Quantifying uncertainty in plasma microturbulence analysis using Multifidelity Monte Sampling

Julia Konrad
Fakultät für Informatik
Technische Universität München
Guided Research

November 6, 2020

Abstract

The simulation of microturbulence in plasma fusion is paramount for understanding the confinement properties of fusion plasmas with magnetic fields. Unfortunately, the measurement of many of the relevant physics parameters for these simulations is subject to uncertainties. Dealing with these problems therefore requires an uncertainty quantification approach. In this work, we employ the Multifidelity Monte Carlo (MFMC) sampling algorithm to plasma microturbulence problems. MFMC improves on the generally slow convergence of standard Monte Carlo (MC) sampling by considering surrogate (low-fidelity) models in addition to the underlying high-fidelity model. For our application, we use the established plasma microturbulence code Gene as a high-fidelity model and construct adaptive sparse grid interpolation approximations to use as data-driven low-fidelity models. While, in general, these have the same stochastic dimensionality as the high-fidelity model, we also make use of the lower intrinsic dimensionality of the considered plasma microturbulence problems and construct reduced-dimension sparse grid interpolation surrogates. We apply the MFMC algorithm with different reduced-dimension low-fidelity models to a high-dimensional real-world test case from plasma microturbulence analysis for which standard MC would be infeasible due to the computational effort required. Using MFMC, we achieve estimators for the high-fidelity model output whose mean squared error is by at least two orders of magnitude lower than that of standard MC estimators for a given computational budget.

1 Introduction

The analysis of plasma microturbulence is an important part of the research and construction of future nuclear fusion devices. Projects like the ITER[1] experiment aim to create a self-sustaining, burning plasma in the laboratory, the accomplishment of which would be a significant step on the way to make fusion power plants possible - which could yield huge amounts of safe, green energy [5]. One obstacle preventing the creation of such a plasma is its efficient and lasting magnetic confinement. This is complicated by microturbulence resulting from steep temperature and density gradients in the plasma [3]. Therefore, it is paramount for plasma physicists to gain a better understanding of the turbulence in order to be able to find ways to counteract it. To this end, simulations of such plasmas are conducted. As the physical measurement of quantities needed for these simulations can be difficult, the input parameters are subject to uncertainties. Therefore, dealing with plasma microturbulence problems requires a UQ approach.

A popular method used to perform uncertainty propagation is Monte Carlo (MC) sampling. Sampling-based approaches are especially attractive for dealing with high-dimensional problems - which plasma physics simulations can be - as they do not suffer from the curse of dimensionality. However, the slow convergence and therefore high required number of evaluations of the underlying model becomes prohibitive very soon if the model is computationally expensive. Again, this is often case in plasma microturbulence simulations. For this reason, we consider Multifidelity Monte Carlo (MFMC) sampling [14; 16] to quantify the uncertainty by computing mean and variance of the model output. Similar to standard MC, MFMC samples the input

[1] `https://www.iter.org/`

domain and evaluates the underlying model for those samples. In addition that, MFMC utilises surrogate models that are less expensive to evaluate but are also generally less accurate. By balancing the evaluations between the different types of models, more efficient estimators can be achieved.

In [11], we applied MFMC to a popular benchmark problem of plasma microturbulence analysis, which was however a simplified test case that does not completely depict reality. In this work, we apply MFMC to a higher-dimensional, more realistic test case. For this purpose, we use the gyrokinetic code GENE as a high-fidelity model and construct surrogate models using an adaptive sparse grid interpolation algorithm presented in [6; 5]. Further, we use the sensitivity information provided by the algorithm to also construct efficient dimension-reduced surrogates, i.e., models that only depend on a subset of the input parameters of the underlying model.

The remainder of this report is organized as follows. In Section 2, we give some detail on the mathematical background of plasma microturbulence analysis and its implementation in GENE, and briefly consider standard MC as a method for uncertainty propagation. Section 3 describes the MFMC algorithm and the construction of the sparse grid surrogates. We apply those techniques to our test case in Section 4. Finally, we conclude this work with a summary and a brief outlook in Section 5.

2 Background

This section gives an overview of the relevant theory behind plasma microturbulence analyis, reviews standard MC sampling and defines the related notations. In this way, we emphasize the motivation behind using MFMC for uncertainty propagation in plasma microturbulence simulation.

2.1 Plasma microturbulence analysis

GENE, the code we use in our simulation of plasma microturbulence, is a so-called *gyrokinetic* solver. Gyrokinetic theory is a popular approach used to model the behavior of the particles in a plasma. In the following, we briefly summarize this method and its implementation in GENE .

In gyrokinetics, the kinetic model using 6D (three positions, three velocties) Vlasov equations is reduced to a 5D state space (three positions, two velocities). This is done by neglecting the resolution of the gyromotion of the charged plasma particles, i.e., their movement around the magnetic field lines. This simplification is valid because the gyromotion happens on a much smaller time-scale than the movement of the center of their orbit along the field lines and only the latter is relevant for the computation of the turbulence. This reduction from a six- to a five-dimensional state space makes the simulation of plasma microturbulence feasible for current day computational resources [5].

Gyrokinetics models the different types of particles of the plasma (e.g. electrons, ions, etc.) by means of their distribution functions

$$\pi_s(\boldsymbol{R}, v_{||}, \mu_m),$$

where s denotes the particle species, $\boldsymbol{R} = (R_x, R_y, R_z)$ is the position of the gyrocenter in real space, $v_{||}$ the velocity parallel to the magnetic field lines, and $\mu_m := m_s v_\perp^2 / 2B$ the magnetic moment with the mass m_s of the species, $v_\perp$ the perpendicular velocity and B the magnetic field.

The behavior of the plasma is then characterized by the evolution of the distribution function for each particle species. This evolution is governed by the gyrokinetic Vlasov equation with collisions,

$$\dot{\pi}_s + \dot{\boldsymbol{R}} \cdot \nabla \pi_s + \dot{v}_{||} \frac{\partial \pi_s}{\partial v_{||}} + \dot{\mu}_m \frac{\partial \pi_s}{\partial \mu_m} = C(\pi_s, \pi_{s'}), \tag{1}$$

where $C(\pi_s, \pi_{s'})$ is a collision operator. Furthermore, we require the gyrokinetic field equations that describe the perturbed electrostatic potential Φ_1 and the perturbed electric potential A_1,

$$\begin{aligned} \text{Poisson's equation:} \quad & -\nabla^2 \Phi_1 = 4\pi \sum_s q_s n_{1s}, \\ \text{Ampère's law:} \quad & -\nabla^2 A_1 = \frac{4\pi}{c} \sum_s n_{1s} \boldsymbol{u}_{1s}, \end{aligned} \tag{2}$$

where c is the speed of light, q_s the charge of species s, n_{1s} describes the 0th space moment and $\boldsymbol{u}_{1s}$ the 1st order velocity moment of π_s. For more detail on the fundamentals of gyrokinetic theory, we refer to [2; 12; 5].

GENE is an Eulerian, i.e., grid-based, code that resolves the described gyrokinetic approach in 5D position-velocity space. The code is heavily parallelized by domain decomposition using MPI [8]. Furthermore, it allows for flux-tube simulations, or radially global and full-flux simulations. In this work, we consider only flux-tube simulations, i.e., thin domains that follow the magnetic field lines and with the method implemented in GENE, can be very efficiently computed (see [8] for more details).

In particular, GENE implements the gyrokinetic equations in the form

$$\dot{\boldsymbol{\pi}}_s = \mathbb{O}_{\text{lin}}(\boldsymbol{\pi}_s) + \mathbb{O}_{\text{nonlin}}(\boldsymbol{\pi}_s), \tag{3}$$

where $\mathbb{O}_{\text{lin}}$ and $\mathbb{O}_{\text{nonlin}}$ are linear and nonlinear operators that describe the respective parts of the Vlasov equation, and $\boldsymbol{\pi}_s$ is a vector containing the linearized values of the transformed discretized distribution function on the grid [12; 5]. Because fully nonlinear simulations can be very computationally expensive, we restrict our work to linear simulations since already the analysis of the linear component can reveal important information about the microinstabilites driving the turbulence. This can be done by simply neglecting the nonlinear part of (3), which then is discretized in GENE as

$$\dot{\boldsymbol{\pi}}_s = O_{\text{lin}}\boldsymbol{\pi}_s.$$

where O_{lin} is a matrix [12; 5]. Typically, O_{lin} has at most about 30% non-zero entries can have a rank of up to a few hundred thousand [5]. More insight into the nature of the microturbulence can be gained from the eigenvalues of the operator O_{lin}. We consider the so-called dominant eigenmode, i.e., the complex eigenvalue of O_{lin} belonging to the first eigenvector [5]. To obtain this quantitiy, GENE's linear eigenvalue solver is used [8].

2.2 Uncertainty propagation using standard Monte Carlo Sampling

This subsection summarizes the mathematical background of standard Monte Carlo sampling and motivates the application of MFMC in the context of plasma microturbulence analysis. In order to quantify the effect of the input uncertainties on the model output, we want to compute estimates of the model output, such as its mean and variance. In the following, we denote the high-fidelity model by $f^{(0)} : \boldsymbol{\mathcal{X}} \to \mathcal{Y}$ with the domain $\boldsymbol{\mathcal{X}} = [\mathcal{X}_1, \dots, \mathcal{X}_d]^T \subseteq \mathbb{R}^d$ of the $d \in \mathbb{N}$ uncertain input parameters and output domain $\mathcal{Y} \in \mathbb{R}$. Furthermore, we define w_0 as the runtime of the high-fidelity model.

Because the uncertainty of the system is modelled in the input parameters, we represent the input as a d-dimensional random vector $\mathbf{X} = (X_1, \dots, X_d) \in \boldsymbol{\mathcal{X}}$ with probability density function (pdf) $\pi(\mathbf{X})$. The goal is then to compute the mean

$$\mu_0 := \mathbb{E}[f^{(0)}(\mathbf{X})] = \int_{\mathbb{R}} f^{(0)}(\mathbf{x})\pi(\mathbf{x})d\mathbf{x} \tag{4}$$

and the variance

$$\sigma_0^2 := \text{Var}[f^{(0)}(\mathbf{X})] = \mathbb{E}[(f^{(0)}(\mathbf{X}) - \mu_0)^2]. \tag{5}$$

of the high-fidelity model output. One very established method to accomplish this is standard MC sampling. The algorithm draws n i.i.d. samples $\mathbf{x_1}, \dots, \mathbf{x_n}$ of the random vector $\mathbf{X}$ and from them, computes the (unbiased) estimators

$$\bar{E}_n^{(0)} = \frac{1}{n}\sum_{i=1}^{n} f^{(0)}(\mathbf{x_i}) \quad \text{and} \quad \bar{V}_n^{(0)} = \frac{1}{n-1}\sum_{i=1}^{n}(f^{(0)}(\mathbf{x_i}) - \bar{E}_n^{(0)})^2 \tag{6}$$

for mean and variance. Because we evaluate the model $f^{(0)}$ at each of the n samples, the computational cost of these estimators is $n \cdot w_0$. For a given computational budget p, standard MC allows for $n = \frac{p}{w_0}$ samples.

The mean squared error (MSE) of the mean estimator for a given p is [17]

$$\mathrm{MSE}[\bar{E}_n^{(0)}] = \frac{\sigma_0^2}{n} = \frac{\sigma_0^2}{p} w_0 \tag{7}$$

which results in a convergence rate of the root mean squared error (RMSE) in $\mathcal{O}(\frac{1}{\sqrt{n}})$. While the convergence of standard MC is independent of the stochastic dimensionality d of the problem and therefore makes this method especially well suited for higher-dimensional problems, the number of samples required to reduce the MSE sufficiently is also typically not feasible. This is in particular the case for our application in plasma microturbulence simulation.

3 Methodology

In this section, we summarize the techniques we apply to our test case in Section 4. We first give an overview of the MFMC algorithm and moreover, describe the construction process of the employed sparse grid surrogate and its dimension-reduced variants.

3.1 Multifidelity Monte Carlo Sampling

The idea of MFMC is, in the construction of estimators for mean and variance, to not just use the high-fidelity but also use evaluations from so-called low-fidelity models, i.e., surrogate models for the high-fidelity model under consideration. Similar to the high-fidelity model $f^{(0)}$, we will denote low-fidelity models by $f^{(1)}, \ldots, f^{(k)} : \boldsymbol{\mathcal{X}} \to \mathcal{Y}$ for $i = 1, \ldots, k$. By distributing model evaluations among high- and low-fidelity models, the MFMC algorithm exploits the often massively lower runtime w_i of the surrogate models $f^{(i)}$, which means that many more low-fidelity than high-fidelity model evaluations can be obtained in a given amount of time. Low-fidelity models typically are, however, not as accurate as the high-fidelity model. As an indicator for the accuracy of a low-fidelity model's output $f^{(i)}(\mathbf{X})$, the Pearson correlation coefficient w.r.t. the high-fidelity model $f^{(0)}$ is used:

$$\rho_{0,i} = \frac{\mathrm{Cov}[f^{(0)}(\mathbf{X}), f^{(i)}(\mathbf{X}))}{\sigma_0 \sigma_i} \tag{8}$$

where σ_i is the standard deviation of the low-fidelity output. By not only invoking the low-fidelity models but also the high-fidelity model, the method ensures that the resulting estimators stay unbiased.

We want to note that MFMC uses all the involved models only as a black box, implying that the way the low-fidelity models were constructed is irrelevant. For this reason, many different kinds of low-fidelity models can be used. For instance, the sparse grid approximation used in our simulations is a so-called *data-driven* surrogate, along with e.g. models obtained from machine learning approaches. Other types include *simplified* low-fidelity models such as reduced-physics models, and *projection-based* models from e.g. proper-orthogonal decomposition [15].

In the following, we describe the computation of MFMC estimators. This section follows [14] closely. The combination of high- and low-fidelity evaluations is done similarly to the construction of control variate estimators (see e.g. [13] for more detail).

Define the number of evaluations of each model $f^{(0)}, f^{(1)}, \ldots, f^{(k)}$ as $0 < m_0 \leq m_1 \leq \ldots \leq m_k$ and draw i.i.d. realizations

$$\mathbf{x_1}, ..., \mathbf{x_{m_k}}$$

of the input random vector $\mathbf{X}$. For $i = 0, 1, \ldots k$, evaluate model $f^{(i)}$ for the first m_i samples and obtain

$$f^{(i)}(\mathbf{x_1}), ..., f^{(i)}(\mathbf{x_{m_i}}).$$

With those evaluations, standard MC estimators for mean and variance are computed. For both the high- and low-fidelity models, compute estimators using all respective m_i model evaluations,

$$\bar{E}_{m_i}^{(i)} = \frac{1}{m_i} \sum_{j=1}^{m_i} f^{(i)}(\mathbf{x_j}) \quad \text{and} \quad \bar{V}_{m_i}^{(i)} = \frac{1}{m_i - 1} \sum_{j=1}^{m_i} (f^{(i)}(\mathbf{x_j}) - \bar{E}_{m_i}^{(i)})^2, \tag{9}$$

for $i = 0, 1, \ldots, k$. Additionally, for the low-fidelity models, we compute standard Monte Carlo estimators from only the first m_{i-1} model evaluations, i.e.,

$$\bar{E}^{(i)}_{m_{i-1}} = \frac{1}{m_{i-1}} \sum_{j=1}^{m_{i-1}} f^{(i)}(\mathbf{x_j}) \quad \text{and} \quad \bar{V}^{(i)}_{m_{i-1}} = \frac{1}{m_{i-1}-1} \sum_{j=1}^{m_{i-1}} (f^{(i)}(\mathbf{x_j}) - \bar{E}^{(i)}_{m_{i-1}})^2 \tag{10}$$

for $i = 1, \ldots, k$. Because the estimators from (10) use a subset of the evaluations used in (9), those estimators are dependent. The multifidelity mean estimator for $\mathbb{E}[f^{(0)}(\mathbf{X})]$ is then given as

$$\hat{E}_{MFMC} = \bar{E}^{(0)}_{m_0} + \sum_{i=1}^{k} \alpha_i (\bar{E}^{(i)}_{m_i} - \bar{E}^{(i)}_{m_{i-1}}), \tag{11}$$

and the estimator for $\mathrm{Var}[f^{(0)}(\mathbf{X})]$ by

$$\hat{V}_{MFMC} = \bar{V}^{(0)}_{m_0} + \sum_{i=1}^{k} \alpha_i (\bar{V}^{(i)}_{m_i} - \bar{V}^{(i)}_{m_{i-1}}), \tag{12}$$

where in both estimators we have coefficients $\alpha_1, \ldots, \alpha_k \in \mathbb{R}$. Let $\mathbf{m} = [m_0, m_1, \ldots, m_k]^T$ describe the number of model evaluations and $\mathbf{w} = [w_0, w_1, \ldots, w_k]^T$ the model runtimes. Then the cost of the MFMC estimators is given by $\mathbf{w}^T\mathbf{m}$, since every model $f^{(i)}$ is evaluated exactly m_i times.

The accuracy of the estimators in (11) and (12) depends on the choice of the number of model evaluations $\mathbf{m}$ and the coefficients $\alpha_1, \ldots, \alpha_k$. It is described in [14] how to choose these parameters in order to minimize the MSE of the mean estimator. Let the low-fidelity models be ordered such that

$$1 > |\rho_{0,1}| > \ldots > |\rho_{0,k}|$$

and their evaluation costs satisfy

$$\frac{w_{i-1}}{w_i} > \frac{\rho^2_{0,i-1} - \rho^2_{0,i}}{\rho^2_{0,i} - \rho^2_{0,i+1}} \quad \text{for} \quad i = 1, \ldots, k$$

with $\rho_{0,k+1} = 0$. Then we compute the optimal coefficients $\alpha^*_1, \ldots, \alpha^*_k$ as

$$\alpha^*_i = \frac{\rho_{0,i}\sigma_0}{\sigma_i} \quad \text{for} \quad i = 1, \ldots, k. \tag{13}$$

Furthermore, we define $\mathbf{r} = [r^*_0, r^*_1, \ldots, r^*_k]^T$ with

$$r^*_i = \sqrt{\frac{w_0(\rho^2_{0,i} - \rho^2_{0,i+1})}{w_i(1 - \rho^2_{0,1})}} \quad \text{for} \quad i = 0, \ldots, k.$$

Then we have as the optimal number of evaluations

$$m^*_0 = \frac{p}{\mathbf{w}^T\mathbf{r}^*} \quad \text{and} \quad m^*_i = m^*_1 r^*_i \quad \text{for} \quad i = 1, \ldots, k \tag{14}$$

for a given computational budget p. The MSE of the mean estimator that uses the optimal $\mathbf{m}$ and $\alpha_1, \ldots, \alpha_k$ can be shown to be [14]

$$\mathrm{MSE}[\hat{E}_{MFMC}] = \frac{\sigma^2_0}{p} \left(\sum_{i=0}^{k} \sqrt{w_i(\rho^2_{0,i} - \rho^2_{0,i+1}))} \right)^2. \tag{15}$$

From (15) it follows that for the MSE of the MFMC mean estimator to be low, the terms $w_i(\rho^2_{0,i} - \rho^2_{0,i+1})$ should be as small as possible. This means that the two factors influencing the accuracy of the estimator are the model runtimes and their interactions with each other. This implies also that in order to decide which models should be employed in computing a multifidelity estimator, (15) can be used to predict the accuracy of an estimator from a given set of models. A model selection process based on this observation is formalized in [14].

Note that we only consider the MSE of the mean estimator $\hat{E}_{MFMC}$. The number of model evaluations and coefficients as described above do not necessarily minimize the MSE of the variance estimator $\hat{V}_{MFMC}$. However, the MSE of a variance estimator using the same parameters will be similar, as shown in [16]. Moreover, determining optimal values for the computation of the variance estimator can be very unstable [16]. For these reasons we use the same parameters to compute both mean and variance estimators, which also means that the same set of model evaluations can be used.

3.2 Adaptive sparse-grid interpolation approximation

In the following, we summarize the construction of the adaptive sparse grid interpolation models that we use as surrogates for the high-fidelity model $f^{(0)}$. For more details, we refer the reader to [6; 5] where this approach was first presented.

Define $\mathcal{U}^i_{\ell_i}$ for $i = 1, \dots, d$ as numerical approximations of one-dimensional continuous linear operators $\mathcal{U}^i$ with the property that

$$\|\mathcal{U}^i - \mathcal{U}^i_{\ell_i}\| \xrightarrow{\ell_i \to \infty} 0$$

for a level $\ell_i \in \mathbb{N}$. From these 1D operators, we obtain a d-dimensional operator on the model $f^{(0)}$ by tensorization:

$$\mathcal{U}^d_{\boldsymbol{\ell}}[f^{(0)}] = \left(\bigotimes_{i=1}^{d} \mathcal{U}^i_{\ell_i} \right) [f^{(0)}].$$

Here, $\boldsymbol{\ell} = (\ell_1, \dots, \ell_d) \in \mathbb{N}^d$ is a multiindex that indicates how far the grid is refined in each of the dimensions. With the full-grid operators $\mathcal{U}^d_{\boldsymbol{\ell}}$, we can compute the hierarchical surpluses

$$\Delta^d_{\boldsymbol{\ell}}[f^{(0)}] = \sum_{\boldsymbol{z} \in \{0,1\}^d} (-1)^{|\boldsymbol{z}|_1} \mathcal{U}^d_{\boldsymbol{\ell}}[f^{(0)}],$$

where $|\boldsymbol{z}|_1 := \sum_{i=1}^d z_i$. Assuming that the $\mathcal{U}^i_{\ell_i}$ signify interpolation operators, the sparse grid interpolation approximation of the high-fidelity model $f^{(0)}$ can be given as

$$\mathcal{U}^d_{\mathcal{L}}[f^{(0)}] = \sum_{\boldsymbol{\ell} \in \mathcal{L}} \Delta^d_{\boldsymbol{\ell}}[f^{(0)}], \tag{16}$$

where $\mathcal{L} \subset \mathbb{R}^d$ is a multiindex set. More specifically, we define $\mathcal{U}^i_{\ell_i}$ as 1D Lagrange interpolation operators (see e.g. [1]) using weighted (L)-Leja points $\{x_n\}_{n=1}^{\ell_i}$ [10]. For a function $f : \mathcal{X}_i \to \mathbb{R}$, univariate Lagrange interpolation reads

$$\mathcal{U}^i_{\ell_i} : C^0(\mathcal{X}_i) \to \mathbb{P}_{P_{\ell_i}}, \quad \mathcal{U}^i_{\ell_i}[f](x) := \sum_{n=1}^{\ell_i} g(x_n) L_n(x) \tag{17}$$

with $\mathbb{P}_{P_{\ell_i}}$ the space of univariate polynomials of degree P_{ℓ_i} and Lagrange polynomials $\{L_n(x)\}_{n=1}^{\ell_i}$ of degree $n-1$ satisfying the interpolation condition $L_n(x_m) = \delta_{nm}$. For numerical stability reasons, (17) is implemented using the barycentric formula [1].

The (L)-Leja points $\{x_n\}_{n=1}^{\ell_i}$ are computed w.r.t. to the density function π_i:

$$\begin{aligned} x_1 &= \operatorname*{argmax}_{x \in \mathcal{X}_i} \pi_i(x) \\ x_n &= \operatorname*{argmax}_{x \in \mathcal{X}_i} \pi_i(x) \prod_{m=1}^{n-1} |(x - x_m)|, \quad n = 2, 3, \dots, \ell_i, \end{aligned}$$

Note that the (L)-Leja points are nested, meaning model evaluations used to compute the approximation $\mathcal{U}^i_{\ell_i}$ can be reused for the computation of $\mathcal{U}^i_{\ell_i+1}$ on the following level.

The composition of the multiindex set $\mathcal{L}$ reflects the adaptivity of the construction process of the model. The algorithm employed in this work is based on the dimension-adaptive algorithm of [7; 9]. The idea behind the algorithm mentioned and the method employed here is to exploit the underlying structure of the high-fidelity model (e.g. a lower intrinsic dimensionality) and not add unnecessary elements $\boldsymbol{\ell} \in \mathcal{L}$, which would make the process of computing the approximation more expensive. Let $\mathcal{L} = \mathcal{O} \dot{\cup} \mathcal{A}$, where $\mathcal{O}$ represents the *old* index set of multiindices that have already been processed, and $\mathcal{A}$ is the set of *active* multiindices that drive the adaptivity. In the beginning, we have $\mathcal{O} = \{\mathbf{1}\}$ and $\mathcal{A} = \{\mathbf{1} + \mathbf{e_i} \mid i = 1, \dots, d\}$. In the adaptive process, a refinement indicator $\epsilon(\boldsymbol{\ell})$ is computed for each $\boldsymbol{\ell} \in \mathcal{A}$ and the multiindex $\boldsymbol{\ell}$ with the largest value $\epsilon(\boldsymbol{\ell})$ is added to the set $\mathcal{O}$. All the so-called *forward neighbors* $\{\boldsymbol{\ell} + \mathbf{e_i} \mid i = 1, \dots, d\}$ are added to the active indices $\mathcal{A}$. The refinement process ends when a specified termination criterion is fulfilled.

In the algorithm of [6; 5], the refinement indicator ϵ is based on the computation of a Sobol' decomposition [19]. In [6; 5], it is shown that the interpolation operators $\mathcal{U}^d_{\boldsymbol{\ell}}[f^{(0)}]$ can be written in the form of a pseudo-spectral projection (PSP):

$$\mathcal{U}^d_{\boldsymbol{\ell}}[f^{(0)}] = \sum_{\boldsymbol{p} \in \mathcal{P}_{\boldsymbol{\ell}}} c_{\boldsymbol{p}} \Phi_{\boldsymbol{p}}(\mathbf{x}),$$

with orthogonal polynomials $\Phi_{\boldsymbol{p}}(\mathbf{x}) := \prod_{i=1}^d \Phi_i(x_i)$ w.r.t. to the pdf π and pseudo-spectral coefficients $c_{\mathbf{p}}$ for $\mathcal{P}_{\boldsymbol{\ell}} := \{\boldsymbol{p} \in \mathbb{N}^d : \mathbf{0} \leq \boldsymbol{p} \leq \boldsymbol{P_\ell} := (\ell_1 - 1, \ell_2 - 1, \dots, \ell_d - 1)\}$. This allows the representation of the hierarchical surpluses using spectral projections:

$$\Delta^d_{\boldsymbol{\ell}}[f^{(0)}] := \sum_{\boldsymbol{p} \in \mathcal{P}_{\boldsymbol{\ell}}} \Delta c_{\boldsymbol{p}} \Phi_{\boldsymbol{p}}(\mathbf{x}), \quad \Delta c_{\boldsymbol{p}} := \sum_{\boldsymbol{z} \in \{0,1\}^d} (-1)^{|\boldsymbol{z}|_1} c_{\boldsymbol{p} - \boldsymbol{z}}, \quad \Delta c_{\mathbf{0}} := c_{\mathbf{0}}.$$

Let now $\Delta \mathrm{Var}^i_{\boldsymbol{\ell}}[f^{(0)}]$ be the variance surplus due to the individual input dimensions and $\Delta \mathrm{Var}^{\mathrm{inter}}_{\boldsymbol{\ell}}[f^{(0)}]$ be the surplus due to the inputs' interaction. [6; 5] show that, using (3.2) and the equivalence of the terms in a spectral projection and a Sobol' decomposition [19], it is possible to compute both types of variance surpluses from the coefficients $\Delta c_{\mathbf{p}}$:

$$\Delta \mathrm{Var}^i_{\boldsymbol{\ell}}[f^{(0)}] := \sum_{\boldsymbol{p} \in \mathcal{J}_i} \Delta c^2_{\boldsymbol{p}}, \quad \mathcal{J}_i := \{\boldsymbol{p} \in \mathcal{P}_{\boldsymbol{\ell}} \setminus \{\mathbf{0}\} : \boldsymbol{p}_i \neq 0 \wedge \boldsymbol{p}_j = 0, \forall j \neq i\}, \tag{18}$$

$$\Delta \mathrm{Var}^{\mathrm{inter}}_{\boldsymbol{\ell}}[f^{(0)}] := \sum_{\boldsymbol{p} \in \mathcal{J}_{\mathrm{inter}}} \Delta c^2_{\boldsymbol{p}}, \quad \mathcal{J}_{\mathrm{inter}} := \{\boldsymbol{p} \in \mathcal{P}_{\boldsymbol{\ell}} \setminus \{\mathbf{0}\} : |\boldsymbol{p}|_1 \geq 2\}. \tag{19}$$

Because these surpluses also correspond to unnormalized Sobol' indices [18], they give an indication for how strong the contribution of an individual input parameter or the parameters' interaction is to the approximation value. Based on this information, the refinement indicator ϵ used here is defined. Let $\epsilon(\boldsymbol{\ell}) = s_{\boldsymbol{\ell}}$ for the so-called *sensitivity index* $s_{\boldsymbol{\ell}}$ for all $\boldsymbol{\ell} \in \mathcal{A}$. The user defines a set of local tolerances $\boldsymbol{\tau} := (\tau_1, \dots, \tau_d, \tau_{d+1})$ where $\tau_1, \dots, \tau_d$ correspond to one direction each and τ_{d+1} to all interactions. At the start, we have $s_{\boldsymbol{\ell}} = 0$. Then, for $i = 1, \dots, d, d+1$, the respective variance surplus (individual or from the interaction) is compared to the tolerance τ_i. If the surplus is larger than τ_i, $s_{\boldsymbol{\ell}}$ is incremented by 1. This means that after considering all surpluses, $s_{\boldsymbol{\ell}}$ will be an integer between 0 and $d+1$ that reflects the underlying structure of the high-fidelity model given by the different inputs. The grid is then refined in the direction of the subspace $\boldsymbol{\ell}$ with the highest sensitivity score. If two subspaces have the same score, the multiindex with the highest total variance surplus $\sum_{i=1}^d \Delta \mathrm{Var}^i_{\boldsymbol{\ell}}[f^{(0)}] + \Delta \mathrm{Var}^{\mathrm{inter}}_{\boldsymbol{\ell}}[f^{(0)}]$ from all inputs and their interaction is chosen. The adaptive refinement based on these sensitivity indices will be terminated if the contributions are too low for all $\boldsymbol{\ell} \in \mathcal{A}$, i.e., $s_{\boldsymbol{\ell}} = 0$, if a maximum level of refinement is reached, or if $\mathcal{A} \neq \emptyset$.

Dimension-reduced models The sparse grid interpolation surrogate constructed by the process described above will result in a model of the same input dimensionality as the underlying high-fidelity model. In our test case, we also use low-fidelity models of the same type that depend only on a subset of the original input parameters. In order to identify which parameters should be used in the construction of such dimension-reduced models, we again employ Sobol' indices.

From the pseudo-spectral coefficients $\{c_{\mathbf{p}}\}_{\mathbf{p}=\mathbf{0}}^{\mathcal{P}_\mathcal{L}}$ with total multivariate degree $\mathcal{P}_\mathcal{L}$ of the resulting approximation after the termination of the refinement process, the total Sobol' indices S_i^T for directions $i = 1, \ldots, d$ can be computed as

$$S_i^T = \frac{\sum_{\|\in\mathcal{K}_i} c_{\mathbf{k}}^2}{\sigma^2[\mathcal{U}_\mathcal{L}^d[f^{(0)}]]} \quad \text{with} \quad \sigma^2[\mathcal{U}_\mathcal{L}^d[f^{(0)}]] = \sqrt{\sum_{\mathbf{p}\neq\mathbf{0}} c_{\mathbf{p}}^2}, \tag{20}$$

where $\mathcal{K}_i = \{\mathbf{0} < \mathbf{k} \leq \mathcal{P}_\mathcal{L} \mid \mathbf{k}_i \neq 0\}$ and $\sigma^2[\mathcal{U}_\mathcal{L}^d[f^{(0)}]]$ is the variance of the computed sparse grid approximation of $f^{(0)}$ that can also be obtained from the PSP coefficients (as can its mean) [19]. Based on the computed total Sobol' indices, which measure the input's individual contributions as well as the inputs' interaction, it can be ascertained which input parameters are the most significant for the output of the model. Therefore, after constructing the fully dimensional surrogate with the described algorithm, we compute the resulting total Sobol' indices and then generate dimension-reduced models that use only a subset of parameters with the highest sensitivity indices. The remaining parameters - which according to the Sobol' indices contribute little to the model output - are fixed to a deterministic value.

This process results in surrogate models that are computationally cheap to produce and evaluate due to the efficient sensitivity scores approach that adapts to the high-fidelity model's structure. Moreover, the additional surrogates of lower input dimension are even less expensive to evaluate and still have high correlation coefficients w.r.t. to the high-fidelity model (see our results in the following section).

4 Numerical Results

In this section we present our results from performing uncertainty propagation using MFMC and adaptive sparse grid low-fidelity models on a realistic test case from plasma microturbulence analysis.

4.1 Realistic test case with 14 uncertain parameters

The test case we consider in this work is based on the scenarios considered in [3; 4] with a setting of parameters similar to a JET L-mode discharge. In [3; 4], the effects of fast ion particles on plasma microturbulence are examined. In many cases, steep ion temperature gradients (ITG) are the main source for microinstabilities that cause the turbulence. It has been shown that even at low densities, the introduction of fast ions to the plasma can suppress the microturbulence stemming from the ITG mode by changing the ions' response to the temperature and density gradients through a wave-particle resonance mechanism. This can already be observed in linear, electrostatic simulations [3; 4].

In this test case, we model five particle species: main ions, electrons, fast deuterium, Helium-3 and carbon (see [3; 4] for more detail on their roles in this scenario). Table 1 lists the 14 parameters we model as uncertain, 12 of which are related to the different types of particles. The other two parameters are the safety factor q and the magnetic shear $\hat{s} = \frac{r}{q}\frac{\delta q}{\delta r}$ with r being the radial coordinate labelling flux surfaces [6]. All parameters are modelled as independent uniform random variables. Other parameters of the simulations, such as the data related to the carbon impurities, are set to deterministic values. As in [3], we perform our simulations for a fixed wave number $k_y\rho_s = 0.5$ for which it is known that the ITG mode drives the microturbulence. Furthermore, we consider a flux-tube domain as described in Subsection 2.1 and employ a linearized Landau-Boltzmann collision operator. The high-fidelity output of interest for our simulations is the growth rate $\gamma[c_s/L_s]$ of the dominant eigenmode (see 2.1) where $c_s = \sqrt{T_e/m_i}$ is the ion sound speed (for ion mass m_i and characteristic length L_s) [6]. We perform both standard Monte Carlo sampling and MFMC to obtain mean and variance estimates of the high-fidelity model output for $\gamma[c_s/L_s]$.

4.2 Low-fidelity models

We apply the adaptive sparse grid interpolation method described in Subsection 3.2 to construct low-fidelity models to be used in the MFMC algorithm alongside the high-fidelity model given by GENE. For our test case with 14 stochastic parameters, we use three different sparse grid surrogates: $f^{(1)}$, which takes the same 14 parameters as the high-fidelity model $f^{(0)}$, and two dimension-reduced sparse grid models $f^{(2)}$ and $f^{(3)}$

	parameter	symbol	probability distribution
p_1	magnetic shear	$\hat{s}$	$\mathcal{U}(0.3920, 0.6534)$
p_2	safety factor	q	$\mathcal{U}(1.3230, 2.1705)$
p_3	ion log temperature gradient	$-L_s\delta_x \ln T_i$	$\mathcal{U}(3.4230, 5.7050)$
p_4	ion log density gradient	$-L_s\delta_x \ln n_i$	$\mathcal{U}(0.0047, 0.0079)$
p_5	ion temperature	T_i	$\mathcal{U}(0.7500, 1.2500)$
p_6	fast deuterium log temperature gradient	$-L_s\delta_x \ln T_D$	$\mathcal{U}(0.7742, 1.2904)$
p_7	fast deuterium log density gradient	$-L_s\delta_x \ln n_D$	$\mathcal{U}(3.5413, 5.9022)$
p_8	fast deuterium density	n_D	$\mathcal{U}(0.0450, 0.0750)$
p_9	fast deuterium temperature	T_D	$\mathcal{U}(7.3500, 12.2500)$
p_{10}	Helium-3 log temperature gradient	$-L_s\delta_x \ln T_{^3\mathrm{He}}$	$\mathcal{U}(5.5544, 9.2573)$
p_{11}	Helium-3 log density gradient	$-L_s\delta_x \ln n_{^3\mathrm{He}}$	$\mathcal{U}(0.3770, 0.6284)$
p_{12}	Helium-3 density	$n_{^3\mathrm{He}}$	$\mathcal{U}(0.0525, 0.0875)$
p_{13}	Helium-3 temperature	$T_{^3\mathrm{He}}$	$\mathcal{U}(9.0000, 15.0000)$
p_{14}	electron log temperature gradient	$-L_s\delta_x \ln T_e$	$\mathcal{U}(1.6695, 2.7825)$

Table 1: Input parameters and their distributions used in the 14-dimensional test case considered in this work. By $\mathcal{U}(a, b)$ we denote a uniform distribution on $[a, b]$.

$f^{(i)}$	d_i	$\mathbf{p_i}$	$\rho_{0,i}$	w_i	σ_i^2
$f^{(0)}$	14D	$\{p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8, p_9, p_{10}, p_{12}, p_{13}, p_{14}\}$	-	11574.86976	0.00397
$f^{(1)}$	14D	$\{p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8, p_9, p_{10}, p_{12}, p_{13}, p_{14}\}$	0.99726	11.67249	0.00404
$f^{(2)}$	9D	$\{p_1, p_3, p_5, p_8, p_9, p_{10}, p_{12}, p_{13}, p_{14}\}$	0.99694	0.38382	0.00405
$f^{(3)}$	6D	$\{p_1, p_3, p_5, p_{10}, p_{12}, p_{14}\}$	0.99007	0.06275	0.00406

Table 2: Overview of the employed low-fidelity with their dimensionality d_i, input parameters $\mathbf{p_i}$, estimated correlation coefficients $\rho_{0,i}$, evaluation costs w_i and variances σ_i^2.

that have nine and six input parameters respectively. The subsets of the original 14 parameters for the lower dimension low-fidelity models have been chosen according to the sensitivity indices obtained from the construction of the 14-dimensional model $f^{(1)}$ as depicted in Figure 1, meaning $f^{(2)}$ and $f^{(3)}$ depend on the respective nine and six parameters that yield the highest total Sobol' indices and are therefore the most important to the the model output. For all surrogates the tolerance parameter of the adaptive algorithm has been set to $\boldsymbol{\tau} = 10^{-6} \cdot \mathbf{1}$.

Table 2 summarizes the models we employ in the MFMC algorithm with their (sub)set $\mathbf{p_i}$ of input parameters. It also lists estimates of the models' correlation coefficient with the high-fidelity model, their runtimes and variance. While the variance can be obtained from the PSP coefficients as in (20), runtime and correlation coefficient w.r.t. $f^{(0)}$ are estimated experimentally. The runtime is averaged from measurements taken from $n = 1000$ model evaluations each, and the sample correlation coefficient is computed from n samples of high- and respective low-fidelity model $i = 1, \ldots, k$ as

$$\bar{\rho}_{0,i} = \frac{1}{\sigma_0 \sigma_i (n-1)} \sum_{j=1}^{n} (f^{(0)}(\mathbf{x_j}) - \bar{E}_n^{(0)})(f^{(i)}(\mathbf{y_j}) - \bar{E}_n^{(i)}) \tag{21}$$

for samples $\mathbf{x_1}, \ldots, \mathbf{x_n}$ of the 14-dimensional input domain, where $\mathbf{y_j} = (\mathbf{x_j})_{\mathbf{p_i}}$, i.e., $\mathbf{y_j}$ contains only the input dimensions relevant for the particular low-fidelity model $f^{(i)}$ of potentially lower dimension. It can be observed that all low-fidelity models have a very high correlation with the high-fidelity even in the case with only six out of the original 14 parameters. This can be explained due to the fact that many of the input dimensions are of little importance according to the sensitivity indices, which for eight of the input parameters are very close to 0 (see Figure 1). Moreover, the reduction of the dimensionality gives a large reduction in the evaluation costs of the resulting models (by a factor 186 from 14D to 6D).

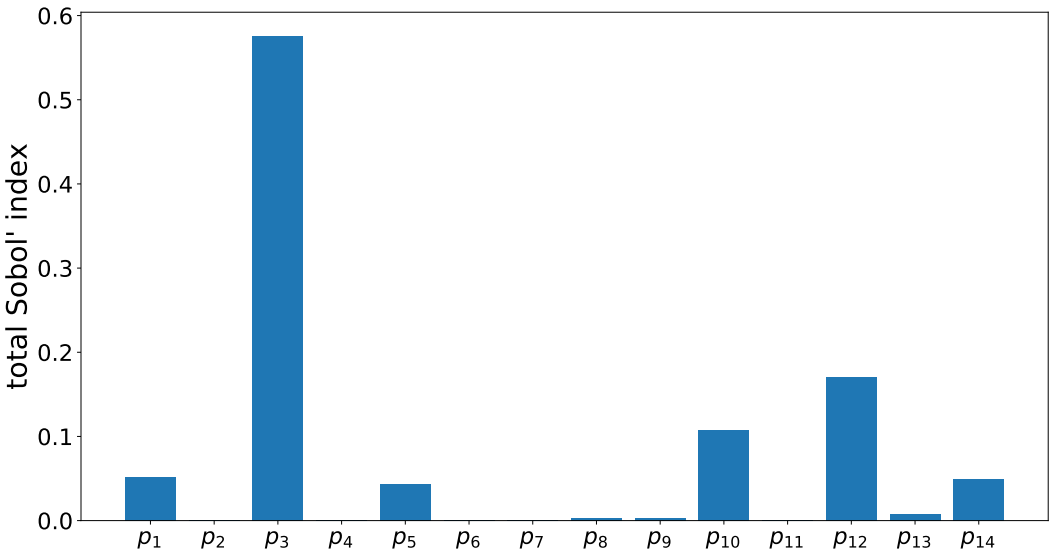

Figure 1: Total Sobol' indices of the 14 stochastic parameters computed by adaptive sparse grid interpolation.

4.3 Results

This section presents our results from performing standard MC and MFMC for the described test case with the listed low-fidelity models and GENE as the high-fidelity model. GENE simulations were conducted on the LRZ Linux cluster CoolMUC-2[2] where a single run is done on 240 cores on nine Intel Xeon E5-2690 nodes.

Reference estimator In order to judge the accuracy of the obtained estimators, we compute the MSE of the mean estimator. Note that we examine the mean estimator as opposed to the variance estimator since the number of model evaluations and coefficients needed to compute the estimators are optimized w.r.t. to the mean (see Subsection 3.1). To compute the MSE, we need a reference mean estimator, which for us, will represent the "true" mean value of the high-fidelity model output. A common approach for obtaining a reference would be to employ standard MC sampling for a very high number of samples. In our case however, the runtime of the high-fidelity model is already large for a single run (≈ 3 CPU hours, see Table 2), which makes computing an accurate reference estimator using standard MC infeasible due to its slow convergence. See Figure 2 for an idea of the convergence of standard MC estimators, the MSE of which decays linearly with the computational budget according to (7). For this reason, we turn again to MFMC and compute a reference using MFMC with a computational budget of $p = 10^7$ seconds and the low-fidelity models we already have. Evaluating the expected MSE of an MFMC mean estimator for all possible model combinations out of our three low-fidelity models according to (15) (as described by the model selection algorithm from [14]), suggests that the most accurate reference will be obtained by combining all three low-fidelity models. With this method, we obtain a reference mean $\mu_{\text{ref}} = 0.1973$ using MFMC with all three low-fidelity models $f^{(1)}, f^{(2)}, f^{(3)}$. This also gives a reference $\sigma^2_{\text{ref}} = 0.003979$ for the variance estimator, though, as described, we do not use it in MSE computations.

MSE of MFMC estimators We perform MFMC and standard MC for a set of computational budgets $p \in \{5 \cdot 10^4, 10^5, 5 \cdot 10^5, 10^6\}$. These were chosen based on the high-fidelity runtime $w_0 = 11574.86976$, which, for a budget of $5 \cdot 10^4$ allows for at most 4 evaluations, while a budget of $p = 10^4$ would be smaller than w_0. Based on (15), we can expect a linear decay of the resulting MSE with increasing budgets. This is depicted

[2] `https://doku.lrz.de/display/PUBLIC/High+Performance+Computing`

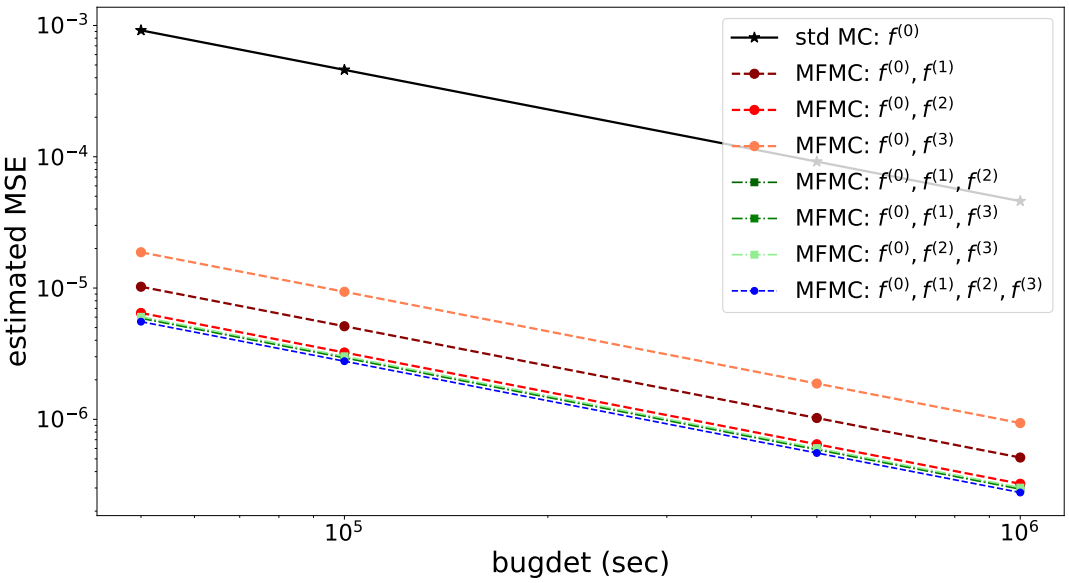

Figure 2: Expected decay of the MSE of the standard MC and MFMC mean estimators.

in Figure 2, which shows already that (a) MFMC with any choice of our low-fidelity models should be by at least two orders of magnitude more efficient than standard MC, and (b) that the different low-fidelity models compliment each other well since the MSE of MFMC with any combination of at least two low-fidelity models is, at least in theory, lower than that from using only one low-fidelity model.

Experimentally, we estimate an MSE as an average over 10 independent mean estimates $\hat{E}_{A,j}, j = 1 \ldots, 10$ obtained from the MC or MFMC algorithms:

$$\mathrm{MSE}[\hat{E}_A] = \frac{1}{10}\sum_{j=1}^{10}(\mu_{\mathrm{ref}} - \hat{E}_{A,j})^2, \tag{22}$$

where $A \in \{MC, MFMC\}$. Figure 3 reports the MSE of the mean estimator computed by standard MC and MFMC with all possible combinations of the available low-fidelity models w.r.t. to the reference μ_{ref}. Note that for MFMC with the 14-dimensional low-fidelity model $f^{(1)}$ not all MFMC estimators could be computed due to time constraints on the cluster used for the high-fidelity simulations. The results we obtain for the estimators' MSE confirm the theoretical considerations from Figure 2 in that the MSE of multifidelity mean estimators is by at least two orders of magnitude lower than that of standard MC estimators. Furthermore, the linear decay of the MSE with increasing budgets can be observed as well. It should be noted that, due to the random nature of the algorithm, we have some estimators that do not completely satisfy the predictions from the theoretical analysis of the MSE, e.g. the MSE using MFMC with only the nine-dimensional low-fidelity model is lower than that of any other model combinations. For this reason, more than 10 runs of the algorithm should be conducted to fully verify the analytical computations. Finally, Tables 3 and 4 list the estimators we computed for mean and variance using the different model combinations and for different computational budgets.

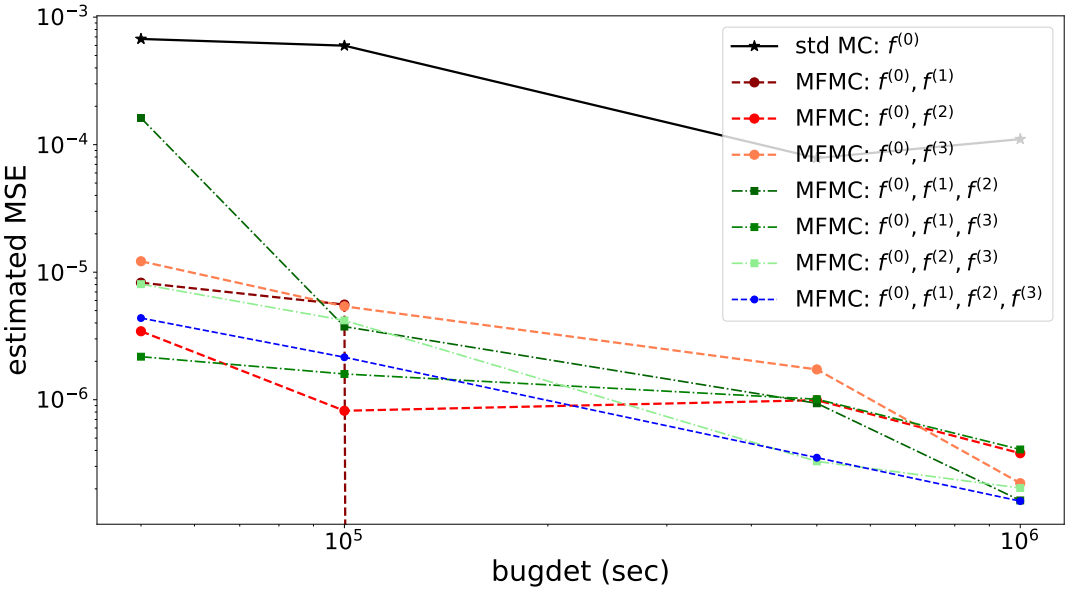

Figure 3: Estimated MSE of standard MC and MFMC mean estimators.

	$p = 5 \cdot 10^4$	$p = 10^5$	$p = 5 \cdot 10^5$	$p = 10^6$
$f^{(0)}$	0.2114	0.2044	0.1955	0.1957
$f^{(0)}, f^{(1)}$	0.1981	0.1960	-	-
$f^{(0)}, f^{(2)}$	0.1970	0.1977	0.1971	0.1972
$f^{(0)}, f^{(3)}$	0.1989	0.1967	0.1975	0.1972
$f^{(0)}, f^{(1)}, f^{(2)}$	0.1989	0.1979	0.1967	0.1972
$f^{(0)}, f^{(1)}, f^{(3)}$	0.1978	0.1977	0.1971	0.1974
$f^{(0)}, f^{(2)}, f^{(3)}$	0.1962	0.1974	0.1972	0.1975
$f^{(0)}, f^{(1)}, f^{(2)}, f^{(3)}$	0.1964	0.1979	0.1971	0.1975

Table 3: **Mean** estimators obtained from standard MC, i.e., only using $f^{(0)}$, and MFMC using the listed sets of low-fidelity models.

	$p = 5 \cdot 10^4$	$p = 10^5$	$p = 5 \cdot 10^5$	$p = 10^6$
$f^{(0)}$	0.002762	0.004253	0.004207	0.004216
$f^{(0)}, f^{(1)}$	0.003921	0.004110	-	-
$f^{(0)}, f^{(2)}$	0.003891	0.003862	0.004040	0.003984
$f^{(0)}, f^{(3)}$	0.003994	0.004003	0.004041	0.003991
$f^{(0)}, f^{(1)}, f^{(2)}$	0.004924	0.003934	0.004007	0.004018
$f^{(0)}, f^{(1)}, f^{(3)}$	0.003681	0.003943	0.004090	0.003943
$f^{(0)}, f^{(2)}, f^{(3)}$	0.003940	0.003916	0.004014	0.003968
$f^{(0)}, f^{(1)}, f^{(2)}, f^{(3)}$	0.004093	0.003940	0.004015	0.003960

Table 4: **Variance** estimators obtained from standard MC, i.e., only using $f^{(0)}$, and MFMC using the listed sets of low-fideliy models.

5 Conclusions and outlook

In this work, we have given an overview of a set-up that can used to perform uncertainty propagation for plasma microturbulence problems using the gyrokinetic code GENE. We used the MFMC technique to counteract the slow convergence of standard MC sampling, which is especially relevant for the typically computationally expensive plasma microturbulence simulations. Furthermore, as standard MC, MFMC being a sampling-based approach makes it particularly interesting for often high-dimensional plasma physics problems. To use as low-fidelity models in MFMC, we constructed sensitivity-driven adaptive sparse grid interpolation approximations of the high-fidelity model with the algorithm presented in [6; 5]. We created dimension-reduced surrogates based on the sensitivity information given by this process. These surrogates proved to be very useful for our application as they are accurate w.r.t. to the high-fidelity model and have a much lower runtime.

Moreover, we applied the MFMC algorithm with the described low-fidelity models to a realistic test case from plasma microturbulence with 14 stochastic parameters. From these test runs we obtained mean estimators with a MSE at least two orders of magnitude lower than standard Monte Carlo estimators for the same computational budget. These results illustrate the suitability of the presented approach using MFMC and sparse grid approximations for performing uncertainty propagation in plasma microturbulence problems.

As mentioned at the end of Section 4.3, the computed MSE does not always completely correspond to what would be expected from analytical computations. For this reason, a part of the future work will include performing more runs of the MC and MFMC algorithms to obtain clearer estimates. Furthermore, a different reference estimator could be used to compute the MSE. As mentioned in the previous section, standard MC is not feasible to accomplish this. However, another MFMC reference could be computed using a higher computational budget than the one we used. Finally, by finding additional or different low-fidelity models for the considered test case, it might be possible to decrease the MSE further.

References

[1] J.-P. Berrut and L. N. Trefethen, *Barycentric lagrange interpolation*, SIAM review, 46 (2004), pp. 501–517, https://doi.org/10.1137/S0036144502417715.

[2] A. Brizard and T. Hahm, *Foundations of nonlinear gyrokinetic theory*, Reviews of Modern Physics, 79 (2007).

[3] A. Di Siena, T. Görler, H. Doerk, E. Poli, and R. Bilato, *Fast-ion stabilization of tokamak plasma turbulence*, Nuclear Fusion, 58 (2018), p. 054002, https://doi.org/10.1088/1741-4326/aaaf26, https://doi.org/10.1088%2F1741-4326%2Faaaf26.

[4] A. Di Siena, T. Görler, E. Poli, R. Bilato, H. Doerk, and A. Zocco, *Resonant interaction of energetic ions with bulk-ion plasma micro-turbulence*, Physics of Plasmas, 26 (2019), p. 052504, https://doi.org/10.1063/1.5087203, https://doi.org/10.1063/1.5087203, https://arxiv.org/abs/https://doi.org/10.1063/1.5087203.

[5] I.-G. Farcaș, *Context-aware Model Hierarchies for Higher-dimensional Uncertainty Quantification*, dissertation, Technische Universität München, München, 2020.

[6] I.-G. Farcaș, T. Görler, H.-J. Bungartz, F. Jenko, and T. Neckel, *Sensitivity-driven adaptive sparse stochastic approximations in plasma microinstability analysis*, Journal of Computational Physics, 410 (2020).

[7] T. Gerstner and M. Griebel, *Dimension–adaptive tensor–product quadrature*, Computing, 71 (2003), pp. 65–87, https://doi.org/10.1007/s00607-003-0015-5.

[8] T. Görler, X. Lapillonne, S. Brunner, T. Dannert, F. Jenko, F. Merz, and D. Told, *The global version of the gyrokinetic turbulence code gene*, Journal of Computational Physics, 230 (2011), pp. 7053 – 7071.

[9] M. Hegland, *Adaptive sparse grids*, in Proc. of 10th Computational Techniques and Applications Conference CTAC-2001, K. Burrage and R. B. Sidje, eds., vol. 44, 2003, pp. C335–C353.

[10] P. Jantsch, C. G. Webster, and G. Zhang, *On the Lebesgue constant of weighted Leja points for Lagrange interpolation on unbounded domains*, IMA J. Numer. Anal., 39 (2018), pp. 1039–1057.

[11] J. Konrad, *Multifidelity Monte Carlo Sampling in Plasma Microturbulence Analysis*, bachelor's thesis, Technische Universität München, 2019.

[12] C. Kowitz and M. Hegland, *The sparse grid combination technique for computing eigenvalues in linear gyrokinetics*, Procedia Computer Science, 18 (2013), pp. 449–458.

[13] B. L. Nelson, *On control variate estimators*, Computers & Operations Research, 14 (1987), pp. 219 – 225, https://doi.org/10.1016/0305-0548(87)90024-4.

[14] B. Peherstorfer, K. Willcox, and M. Gunzburger, *Optimal model management for multifidelity Monte Carlo estimation*, SIAM Journal on Scientific Computing, 38 (2016), pp. A3163–A3194, https://doi.org/10.1137/15M1046472.

[15] B. Peherstorfer, K. Willcox, and M. Gunzburger, *Survey of multifidelity methods in uncertainty propagation, inference, and optimization*, SIAM Review, 60 (2018), pp. 550–591, https://doi.org/10.1137/16M1082469.

[16] E. Qian, B. Peherstorfer, D. O'Malley, V. Vesselinov, and K. Willcox, *Multifidelity Monte Carlo estimation of variance and sensitivity indices*, SIAM/ASA Journal on Uncertainty Quantification, 6 (2018), pp. 683–706, https://doi.org/10.1137/17M1151006.

[17] R. C. Smith, *Uncertainty Quantification: Theory, Implementation, and Applications*, Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2013.

[18] I. M. Sobol, *Global sensitivity indices for nonlinear mathematical models and their monte carlo estimates*, Math. Comput. Simul., 55 (2001), pp. 271–280, https://doi.org/https://doi.org/10.1016/S0378-4754(00)00270-6.

[19] B. Sudret, *Global sensitivity analysis using polynomial chaos expansions*, Reliab. Eng. Syst. Safe., 93 (2008), pp. 964–979, https://doi.org/https://doi.org/10.1016/j.ress.2007.04.002.