MATHICSE
Mathematics Institute of Computational Science and Engineering
School of Basic Sciences - Section of Mathematics

MATHICSE Technical Report
Nr. 41.2014
September 2014

Comparison of Clenshaw-Curtis and Leja quasi-optimal sparse grids for the approximation of random PDEs

Fabio Nobile, Lorenzo Tamellini, Raúl Tempone

http://mathicse.epfl.ch
Phone: +41 21 69 37648
Address:
EPFL - SB - MATHICSE (Bâtiment MA)
Station 8 - CH-1015 - Lausanne - Switzerland
Fax: +41 21 69 32545

Comparison of Clenshaw–Curtis and Leja quasi-optimal sparse grids for the approximation of random PDEs

Fabio Nobile, Lorenzo Tamellini, Raul Tempone

Abstract In this work we compare numerically different families of nested quadrature points, i.e. the classic Clenshaw–Curtis and various kinds of Leja points, in the context of the quasi-optimal sparse grid approximation of random elliptic PDEs. Numerical evidence suggests that the performances of both families are essentially comparable within such framework.
Key words: Uncertainty Quantification; PDEs with random data; linear elliptic equations; Stochastic Collocation methods; Sparse grids approximation; Leja points; Clenshaw–Curtis points.
AMS Subject Classification: 41A10, 65C20, 65N12, 65N30, 65N35.

1 Introduction

While it is nowadays widely acknowledged that Uncertainty Quantification problems can be conveniently tackled with polynomial approximation schemes whenever the output quantities of interest depend smoothly on a moderate number of random parameters, the search for algorithms whose performance is resilient with respect to the number of such random parameters is a very active research area.

Fabio Nobile
CSQI - MATHICSE, Ecole Politechnique Fédérale Lausanne, Switzerland, e-mail: fabio.nobile@epfl.ch

Lorenzo Tamellini
CSQI - MATHICSE, Ecole Politechnique Fédérale Lausanne, Switzerland, e-mail: lorenzo.tamellini@epfl.ch

Raul Tempone
Applied Mathematics and Computational Science, King Abdullah University of Science and Technology, Saudi Arabia, e-mail: raul.tempone@kaust.edu.sa

In the context of sparse grids approximation [1, 4, 13], this has lead on the one hand to the development of more efficient sparse grid algorithms, which exploit the anisotropic structure of the problem at hand (either via an "a-priori" analysis, see e.g. [2, 3, 11], or with an "a-posteriori" adaptation, see [6, 8, 12]), and on the other hand to the study of appropriate univariate collocation points to be used as a basis for the sparse grids construction.

To maximize the efficiency of the sparse grids, such collocation points are typically chosen to be nested. Clenshaw–Curtis points are a classical choice in this sense; more recently, an increasing attention has been devoted to the study of the performance of the so-called Leja points (see [5, 6, 10, 12]), which are promising since the cardinality of Leja quadrature rules grows slower than that of Clenshaw–Curtis rules when increasing the approximation level. In the literature, Leja points have only been applied to "a-posteriori" adaptive sparse grids [6, 12]: the aim of this work is to test their performance in the context of the quasi-optimal "a-priori/a-posteriori" sparse grids that we have proposed in a series of previous papers [2, 3, 11], focusing on the case of elliptic PDEs with diffusion coefficients parametrized by uniform random variables.

The rest of this work is organized as follows. The general problem setting will be introduced in Section 2, and quasi-optimal sparse grids in Section 3. Clenshaw–Curtis and Leja points will be discussed in Section 4, while numerical tests and some conclusions will be presented in Section 5.

2 Problem setting

Let $N \in \mathbb{N}$ and $\Gamma \subset \mathbb{R}^N$ be an N-variate hyper-rectangle $\Gamma = \Gamma_1 \times \ldots \times \Gamma_N$, and assume that each Γ_n is endowed with a uniform probability measure $\varrho_n(y_n)dy_n = \frac{1}{|\Gamma_n|}dy_n$, so that $\varrho(\mathbf{y})d\mathbf{y} = \prod_{n=1}^N \varrho_n(y_n)dy_n$ is a uniform probability measure on Γ and $(\Gamma, B(\Gamma), \varrho(\mathbf{y})d\mathbf{y})$ is a probability space, $B(\Gamma)$ being the Borel σ-algebra on Γ. Given a convex polygonal domain D in $\mathbb{R}^d$, $d = 1, 2, 3$, we consider the following problem:

Problem 1. Find a real-valued function $u : \overline{D} \times \Gamma \to \mathbb{R}$, such that $\varrho(\mathbf{y})d\mathbf{y}$-almost everywhere there holds:

$$\begin{cases} -\operatorname{div}(a(\mathbf{x},\mathbf{y})\nabla u(\mathbf{x},\mathbf{y})) = f(\mathbf{x}) & \mathbf{x} \in D, \\ u(\mathbf{x},\mathbf{y}) = 0 & \mathbf{x} \in \partial D, \end{cases}$$

where the operators div and ∇ imply differentiation with respect to the physical coordinate only, and $a : \overline{D} \times \Gamma \to \mathbb{R}$ is such that

$$0 < a_{min} \leq a(\mathbf{x},\mathbf{y}) \leq a_{max} < \infty \qquad (1)$$

for some positive and bounded constants a_{min}, a_{max}.

By introducing the Hilbert space $V = H_0^1(D)$, the above problem is shown to be well-posed in the Bochner space $L^2_\varrho(\Gamma; V) = \left\{ u : \Gamma \to V \text{ s.t. } \int_\Gamma \|u(\mathbf{y})\|_V^2 \, \varrho(\mathbf{y}) d\mathbf{y} < \infty \right\}$, due to the boundedness assumption (1). Moreover, under additional assumptions on a (e.g. $\mathbf{y}$-linearity or mild assumptions on the growth of its $\mathbf{y}$-derivatives), it can be shown that the map $\mathbf{y} \to u(\cdot, \mathbf{y})$ is analytic, see e.g. [2, 7].

3 Quasi-optimal sparse grid approximation

Let $\mathbb{P}_r(\Gamma_n)$ be the set of polynomials of degree at most r over Γ_n, $C^0(\Gamma_n)$ the set of continuous functions over Γ_n, and for a given interpolation level i_n let $\mathcal{U}_n^{m(i_n)} : C^0(\Gamma_n) \to \mathbb{P}_{m(i_n)-1}(\Gamma_n)$ be the lagrangian interpolant operator over $m(i_n)$ points, with $m : \mathbb{N} \to \mathbb{N}$ a non-decreasing function, the so-called "level-to-nodes" function. Next, for any multi-index with non-zero components $\mathbf{i} \in \mathbb{N}_+^N$ let us define the "hierarchical surplus" operator $\boldsymbol{\Delta}^{m(\mathbf{i})} = \bigotimes_{n=1}^N (\mathcal{U}_n^{m(i_n)} - \mathcal{U}_n^{m(i_n-1)})$, and let $\{\mathcal{I}(\mathrm{w})\}_{\mathrm{w}\in\mathbb{N}}$ denote a nested sequence of index sets with non-zero components, $\mathcal{I}(\mathrm{w}) \subset \mathbb{N}_+^N$, with $\mathcal{I}(0) = [1, 1, \dots, 1]$ and $\bigcup_{\mathrm{w}\in\mathbb{N}} \mathcal{I}(\mathrm{w}) = \mathbb{N}_+^N$. The sparse grid approximation of u is then written as

$$\mathcal{S}^m_{\mathcal{I}(\mathrm{w})}[u](\mathbf{y}) = \sum_{\mathbf{i}\in\mathcal{I}(\mathrm{w})} \boldsymbol{\Delta}^{m(\mathbf{i})}[u](\mathbf{y}), \tag{2}$$

where one usually requires the sets $\mathcal{I}(\mathrm{w})$ to be *lower sets*[1], see e.g. [8]. In practice, to build a sparse grid one has to specify (i) the family of interpolation nodes, that should be chosen according to the probability measure over Γ (as previously mentioned, in this work we will use Leja and Clenshaw–Curtis points, which are suitable for uniform measures), (ii) the function $m(\cdot)$, and (iii) the sequence of index sets $\mathcal{I}(\mathrm{w})$.

To detail the choice of the sequence $\mathcal{I}(\mathrm{w})$, let us now denote by $\Delta E(\mathbf{i})$ the error decrease obtained by adding a given hierarchical surplus $\boldsymbol{\Delta}^{m(\mathbf{i})}$ to the sparse grid approximation of u and by $\Delta W(\mathbf{i})$ its associated cost, i.e. the number of interpolation points added to the sparse grid by $\boldsymbol{\Delta}^{m(\mathbf{i})}$, and let us define the *profit* $P(\mathbf{i})$ of each $\boldsymbol{\Delta}^{m(\mathbf{i})}$ as the ratio $P(\mathbf{i}) = \frac{\Delta E(\mathbf{i})}{\Delta W(\mathbf{i})}$. The optimal sequence $\mathcal{I}(\mathrm{w})$ should then progressively add to the sparse grid approximation of u the hierarchical surpluses $\boldsymbol{\Delta}^{m(\mathbf{i})}$ ordered by decreasing profits, see [2, 8, 9, 11],

$$\mathcal{I}(\mathrm{w}) = \left\{ \mathbf{i} \in \mathbb{N}_+^N : P(\mathbf{i}) \geq \epsilon_{\mathrm{w}} \right\}, \tag{3}$$

[1] Also known as *admissible sets* or *downward closed sets*, i.e. such that $\forall\, \mathbf{i} \in \mathcal{I}(\mathrm{w})$ and $\forall\, \mathbf{j} \in \mathbb{N}_+^N$ s.t. $\mathbf{j} \leq \mathbf{i}$, there holds $\mathbf{j} \in \mathcal{I}(\mathrm{w})$, where the inequality is to be understood component-wise.

with $\{\epsilon_{\mathrm{w}}\}_{\mathrm{w}\in\mathbb{N}}$ positive sequence decreasing to 0. Note that $\mathcal{I}(\mathrm{w})$ in (3) may not be a lower set, and this condition will have to be explicitly enforced.

The above criterion (3) can be implemented either by an "a-posteriori" adaptive procedure (see e.g. [6, 8, 12]) that explores the space of hierarchical surpluses and adds to $\mathcal{I}(\mathrm{w})$ the most profitable one, or, as we have previously detailed in [2, 3, 11], with a procedure based on a-priori estimates of $\Delta E(\mathbf{i})$ and $\Delta W(\mathbf{i})$, tuned to the problem at hand by some cheap preliminary computations ("a-priori/a-posteriori" approach); in this work, we consider this latter approach. Of course, if on the one hand the "a-priori/a-posteriori" approach saves the computational cost of the exploration of the space of hierarchical surpluses, on the other hand it will be effective only if the estimates of $\Delta E(\mathbf{i})$ and $\Delta W(\mathbf{i})$ are sufficiently sharp.

The work contribution $\Delta W(\mathbf{i})$ can actually be computed exactly if the points used in the sparse grid construction are nested (as it is the case in this work) and $\mathcal{I}(\mathrm{w})$ is a lower set:

$$\Delta W(\mathbf{i}) = \prod_{n=1}^{N} (m(i_n) - m(i_n - 1)). \tag{4}$$

As for the error contribution $\Delta E(\mathbf{i})$, we propose to use certain problem-dependent estimates, that we will specify later on.

4 Leja and Clenshaw–Curtis quadrature rules

A Leja sequence on a generic compact set X is defined recursively, by first choosing $x_1 \in X$ and then letting $x_n = \mathrm{argmin}_{x\in X} \prod_{k=1}^{n-1}(x - x_k)$, see e.g. [5, 6, 10, 12], while the corresponding quadrature weights are computed by enforcing the maximal degree of polynomial exactness. More specifically, we will consider the following families of Leja points:

Line Leja: Let $X = [-1, 1]$ and $x_1 = -1$. Then $x_2 = 1$, $x_3 = 0$, and $x_n = \mathrm{argmin}_{(-1,1)} \prod_{k=1}^{n-1}(x - x_k)$.

Sym-Line Leja: Let $x_1 = 0, x_2 = 1, x_3 = -1, x_n = \mathrm{argmin}_{(-1,1)} \prod_{k=1}^{n-1}(x-x_k)$ for n even, and x_{n+1} be the symmetric point of x_n with respect to 0. Observe that this is *not* a Leja sequence according to the definition above.

P-Disk Leja: Let $x_k = \cos\phi_k$, with $\phi_1 = 0$, $\phi_2 = \pi$, $\phi_3 = \pi/2$, $\phi_{2k+2} = \frac{\phi_{k+2}}{2}$, and $\phi_{2k+3} = \phi_{2k+2} + \pi$. These points correspond to the projection on the real axis (with no repetitions) of the Leja sequence obtained with $x_1 = 1$ and X the complex unit ball (see [5]), and are *not* a Leja sequence.

We will test the Leja families above with two different level-to-nodes functions, i.e. $m_s(i_n) = i_n$ and $m_t(i_n) = 2(i_n - 1) + 1$. This latter "two-stepping" rule has been introduced in the adaptive context (see e.g. [12]), where the error contributions $\Delta E(\mathbf{i})$ are estimated via successive differences of the integral of u (or an approximation of u by e.g. finite elements) over the parameter space: indeed, observe that whenever one point is added to a symmetric quadrature rule, the corresponding quadrature weight will be zero, by symmetry; hence if one were using the "single-stepping" rule $m_s(i_n)$, two consecutive integrals may be equal (up to numerical noise) and the algorithm might prematurely stop.

Finally, the Clenshaw–Curtis points (cf. e.g. [11]) are defined as

$$x_j = \cos\left(\frac{(j-1)\pi}{m(i_n)-1}\right), 1 \le j \le m(i_n),$$

together with the following level-to-nodes relation $m_d(i_n)$, that ensures their nestedness[2]: $m_d(0) = 0$, $m_d(1) = 1$, $m_d(i_n) = 2^{i_n - 1} + 1$. Observe $m_d(i_n)$ grows exponentially in i_n, while $m_s(i_n)$ and $m_t(i_n)$ grow linearly; quoting [10], we say that Leja points have a much finer "granularity".

5 Numerical tests

In this section we consider two different examples of Problem 1; in both cases, we will introduce a bounded linear functional $\Theta : V \to \mathbb{R}$, and monitor the convergence of the quantity

$$\varepsilon = \sqrt{\mathbb{E}\left[\left(\Theta(\mathcal{S}^m_{\mathcal{I}(\mathrm{w})}[u]) - \Theta(u)\right)^2\right]}, \tag{5}$$

with respect to the number of sparse grid points, that will converge with the same rate as the full error $\mathbb{E}\left[\left(\mathcal{S}^m_{\mathcal{I}(\mathrm{w})}[u] - u\right)^2\right]^{1/2}$, given the linearity of Θ. In practice, we have estimated (5) with a Monte Carlo sampling (see Figure 1 for the sample size for each test); we underline that the sample sizes have been verified to be sufficient for our purposes.

In the first test, we consider $\Gamma_n = [-1, 1]$, $D = (0, 1)$ and two different expressions of $a(\mathbf{x}, \mathbf{y})$, both complying with equation (1), that is $a_1(x, \mathbf{y}) = 4 + y_1 + 0.2\sin(\pi x)y_2 + 0.04\sin(2\pi x)y_3 + 0.008\sin(3\pi x)y_4$, and $a_2(\mathbf{x}, \mathbf{y}) = \exp\left(a_1(x, \mathbf{y})\right)$. We also set $f(\mathbf{x}) = 1$ and $\Theta(u) = u(0.7)$. For this case, the estimate for the error contribution $\Delta E(\mathbf{i})$ in (3) is (cf. [2])

$$\Delta E(\mathbf{i}) \le Ce^{-\sum_{n=1}^N g_n m(i_n - 1)} \left(\prod_{n=1}^N \mathbb{L}_n^{m(i_n)}\right) \frac{|m(\mathbf{i})|!}{m(\mathbf{i})!},$$

[2] When $2^m + 1$ p-Disk Leja points are computed, they coincide with the Clenshaw–Curtis points.

where C is a positive constant, $\mathbb{L}_n^{m(i_n)}$ is the Lebesgue constant associated to the interpolation scheme $\mathcal{U}_n^{m(i_n)}$ that can either be computed numerically or estimated a-priori (cf. [11]), $m(\mathbf{i})! = \prod_n m(i_n)!$, $|m(\mathbf{i})| = \left(\sum_n m(i_n)\right)!$, and g_n can be tuned with cheap preliminary computations, see e.g. [2].

In the second test, we consider instead $\Gamma_n = [-0.99, 0.99]$, $D = [0,1]^2$ and $a(\mathbf{x},\mathbf{y}) = 1 + \sum_{n=1}^N \gamma_n \chi_n(\mathbf{x}) y_n$, for $N = 4, 8$. Here $\chi_n(\mathbf{x})$ are the indicator functions of the disjoint circular sub-domains $D_n \subset D$ as in Figure 1, and γ_n are real coefficients such that (1) holds true; more specifically, we consider both an isotropic setting, $\gamma_n = 1$ for each subdomain, and an anisotropic setting, see Figure 1 for the values of γ_n in this latter setting. Finally, we set $f(\mathbf{x}) = 100\chi_F(\mathbf{x})$ and $\Theta(u) = \int_F u(\mathbf{x}) d\mathbf{x}$. In this case, the estimate for the error contribution $\Delta E(\mathbf{i})$ in (3) is

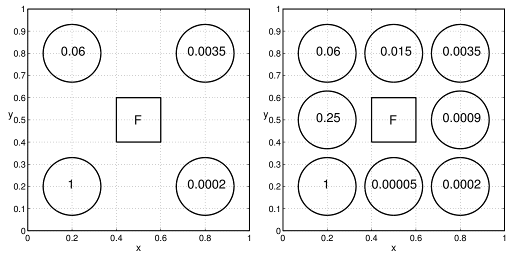
$$\Delta E(\mathbf{i}) = C e^{-\sum_{n=1}^N g_n m(i_n - 1)} \left(\prod_{n=1}^N \mathbb{L}_n^{m(i_n)} \right),$$

see [11], where we also provide a convergence estimate for the resulting sparse grid.

Numerical results are shown in Figure 2. It can be seen that sym-line Leja points with "two-stepping" seems to have the same (or slightly better) performance than Clenshaw–Curtis points, while the other families present non-negligible improvements in some cases but underperform in other tests. Additional tests carried on monitoring the quadrature error for $\Theta(u)$ rather than the interpolation error (5) (see Figure 3), show again that the performance of sym-Leja points with two-stepping is comparable to that of Clenshaw–Curtis, while other families this time always show a slight performance deterioration. This is likely due to the fact that Leja points are designed to minimize the Lebesgue constant, hence more suited for interpolation than for quadrature.

In conclusion, these tests seem to suggest that Leja points do not exhibit significative advantages over Clenshaw–Curtis points in the framework of the quasi-optimal sparse grids; also, in both cases the resulting sparse grids do not suffer from an excessive increase of points as we raise the level (at least for moderate dimensions) despite the improved granularity of Leja points, likely due to the fact the quasi-optimal construction adds only one or few hierarchical surpluses per level.

Acknowledgements F. Nobile and L. Tamellini have been partially supported by the Swiss National Science Foundation under the Project No. 140574 "Efficient numerical methods for flow and transport phenomena in heterogeneous random porous media" and by the Center for ADvanced MOdeling Science (CADMOS). R. Tempone is a member of the KAUST SRI Center for Uncertainty Quantification in Computational Science and Engineering.



Test	**Sample size**
Test 1, case a_1	1000
Test 1, case a_2	1000
Test 2, $N = 4$, iso.	5000
Test 2, $N = 4$, aniso.	5000
Test 2, $N = 8$, iso.	50000
Test 2, $N = 8$, aniso.	25000

Fig. 1: Left: domains for test 2 with $N = 4$ and $N = 8$, with values of the coefficients γ_n for the anisotropic settings. Right: sample size for the Monte Carlo estimate of (5) for each test.

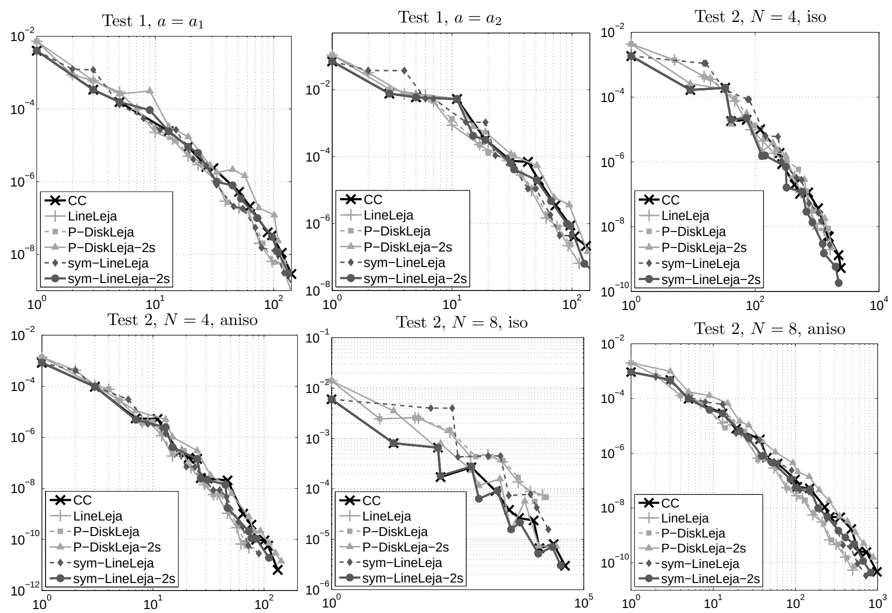


Fig. 2: Convergence of error (5) vs. sparse grids cardinality. The suffix "2s" refers to the "two-stepping" function for Leja points.

References

1. Babuška, I., Nobile, F., Tempone, R.: A stochastic collocation method for elliptic partial differential equations with random input data. SIAM Review **52**(2), 317–355 (2010)
2. Beck, J., Nobile, F., Tamellini, L., Tempone, R.: On the optimal polynomial approximation of stochastic PDEs by Galerkin and collocation methods. Mathematical Models and Methods in Applied Sciences **22**(09) (2012)
3. Beck, J., Nobile, F., Tamellini, L., Tempone, R.: A Quasi-optimal Sparse Grids Procedure for Groundwater Flows. In: Spectral and High Order Methods for Partial

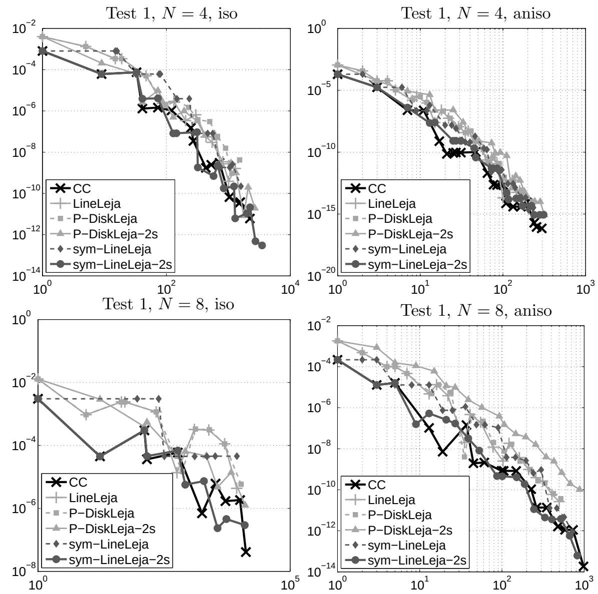


Fig. 3: Convergence of quadrature error for $\Theta(u)$ vs. sparse grids cardinality. The suffix "2s" refers to the "two-stepping" function for Leja points.

Differential Equations - ICOSAHOM 2012, *Lecture Notes in Computational Science and Engineering*, vol. 95, pp. 1–16. Springer (2014)
4. Bungartz, H., Griebel, M.: Sparse grids. Acta Numer. **13**, 147–269 (2004)
5. Chkifa, A.: On the Lebesgue constant of Leja sequences for the complex unit disk and of their real projection. Journal of Approximation Theory **166**(0), 176 – 200 (2013)
6. Chkifa, A., Cohen, A., Schwab, C.: High-dimensional adaptive sparse polynomial interpolation and applications to parametric pdes. Foundations of Computational Mathematics pp. 1–33 (2013)
7. Cohen, A., Devore, R., Schwab, C.: Analytic regularity and polynomial approximation of parametric and stochastic elliptic PDE'S. Anal. Appl. (Singap.) **9**(1), 11–47 (2011)
8. Gerstner, T., Griebel, M.: Dimension-adaptive tensor-product quadrature. Computing **71**(1), 65–87 (2003)
9. Griebel, M., Knapek, S.: Optimized general sparse grid approximation spaces for operator equations. Math. Comp. **78**(268), 2223–2257 (2009)
10. Narayan, A., Jakeman, J.: Adaptive Leja sparse grid constructions for stochastic collocation and high-dimensional approximation. arXiv arXiv:1404.5663, e-print (2014)
11. Nobile, F., Tamellini, L., Tempone, R.: Convergence of quasi-optimal sparse grids approximation of Hilbert-valued functions: application to random elliptic PDEs. Mathicse report 12/2014, EPFL (2014). Submitted
12. Schillings, C., Schwab, C.: Sparse, adaptive Smolyak quadratures for Bayesian inverse problems. Inverse Problems **29**(6) (2013)
13. Xiu, D., Hesthaven, J.: High-order collocation methods for differential equations with random inputs. SIAM J. Sci. Comput. **27**(3), 1118–1139 (2005)

Recent publications:

MATHEMATICS INSTITUTE OF COMPUTATIONAL SCIENCE AND ENGINEERING
Section of Mathematics
Ecole Polytechnique Fédérale
CH-1015 Lausanne

29.2014 ASSYR ABDULLE, PATRICK HENNING:
Localized orthogonal decomposition method for the wave equation with a continuum of scales

30.2014 DANIEL KRESSNER, ANDRÉ USCHMAJEW:
On low-rank approximability of solutions to high-dimensional operator equations and eigenvalue problems

31.2014 ASSYR ABDULLE, MARTIN HUBER:
Finite element heterogeneous multiscale method for nonlinear monotone parabolic homogenization problems

32.2014 ASSYR ABDULLE, MARTIN HUBER, GILLES VILMART:
Linearized numerical homogenization method for nonlinear monotone parabolic multiscale problems

33.2014 MARCO DISCACCIATI, PAOLA GERVASIO, ALFIO QUARTERONI:
Interface control domain decomposition (ICDD) methods for heterogeneous problems

34.2014 ANDRE USCHMAJEW:
A new convergence proof for the high-order power method and generalizations

35.2014 ASSYR ABDULLE, ONDREJ BUDÁC:
A Petrov-Galerkin reduced basis approximation of the Stokes equation in parametrized geometries

36.2014 ASSYR ABDULLE, MARTIN E. HUBER:
Error estimates for finite element approximations of nonlinear monotone elliptic problems with application to numerical homogenization

37.2014 LARS KARLSSON, DANIEL KRESSNER, ANDRÉ USCHMAJEW:
Parallel algorithms for tensor completion in the CP format

38.2014 PAOLO TRICERRI, LUCA DEDÈ, SIMONE DEPARIS, ALFIO QUARTERONI, ANNE M. ROBERTSON, ADÉLIA SEQUEIRA:
Fluid-structure interaction siumlations of cerebral arteries modeled by isotropic and anistropic constitutive laws

39.2014 FRANCESCA BONIZZONI, FABIO NOBILE, DANIEL KRESSNER:
Tensor train approximation of moment equations for the log-normal Darcy problem

40.2014 DIANE GUIGNARD, FABIO NOBILE, MARCO PICASSO:
A posteriori error estimations for elliptic partial differential equations with small uncertainties

41.2014 FABIO NOBILE, LORENZO TAMELLINI, RAÚL TEMPONE:
Comparison of Clenshaw-Curtis and Leja quasi-optimal sparse grids for the approximation of random PDEs