. Attempts to map the irregular elements into regular elements will be difficult, if not impossible, especially in high dimensional spaces. Our proposed strategy is to use the stochastic collocation (SC) method directly on the irregular elements. In particular, since we have already computed the solution ensembles in the discontinuity detection step, we will not conduct further SC simulations. Instead, we will seek to construct gPC approximations in each element using the existing simulation results on the sparse grid generated by the discontinuity detector. The difficulty is that now the collocation points do not possess any structure because of the arbitrary boundaries imposed by the discontinuities. Also the number of collocation points in each element can be arbitrary. In order to construct high-order gPC polynomial approximations in each element, we employ the "least orthogonal interpolation" method developed in [8]. This is another important development of this project and is described in detail in Section 3. This method allows one to construct high order polynomial approximations in high dimensions based on arbitrary number of collocation nodes located at arbitrary locations. In our new mE-gPC method, we further improve the performance of the least orthogonal interpolation by adaptively choosing subsets of the sparse grids, from the discontinuity detection step, and constructing an interpolant that minimizes oscillations. The result is a nonlinear interpolation method, robust for smooth functions, that can perform post-processing polynomial construction regardless of nodal distribution or Euclidean dimension. We show that this method performs well when applied to point sets given by the discontinuity detector. We remark that the least orthogonal interpolation is merely a choice we make here. One is free to use other technique for the polynomial approximation in each element. For example, one can employ a least-square type polynomial regression.

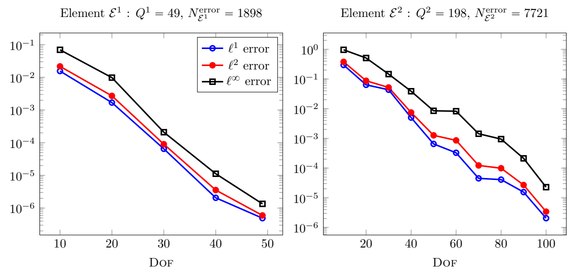
Figure 5 show the numerical errors in the function approximation of the same examples in Figure 1. We clearly observe exponential decay of the numerical errors. For all of these functions with discontinuities, the exponential decay of errors is only made possible because of the use of mE-gPC. This is because the mE-gPC constructs the gPC approximations in each smooth subdomains separately and therefore is immune to the discontinuities. Again, the number of smooth subdomain is determined by the actual number of smooth subdomains of the problem, and is at absolute minimum.

The efficacy of the mE-gPC method is further demonstrated in the following example containing multiple discontinuities.

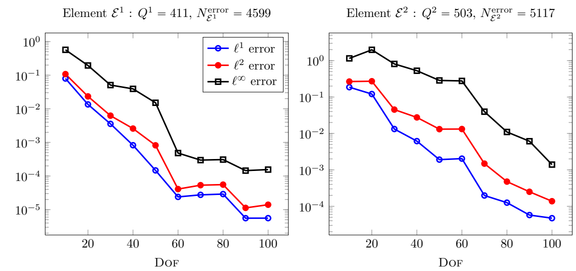
$$f_d^{\text{multi}}(x) = \begin{cases} f^1(x) - 2, & 3x_1 + 2x_2 \geq 0 \text{ and } -x_1 + 0.3x_2 < 0, \\ 2f_d^2(x), & 3x_1 + 2x_2 \geq 0 \text{ and } -x_1 + 0.3x_2 \geq 0, \\ 2f^1(x) + 4, & (x_1+1)^2 + (x_2+1)^2 < 0.95^2 \text{ and } d = 2, \\ f^1(x), & \text{otherwise.} \end{cases} \tag{3}$$

The function surface is shown in Figure 6 (a). When applied to this function, the minimal element method splits the input domain $[-1, 1]^2$ into four elements. The collocation nodes generated by the discontinuity detector are shown in Figure 6 (b) and the classification of $10,000$ random Monte Carlo samples is shown in Figure 6 (c).

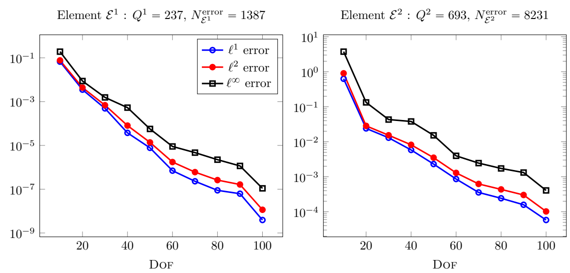
Figure 7 displays the error in each of the four elements obtained when the adaptive least orthogonal method is used to construct the interpolants. In all regions a very high level of accuracy is achieved using only a small number of points. Again, we observe exponential decay of the numerical errors – a unique and most desirable property achieved by mE-gPC.



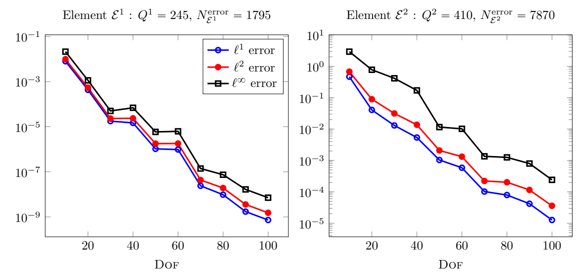
(a) Squared-shape discontinuity.



(b) linear-cut discontinuity.



(c) Triangle discontinuity.



(d) Circle discontinuity.

Figure 5: Numerical errors of mE-gPC method for the four examples in two-dimension. Each two plot show errors in the two smooth subdomains.

The mE-gPC method is presented in a paper [7].

- J. Jakeman, A. Narayan and D. Xiu, Minimal Multi-element Stochastic Collocation for Uncertainty Quantification of Discontinuous Functions, Journal of Computational Physics, Volume 242, 790-808, 2013.

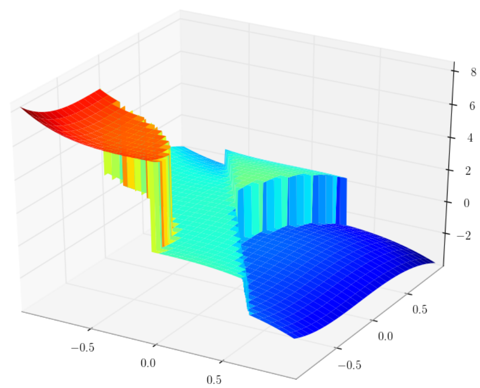
5 Personnel

The project has produced fundamentally new methods for high performance stochastic computing and uncertainty quantification. It has provided support for the PI, Prof. Dongbin Xiu at The University of Utah, to conduct this line of research. Meanwhile, during the course of the project, it has also provided support for Dr. Xueyu Zhu, who conducted post-doctoral research at University of Utah. Meanwhile, the team at Utah collaborated with Dr. Richard Archibald at Oak Ridge National Laboratory, Dr. John Jakeman at Sandia National Laboratory, and Prof Akil Narayan at University of Massachusetts at Dartmouth. The collaboration resulted in the high performance adaptive edge detection method (described in Section 2).

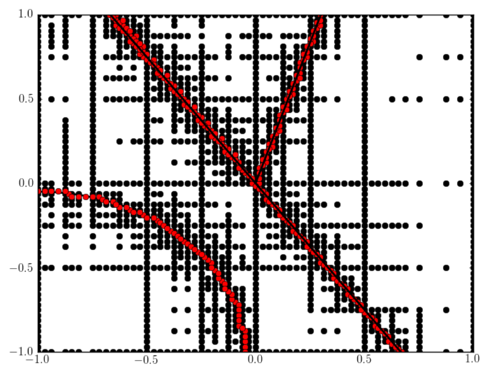
References

[1] R. Archibald, A. Gelb, R. Saxena, and D. Xiu. Discontinuity detection in multivariate space for stochastic simulations. *J. Comput. Phys.*, 228(7):2676–2689, 2009.

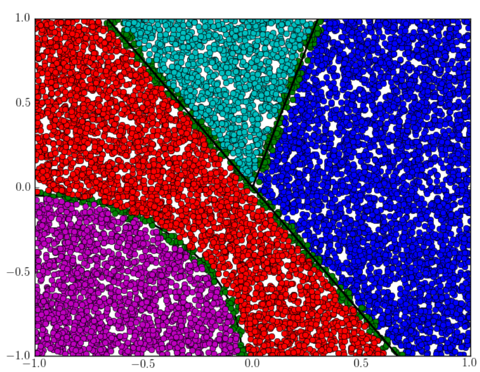
[2] R. Archibald, A. Gelb, and J. Yoon. Polynomial fitting for edge detection in irregularly sampled signals and images. *SIAM J. Numer. Anal.*, 43(1):259–279, 2005.



(a) $f_2^{\text{multi}}(x)$



(b) Collocation nodes



(c) MC sample classification

Figure 6: Discontinuity detection applied to a two dimensional function f_2^{multi} with multiple discontinuities. The true function is shown in (a), the points generated by the discontinuity detection algorithm are shown in (b), and 10000 randomly classified points are shown in (c). Four smooth regions are identified. Green points represent the subset of random points that cannot be classified with certainty.

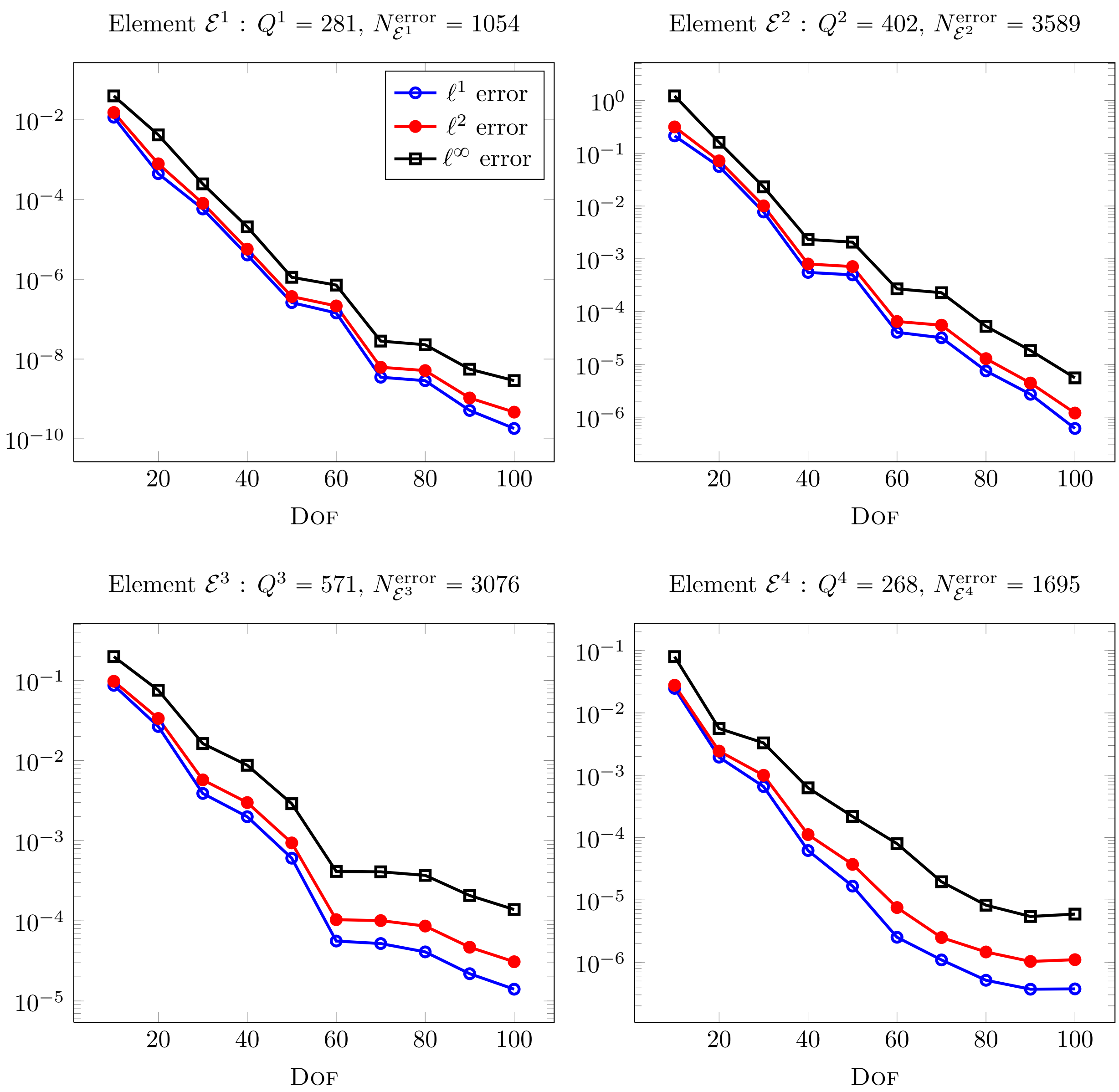


Figure 7: Errors for the minimal element method on test function f_2^{multi}.

[3] C. de Boor and A. Ron. On multivariate polynomial interpolation. *Constr. Approx.*, 6:287–302, 1990.

[4] C. de Boor and A. Ron. Computational aspects of polynomial interpolatoin in several variables. *Math. Comput.*, 58(198):705–727, 1992.

[5] R.G. Ghanem and P. Spanos. *Stochastic Finite Elements: a Spectral Approach.* Springer-Verlag, 1991.

[6] J. Jakeman, R. Archibald, and D. Xiu. Characterization of discontinuities in high-dimensional stochastic problmes on adaptive sparse grids. *J. Comput. Phys.*, 230:3977–3997, 2011.

[7] J.D. Jakeman, A. Narayan, and D. Xiu. Minimal multi-element stochastic collocation for uncertainty quantification of discontinuous functions. *J. Comput. Phys.*, 242, 2013.

[8] A. Narayan and D. Xiu. Stochastic collocation methods on unstructured grids in high dimensions via interpolation. *SIAM J. Sci. Comput.*, 34:A1729–A1752, 2012.

[9] X. Wan and G.E. Karniadakis. An adaptive multi-element generalized polynomial chaos method for stochastic differential equations. *J. Comput. Phys.*, 209(2):617–642, 2005.

[10] X. Wan and G.E. Karniadakis. Multi-element generalized polynomial chaos for arbitrary probability measures. *SIAM J. Sci. Comput.*, 28:901–928, 2006.

[11] D. Xiu and J.S. Hesthaven. High-order collocation methods for differential equations with random inputs. *SIAM J. Sci. Comput.*, 27(3):1118–1139, 2005.

[12] D. Xiu and G.E. Karniadakis. The Wiener-Askey polynomial chaos for stochastic differential equations. *SIAM J. Sci. Comput.*, 24(2):619–644, 2002.