Abstract

In this paper, we propose and analyse a new adaptive multilevel stochastic collocation method for randomized elliptic PDEs. A hierarchical sequence of adaptive mesh refinements for the spatial approximation is combined with adaptive anisotropic sparse Smolyak grids in the stochastic space in such a way as to minimize computational cost. We provide a rigorous analysis for the convergence and computational complexity of the adaptive multilevel algorithm. Two numerical examples demonstrate the reliability of an error control by adaptive methods and the significant decrease in complexity versus uniform spatial refinements and single-level stochastic sampling methods.

Keywords. multilevel methods, hierarchical methods, adaptivity, stochastic collocation, PDEs with random input data, sparse grids, uncertainty quantification, high-dimensional approximation

AMS subject classifications. 65C20, 65C30, 65N35, 65M75

1 Introduction

Multilevel methods have made their way into Monte Carlo and stochastic collocation methods [4, 20]. They were originally proposed by Giles [13]. Recently, also combinations with spatial adaptivity have been investigated [5, 7, 8, 14, 18]. Here, we formulate first ideas to include
rigorous error control in both the solution of the physical PDE and the stochastic collocation method. We will follow the paper by Teckentrup, Jantsch, Webster, and Gunzburger \[20\] and concentrate on elliptic partial differential equations with random input data in the standard setting: Find \( u(x, y) : D \times \Gamma \rightarrow \mathbb{R} \) such that almost surely

\[
A(x, y) u(x, y) = f(x, y), \quad (x, y) \in D \times \Gamma, \\
u(x, y) = g(x, y), \quad (x, y) \in \partial D \times \Gamma,
\]

where \( \Gamma = \prod_{n=1}^{N} \Gamma_n \) with bounded \( \Gamma_n \) (boundedness assumption) is a stochastic parameter space of finite dimension \( N \) (finite noise assumption) and \( D \subset \mathbb{R}^d, d = 1, 2, 3, \) is the deterministic (physical) space. The random variables \( y_1, \ldots, y_n \) have a joint probability density function \( \rho(y) = \prod_{n=1}^{N} \hat{\rho}_n \in L^\infty(\Gamma), \hat{\rho}_n : [-1, 1] \rightarrow \mathbb{R} \). We assume that problem (1)-(2) admits a unique solution \( u(x, y) \in L^2_{\rho}(\Gamma; H^1(D)) \) in the weighted Bochner space

\[
L^2_{\rho}(\Gamma; H^1(D)) = \{ v : \Gamma \rightarrow H^1(D) \text{ measurable: } \int_{\Gamma} \| v(y, \cdot) \|_{H^1(D)}^2 \rho(y) dy < \infty \} (3)
\]

with corresponding norm

\[
\| v \|_{L^2_{\rho}(\Gamma; H^1(D))}^2 = \int_{\Gamma} \| v(y, \cdot) \|_{H^1(D)}^2 \rho(y) dy =: E \left[ \| v(y, \cdot) \|_{H^1(D)}^2 \right]. (4)
\]

The standard, single-level stochastic collocation method uses a set of sampling points \( \{y_m\}_{m=1,\ldots,M} \) in \( \Gamma \) and finite-dimensional spatial approximations \( u_h(y) \in V_h \subset H^1(D) \) to construct an interpolant

\[
u^{(SL)}_{M,h}(x, y) = \mathcal{I}[u_h](x, y) = \sum_{m=1}^{M} u_m(x) \phi_m(y), (5)\]

in the polynomial space \( P_M = \text{span}\{\phi_m\}_{m=1,\ldots,M} \subset L^2_{\rho}(\Gamma) \) with basis functions \( \{\phi_m\}_{m=1,\ldots,M} \). The coefficients \( u_m(x) \) are determined by the interpolating condition \( \mathcal{I}[u_h](x, y_m) = u_h(x, y_m) \) for \( m = 1, \ldots, M \). The quality of the interpolation process depends on the accuracy of the spatial approximations \( u_h(y) \) and the number of collocation points \( M \), which in practice can quickly grow with increasing stochastic dimension \( N \). Multilevel methods aim at reducing the overall computational cost through exploiting hierarchies of different resolutions in both spatial and stochastic approximations. Examples are uniform mesh refinement and generalized sparse grids as used in \[20\]. In what follows, we will combine an adaptive mesh refinement for the approximation of the spatial approximations \( u_h(y) \) with an adaptive (anisotropic) sparse Smolyak grid in order to improve the multilevel method and to reach a user-prescribed tolerance for the accuracy of the multilevel interpolant.

2 Adaptive Spatial Approximation

Let \( \{TolX_k\}_{k=0,\ldots,K} \) be a decreasing sequence of tolerances with

\[
1 \geq TolX_0 > TolX_1 > \cdots > TolX_k > \cdots > TolX_K > 0. \quad (6)
\]
Then, for each fixed $y \in \Gamma$, we adaptively compute approximate spatial solutions $u_k(y) \in V_k(y)$ on nested subspaces
\[ V_0(y) \subset V_1(y) \subset \cdots \subset V_K(y) \subset H^1(D) \] (7)
with the pointwise error estimates
\[ \|u(\cdot, y) - u_k(y)\|_{H^1(D)} \leq C_k(y) \cdot TOL X_k, \quad k = 0, \ldots, K. \] (8)
Supposing measurability of the discrete spaces $V_k(y)$ and bounded second moments of $C_k(y)$, we directly get
\[ \|u - u_k\|_{L_2^2(\Gamma; H^1(D))} \leq C_x \cdot TOL X_k, \quad k = 0, \ldots, K, \] (9)
with a constant
\[ C_x := \max_{k = 0, \ldots, K} \left( \int_{\Gamma} C^2_k(y) \rho(y) dy \right)^{1/2} \] (10)
that does not depend on $y$ and $k$. Adaptive algorithms as proposed in [6, 15] converge for fixed $y \in \Gamma$ and $TOL X_k \to 0$. The constant $C_x$ depends on the quality of the a posteriori error estimator used. Values close to one can be obtained by hierarchical error estimators and gradient recovery techniques in the asymptotic regime.

3 Adaptive Stochastic Interpolation

Let us assume $u \in C^0(\Gamma; H^1(D))$ and denote by $\{I_{M_k}\}_{k=0,1,\ldots}$ a sequence of interpolation operators
\[ I_{M_k} : C^0(\Gamma) \to L_2^2(\Gamma) \] (11)
with $M_k$ points from the $N$-dimensional space $\Gamma = \Pi_{n=1,\ldots,N} \Gamma_n$. We construct each of these operators by a hierarchical sequence of one-dimensional Lagrange interpolation operators with the anisotropic Smolyak algorithm, which was introduced by Gerstner and Griebel [12]. The method is dimension-adaptive, using the individual surplus spaces in the multi-dimensional hierarchy as natural error indicators.

Let $\{TOL Y_k\}_{k=0,\ldots,K}$ be a second sequence of tolerances, a priori not necessarily decreasing. Under suitable regularity assumptions for the uncertain data (see e.g. Babuška, Nobile, Tempone [1, Lem. 3.1, 3.2]), we can assume that there exist numbers $M_k$, $k = 0, 1, \ldots, K$, and a constant $C_y > 0$ not depending on $k$ such that
\[ \|(u_k - u_{k-1}) - I_{M_{K-k}}[u_k - u_{k-1}]\|_{L_2^2(\Gamma; H^1(D))} \leq C_y \cdot TOL Y_{K-k}, \quad k = 0, \ldots, K, \] (12)
where, for simplicity, we set $u_{-1} = 0$.

Since (8) implies $\|u_k - u_{k-1}\|_{L_2^2(\Gamma; H^1(D))} \leq C \cdot TOL X_{k-1}$, which is decreasing as $k \to \infty$, we can expect that for higher $k$ it suffices to use less accurate interpolation operators, i.e., smaller numbers $M_{K-k}$, to achieve a required accuracy. Indeed this is the main motivation to set up a multilevel interpolation approximation. Note that in this way, the tolerances $TOL Y_{K-k}$ are strongly linked to the spatial tolerances $TOL X_k$. We will make this connection more precise and give suitable values for the sequence of tolerances $\{TOL Y_k\}_{k=0,\ldots,K}$ in the next section.
4 Adaptive Multilevel Method

Given the sequences \( \{u_k\} \) and \( \{I_{M_k}\} \), we define the multilevel interpolation approximation in the usual way (as in [20]) by

\[
u^{(ML)}_K = \sum_{k=0}^{K} I_{M_k} [u_k - u_{k-1}] = \sum_{k=0}^{K} \left( u^{(SL)}_{M_{K-k},k} - u^{(SL)}_{M_{K-k},k-1} \right).
\]

(13)

Observe that the most accurate interpolation operator \( I_{M_K} \) is used on the coarsest spatial approximation \( u_0 \) whereas the least accurate interpolation operator \( I_{M_0} \) is applied to the difference of the finest spatial approximations \( u_K - u_{K-1} \). From (13), the close relations of the single index \( k \) for the spatial and stochastic approximations are clearly visible.

Convergence analysis

To show the convergence of the multilevel approximation \( u^{(ML)}_K \) to the true solution \( u \), we split the error into the sum of a spatial discretization error and a stochastic interpolation error. This yields with the triangle inequality

\[
\|u - u^{(ML)}_K\|_{L^2(\Gamma;H^1(D))} \leq \|u - u_K\|_{L^2(\Gamma;H^1(D))} + \|u_K - u^{(ML)}_K\|_{L^2(\Gamma;H^1(D))}.
\]

(14)

Due to (9) the first term on the right hand side of (14) is bounded by \( C_x \cdot TolX_K \). The aim is now to choose the tolerances \( TolY_k \) in an appropriate way to reach the same accuracy for the second term. From (12), we estimate the stochastic interpolation error as follows:

\[
\|u_K - u^{(ML)}_K\|_{L^2(\Gamma;H^1(D))} \leq \sum_{k=0}^{K} \left\| (u_k - u_{k-1}) - I_{M_k} [u_k - u_{k-1}] \right\|_{L^2(\Gamma;H^1(D))} \leq \sum_{k=0}^{K} C_y \cdot TolY_{K-k}.
\]

(15)

Let us first discuss a first choice for the tolerances \( TolY_k \). To obtain an accuracy of the same size as for the spatial discretization error, we could simply require \( TolY_k \leq C_x \cdot TolX_K / ((K + 1) C_y) \) for \( k = 0, \ldots, K \). It would follow

\[
\|u - u^{(ML)}_K\|_{L^2(\Gamma;H^1(D))} \leq 2 C_x \cdot TolX_K,
\]

(16)

and thus convergence of the adaptive multilevel method for \( TolX_K \to 0 \). However, the values for \( TolY_k \) can be optimized by minimizing the computational costs while keeping the desired accuracy. This is considered next.

Cost analysis

We will analyse the computational cost, \( C^{(ML)}_\epsilon \), of the multilevel stochastic collocation estimator \( u^{(ML)}_K \), required to achieve an accuracy \( \epsilon \). In order to quantify the contributions from the
spatial discretization and the stochastic collocation, we make two assumptions to link the cost with the accuracies in (9) and (12). Let $A_k$ denote a lower bound for the cost to solve the deterministic PDE for one sample point $y \in \Gamma$ with accuracy $TolX_k$. Then, we assume that for all $k = 0, \ldots, K$, 

\begin{align}
(A1) \quad A_k & \leq C_c \cdot TolX_k^{-s}, \\
(A2) \quad C_y \cdot TolY_{K-k} = C_I(N) M_{K-k}^{-\mu} TolX_{K-1}
\end{align}

as well as the special case

$$
\|u_0\|_{L^2(\Gamma; H^1(D))} \leq TolX_0 := \text{const}.
$$

Here, $C_I(N) > 0$, $C_c > 0$ does not depend on $k$, and $s, \mu > 0$ are two real numbers.

Assumption (A1) usually holds for first-order adaptive spatial discretization methods with $s = d$, when they are coupled with optimal linear solvers such as multigrid. The factors on the right-hand side in (A2) reflect best the convergence of the sparse grid approximations in (12) with respect to the total number $M_{K-k}$ of collocation points, see [16, 17] or [20, Theorem 5.5]. To estimate the difference $u_k - u_{k-1}$, we have used the fact that

$$
\|u_k - u_{k-1}\|_{L^2(\Gamma; H^1(D))} \leq C \cdot TolX_{k-1}
$$

with a constant $C > 0$ close to $C_x$. It follows from $\|u - u_{k-1}\|_{H^1(D)} \approx \|u_k - u_{k-1}\|_{H^1(D)}$, which is the basis for the very good performance of hierarchical error estimators. We absorb $C$ in $C_I$. The factor $\mu$ strongly depends on the dimension $N$. Theoretical results for the anisotropic classical Smolyak algorithm are given in [16, Thm. 3.8].

The total computational cost of the approximation $u^{(ML)}_K$ can be defined as

$$
C^{(ML)} = \sum_{k=0}^K M_{K-k} (A_k + A_{k-1}),
$$

with $A_{-1} := 0$. In a first step, we will consider a general sequence $\{TolX_k\}_{k=0,\ldots,K}$ without defining a decay rate a priori. We have the following result for the $\epsilon$-cost $C^{(ML)}_\epsilon$ and the optimal choice of the tolerances $TolY_k$ in (12).

**Theorem 4.1.** Let $TolX_0, TolX_1, \ldots,$ be a decreasing sequence of spatial tolerances satisfying (6). Suppose assumptions (A1) and (A2) hold. Then, for any $\epsilon$, there exist an integer $K = K(\epsilon)$ and a sequence of tolerances $\{TolY_k\}_{k=0,\ldots,K}$ in (12) such that

$$
\|u - u^{(ML)}_K\|_{L^2(\Gamma; H^1(D))} \leq \epsilon
$$

and

$$
C^{(ML)}_\epsilon \leq C \cdot G_K(\mu) \frac{\mu+1}{\mu} \epsilon^{-\frac{1}{\mu}} + C_c \sum_{k=0}^K F_k TolX_{k-1}
$$

with $C = C_c (2 C_I)^{\frac{1}{\mu}}$ and

$$
F_k = (TolX_k^{-s} + (1 - \delta_{k0}) TolX_{k-1}^{-s}) TolX_{k-1}^{-1}, \quad k = 0, \ldots, K,
$$

$$
G_K(\mu) = \sum_{k=0}^K F_k^{\frac{\mu}{\mu+1}} TolX_{k-1}.
$$
Here, as usual, $\delta_{k0} = 1$ for $k = 0$ and zero otherwise. The optimal choice for the tolerances $TolY_k$ is given by

$$TolY_{K-k} = (2 C_y G_K(\mu))^{-1} F_k^{\mu+1} TolX_{k-1} \epsilon. \quad (22)$$

Proof: As in the convergence analysis above, we split the error and make sure that both the spatial discretization error and the stochastic interpolation error are bounded by $\epsilon/2$. First, we choose an appropriate $K \geq 0$ and $TolX_K$ such that $C_x \cdot TolX_K < \epsilon/2$. This is, of course, always possible and fixes the number $K = K(\epsilon)$ as a function of $\epsilon$. Next we determine the set $\{M_k\}_{k=0, \ldots, K}$ so that the computational cost in (17) is minimized subject to the requirement that the stochastic interpolation error is bounded by $\epsilon/2$. Using assumptions (A1) and (A2), this reads

$$\min_{M_0, \ldots, M_K} \sum_{k=0}^{K} C_c \cdot M_{K-k} \left( TolX_k^{-s} + (1 - \delta_{k0})TolX_{k-1}^{-s} \right)$$

$$s.t. \sum_{k=0}^{K} C_I \cdot M_{K-k}^{-\mu} TolX_{k-1} = \frac{\epsilon}{2}. \quad (23)$$

Application of the Lagrange multiplier method with all $M_k$ treated as continuous variables gives the optimal choice for the number of samples

$$M_{K-k} = (2 C_I G_K(\mu))^{1/\mu} F_k^{\frac{1}{\mu+1}} \epsilon^{-\frac{1}{\mu}} \quad (24)$$

with

$$F_k = \left( TolX_k^{-s} + (1 - \delta_{k0})TolX_{k-1}^{-s} \right) TolX_{k-1}^{-1}, \quad k = 0, \ldots, K, \quad (25)$$

$$G_K(\mu) = \sum_{k=0}^{K} F_k^{\mu+1} TolX_{k-1}. \quad (26)$$

To ensure that $M_{K-k}$ is an integer, we round up to the next integer. The complexity of the multilevel approximation can now be estimated:

$$C_{\epsilon}^{(ML)} \leq \sum_{k=0}^{K} C_c \cdot (M_{K-k} + 1) \left( TolX_k^{-s} + (1 - \delta_{k0})TolX_{k-1}^{-s} \right)$$

$$= \sum_{k=0}^{K} C_c \cdot \left( (2 C_I G_K(\mu))^{1/\mu} F_k^{\frac{1}{\mu+1}} \epsilon^{-\frac{1}{\mu}} + 1 \right) F_k TolX_{k-1}$$

$$\leq C \cdot G_K(\mu)^{1/\mu} \epsilon^{-\frac{1}{\mu}} \sum_{k=0}^{K} F_k^{\frac{\mu+1}{\mu}} TolX_{k-1} + C_c \sum_{k=0}^{K} F_k TolX_{k-1}$$

$$= C \cdot G_K(\mu)^{\frac{\mu+1}{\mu}} \epsilon^{-\frac{1}{\mu}} + C_c \sum_{k=0}^{K} F_k TolX_{k-1}$$
with $C = C_c (2 C_f)^{\frac{1}{2}}$.

The optimal tolerances $TolY_k$ can be directly determined from assumption (A2). We get

$$TolY_{K-k} = (2 C_y G_K(\mu))^{-1} F_k^{\mu} TolX_{k-1}.$$

(27)

Note that with these values $\sum_{k=0}^K C_y \cdot TolY_k = \epsilon/2$, which gives the desired accuracy in (15).

Observe that the function $G_K(\mu)$ as well as the second term in (19) still depend on $\epsilon$, because $K$ is a function of $\epsilon$. In this way, the choice of the tolerances $TolX_k$ has an influence on the rate $-1/\mu$, which could be further optimized. However, in the following we restrict to a typical geometric design with $TolX_k = q^k TolX_0$, $k = 1, 2, \ldots$, with a positive reduction factor $q < 1$.

In the following, we use the notation $a \lesssim b$ to denote $a \leq C b$ with a generic constant $C$ that does not depend on the dimension $N$ and the number of samples $M_k$. Furthermore, we write $a \同等 b$ for ($a \lesssim b$ and $b \lesssim a$).

**Theorem 4.2.** Let the sequence of spatial tolerances $\{TolX_k\}_{k=0,\ldots,K}$ in (6) be defined by

$$TolX_k = q^k TolX_0$$

with a reduction factor $q < 1$. Suppose assumptions (A1) and (A2) hold.

Then, for any $\epsilon < 1$, there exists an integer $K(\epsilon)$ such that

$$\|u - u_K^{(ML)}\|_{L^2_\rho(\Gamma; H^1(D))} \leq \epsilon$$

(28)

and

$$C^{(ML)}(\epsilon) \lesssim \begin{cases} \epsilon^{-\frac{1}{\mu}} & \text{if } s\mu < 1 \\ \epsilon^{-\frac{1}{\mu}} \log \epsilon^{1 + \frac{1}{\mu}} & \text{if } s\mu = 1 \\ \epsilon^{-s} & \text{if } s\mu > 1. \end{cases}$$

(29)

**Proof:** We start with identifying the number $K$. From the accuracy requirement $C_x \cdot TolX_K = q^K C_x \cdot TolX_0 < \epsilon/2$ we deduce

$$K = \left\lfloor \log_q \left( \frac{\epsilon}{2 C_x \cdot TolX_0} \right) \right\rfloor.$$

(30)

This gives $K \leq \log_q (\epsilon/(2 C_x \cdot TolX_0)) + 1$ and the estimate

$$C_c \sum_{k=0}^K (TolX_k^{-s} + (1 - \delta_{k0})TolX_{k-1}^{-s}) \lesssim C_c \sum_{k=0}^K q^{-ks} \lesssim \frac{q^{-sK}}{1 - q^s} \lesssim \epsilon^{-s}.$$

(31)

To estimate the term $G_K(\mu)$ in (19), we first recall that in Assumption (A2) we assumed $TolX_{-1} = \text{const.}$, which yields

$$G_K(\mu) \lesssim \sum_{k=0}^{K-1} q^{k(1-s\mu)/(\mu+1)}.$$

(32)
The behaviour of this geometric sum depends on the sign of $1 - s\mu$. When $1 - s\mu > 0$, the sum converges to a limit independent of $K$. Since in this case $e^{-s} < e^{-1/\mu}$ for $\epsilon < 1$, we end up with $C_c^{(ML)} \lesssim e^{-1/\mu}$. When $1 - s\mu = 0$, we have $G_K(\mu) \lesssim K$, which gives together with $K$ as defined in (30) the additional logarithmic term, i.e., $C_c^{(ML)} \lesssim e^{-1/\mu} \log e^{1+1/\mu}$. Eventually, when $1 - s\mu < 0$, we estimate

$$G_K(\mu) \lesssim q \left( \frac{1}{\mu+1} \right)^{K-1/\mu} \lesssim q \left( \frac{1}{\mu+1} \right)^{\frac{1}{\mu+1}} \lesssim e^{1/\mu}$$

and find

$$C_c^{(ML)} \lesssim e^{-\frac{1}{\mu} (1 - s\mu) \frac{1}{\mu} + \epsilon^{-s} \approx \epsilon^{-s}}. \quad (34)$$

This completes the proof.

For a comparison with the standard single-level stochastic collocation method, we set $K = 0$ and use the estimate

$$\|u - u_0^{(SL)}\|_{L^2_\beta(T;H^1(D))} \leq C_x \cdot TolX_0 + C_I \|u_0\|_{H^1(D)} M_0^{-\mu}. \quad (35)$$

Balancing both contributions with $\epsilon/2$, requests $TolX_0 \approx \epsilon$ and $M_0 \approx e^{-1/\mu}$. The computational $\epsilon$-cost is then bounded by

$$C_{\epsilon}^{(SL)} \approx M_0 \cdot TolX_0^{-s} \approx e^{-s - \frac{1}{\mu}}. \quad (36)$$

In terms of savings, we find a reduction factor $\Theta := C_{\epsilon}^{(ML)}/C_{\epsilon}^{(SL)} \approx e^s$ for the case $s\mu < 1$, which also holds, up to a log factor, for $s\mu = 1$. When $s\mu > 1$, we have $\Theta \approx e^{1/\mu}$. Thus, the advantage of the multilevel method is obvious.

**Multilevel Approximations of Functionals**

In many applications, it is more natural to consider a functional $\psi$ of the solution $u$ instead of the solution itself. Suppose a (possibly nonlinear) functional $\psi : H^1(D) \to \mathbb{R}$ with $\psi(0) = 0$ is given. Then, we define the following single-level and multi-level stochastic collocation approximations:

$$\psi_K^{(SL)} := \mathcal{I}_{MK} \left[ \psi(u_K) \right], \quad (37)$$

$$\psi_K^{(ML)} := \sum_{k=0}^K \mathcal{I}_{MK-k} \left[ \psi(u_k) - \psi(u_{k-1}) \right] \quad (38)$$

with $u_{-1} := 0$. As in (9) and (12), we ensure that for the adaptive error control of the expected values, for all $k = 0, \ldots, K$, we have

$$|E[\psi(u) - \psi(u_k)]| \leq C_x \cdot TolX_k, \quad (39)$$

$$|E[\psi(u_k) - \psi(u_{k-1}) - \mathcal{I}_{MK-k} \left[ \psi(u_k) - \psi(u_{k-1}) \right]]| \leq C_Y \cdot TolY_{K-k}. \quad (40)$$

Then, we have the following analogous result to Theorem 4.2 for the expected value of the error of the multilevel approximation of functionals:
Theorem 4.3. Let the sequence of spatial tolerances \( \{\text{TolX}_k\}_{k=0,1,...,K} \) in (6) be defined by \( \text{TolX}_k = q^k \text{TolX}_0 \) with a reduction factor \( q < 1 \). Suppose assumptions (A1) and (A2) hold. Then, for any \( \epsilon < 1 \), there exists an integer \( K(\epsilon) \) such that

\[
\mathbb{E} \left[ \psi(u) - \psi^{(ML)}_K \right] \leq \epsilon
\]

and

\[
C^{(ML)}_\epsilon \lesssim \begin{cases} 
\epsilon^{-\frac{1}{\mu}} & \text{if } \mu \leq 1 \\
\epsilon^{-\frac{1}{\mu}} |\log \epsilon|^{1 + \frac{1}{\mu}} & \text{if } \mu = 1 \\
\epsilon^{-s} & \text{if } \mu > 1
\end{cases}
\]  

(42)

Proof: All steps can be done as in the proof of Theorem 4.2. □

We would like to mention that the convergence rate of the error in the functional \( \psi(u_k) \) in general larger than the convergence rate of the error in the \( H_1 \)-norm, which gives a smaller value for \( s \) in assumption (A1). We will see in Section 6 that this super-convergence property depends on the regularities of the primal and dual solutions (cf. formula (44) below). In the special case of quadratic convergence, we have \( s = d/2 \) if an optimal linear solver such as multigrid is used.

5 Numerical Algorithm

In what follows, we will describe the algorithmic procedure of our adaptive multilevel stochastic collocation method. The proposed method has a self-adaptive nature. Once the tolerances \( \{\text{TolX}_k\}_{k=0,1,...,K} \) and \( \{\text{TolY}_k\}_{k=0,1,...,K} \) are set, the algorithm delivers a numerical solution \( u^{(ML)}_K \) or an approximate functional \( \psi^{(ML)}_K \) with an accuracy close to a user-prescribed tolerance \( \epsilon \).

First, the reliability of the estimation for the adaptive spatial discretization and the adaptive Smolyak algorithm has to be studied in order to provide values for the constants \( C_x \) and \( C_y \). A brief discussion is given in the next section. While the spatial tolerances \( \text{TolX}_k \) can be freely chosen by setting \( K \) and \( q \), the optimal choice of the tolerances \( \text{TolY}_k \) in (22) requests the knowledge of the parameters \( s \) and \( \mu \). They have to be determined in advance through an appropriate number of samples. Note that the adaptive anisotropic Smolyak algorithm will automatically detect the importance of various directions in the parameter space \( \Gamma \subset \mathbb{R}^N \). Tab. 1 illustrates the main steps of the whole algorithm for the approximation of functionals.

A crucial point already mentioned in [20] is that the optimal rounded values for the number of samples, \( M_k \), will not be used by the algorithm, because they do not necessarily correspond to an adaptive sparse grid level. However, for each level \( k \), the tolerance \( \text{TolY}_k \) can be ensured by choosing \( \bar{M}_k \geq M_k \) slightly larger, resulting in a slight inefficiency of the sparse grid approximation. Note that, in practice, the same behaviour is observed for adaptive spatial discretizations. In any case, there is no restart necessary as used in [20, Section 6.3].
Algorithm: Adaptive Multilevel Stochastic Collocation Method

1. Given $\varepsilon$, $K$ and $q$.
2. Estimate $C_x$, $C_y$, $s$ and $\mu$.
3. Compute $TolX_0 := \varepsilon/(2C_x q^K)$.
4. Set spatial tolerances:
   $TolX_k := q^k TolX_0$, $k = 1, \ldots, K$.
5. Compute $TolX_{-1} := E[I_{M_0}[\psi(u_0)]]$ with $TolX_0$ and $TolY_0 = TolX_0$.
6. Set stochastic tolerances:
   $F_k := (TolX_k^{-s} + (1 - \delta_k) TolX_{k-1}^{-s}) TolX_{k-1}^{-1}$, $k = 0, \ldots, K$,
   $G_K := \sum_{k=0}^{K} F_k^{\mu/(\mu+1)} TolX_{k-1}$,
   $TolY_{K-k} := (2C_y G_K)^{-1} F_k^{\mu/(\mu+1)} TolX_{k-1} \varepsilon$, $k = 0, \ldots, K$.
7. Initial step, $k = 0$ (reuse samples from step 5):
   Compute $E_0 := E[I_{M_K}[\psi(u_0)]]$ with $TolX_0$ and $TolY_K$.
8. Multilevel steps, $k = 1, \ldots, K$ (reuse samples from level $k-1$):
   Compute $E_k := E[I_{M_{K-k}}[\psi(u_k) - \psi(u_{k-1})]]$ with $TolX_k$, $TolX_{k-1}$, $TolY_{K-k}$.
9. Compute $E[\psi^{(ML)}_K] := \sum_{k=0}^{K} E_k$.

Table 1: Numerical algorithm to approximate solution functionals $\psi(u)$ by an adaptive multilevel stochastic collocation method.

6 Numerical Examples

First, we give general informations on the adaptive solvers used. Then, numerical examples are presented for two Poisson problems with uncertain source term and uncertain geometry, respectively. All calculations have been done with Matlab-Version R2017a on a Latitude 7280 with an i5-7300U Intel processor at 2.7 GHz.

Adaptive Finite Element Method

For the solution of second order elliptic partial differential equations with the Laplace operator $A(x, y) = -\Delta$ in two spatial dimensions, which will be considered in our numerical examples, we use an adaptive $P1$-finite element method (AFEM) implemented by Funken, Praetorius, and Wissgott in the Matlab package $p1afem$. Using Matlab built-in functions and vectorization for an efficient realization, the code performs with almost linear complexity in terms of degrees of freedom with respect to the runtime. The complete Matlab code of $p1afem$ can be downloaded from one of the authors webpage [9]. A general description of the underlying ideas can be found in [11] and the technical report [10] provides a detailed documentation. The code is easy to modify. In order to control the accuracy of solution functionals, we have implemented the dual weighted residual method (DWRM) introduced by Becker and Rannacher [3]. In what follows, we will give a short summary of the underlying principles that are relevant for our
Let $y \in \Gamma$ be fixed and suppose the solution functional $\psi(u)$ has the special form

$$
\psi(u) = \int_D N(u) \, dx
$$

with a possibly nonlinear function $N(u)$. Let $u_k \in V_k$ be the finite element solution computed on the (adaptive) mesh $T_k$. Then the DWRM provides a representation of the error in the solution functional in the form

$$
\psi(u) - \psi(u_k) = \sum_{T \in T_k} \left\{ \int_T f(x,y)(w - v_k) - \nabla u_k \cdot \nabla (w - v_k) \, dx \right\} + H.O.T.
$$

with an arbitrary $v_k \in V_k$ and $w$ being the exact solution of the dual problem

$$
A^*(x,y) w(x,y) = N'(u_k), \quad (x,y) \in D \times \Gamma,
$$

$$
w(x,y) = 0, \quad (x,y) \in \partial D \times \Gamma.
$$

The higher order terms, H.O.T., in (44) arise from the linearization of $N(u)$. They are of order $O(\|u - u_k\|_{L^2(D)})$ and will be neglected. Next we replace $v_k$ by the linear finite element approximation $w_k \in V_k$ of the dual problem and compute an approximation $\phi_k \approx w - w_k$ of its error in the hierarchical surplus space of quadratic finite elements. Such kind of hierarchical error estimators are already implemented in $p1afem$ for the primal solution $u_k$. Eventually, we obtain

$$
\psi(u) - \psi(u_k) = \sum_{T \in T_k} \eta_T + H.O.T.
$$

with

$$
\eta_T = \int_T f(x,y)\phi_k - \nabla u_k \cdot \nabla \phi_k \, dx.
$$

We use $|\eta_T|$ as refinement indicators and mark elements $T \in T_k$ for refinement by the DÖRFLE criterion [6], which determines the minimal set $\mathcal{M} \subset T_k$ such that $\theta \sum_{T \in T_k} |\eta_T| \leq \sum_{T \in \mathcal{M}} |\eta_T|$. We set $\theta = 0.6$ in our calculations. Refinement by newest vertex bisection is applied to guarantee nested finite element spaces and the optimal convergence of the adaptive finite element method, see [15]. We stop the adaptation if the absolute value of $\sum_{T \in T_k} \eta_T$ is less than a prescribed tolerance. This error estimator is quite accurate and has been successfully applied in the solution of various problems.

**Adaptive Anisotropic Smolyak Algorithm**

The main idea for the construction of the sparse grid interpolation operators in (11) is to use the hierarchical decomposition

$$
\mathcal{I}_M[u_h](y) = \sum_{i \in I} \bigtriangleup^m[i]u_h](y) := \sum_{i \in I} \bigotimes_{n=1}^{N} \left( \mathcal{I}^m[i_n]u_h](y) - \mathcal{I}^m[i_{n-1}]u_h](y) \right)
$$
with multi-indices \( \mathbf{i} = (i_1, \ldots, i_N) \in I \subset \mathbb{N}_+^N \), \( m(\mathbf{i}) = (m(i_1), \ldots, m(i_N)) \), and univariate polynomial interpolation operators \( I_{\mathbf{n}}^{m(\mathbf{i})} : C^0(\Gamma_n) \to \mathbb{P}_{m(\mathbf{i})-1} \), which use \( m(\mathbf{i}_n) \) collocation points to construct a polynomial interpolant in \( y_n \in \Gamma_n \) of degree at most \( m(\mathbf{i}_n) - 1 \). The operators \( \Delta^{m(\mathbf{i})} \) are often referred to as hierarchical surplus operators. The function \( m \) has to satisfy \( m(0) = 0 \), \( m(1) = 1 \), and \( m(i) < m(i+1) \). We set \( I_0 = 0 \) for all \( n = 1, \ldots, N \) and use the nested sequence of univariate Clenshaw-Curtis nodes with \( m(\mathbf{i}_n) = 2^{i-1} + 1 \) if \( i > 1 \). In (49), \( M_k \) is then the number of all explored quadrature points in \( \Gamma \) determined by \( m(\mathbf{i}) \). To get good approximation properties, the index set \( I \) should satisfy the downward closed set property, i.e.,

\[
\text{if } \mathbf{i} \in I, \text{ then } \mathbf{i} - \mathbf{e}_j \in I \text{ for all } j = 1, \ldots, N \text{ such that } i_j > 1.
\]  

(50)

As usual, we ensure that \( 1 \in I \) to also recover constant functions.

The hierarchical structure in (49) allows to interpret updates that are derived by adding further differences \( \Delta^{m(\mathbf{i}_a)} \), i.e., enhancing the index set \( I \) by an admissible index \( \mathbf{i}_a \), as error indicators for already computed approximations. There are several adaptive strategies available. One could explore the whole margin of \( I \) defined by

\[
M_I := \{ \mathbf{i} \in \mathbb{N}_+^N \setminus I : \mathbf{i} - \mathbf{e}_n \in I \text{ for some } n \in \{1, \ldots, N\} \}.
\]

(51)

Generally, this approach is computationally challenging and yields a fast increase of quadrature points. Instead, as originally suggested by Gerstner and Griebel in [12], the margin is reduced to the set

\[
R_I := \{ \mathbf{i} \in M_I : \mathbf{i} - \mathbf{e}_n \in I \text{ for all } n = 1, \ldots, N \text{ with } i_n > 1 \}.
\]

(52)

In each iteration step, the adaptive Smolyak algorithm computes the profits \( \Delta^{m(\mathbf{i}_a)} \) for all \( \mathbf{i}_a \in R_I \) (already computed profits are reused) and replaces the index in \( R_I \) with the highest profit, say \( \mathbf{i}_{\text{max}} \), by its admissible neighbours taken from the set \( \{ \mathbf{i}_{\text{max}} + \mathbf{e}_j, j = 1, \ldots, N \} \). These neighbours are then explored next. The algorithm stops if the absolute value of the highest profit is less than a prescribed tolerance. This adaptive strategy has been implemented in the Sparse Grid Matlab Kit, version 17-5, which can be downloaded from [19]. Numerical comparisons are presented by Bäck, Nobile, Tamellini, and Tempone in [2].

In our numerical experiments, we follow this main line with one small change: After reaching the final iteration step, we do not add the new profits to the interpolant and deliver the previous approximate value instead. Hence, the last highest profit can then be considered as a more realistic error indicator. This adaptation allows a better understanding of the convergence behaviour of the multilevel approach. Clearly, in practical calculations one would proceed with the better values.

6.1 Uncertain Source and Boundary Conditions: \( d=2, N=2 \)

We consider the Poisson problem with random right-hand side proposed in [14, Sect. 5.1] to study adaptive multilevel Monte Carlo methods. The problem is slightly modified by an anisotropic
factor. We choose the two-dimensional computational domain $D = (-1, 1)^2 \subset \mathbb{R}^2$ and set $A(x, y) = -(\partial_{x_1} + \partial_{x_2})$. The uncertain source term in (1) and the uncertain boundary conditions in (2) are given through

\begin{align*}
    f(x, y) &= d(x, y, \alpha(y)) \exp \left( -50 (\alpha(y)(x_1 - y_1)^2 + (x_2 - y_2)^2) \right), \ x \in D, \\
    g(x, y) &= \exp \left( -50 (\alpha_1(y_1)(x_1 - y_1)^2 + (x_2 - y_2)^2) \right), \ x \in \partial D,
\end{align*}

with $\alpha(y) = (36y_1 + 11)/2$, $\beta = 50$ and

\begin{align*}
    d(x, y, \alpha) &= -4\alpha^2\beta^2(x_1 - y_1)^2 - 4\beta^2(x_2 - y_2)^2 + 2\beta(1 + \alpha).
\end{align*}

Here, the random vector $y = (y_1, y_2)^T$ consists of two uniformly distributed random variables $y_1, y_2 \sim \mathcal{U}[-0.25, 0.25]$. Observe that $\alpha(y_1)$ varies over the interval $[1, 10]$ and hence yields an anisotropic solution behaviour. The data are chosen in such a way that the unique pathwise solution $u(x, y)$ of (1)-(2) is determined by

\begin{align*}
    u(x, y) &= \exp \left( -50 (\alpha(y)(x_1 - y_1)^2 + (x_2 - y_2)^2) \right), \ x \in D.
\end{align*}

The goal is to approximate the quantity of interest

\begin{align*}
    \mathbb{E}[\psi(u)] &= \mathbb{E} \left[ \int_D u^2(x, y) \, dx \right].
\end{align*}

The initial mesh is obtained by three uniform refinement steps, starting from two triangles with common edge from $(-1, -1)$ to $(1, 1)$. For the specific parameter choice $y = (-0.22, -0.22)^T$ and spatial tolerance $TolX = 5 \times 10^{-6}$, the numerical solution and its corresponding adaptive mesh, consisting of 158734 points, are shown in Fig. 1. The requested refinement steps are 13, which, if applied globally, would correspond to nearly $4.3 \times 10^3$ uniform mesh points. The quality of the error estimation based on the dual weighted residual method is illustrated in the left part of Fig. 2. The effectivity index, i.e., the ratio between estimator and true error, tends asymptotically to 1 for both uniform and adaptive refinement. For higher tolerances, the gain of efficiency by the use of adaptive meshes is nearly a factor 10. The very good quality of the estimation process is still valid if we sample over the whole parameter space exemplified by means of an isotropic Smolyak algorithm with 145 collocation points. The resulting expected values for adaptive meshes are plotted in the right part of Fig. 2. Observe that the estimators deliver upper bounds for the numerical errors and the tolerances are always satisfied. As expected from the theory, the convergence rates for $\psi(u)$ and $\mathbb{E}[\psi(u)]$ in terms of computing time are both close to $-1$. So we have $C_x = 1$ in (39) and $s = 1$ in our assumption (A1) above.

Next we consider the convergence behaviour and the quality of the error estimates for the adaptive anisotropic Smolyak algorithm. For comparison, the spatial tolerance $TolX = 5 \times 10^{-7}$ is fixed and a reference value $\mathbb{E}[\psi(u)] = 0.015096132160028553$ with 165 adaptive collocation points is computed. The employed polynomial degrees are plotted in the right part of Fig. 3. The anisotropic character of the problem is obvious. The left diagram of Fig. 3 shows the results
for stochastic tolerances $TolY = 10^{-3}, \ldots, 10^{-8}$. The corresponding numbers of collocation points are 5, 7, 11, 11, 19, and 101. The quality of the estimator is fine for lower tolerances, but reduces when the number of grid points increases. However, the prescribed tolerances are always satisfied and the observed errors (with respect to the reference solution) are at most a factor of around 10 smaller than the requested tolerances. A least-squares fit gives an averaged value of $p_c = -4.35$ for the convergence order. Therefore, we will use $C_y = 0.1$ and $\mu = -p_c$ in our assumption (A2) above.

We now apply the adaptive multi-level approach with overall accuracy requirements $\epsilon = 10^{-5}, 5 \times 10^{-6}, 2.5 \times 10^{-6}, 10^{-6}$. We set $K = 2$, i.e., three levels are used, and choose the reduction factor $q = 0.2$. The spatial tolerances are then given by $TolX_k = \epsilon q^{k-2}/2$ with $k = 0, 1, 2$. To calculate, in a first step, a sufficiently accurate approximation for the tolerance $TolX_{-1} = E[\psi(u_0)]$ at reasonable cost, we apply the anisotropic Smolyak algorithm with $TolX = TolY = TolX_0$. Note that the samples can be reused in the first level of the multi-level scheme. Eventually, the stochastic tolerances are derived from (22) with $C_y = 0.1$, $\mu = 4.35$, and $s = 1$.

Results for the three-level and single-level approach are summarized in Fig. 4. For adaptive spatial meshes, we observe that the errors of the expected values $E[\psi_2^{(ML)}]$ are very close to the prescribed tolerances. This is not always the case for the values $E[\psi_2^{(SL)}]$ computed by the single-level approach. The three-level approach performs reliably and outperforms the single-level
version clearly. The orders of convergence, \( p_{ML} = -1/s = -1 \) and \( p_{SL} = -1/(s + 1/\mu) \approx -0.81 \), predicted by Theorem 4.3 for the accuracy in terms of computational complexity are also visible. The three-level approach with uniform spatial meshes is not competitive and exceeds memory restrictions for higher tolerances.

We have also applied an adaptive multilevel Monte Carlo method [4, 14]. For the lowest tolerance, \( \epsilon = 10^{-5} \), and an average over 5 independent realizations, the three-level algorithm achieves an accuracy of \( 2.77 \times 10^{-6} \) in \( 7.76 \times 10^{4} \) sec. The numbers of averaged samples for each level are \( M_0 = 519634 \), \( M_1 = 6153 \), and \( M_2 = 243 \). Obviously, the slow convergence rate \( \mu = 0.5 \) of the Monte Carlo method is prohibitive for higher tolerances.

6.2 Uncertain Geometry: \( d=2, N=16 \)

In a second example, we again consider the two-dimensional Poisson problem with \( A = -(\partial_{x_1^2} + \partial_{x_2^2}) \), but now with uncertain geometry of the computational domain to study, e.g., the impact of imprecise manufacturing machines. The uncertain domain is defined by

\[
D(y) = (0, 6) \times (0, 6) \setminus (D_1(y) \cup D_2(y)),
\]

(58)
Figure 3: One peak problem: History of error estimators and true errors for anisotropic Smolyak approximations with stochastic tolerances TolY = 10^{-(i+3)}, i = 0,\ldots,5 (left); employed multi-index set G for the reference solution with TolX = 5 \times 10^{-7} and 165 adaptive collocation points (right). The numerically observed, averaged convergence order for $E\left[\psi(u)\right]$ in terms of collocation points is -4.35.

where the holes $D_1(y) = P_1P_2P_3P_4$ and $D_2(y) = P_5P_6P_7P_8$ are taken as quadrilaterals with the uncertain vertices

\begin{align*}
P_1 &= (1 + y_1/a_1, 1 + y_2/a_2), & P_2 &= (2 + y_3/a_3, 1 + y_4/a_4), \\
P_3 &= (2 + y_5/a_5, 3 + y_6/a_6), & P_4 &= (1 + y_7/a_7, 3 + y_8/a_8), \\
P_5 &= (4 + y_9/a_9, 1 + y_{10}/a_{10}), & P_6 &= (5 + y_{11}/a_{11}, 1 + y_{12}/a_{12}), \\
P_7 &= (5 + y_{13}/a_{13}, 5 + y_{14}/a_{14}), & P_8 &= (4 + y_{15}/a_{15}, 5 + y_{16}/a_{16}).
\end{align*}

(59)

Here, the random vector $y = (y_1,\ldots,y_{16})^T$ consists of sixteen uniformly distributed random variables $y_i \sim U[-1,1], i = 1,\ldots,16$. We set

\begin{equation}
(a_1,\ldots,a_{16}) = (5, 5, 5, 5, 5, 5, 5, 5, 10, 10, 10, 10, 10, 10, 10, 20, 20, 20, 20)
\end{equation}

(60)

to represent different strength of uncertainty and hence anisotropy in the stochastic space. We impose a constant volume force $f \equiv 1$ and fix the component by homogeneous Dirichlet boundary conditions on the whole boundary $\partial D(y)$ including $\partial D_1(y) \cup \partial D_2(y)$ with stochastically varying positions in space. Our goal is then to study the effect of this uncertainty on the expectation of the overall displacement calculated by

\begin{equation}
E[\psi(u,y)] = E \left[ \int_{D(y)} u^2(x,y) \, dx \right].
\end{equation}

(61)

This allows an assessment of the desired averaged load capacity of the component, which takes into account uncertainties in the manufacturing process.
Applying a parameter-dependent map, each domain $D(y)$ can be mapped to the fixed nominal domain $D_0 = D(0)$ with $0 \in \mathbb{R}^{16}$. Such a domain mapping approach was first introduced by Xiu and Tartakovsky [21] and allows to reformulate the problem in the form (1) with parameter-dependent coefficients on $D_0$. The well established theory for elliptic partial differential equations with random input data can then be applied without modifications to show the well posedness of the setting with random domains.

All calculations start with an initial criss-cross structured mesh consisting of 1920 triangles, which are adjusted to the random holes. In Fig. 5, the numerical solution and its corresponding adaptive mesh for a spatial tolerance $TolX = 10^{-1}$ and the random vector $y = (0.5, 0.5, -0.5, -0.5, -0.5, -0.5, 1, -1, -1, 1, 1, -1, 1, -1, 1, -1, 1, -1, -1)^T$ are shown. The adaptive algorithm refines the mesh at the eight reentrant corners due to the fact that the exact solution contains a loss of regularity there. Exemplarily, we study the convergence rates of adaptive and uniform refinements for the nominal domain $D_0$, where $u(0)$ is contained in $H^{5/3}(D_0)$. While the correctly adapted grids still recover the optimal order $-1$ for the approximation of $\psi(u(0))$ in terms of CPU time, the use of a sequence of uniform meshes lets the order drop down to the
Figure 5: Two hole problem: Solution for the special random domain $D(y)$ with $y = [0.5, 0.5, -0.5, -0.5, -0.5, -0.5, 1, -1, -1, 1, 1, -1, -1, -1, -1]$ (left) and corresponding adaptive mesh based on the dual weighted residual approach (right). A spatial tolerance $TolX = 10^{-1}$ yields 2650 mesh points to efficiently resolve the corner singularities.

Theoretical value $-0.66$, as can be seen in the left part of Fig. 6. The DWR-estimators perform quite well and deliver accurate upper bounds of the error. So we again set $C_x = 1$ in (39) and use $s = 1$ for adaptive meshes and $s = 1.6$ for uniform meshes in our assumption (A1) above. Note that the latter choice takes into account that due to the larger interior angle at the random holes for some parameter values $y \neq 0$, the regularity of the solution, and thus also the spatial convergence rate, is slightly lower.

In order to estimate the parameter $\mu$, we apply the anisotropic Smolyak algorithm and calculate samples at reasonable costs with tolerances $TolX = TolY = 10^{-2}/2^i, i = 0, \ldots, 3$. The value of $\mathbb{E}[\psi(u)]$ for $i = 3$ is taken as the reference value, leading to an approximate convergence order of $p_c = -3.4$. From Fig. 6, we detect that the errors are significantly smaller than the prescribed tolerances, as it was also the case in the first test example. Therefore, we again choose $C_y = 0.1$ and set $\mu = 3.4$ in our assumption (A2) above. For later comparison, we compute a reference value $\mathbb{E}[\psi(u)] = 4.7580572144972830$ with $TolX = 10^{-4}$ and $TolY = 5 \times 10^{-4}$. The required number of collocation points is 233 and the vector of maximum polynomial degree is $(3233333332232322)^T$. 

18
For the adaptive multilevel approach, we use again three levels, i.e., $K = 2$, and set the reduction factor to $q = 0.2$. The spatial tolerances are then computed from $TolX_k = \epsilon q^{k-2}/2$, $k = 0, 1, 2$, with the six overall accuracy requirements $\epsilon = 10^{-2}, 5 \times 10^{-3}, 2.5 \times 10^{-3}, 10^{-3}, 5 \times 10^{-4}, 2.5 \times 10^{-4}$. Following the algorithm given in Tab. 1, we first compute $TolX_{-1} = E[\psi(u_0)]$ with $TolX = TolY = TolX_0$ and then derive the stochastic tolerances $TolY_k$ from (22) with $C_y = 0.1$, $\mu = 3.4$, and $s = 1$. The corresponding values are given in Tab. 2. For the multilevel approach with uniform spatial meshes, we use $s = 1.6$ to determine the stochastic tolerances.

In Fig. 7, errors for the expected value $E[\psi^{(ML)}]_2$ versus computing time are plotted for the three-level approach with adaptive and uniform spatial meshes. For comparison, we also show results for the single-level approach with adaptive spatial meshes, computed with $TolX = \epsilon/2$ and $TolY = \epsilon/(2C_y)$. In all cases, except the first two for the single-level method, the overall tolerances are satisfied. The schemes work remarkably reliably. However, the achieved accuracy is often much better than the prescribed one. One possible reason could be that cancellation effects are typically overlooked if the overall error is split into two parts, which are then individually controlled.

The multilevel method with adaptive spatial meshes outperforms the single-level approach. The number of collocation points taken by the anisotropic Smolyak algorithm are listed in Tab. 2. We observe that the number of samples for the differences keeps constant, showing that the increasing samples in the zeroth level always catch enough information to eventually reach the tolerance. The corresponding numbers of collocation points for the single-level method are...
(1, 1, 33, 41, 51, 105). Note that the computing time also includes the effort for the estimation process, which is not visible in these numbers; see also the discussion for our realization of the adaptive anisotropic Smolyak algorithm above. It is also clear that the averaged constant $C_y = 0.1$ is too optimistic for the first runs with $\epsilon = 10^{-2}, 5 \times 10^{-3}$, which leads to the algorithm taking only one collocation point. The approximate orders of convergence for both methods, $p_{ML} = -1/s = -1$ and $p_{SL} = -1/(s + 1/\mu) \approx -0.77$, predict the observed asymptotic rates for the computing times quite well. However, the actual estimates can only serve as rough indicators for the achieved accuracy.

The multilevel method with uniform spatial meshes performs better than the single-level one for the first two tolerances, but becomes quickly inefficient for higher tolerances. Due to the larger value $s = 1.6$, it needs significantly more samples for coarser meshes: $M_0 = (65, 131, 567)$ and $M_1 = (33, 33, 105)$, compared to the multilevel approach with adaptive spatial meshes, see Tab. 2, $M_2$ remains the same. The observed convergence order $-0.54$ is close to the predicted value $-1/s = -0.625$. Also for this example, it becomes obvious that uniform meshes cannot compete with adaptive meshes for higher tolerances.

This is also valid for an adaptive multilevel Monte Carlo method implemented along the principles given in [4] [14]. For accuracy requirements $\epsilon = 10^{-2}, 5 \times 10^{-3}$, results are shown in Fig. 7 for three levels. The numbers of optimized samples are $M_0 = (4140, 17013), M_1 = (48, 108), M_2 = (5, 6)$, respectively. All values are calculated from an average over 5 independent realizations. Although the variance reduction is quite high, leading to surprisingly small numbers $M_2$, the fast increasing numbers for $M_0$ are still too challenging, especially for higher tolerances.

### 7 Conclusion

Drawing inspiration from [14] [20] and our own work on adaptive schemes for PDEs with random input data [5] [18], we have developed an adaptive multilevel collocation method based on a
hierarchy of adaptive spatial and stochastic approximation spaces. These spaces are constructed such that individual sequences of tolerances are satisfied in the spatial and stochastic space. A decoupling of both spaces allows a detailed computational cost analysis of the multilevel method and, given a decreasing sequence of spatial tolerances, the determination of optimal tolerances for the stochastic degrees of freedom.

Adaptive methods have the potential to drastically reduce the number of degrees of freedom and to reach optimal complexity in terms of computing time, even in the case of solutions exhibiting singularities. They usually outperform methods based on uniform mesh refinement, and provide reliable a posteriori error estimators. Coupled with multilevel stochastic algorithms, they can increase the computational efficiency in the sampling process when the stochastic dimension increases and hence provide a general tool to further delay the curse of dimensionality, which is inherent for larger stochastic dimensions.

In two numerical examples of different complexity, we have applied up-to-date adaptive
methods implemented in the open-source MATLAB packages *p1afem* and *Sparse Grid Matlab Kit*. An adaptive P1-finite element method extended with a dual weighted residual error estimator for solution functionals has been coupled with an anisotropic Smolyak algorithm exploiting sparse grid approximations. The numerical results for the adaptive multilevel collocation method demonstrate the reliability of an error control by adaptive methods and the significant decrease in complexity versus uniform spatial refinements, single-level stochastic sampling methods and even adaptive multilevel Monte Carlo methods. Though these advantages are already known for the single methods involved, they can be efficiently combined to solve PDEs with random input data, applying them with appropriately chosen particular accuracy requirements.

8 Acknowledgement

The first author is supported by the German Research Foundation within the collaborative research center TRR154 “Mathematical Modeling, Simulation and Optimisation Using the Example of Gas Networks” (DFG-SFB TRR154/2-2018, TP B01), the Graduate School of Excellence Computational Engineering (DFG GSC233), and the Graduate School of Excellence Energy Science and Engineering (DFG GSC1070). The second author is supported by the UK Engineering and Physical Sciences Research Council (EPSRC) under grant number EP/K031368/1.

The authors would also like to thank the Isaac Newton Institute for Mathematical Sciences, Cambridge, for support and hospitality during the programme “Uncertainty Quantification for Complex Systems: Theory and Applications” (January - July 2018), where work on this paper was partially undertaken. This work was supported by EPSRC grant no EP/K032208/1.

References

[1] I. Babuška, F. Nobile, and R. Tempone. A stochastic collocation method for elliptic partial differential equations with random input data. *SIAM J. Numer. Anal.*, 45:1005–1034, 2007.

[2] J. Bäck, F. Nobile, L. Tamellini, and R. Tempone. Stochastic spectral Galerkin and collocation methods for PDEs with random coefficients: a numerical comparison. In J.S. Hesthaven and E.M. Ronquist, editors, *Spectral and High Order Methods for Partial Differential Equations*, volume 76 of *Lecture Notes in Computational Science and Engineering*, pages 43–62. Springer, 2011.

[3] R. Becker and R. Rannacher. An optimal control approach to a posteriori error estimation in finite element methods. *Acta Numerica*, 10:1–102, 2001.

[4] K.A. Cliffe, M.B. Giles, R. Scheichl, and A.L. Teckentrup. Multilevel Monte Carlo methods and applications to elliptic PDEs with random coefficients. *Comput. Visual. Sci.*, 14:3–15, 2011.

[5] G. Detommaso, T. Dodwell, and R. Scheichl. Continuous level Monte Carlo and sample-adaptive model hierarchies. *SIAM/ASA J. Uncertainty Quantification*, 7:93–116, 2019.
[6] W. Dörfler. A convergent adaptive algorithm for Poisson’s equation. *SIAM J. Numer. Anal.*, 33:1106–1124, 1996.

[7] M. Eigel, Ch. Merdon, and J. Neumann. An adaptive Monte Carlo method with stochastic bounds for quantities of interest with uncertain data. *SIAM/ASA J. Uncertainty Quantification*, 4:1219–1245, 2016.

[8] D. Elfverson, F. Hellman, and A. Målqvist. A multilevel Monte Carlo method for computing failure probabilities. *SIAM/ASA J. Uncertainty Quantification*, 4:312–330, 2016.

[9] S. Funken, D. Praetorius, and P. Wissgott. Efficient implementation of adaptive P1-FEM in Matlab, software available at http://www.asc.tuwien.ac.at/praetorius/matlab.

[10] S. Funken, D. Praetorius, and P. Wissgott. Efficient implementation of adaptive P1-FEM in Matlab (extended preprint). ASC Report 19/2008, Vienna University of Technology, 2008.

[11] S. Funken, D. Praetorius, and P. Wissgott. Efficient implementation of adaptive P1-FEM in Matlab. *Comput. Meth. Appl. Math.*, 11:460–490, 2011.

[12] T. Gerstner and M. Griebel. Dimension-adaptive tensor-product quadrature. *Computing*, 71:65–87, 2003.

[13] M. Giles. Multilevel Monte Carlo path simulation. *Operations Res.*, 56:607–617, 2008.

[14] R. Kornhuber and E. Youett. Adaptive multilevel Monte Carlo methods for stochastic variational inequalities. *SIAM J. Numer. Anal.*, 56:1987–2007, 2018.

[15] C. Kreuzer and K.G. Siebert. Decay rates of adaptive finite elements with Dörfler marking. *Numer. Math.*, 1147:679–716, 2011.

[16] 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.*, 46:2411–2442, 2008.

[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.*, 46:2309–2345, 2008.

[18] B. Schieche and J. Lang. Adjoint error estimation for stochastic collocation methods. In J. Garcke and D. Pflüger, editors, *Sparse Grids and Applications - Munich 2012*, volume 97 of *Lecture Notes on Computational Science and Engineering*, pages 271–294. Springer, 2014.

[19] L. Tamellini, F. Nobile, D. Guignard, F. Tesei, and B. Sprungk. Sparse grid matlab kit, version 17-5, software available at https://csqi.epfl.ch.
[20] A.L. Teckentrup, P. Jantsch, C.G. Webster, and M. Gunzburger. A multilevel stochastic collocation method for partial differential equations with random input data. *SIAM/ASA J. Uncertainty Quantification*, 3:1046–1074, 2015.

[21] D. Xiu and D. Tartakovsky. Numerical methods for differential equations in random domains. *SIAM J. Sci. Comput.*, 28:1167–1185, 2007.