Commun. Comput. Phys.
doi: 10.4208/cicp.130813.060214a

Vol. **16**, No. 2, pp. 365-381
August 2014

On the Choice of Design Points for Least Square Polynomial Approximations with Application to Uncertainty Quantification

Zhen Gao[1] and Tao Zhou[2,*]

[1] *School of Mathematical Sciences, Ocean University of China, Qingdao, China.*
[2] *Institute of Computational Mathematics, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China.*

Received 13 August 2013; Accepted (in revised version) 6 February 2014

Available online 9 May 2014

Abstract. In this work, we concern with the numerical comparison between different kinds of design points in least square (LS) approach on polynomial spaces. Such a topic is motivated by uncertainty quantification (UQ). Three kinds of design points are considered, which are the Sparse Grid (SG) points, the Monte Carlo (MC) points and the Quasi Monte Carlo (QMC) points. We focus on three aspects during the comparison: (i) the convergence properties; (ii) the stability, i.e. the properties of the resulting condition number of the design matrix; (iii) the robustness when numerical noises are present in function values. Several classical high dimensional functions together with a random ODE model are tested. It is shown numerically that (i) neither the MC sampling nor the QMC sampling introduce the low convergence rate, namely, the approach achieves high order convergence rate for all cases provided that the underlying functions admit certain regularity and enough design points are used; (ii)The use of SG points admits better convergence properties only for very low dimensional problems (say $d \leq 2$); (iii)The QMC points, being deterministic, seem to be a good choice for higher dimensional problems not only for better convergence properties but also in the stability point of view.

AMS subject classifications: 65C20, 65M70

Key words: Least square, polynomial approximations, uncertainty quantification, condition number.

1 Introduction

In recent years, there has been a growing need for including uncertainty in mathematical models and quantify its effect on given outputs of interest used in decision making. In

*Corresponding author. *Email addresses:* zhengao@ouc.edu.cn (Z. Gao), tzhou@lsec.cc.ac.cn (T. Zhou)

general, a probabilistic setting can be used to include these uncertainties in mathematical models. In such framework, the input data are modeled as random variables, or more generally, as random fields with a given correlation structure. Thus, the goal of the mathematical and computational analysis becomes the prediction of statistical moments of the solution or statistics of some quantities of physical interest of the solution, given the probability distribution of the input random data. Examples of quantities of interest could be the mean or the variance of the exact solution, the solution values in a given region, etc. This is the so called Uncertainty Quantification (UQ).

A fundamental problems in UQ is to approximate a multivariate function $Z = f(x,Y_1,Y_2,\cdots,Y_N)$ with random parameters $\{Y_i\}_{i=1}^N$, which might be a solution resulting from a stochastic PDE problem or other complex models. Stochastic modeling methods for uncertainty quantification are being well developed in recent years. A traditional approach is the Monte Carlo (MC) method [7]. In MC method, one first generates a number of random realizations for the prescribed random inputs and then utilizes existing deterministic solvers for each realization. Although the convergence rate of Monte Carlo method is relatively slow (converges asymptotically at a rate of $\frac{1}{\sqrt{K}}$ with K realizations), it is independent of the dimensionality of the random space, i.e., independent of the number of random variables used to characterize the random inputs. Significant advances have been made in improving the efficiency of Monte Carlo schemes during the past years. Stochastic collocation (SC) [17,18,23,24] method is another non-intrusive method. Like MC method, the SC method can be easily implemented and leads naturally to the solution of uncoupled deterministic problems, even in presence of input data which depend nonlinearly on the driving random variables. When the number of input random variables is small, the SC method is a very effective numerical tool. However, in many cases, a large number of collocation points are still needed to get a good convergence rate.

One of the most popular intrusive methods is the generalized polynomial chaos (gPC) methods [10,25], which are the generalizations of the Wiener-Hermite polynomial chaos expansion developed in [22]. Compared to the SC method, the gPC methods need relatively smaller number of degree of freedom, and such methods also exhibit fast convergence rates with increasing order of the expansions, provided that solutions are sufficiently smooth with respect to the random variables. However, the resulting set of deterministic equations is often coupled, thus, care is needed to design efficient and robust solver, and furthermore, the form of the resulting equations can become very complicated if the underlying differential equations have nontrivial and nonlinear forms [2,28].

To efficiently build such a gPC approximation, we consider in this work the lease square projection onto the polynomial spaces. The LS approach is actually a combination of the SC method and gPC method. It sakes a polynomial approximation for the unknown solution while using collocation points to evaluate the expansion coefficients. Such an idea has been used by many researchers and been given several different names [1,5,6,12,13], to name a few. To assure the well-posedness, the number of samples drawn from the input distribution is typically taken to be 2 to 3 times the dimension of the polynomial space.

We will focus on the choice of the design points used in the LS approach. Three kinds of points will be considered, namely, the sparse grid points, the low-discrepancy Hammersley/Halton (QMC) points and the random points. The comparison between these choices will be done in three aspects: (i) we investigate the efficiency and convergence properties; (ii) we investigate the stability; (iii) we assess the robustness when numerical noises are present in function values. Despite the large amount of the literature, such problems have not been well studied.

The rest of the paper are organized as following: we give a brief introduce of the LS approach in Section 2. Different choices of design points are introduced in Section 3. Detail numerical comparisons are provided in Section 4, and we finally give some conclusions in Section 5.

2 The least square approach

In this section, we follow closely the work [4, 15, 27] to give a basic introduction for the LS approach.

Let $\mathbf{Y}=(Y_1,\cdots,Y_N)^T$ be a vector with N random variables, taking values in $\Gamma\subset\mathbb{R}^N$. For simplicity, we assume that $\Gamma\equiv[-1,1]^d$. We assume that the random variables $\{Y_i\}_{i=1}^N$ are independent with each other and there is a probability density function ρ_i for each random variable Y_i. We denote with $\rho(\mathbf{Y})=\prod_{i=1}^N\rho_i(Y_i):\Gamma\to\mathbb{R}^+$ the probability density function of $\mathbf{Y}$. The problem considered here is to approximate a function $Z=f(\mathbf{Y})$ taking values in $\mathbb{R}$. The case for functions with values in some Banach space, which representing the solution of a (possibly non-linear) differential or integral problem ($Z=f(x,\mathbf{Y})$) can be discussed in a similar way.

The mean of $Z=f(\mathbf{Y})$ is defined as

$$\mathbb{E}[Z]:=\int_\Gamma f(\mathbf{Y})\rho(\mathbf{Y})d\mathbf{Y}. \tag{2.1}$$

We assume that the functions considered in this paper are in the space L^2_ρ endowed with the norm

$$||f||_{L^2_\rho}=\Big(\int_\Gamma f^2(\mathbf{Y})\rho(\mathbf{Y})d\mathbf{Y}\Big)^{1/2}. \tag{2.2}$$

Considering problems with only one stochastic dimension, the `best` type of approximation polynomial can be chosen according to the measure of the random variable, for example, Legendre polynomials for uniform distribution, Hermite polynomials for Gaussian distribution and so on [25]. For higher dimensional case, one can construct the multivariate polynomial basis by tensorizing 1D orthogonal polynomials with respect to each weight $\{\rho_n\}_{n=1}^N$ separately. To do this, let us first define the following multi-index:

$$\mathbf{n}=(n_1,\cdots,n_d)\in\mathbb{N}^d,\quad\text{with}\quad|\mathbf{n}|=n_1+\cdots+n_d.$$

Then, every d dimensional polynomial can be written as

$$\mathbf{\Phi}_{\mathbf{n}}(\mathbf{Y})=\prod_{i=1}^{d}\phi_{n_i}^{i}(Y_i), \tag{2.3}$$

where $\{\phi_k^i\}_{k=1}^{\infty}$ are one dimension orthonormal polynomials with respect to Y_i.

Let $\Lambda\subset\mathbb{N}^d$ be an finite multi-index set which is monotonous [15]. The dimension of an index set Λ will be denoted by $\#\Lambda$. A finite dimension polynomial spaces introduced by Λ can be defined by

$$\mathbf{P}^{\Lambda}:=\text{span}\{\mathbf{\Phi}_{\mathbf{n}}(\mathbf{Y}),\mathbf{n}\in\Lambda\}. \tag{2.4}$$

For convenience, we arrange the index set Λ in the lexicographical order, namely, given $\mathbf{n}',\mathbf{n}''\in\Lambda$

$$\mathbf{n}'<\mathbf{n}''\ \Leftrightarrow\ \exists j: n_j'<n_j''\wedge(n_i'=n_i'',\forall i\leq j). \tag{2.5}$$

According to this order, we can denote by $\mathbf{n}_j$ the jth multi-index of Λ, and then, the bases in $\mathbf{P}^{\Lambda}$ can be denoted as $\{\mathbf{\Phi}_j\}_{j=1}^{\#\Lambda}$.

To construct a polynomial approximation $f^{\Lambda}\in\mathbf{P}^{\Lambda}$ for the function $Z=f(\mathbf{Y})$ by LS approach, we first compute the exact function $f(\mathbf{Y})$ in $\mathcal{M}>\#\Lambda$ collocation points $\mathbf{y}_1,\cdots,\mathbf{y}_{\mathcal{M}}$. Then, we find a discrete least square approximation ϕ_{Λ} by requiring

$$f^{\Lambda}=P_{\mathcal{M}}^{\Lambda}f=\underset{v\in\mathbb{P}^{\Lambda}}{\text{argmin}}\frac{1}{\mathcal{M}}\sum_{k=1}^{\mathcal{M}}(f(\mathbf{y}_i)-v(\mathbf{y}_i))^2. \tag{2.6}$$

We introduce the discrete inner product

$$(u,v)_{\mathcal{M}}=\frac{1}{\mathcal{M}}\sum_{k=1}^{\mathcal{M}}u(\mathbf{y}_i)v(\mathbf{y}_i) \tag{2.7}$$

and the corresponding discrete norm $||u||_{\mathcal{M}}=(u,u)_{\mathcal{M}}^{1/2}$.

Remark 2.1. We remark that we are considering here the standard least square approach (2.6). While it will be interesting if one considers the weighted LS approach in the following sense

$$f^{\Lambda}=P_{\mathcal{M}}^{\Lambda}f=\underset{v\in\mathbb{P}^{\Lambda}}{\text{argmin}}\frac{1}{\mathcal{M}}\sum_{k=1}^{\mathcal{M}}\omega_i(f(\mathbf{y}_i)-v(\mathbf{y}_i))^2. \tag{2.8}$$

When dealing with different original weight functions ρ, adding some suitable weights into the LS approach can indeed make the approach more stable. We will report such weighted LS approaches in our forthcoming paper [27]. However, in this paper, we only consider the uniform least square approach (2.6).

2.1 High dimensional polynomial spaces

One traditional high dimensional polynomial space is the full tensor product (TP) space

$$\mathbf{P}_q^d := \text{span}\Big\{ \mathbf{\Phi}_{\mathbf{n}}(\mathbf{x}) : \mathbf{n} \in \mathbb{N}^d, \max_{j=1,\cdots,d} n_j \le q \Big\}. \tag{2.9}$$

The dimension of such TP space is

$$\#\Lambda = \dim(\mathbf{P}_q^d) = (q+1)^d. \tag{2.10}$$

Note that when $d \gg 1$ the cardinality of TP polynomial spaces grows exponentially with the polynomial degree q. This is the so-called *curse of dimensionality*, and this is also the reason why the full TP spaces are rarely used in real applications when $d > 5$.

In real applications, the following total degree (TD) polynomial space is often used instead of the TP spaces [17, 25]

$$\mathbf{W}_q^d := \text{span}\Big\{ \mathbf{\Phi}_{\mathbf{n}}(\mathbf{x}) : \mathbf{n} \in \mathbb{N}^d, |\mathbf{n}| = \sum_{j=1}^{d} n_j \le q \Big\}. \tag{2.11}$$

The degree of freedom (DoF) of the TD space is

$$\#\Lambda' = \dim(\mathbf{W}_q^d) = \binom{q+d}{q}. \tag{2.12}$$

It is noted that the Dof of the TD spaces still increase when the degree q is increased, however, such grows property is more slowly than that of the TP spaces. The differences between the TP spaces and the TD spaces in 3 dimensional case are shown in Fig. 1.

Motivated by the above discussions, we will consider in this work the high dimensional TD spaces.

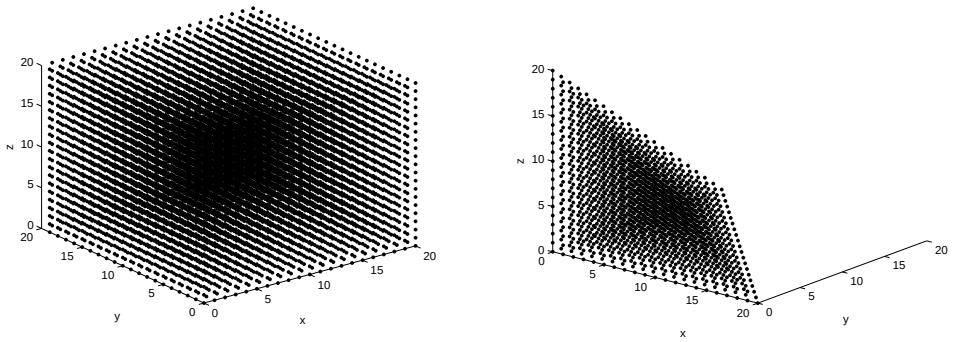

Figure 1: Degree of freedom of the polynomial spaces. TP space(left), TD space (right).

2.2 Algebraic formulation

Consider the approximation in the space $\mathbf{P}^\Lambda = \text{span}\{\mathbf{\Phi}_j\}_{j=1}^{\#\Lambda}$ with collocation points $\{\mathbf{y}_k\}_{k=1}^{\mathcal{M}}$, the least square solution can be written in

$$f^\Lambda = \sum_{j=1}^{\#\Lambda} c_j \mathbf{\Phi}_j, \tag{2.13}$$

where $\mathbf{c} = (c_1, \cdots, c_{\#\Lambda})^T$ is the coefficient vector. By using the collocation points $\{\mathbf{y}_k\}_{k=1}^{\mathcal{M}}$, we get the following design matrix

$$\mathbf{D} = \big(\mathbf{\Phi}_j(\mathbf{y}_k)\big), \quad j = 1, \cdots, \#\Lambda, \quad k = 1, \cdots, \mathcal{M}. \tag{2.14}$$

Then the algebraic problem to determine the unknown coefficient $\mathbf{c}$ can be formulated as:

$$\mathbf{c} = \underset{\mathbf{c}\in\mathbb{R}^{\#\Lambda}}{\arg\min} ||\mathbf{D}\mathbf{c} - \mathbf{b}||_2, \tag{2.15}$$

where $\mathbf{b}$ contains the evaluations of the target function f in the collocation points. In what follows, for notation simplicity, we denote with $\mathcal{N}$ the degree of freedom of the polynomial space $\mathbf{P}^\Lambda$, e.g. $\mathcal{N} = \#\Lambda$.

The solution of the least square problem can also be computed by solving an $\mathcal{N} \times \mathcal{N}$ system. Namely, one can solve the following system

$$\mathbf{A}\mathbf{c} = \mathbf{f}, \tag{2.16}$$

with $\mathbf{A} = \mathbf{D}^T\mathbf{D}, \mathbf{f} = \mathbf{D}^T\mathbf{b}$, and furthermore, we have

$$\mathbf{f} := \big((f, \mathbf{\Phi}_j)_{\mathcal{M}}\big)_{j=1,\cdots,\mathcal{N}}, \quad \mathbf{A} := \big((\mathbf{\Phi}_i, \mathbf{\Phi}_j)_{\mathcal{M}}\big)_{i,j=1,\cdots,\mathcal{N}}. \tag{2.17}$$

For the computation point of view, we can solve problem (2.15) by using QR factorization. Alternatively, we can also solve (2.16) by Cholesky factorization.

3 Choices of nodal sets

As we mainly focus on the choices of collocation points in the LS approach, we introduce in this section some nature choices of the collocation points. Various kinds of nodal sets have been investigated by researchers when dealing with different problems, for example, the random sampling method [15], the Hammersley/Halton dataset [11]. In [6], sparse grid points set is also introduced. During the review procedure of this paper, Zhou et al. also introduce a new type of specially designed points in [27]. However, as far as we know, there is no detailed comparison between these choices. In this paper, we will concern with such numerical comparison between these choices. To do this, we give in the following a quickly review for the three kinds of design points mentioned above.

3.1 Random sampling

A nature choice is the random sampling according to the measure ρ, which is well known as Monte Carlo (MC) method. MC methods have been applied to many applications and their implementations are straightforward. Although the convergence rate for the ensemble mean of the direct MC simulation with M realizations converges asymptotically at a rate $\frac{1}{\sqrt{M}}$ is relatively slow, it is independent of the dimensionality of the random space, i.e., independent of the number of random variables used to characterize the random inputs. In LS framework, the use of MC points has been investigated in [4, 15] both numerically and theoretically. For bounded measures ρ, it is shown in [15] that the least square projection on a polynomial space with random samples leads to quasi optimal convergence rates (up to logarithmic factors) provided that the number of samples scales as $M \sim (\#\Lambda)^2$ (Note this is an one dimensional result). Using different techniques, the authors in [4] proved a more general result. In particular, The approach in [4] is not limited to polynomial spaces. We remark that in [4,15], the convergence results are in probability, e.g, convergent with high probability or convergent in the expectation sense, and this is nature when random samples are used.

3.2 Sparse grid

The sparse grid quadrature was first proposed by Smolyak [21], and further development has been done by researchers see, e.g., [19, 20]. It is noticed that the sparse grid reduce dramatically the number of collocation points compared to the full tensor grid, while preserving a high level of accuracy. The sparse grid quadrature points are introduced in the LS approach in [6]. Unlike the tensor produce formula, the Smolyak algorithm is a linear combination of product formulas, and the linear combination is chosen in such a way that an interpolation property for $d=1$ is preserved for $d>1$. Compared to the tensor product rule, only products with a relatively small number of points are used and the resulting nodal set has significantly less number of nodes starting with the one-dimensional interpolation formula, the Smolyak algorithm is given by

$$A(q,d)=\sum_{q-d+1\le|\mathbf{i}|\le q}(-1)^{l-|\mathbf{i}|}\binom{d-1}{l-|\mathbf{i}|}\left(\mathcal{U}^{i_1}\otimes\cdots\otimes\mathcal{U}^{i_d}\right),$$

where $\mathbf{i}=(i_1,\cdots,i_d)$ and $\mathcal{U}^{i_k}$ is the traditional one dimensional interpolation operator using i_k collocation points. To compute $A(q,d)$, we only need to evaluate function on the `sparse grid`

$$\Theta=\mathcal{H}(q,d)=\bigcup_{q-d+1\le|\mathbf{i}|\le q}\left(\Theta_1^{i_1}\otimes\cdots\otimes\Theta_1^{i_d}\right).$$

There are mainly two kinds of abscissas in the construction of the Smolyak formula. One is the `Clenshaw-Curtis` abscissas. These abscissas are the extrema of Chebyshev polynomials and can be constructed in a nested way. The other frequently used abscissas

is the `Gaussian abscissas`, in which one uses the zeros of the orthogonal polynomials with respect to some positive weight. But, these Gaussian abscissas are in general not nested. The natural choice of the weight should be the probability density function ρ_i of the random variables ξ_i. For details on the sparse grid, please refer to [18, 24] and references therein. It has been shown that if we set $q=d+k$ (the index k is called `level`), then $A(d+k,d)$ is exact for all polynomials in $\mathbf{W}_k^d$ [24].

3.3 Hammersley/Halton points

The Hammersley dataset and the Halton dataset, which belong to the low discrepancy sequences [11] have been investigated in the area of the quasi-Monte Carlo integration with large number of variables. The Hammersley and Halton points are closely related, in the sense that if one drops the first dimension of the standard d-dimensional Hammersley sequence, you get the standard Halton sequence of dimension $d-1$. Thus, they are usually referred to as the Hammersley/Halton points. One can refer to [16] for a comparison work between these datasets.

The standard Hammersley sequence has slightly better dispersion properties than the standard Halton sequence. However, it suffers from the problem that you must know, beforehand, the number of points you are going to generate. Namely, if you have computed a Hammersley sequence of length M, and you want to compute a Hammersley sequence of length $2M$, you have to discard your current values and start over. By contrast, you can compute M points of a Halton sequence, and then M more, and this will be the same as computing the first $2M$ points of the Halton sequence in one calculation.

In general, for the same sample size, Hammersley/Halton sample is more representative of the random distribution than the random generated samples. In low dimensions, both the multidimensional Hammersley sequence and Halton sequence quickly fills up the space in a well-distributed pattern. However, for higher dimensions (such as $N=40$), the initial elements of the sequences can be very poorly distributed; it is only when M, the number of sequence elements, is large enough relative to the spatial dimension, that the sequence is properly behaved. Remedies for this problem were introduced by researchers, see e.g., Kocis and Whiten [14].

We show in Fig. 2 the comparison of points distribution of the above three kinds of datasets in the uniformly distribution case.

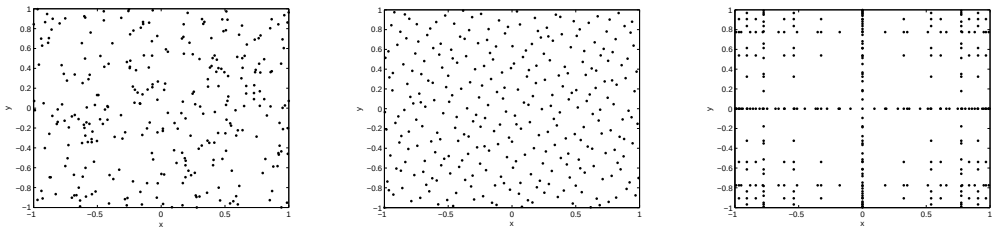

Figure 2: Points distribution: MC (left), QMC (middle) and SG with *level* $k=5$ (right).

Remark 3.1. We remark that the use of deterministic points in the LS approach is of independent interest. In [27], specially designed deterministic points are used in the LS framework, and similar analysis results as in [4,15] are also provided in [27].

4 Numerical examples

In this section, we provide with several numerical examples to show the performance of the datasets mentioned above in the LS approach. Throughout this section, the random variables are set to be uniformly distributed random variables and the SG points are set up based on the one dimensional Gauss-Legendre points. The main purpose is to investigate the numerical performance of different points. It is well know that, unlike the MC and QMC points, the number of the SG points is restricted by the level k, and thus not continuous. To make a fair comparison, we thus adjust the number of MC and QMC points according to the number of SG points. For one dimensional case, the number of points with level k is defined by $M=2^k+1$. The following classic test functions are considered:

- Gaussian function: $f_1(\mathbf{Y})=\exp^{-\sum_{i=1}^d c_i^2(Y_i-w_i)^2}$;
- Meromorphic function: $f_2(\mathbf{Y})=\frac{1}{\sum_{i=1}^d w_i+c_iY_i}$;
- A function with lower regularity: $f_3(\mathbf{Y})=\sum_{i=1}^d \left|Y_i\right|^3$,

where w_i,c_i are randomly chosen constants.

4.1 One-dimensional tests

We begin with the one dimensional tests to show the basic properties of the LS approach.

In Fig. 3, we plot the convergence properties with respect to the polynomial order of LS approach for the test function f_1. The left plots are for MC, the middle plots are for QMC and the right plots are for SG. The affect of noises in function values to the convergence rate is also given, more precisely, no noises are introduced in top (a) plots, 0.01% Gaussian noises (e.g., $\tilde{f}=f*(1+0.01*N(0,1))$) are introduced into the function values in the bottom (b) plots. It is shown in the pictures that: (i) neither the MC sampling nor the QMC sampling introduce the low convergence rate, namely, the LS approach can achieve high order convergence rate (with high level grids) using any kind of the three points; (ii) the use of SG points results in good convergence properties than the other two choices, and it is also noted that the LS approach using SG points is more stable than the other choices in the one dimensional case; (iii) in presence of noise values, the convergence rate is deteriorated for all the three cases, however, the LS approach with SG points is more robust than the other two.

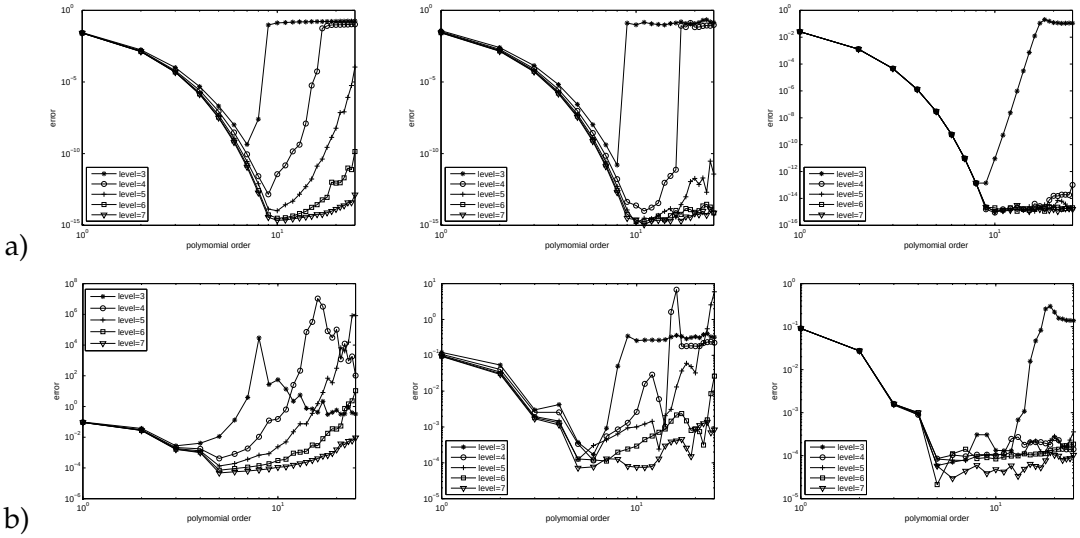

Figure 3: One dimensional test: error decay for f_1 with respect to polynomial order. a) noise free, b) 0.01% Gaussian noise. MC points (left), QMC points (middle), and SG points (right).

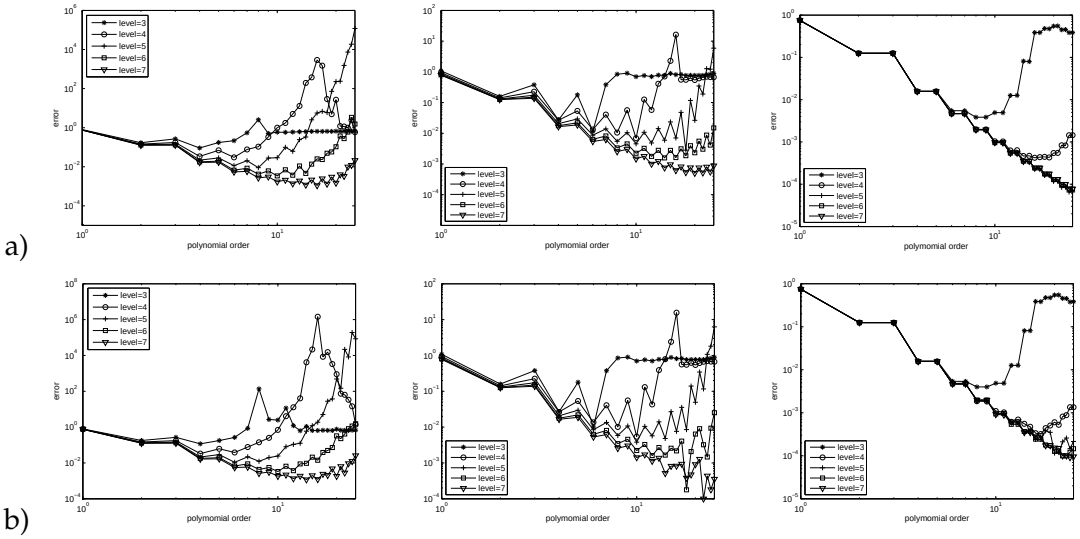

Figure 4: One dimensional tests: error decay for f_3 with respect to polynomial order. a) noise free, b) 0.01% Gaussian noise. MC points (left), QMC points (middle), and SG points (right).

Analogous plots are done for the test function f_3 in Fig. 4. Similar conclusions can be made as above discussion, nevertheless, the convergence rate for f_3 is lower than that of for f_1, and this is due to the low regularity of the test function f_3.

Let us now consider the properties of the resulting condition numbers of the LS approach, namely, $\text{cond}(\mathbf{A}) = \frac{\sigma_{\max}(\mathbf{A})}{\sigma_{\min}(\mathbf{A})}$. It is clear that the resulting condition number using

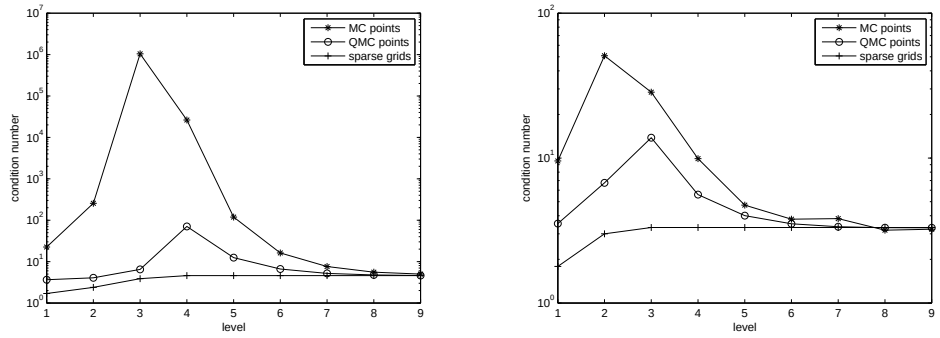

Figure 5: One dimension tests: condition numbers for different points sets. PC order $q=9$ (left) and PC Order $q=4$ (right).

MC points will be `random`. To deal with such randomness, it is nature to report the mean condition number, namely, we repeat the computation M times and get the mean of condition numbers. In the following, whenever the resulting condition number of MC points is considered, it refers to the mean condition number of 1000 repeats. The comparison between three choices are shown in Fig. 5, it is noticed that the SG points is more stable in the stability point of view (smaller condition number). However, each choice of nodal points result in small condition number when large number of points (or, high level grids) are used.

Remark 4.1. Concerning the discussions above, we would like to make the following remarks:

- For the one dimension case, the SG points are just the Gauss points, which is known to be the best quadrature points. In fact, $M>2q$ points are enough to guarantee the optimal convergence when approximating functions in the polynomial space P^q. Thus, the SG points works very well. However, one is more interested in higher dimensional cases.

- Note that in Figs. 3-5, few points are used for low level grids. This is not the standard least square approach, as the number of points are less than the DoF of the approximated spaces. However, it is not an issue in the computation point of view, as one can solve the `minimum norm` least square solution. Alternatively, a standard way to use less points is to use the L^1 minimization approach [26]. In standard least square approach, one is interested in the case where the number of points are larger than the DoF. Typically, one may interested in the linear rule $(M=c\#\Lambda)$ and/or the quadratic rule $(M=c(\#\Lambda)^2)$.

- We noticed that for the MC points, there are peaks in the condition number plots. The peaks happen when the number of points is the same as the number of the DoF, and in such cases, the problem reduced to the interpolation problem. The peaks are probably due to the bad distribution of the points.

4.2 Two-dimensional tests

Let us now consider the two dimensional case. Again, we adjust the number of points of MC and QMC according to SG. We remark that for level k, the number of SG points behaves approximately as $M \sim \frac{2^k}{k!} d^k$. In Fig. 6, we plot the convergence properties of LS approach for f_3. The left plots are for MC, the middle plots are for QMC and the right plots are for SG. The affect of noises in function values to the convergence rate is also given, more precisely, no noises are introduced in the top (a) plots, and 0.01% Gaussian noises are introduced into the function values in the middle (b) plots. It is shown that the SG points lost the dominant superiority in view of convergence rate, and the convergence

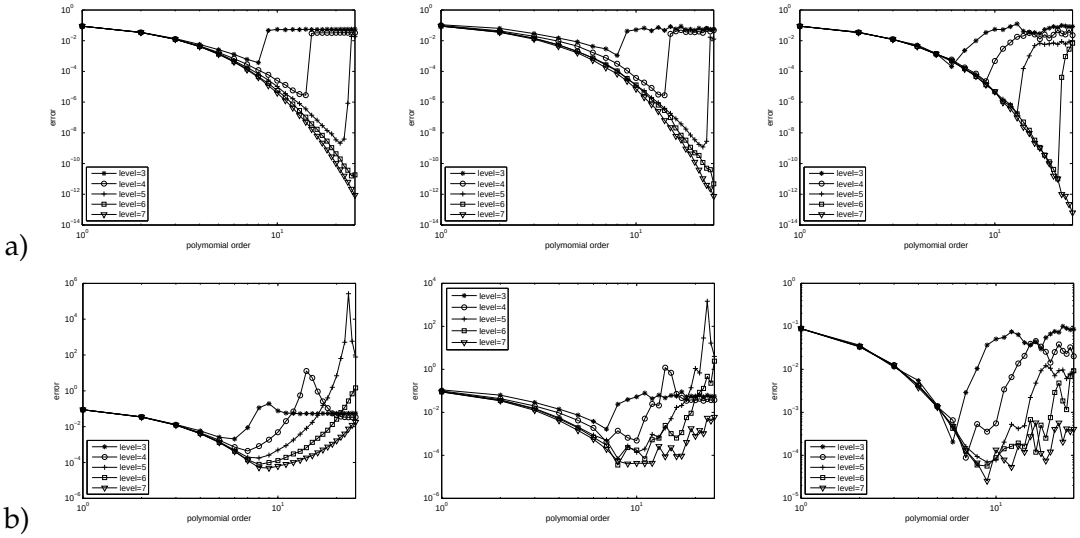

Figure 6: Two dimension tests: error decay for f_2 with respect to polynomial order. a) noise free, b) 0.01% Gaussian noise. MC points (left), QMC points (middle), and SG points (right).

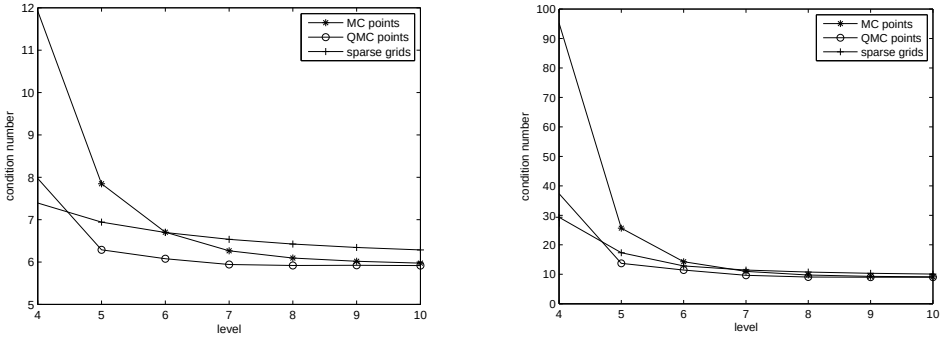

Figure 7: Two dimension tests: condition number with respect to the number of points. Left: polynomial order $q=5$. Right: polynomial order $q=8$.

rate of MC and QMC becomes comparable with SG. Properties of the resulting condition numbers is provided in Fig. 7. It can be seen the condition number for three kinds of points are comparable.

4.3 High dimensional tests

Five dimensional problems are also tested. The tests are done in the same way as the above sections (Figs. 8-9). It is noticed that the QMC points works better than the other two both in the convergence (Fig. 8, for f_3) and in the stability (Fig. 9) point of view.

As it has been shown in [15] that the quadratic rule $M = c(\#\Lambda)^2$ results in stable LS approach when MC points are used. We would like to investigate this property here for

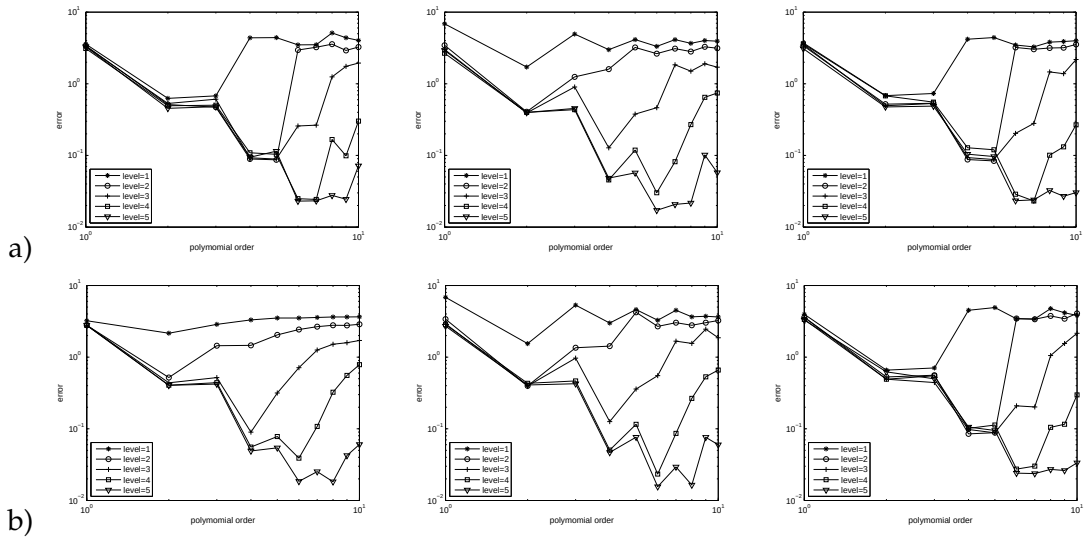

Figure 8: Five dimension tests: error decay for f_3 with respect to the expansion term. a) noise free, b) with 0.01% Gauss distributed noise. (Left) MC (100 repetitions), (middle) QMC, (right) SG.

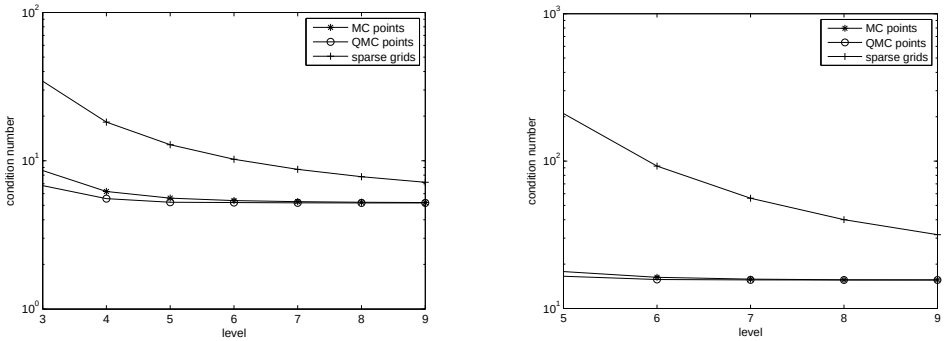

Figure 9: Five dimension test: condition numbers. Left: $q=3$ and right: $q=5$.

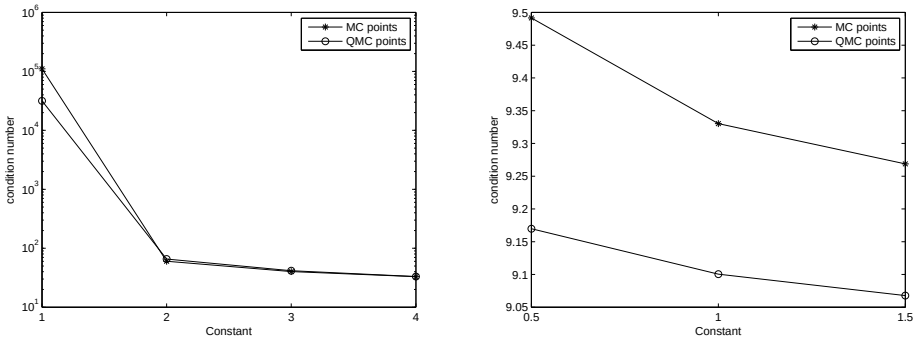

Figure 10: Eight dimension tests: condition numbers for q=5. Left: $M=c\#\Lambda$. Right: $M=c\#\Lambda^2$.

QMC points. In Fig. 10, condition numbers are reported for MC and QMC points for the eight dimensional TD space. The left plot is for the linear rule $c(\#\Lambda)$ and the right plot is for the quadratic rule $c(\#\Lambda)^2$. It is noticed that the using of $c(\#\Lambda)^2$ points (right plot) results in smaller condition numbers compared to the linear rule (left plot). In other words, the QMC points share similar properties with the MC points, however, we do not have corresponding proofs as in [4, 15].

4.4 A random ODE model: Genetic toggle switch

We now consider a random ODE model: the genetic toggle switch. Such model is used to describe a genetic switch in EScherichia coli [9]. The system reads

$$\frac{du}{dt}=\frac{\alpha_1}{1+v^\beta}-u,\quad \frac{dv}{dt}=\frac{\alpha_2}{1+\omega^\gamma}-v,\quad \omega=\frac{u}{(1+[IPTG]/\mathcal{K})^\eta},$$

where α_1, α_2, β, γ, η, $\mathcal{K}$ are parameters and $[IPTG]$ is a system input that controls the behavior of the steady state solution.

We first choose (randomly) parameters α_2,β as random variables on the form $\boldsymbol{\alpha}=\langle\boldsymbol{\alpha}\rangle(1+\sigma x)$, where $\langle\boldsymbol{\alpha}\rangle=(15.6,2.5)$ are the expectation values, and $x=(x_1,x_2)$ are uniformly distributed random variables with $x_i\in[-1,1]$, $i=1,2$. other parameters are fixed as the mean values, One can refer to [8,23] for more details. A 10% variation is introduced during the computation, that is, $\sigma=0.1$. We employ a 4th-order Runge-Kutta solver as the deterministic solver and the Monte Carlo method with 10000 simples are used to get the *exact solutions*. In Fig. 11, we show the convergence properties of LS approach with respect to the polynomial order for different number of points. High order convergence rate is obtained.

We then randomly choose parameters α_2, β, γ, η as random variables with the vector form $\boldsymbol{\alpha}=\langle\boldsymbol{\alpha}\rangle(1+\sigma x)$, where $\langle\boldsymbol{\alpha}\rangle=(15.6,2.5,1,2.0015)$ are the expectation values, and $x=(x_1,x_2,x_3,x_4)$ are uniformly distributed random variables with $x_i\in[-1,1]$, $i=1,\cdots,4$. We set again a 10% variation. The convergence rates are shown in Fig. 12. It is shown

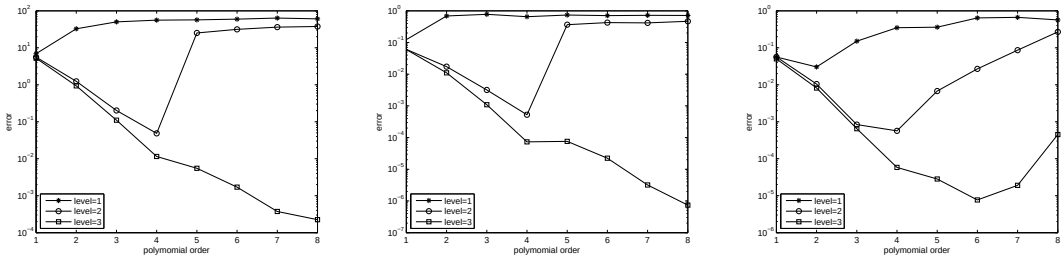

Figure 11: Two dimensional random ODE: error decay with respect to the polynomial order ($T=2$). MC (left), QMC (middle) and SG (right).

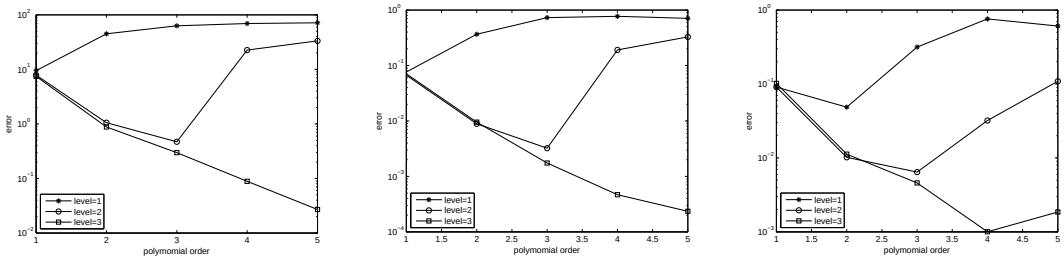

Figure 12: Four dimensional random ODE: error decay with respect to the PC order ($T=1$). MC (left), QMC (middle) and SG (right).

that the LS approach with suitable design points can be a efficient way to approximate the solution.

5 Conclusions

In this work, the numerical comparison between different kinds of nodal points for the LS approach is considered. Three kinds of nodal points are introduced, namely, the Sparse Grid points, the Monte Carlo points and the Quasi Monte Carlo points. It is shown numerically that neither the MC sampling nor the QMC sampling introduce the low convergence rate, namely, the LS approach can achieve high order convergence rate for all the three kinds of nodal points. Furthermore, the use of SG points has better convergence properties only for very low dimensional problems (say 1-2). The QMC points seems to be a good choice for higher dimensional problems not only for better convergence properties but also in the stability point of view.

We remark that our setting here is quite standard. It would be more efficient if weighted LS approach is considered, and we will report such results in our forthcoming work [27]. Furthermore, one may also interested in testing other kinds of sparse grid points such as Clenshaw-Curtis. We also remark that only one kind of QMC points is tested, however, there are many other kinds of QMC points, and we would expect that all the low discrepancy QMC points work well in the LS approach.

Acknowledgments

The first author is supported by National Natural Science Foundation of China (11201441), the Natural Science Foundation of Shandong Province (ZR2012AQ003), and China Postdoctoral Science Foundation (2012M521374/2013T60684). The second author is supported by the National Natural Science Foundation of China (No. 91130003 and No. 11201461).

References

[1] G. Blatman and B. Sudret. Sparse polynomial chaos expansions and adaptive stochastic finite elements using regression approach. C.R.Mechanique, 336:518-523, 2008.
[2] Q.-Y. Chen, D. Gottlieb and J. Hesthaven, Uncertainty analysis for the steady-state flows in a dual throat nozzle, J. Comput. Phys., 204 (2005), pp. 387-398.
[3] H. Cheng and A. Sandu, Collocation least-squares polynomial chaos method. In Proceedings of the 2010 Spring Simulation Multiconference, pages 94-99, April, 2010.
[4] A. Cohen, M.A. Davenport, and D. Leviatan, On the stability and accuracy of Least Squares approximations, Found. Comput. Math., (2013) 13:819-834.
[5] A. Doostan and G. Iaccarino, A least-squares approximation of partial differential equations with high-dimensional random inputs, J. Comput. Phys., 228 (2009) 4332-4345.
[6] M. S. Eldred, Recent Advances in Non-Intrusive Polynomial Chaos and Stochastic Collocation Methods for Uncertainty Analysis and Design, Proceedings of the 11th AIAA Nondeterministic Approaches Conference, No. AIAA-2009-2274, Palm Springs, CA, May 4-7 2009.
[7] G. Fishman, Monte Carlo: Concepts, Algorithms, and Applications, Springer-Verlag, New York, 1996.
[8] Zhen Gao and Jan S. Hesthaven, Efficient solution of ordinary differential equations with high-dimensional parametrized uncertainty, Commun. Comput. Phys. 10, 253-278, 2011.
[9] T. Gardner, C. Cantor and J. Collins, Construction of a Genetic Toggle Switch in Escherichia coli, Nature 403, 339-342, 2000.
[10] R. Ghanem and P. Spanos, Stochastic Finite Elements: A Spectral Approach, Springer-Verlag, New York, 1991.
[11] John Hammersley, Monte Carlo methods for solving multivariable problems, Proceedings of the New York Academy of Science, Volume 86, 1960, pages 844-874.
[12] S. Hosder, R. W. Walters, and M. Balch, Efficient Sampling for Non-Intrusive Polynomial Chaos Applications with Multiple Uncertain Input Variables, Proceedings of the 48th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, No. AIAA-2007-1939, Honolulu, HI, April 23-26, 2007.
[13] S. Hosder, R. W. Walters, and M. Balch. Point-collocation nonintrusive polynomial chaos method for stochastic computational fluid dynamics. AIAA Journal, 48:2721-2730, 2010.
[14] Ladislav Kocis and William Whiten, Computational Investigations of Low-Discrepancy Sequences, ACM Transactions on Mathematical Software, Volume 23, Number 2, 1997, pages 266-294.
[15] G. Migliorati, F. Nobile, E. Schwerin, and R. Tempone, Analysis of the discrete L^2 projection on polynomial spaces with random evaluations, Foundations of Computational Mathematics, DOI:10.1007/s10208-013-9186-4, 2014.

[16] W. J. Morokoff and R. E. Caflisch, Quasi-random sequences and their discrepancies. SIAM J. Sci. Comput., 15(6):1251-1279, November 1994.
[17] F. Nobile, R. Tempone and C. Webster, A sparse grid stochastic collocation method for partial differential equations with random input data, SIAM J. Numer. Anal., 2008, vol. 46/5, pp. 2309–2345.
[18] F. Nobile, R. Tempone and C. Webster, An anisotropic sparse grid stochastic collocation method for partial differential equations with random input data, SIAM J. Numer. Anal., 2008, vol. 46/5, pp. 2411-2442.
[19] E. Novak and K. Ritter, High dimensional integration of smooth functions over cubes, Numer. Math., 75 (1996), pp. 79-97.
[20] E. Novak and K. Ritter, Simple cubature formulas with high polynomial exactness, Constructive Approx., 15 (1999), pp. 499-522.
[21] S. Smolyak, Quadrature and interpolation formulas for tensor products of certain classes of functions, Soviet Math. Dokl., 4 (1963), pp. 240-243.
[22] N. Wiener, The homogeneous chaos, Am. J. Math., 60 (1938), 897-936.
[23] D. Xiu, Efficient collocational approach for parametric uncertainty analysis, Commun. Comput. Phys., 2 (2007), 293-309.
[24] D. Xiu and J.S. Hesthaven, High-order collocation methods for differential equations with random inputs, SIAM J. Sci. Comput., 27 (2005), pp. 1118-1139.
[25] D. Xiu and G.E. Karniadakis, The Wiener-Askey polynomial chaos for stochastic differential equations. SIAM J. Sci. Comput., Vol. 24, No. 2, pp. 619-644.
[26] Z. Xu and T. Zhou, On sparse interpolation and the design of deterministic interpolation points, arxiv:1308.6038, Aug. 2013.
[27] T. Zhou, A. Narayan and Z. Xu, Multivariate discrete least-squares approximations with a new type of collocation grid, available online: `http://arxiv.org/abs/1401.0894`, Jan. 2014.
[28] T. Zhou and T. Tang, Galerkin Methods for Stochastic Hyperbolic Problems Using Bi-Orthogonal Polynomials, J. Sci. Comput., (2012)51:274-292.