Efficient estimation of polynomial chaos proxies using generalized sparse quadrature
M.J. Tompkins*, Schlumberger-EMI, and M. Prange, Schlumberger-Doll Research

Summary
We investigate the use of sparse grid methods in computing polynomial chaos (PC) proxies for forward stochastic problems associated with numerically-expensive simulators. These are problems where some input parameters are random with known distributions, and stochastic properties of the simulator output are desired. The bottleneck for PC proxy construction is the estimation of the coefficients, which typically require computationally intensive forward simulations and multi-dimensional integration. To minimize the number of simulations, we compare two methods for computing polynomial coefficients using sparse quadrature integration: generalized Fejér quadrature (FQ), and sparse reduced quadrature (RQ). We compare the efficiency (as determined by the number of quadrature points needed to accurately estimate coefficients) of these methods for a 5-dimensional stochastic electromagnetic problem. Paradoxically, we find that for general weight functions, sparse FQ requires very high degree exactness to accurately estimate proxy coefficients, which makes this scheme very inefficient. In contrast, RQ requires the minimum number of quadrature points for a pre-defined polynomial exactness. By using the sparse reduced quadrature approach, PC can apply to problems with arbitrary input PDFs and high-dimensional spaces. The trade-off is that sparse FQ has nested abscissae allowing for adaptive refinement of integration degree, while RQ does not.

Introduction
Recently, uncertainty quantification (UQ) has seen increasing interest in academic and industrial fields such as climate modeling, acoustics, hydrology, and decision analysis. This type of analysis is important, because no forward or inverse solution can be interpreted fully without an appreciation for its robustness or statistical significance. In the geosciences, quantitative uncertainty solutions have been implemented (Tompkins et al., 2011; Osypov et al., 2008; Sarma, 2006), but scalability is always an issue. This paper is an attempt to address scalability for a particular type of problem: forward stochastic problems whose input parameters have known random behavior. Such problems are commonly seen in the reservoir simulation literature (e.g., Li et al., 2009; Sarma, 2006) and to a lesser extent in geophysics (Tatang et al., 1997).

The topic of polynomial chaos has received some attention in the last twenty years as a means to efficiently estimate model outcomes based on known stochastic processes (e.g., Chauvière et al., 2006; Ghanem and Spanos, 1991). A PC proxy represents a stochastic process as a polynomial expansion, with corresponding expansion coefficients, over random variables of known distributions. This approach can exhibit very fast convergence if the observables depend smoothly on the random parameters. The corresponding qth degree polynomial chaos approximation for a function \( f(x) \) of \( d \) random variables \( x = (x_1, x_2, \ldots, x_d) \) with PDF \( \pi(x) \) is

\[
I_d^q(x) = \sum_{m=1}^{Q} a_m P_m(x), \quad Q = \left\{ \frac{q + d}{q} \right\},
\]

with expansion coefficients \( a_m \) and multivariate orthonormal polynomials \( P_m(x) \). Here, the summations are over all possible combinations of the multi-index \( i = (i_1, i_2, \ldots, i_d) \) such that \( \|i\| = \sum_{j=1}^{d} i_j \leq q \). The \( P_m(x) \) are selected such that

\[
\int_{\Omega} P_j(x) P_j(x) \pi(x) dx = \delta_{ij},
\]

where \( \delta_{ij} \) is the Kronecker delta. This orthonormality allows the expansion coefficients to be determined by

\[
a_m = \int_{\Omega} f(x) P_m(x) \pi(x) dx .
\]

The number of summands in (1) grows as

\[
dim(Q_d^q) = \left\{ \frac{q + d}{q} \right\} \sim \frac{d^q}{q!} \text{ for } d \gg 1 .
\]

For generalized polynomial chaos (gPC), polynomials are determined which best match the distribution of the non-Gaussian input random variables (Xiu and Karniadakis, 2003). Determining the expansion coefficients in (1) involves the potentially time-consuming task of evaluating sufficient forward problems, \( f(x) \), to accurately estimate these coefficients. However, once these coefficients have been computed, any future function estimations only require the evaluation of (1). Thus, the potentially large
Polynomial Chaos Using Generalized Sparse Quadrature

The number of forward-problem evaluations needed to quantify the distribution of \( f(x) \) via Monte Carlo sampling is replaced by the potentially much smaller number of evaluations needed to pre-compute the coefficients for the PC expansion, from which a Monte Carlo analysis can be inexpensively performed. In order for this to pay off, we need a low-degree expansion and accurate computation of the coefficients in (3) with minimal forward evaluations.

Generalized Sparse Fejér Quadrature

We need to have an efficient integration scheme, but standard tensor-product quadrature is not appropriate for higher-dimensions. First, we consider Smolyak’s sparse grid method (See Barthelmann et al., 2000; Smolyak, 1963), where well-established univariate integration formulas are extended to the multivariate case by using a subset of the complete tensor product set of abscissae. Ideally, we can perform accurate integrations that require orders of magnitude fewer forward model evaluations than for uniform multi-dimensional quadrature. For a given \( d \)-dimensional rule, \( A(k+d,d) \), the rule’s sparsity depends directly on the degree exactness, \( k \), required for the integration in (3):

\[
\dim(A(k+d,d)) \sim \frac{2^k}{k!} d^k d > 1 \quad (5)
\]

The number of points is much less than that given by full tensor products, \( \sim k^d \), when the degree of integration is small \( (k \ll d) \), which is why sparse grids are commonly used for problems of high stochastic dimension. Unfortunately, Smolyak rules grow exponentially with degree \( k \), so if high degrees are required for (3), these rules can be less sparse than their tensor product counterparts.

The 1D quadrature rule we consider here is Fejér Type-2 which is defined over the open interval \((-1,1)\). The abscissae are defined by the Chebyshev formula with endpoints removed (Waldvogel, 2003):

\[
Y_j = -\cos\left(\frac{\pi(j-1)}{m_t}\right), \quad j = 1, \ldots, m_t \quad (6)
\]

And the weights are explicitly defined as

\[
w_j = \frac{2}{m_t} \left[ 1 - \sum_{r=1}^{m_t} \frac{1}{2r+1} \cos\left( \frac{r(2j+1)\pi}{m_t} \right) \right],
\]

\[ j = 0, 1, \ldots, m_t - 1 \quad (7)
\]

The univariate rules, \( Y^i \), together with weights, \( W^i \), constitute the 1D quadrature that is used with the sparse grid scheme to generate the multivariate quadrature that will evaluate integrals of the form

\[
a_i = \frac{1}{2^d} \int_{-1}^{1} f(x)p_i(x)dx \quad (8)
\]

where the uniform weight function is taken as \( 1/2^d \). However, in a generalized case, the coefficient integral has the form of (3), requiring a generalization of the sparse grid scheme. This may be accomplished for each of 1D rule via a change of variable in terms of the cumulative distribution function

\[
c(x_j) = \int_{-\infty}^{x_j} \pi(\xi)d\xi, \quad (9)
\]

where \( \pi(x) \) is assumed to be separable in the form

\[
\pi(x) = \prod_{i=1}^{d} \pi_i(x_i) \quad (10)
\]

Since \( c(x) \in [0,1] \), defining \( y = 2c(x)-1 \) maps \( x \) into the Fejér domain \([-1,1]\). Using this change of variable to transform (3) into the form of a Fejér integral (noting that \( dy = 2\pi(x)dx \)),

\[
a_i = \frac{1}{2^d} \int_{-1}^{1} f(x(y))p_i(x(y))dy \quad (11)
\]

The Fejér points, \( Y^i \in (-1,1) \), are defined by (6) in the \( y \) domain and then mapped into the \( x \) domain before being used in the evaluation \( f(x) \) and \( p_i(x) \) in the quadrature rule. With this transformation, the weights associated with the Fejér quadrature rule remain the same as in (7), and are used in the evaluation of (11). Figure 1 provides an example of the 1D transformation of the quadrature abscissa for a lognormal(0,1) distribution, along with associated low-order polynomials.

Sparse Reduced Quadrature

Another approach to obtaining a sparse quadrature formula is due to Davis (1967) and Wilson (1969), who base their work on a theorem by Tchakaloff (1957) stating that if \( p_i(x), i \in \{1, \ldots, Q\} \), are linearly independent and continuous functions in \( \Omega \), of which one does not vanish in \( \Omega \), then one can find \( Q \) points \( x_j \) lying in \( \Omega \) with \( Q \) weights \( a_i \geq 0 \) such that:
Polynomial Chaos Using Generalized Sparse Quadrature

\[ \int_{\Omega} p_i(x) dx = \sum_{j=1}^{Q} a_j p_i(x_j). \] (12)

This means that a quadrature rule that exactly integrates polynomials up to total degree \( q \) in dimension \( d \) requires only \( Q = \frac{(q + d)}{q} \) quadrature points. The Wilson (1969) algorithm can be applied to our case as follows:

1) Form the set of polynomials of total degree \( q \)
   as \( F = \{ P_0(x), ..., P_Q(x) \} \).

2) Form the vector \( c_i = \int_{\Omega} P_i(x) \pi(x) dx \). Note that since the \( P_i(x) \) are orthonormal with respect to the weight function \( \pi \) and \( P_0 = 1 \), then \( c_i = [1, 0, ..., 0] \).

3) Form matrix \( A \) using \( A_{ij} = P_i(x_j) \), \( i = 1, ..., Q, j = 1, ..., n \).

By choosing the \( n \) quadrature points \( x_j \) as the Gauss quadrature tensor product rule appropriate for \( \pi \), we have \( A \lambda = c \), where \( \lambda \) is the vector of Gauss quadrature weights.

This is because the original rule exactly integrates the reduced set of polynomials. Note that \( n \) is much larger than \( Q \), so the goal of the next two steps is to reduce the quadrature rule to at most \( Q \) points.

4) Use a linear programming algorithm to solve:

\[ \min \lambda \cdot 1, \lambda \cdot \pi \quad \text{s.t.} \quad A \lambda = c, \lambda \geq 0. \] (13)

Since \( A \lambda = c \) is underdetermined, we can expect as many zeros in the solution \( \lambda \) as there are dimensions to the nullspace of \( A \).

5) If the resulting \( \lambda \) has any zeros, the corresponding superfluous quadrature points can be removed from the quadrature rule. The non-zero \( \lambda \) values are the new quadrature weights corresponding to the remaining quadrature abscissa.

5D Electromagnetic Example

We now test the two methods using a 5-layer resistivity earth model, shown in Figure 2, and simulating the 1D electromagnetic response function based on a semi-analytical solution for dipoles in transversely isotropic 5-Layer Earth Model

![Figure 2: 5-layer resistivity earth model used for simulating EM responses.](image)

stratified media (Tompkins, 2003). For the data, we simulated a surface controlled-source electric field experiment using a horizontal dipole source at \((x,y,z) = (0,0,0)\) and a horizontal surface receiver array Rx(x) = [0.5, 3, 8, and 12]km. The horizontally polarized electric field data were simulated at 0.25Hz, and amplitudes were computed from real and imaginary components. The lognormal probability distributions for the 5 layer resistivities are shown in Figure 3 (\((1:5) = [10, 1, 5, 1, 100]\) and \((\alpha(1:5) = [4.47, 1.00, 4.47, 1.00, 44.87])\). For comparison with our PC results, E-field amplitude response histograms for a quasi-random sampling of 50,000 points drawn from the prior resistivity distributions are shown in Figures 4 and 5.

Sparse Fejér Quadrature

We first test the efficiency of our sparse FQ rule for our 5D EM problem. From Figure 4, it is clear that the EM responses are nonlinear and that the degree of nonlinearity changes with receiver position (Tx-Rx offset). This is a well-known property of low-frequency electric fields propagating in lossy media. Stated simply, we expect the near offset (<4km or so in this case) amplitudes to fall off \( \propto 1/\text{distance}^3 \), while the long offset amplitudes will follow a less severe decay. The implication of this is that the near offset EM responses will require very high degree integrations (at least 2x2x2=8) even for a degree-2 PC proxy (because we must integrate the function over the non-polynomial weights \( \pi(x) \)). The troubling point is that this causes the sparse FQ in 5-D to actually involve many more points than the equivalent full tensor product grids (187,903 vs. 59,049).
Polynomial Chaos Using Generalized Sparse Quadrature

Therefore, we test the 5D EM amplitude response estimation using a degree-2 proxy and only a degree-6 sparse FQ integration (18,943 points). The distributions estimated from the PC proxy at 4 offsets are shown overlain on Monte-Carlo results in Figure 4. As expected, the proxy provides poor estimates of the distributions at most offsets (with some improvement for larger offsets) given that the degree of integration is too low.

Figure 4: Comparison of Monte-Carlo (bold) 1D E-Field amplitude response histograms (10,000 samples) for the 5-parameter stochastic problem and degree-2 PC proxy estimates (dashed) performed using degree-6 sparse FQ for 18,943 points.

While more accurate estimations are possible for degree-8 integration, the number of points required is prohibitive (187,903) even in 5D and far more than that required for the Monte-Carlo method.

Sparse Reduced Quadrature

Here we repeat the PC proxy estimation only now we use the sparse reduced quadrature method for the same degree-2 proxy, which requires at most a degree-4 quadrature rule (because we optimize for the weights). When applying the reduced quadrature method to a Gauss-Quadrature rule, we found that of the 3,125 points used in the full tensor-product 5-D rule, only 135 points are needed when integrating to maximum of degree-4 (tensor-product quadrature “over” integrates in certain projections of the 5D space). This means that our integration requires 135 forward simulations as opposed to the 18,943 required for degree-6 sparse FQ.

The results are presented in Figure 5 along with the Monte-Carlo distributions. It is clear that these results are not only more accurate than the FQ, but are far more efficient. Again, the one drawback is that these RQ rules cannot be reused to increase the accuracy of the integration in (3) if desired. In the case that we have to increase our exactness, a complete new multi-dimensional rule would be required with corresponding forward simulations.

Figure 5: Comparison of Monte-Carlo (bold) 1D E-Field amplitude response distribution (10,000 samples) for our 5-D stochastic problem with degree-2 PC proxy estimates (dashed) performed using degree-4 sparse RQ for 135 points.

Conclusions

We have presented two methods for the explicit estimation of polynomial chaos proxy coefficients for arbitrary PDFs and arbitrary support using sparse quadrature integration. By comparing standard Smolyak sparse quadrature, using generalized Fejér Type-2 rules, with Monte-Carlo integration, we find that this method is far too inefficient when integrating in high degrees. As an alternative, we present a sparse reduced quadrature method that is more than two orders of magnitude more efficient in our 5D example without loss of accuracy.
EDITED REFERENCES
Note: This reference list is a copy-edited version of the reference list submitted by the author. Reference lists for the 2012 SEG Technical Program Expanded Abstracts have been copy edited so that references provided with the online metadata for each paper will achieve a high degree of linking to cited sources that appear on the Web.

REFERENCES


