

Stochastic finite element approaches using derivative information for uncertainty quantification

Oleg Roderick,^{*} Mihai Anitescu,[†] and Paul Fischer[‡]

Mathematics and Computer Science Division, Argonne National Laboratory

(Dated: October, 2008)

In this work we describe a stochastic finite-element-based approach for analyzing the performance of a complex system that is described by a mathematical model depending on several stochastic parameters.

We construct a surrogate model as a goal-oriented projection onto an incomplete space of polynomials; find coordinates of the projection by collocation; and use derivative information to significantly reduce the number of the required collocation sample points. The simplified model can be used as a control variate to significantly reduce the sample variance of the estimate of the goal.

For our test model, we take a steady-state description of heat distribution in the core of the nuclear reactor core, and as our goal we take the maximum centerline temperature in a fuel pin. For this case, the resulting surrogate model is substantially more computationally efficient than random sampling or approaches that do not use derivative information, and it has greater precision than linear models.

Keywords: uncertainty; stochastic finite element; derivative; control variate

I. INTRODUCTION

The quantitative characterization and reduction of uncertainties in large models is an important area of research, related to statistical analysis of random phenomena and to physical study of complex systems. One of its most general tasks is to describe the influence of random inputs on an arbitrary but given output. Any improvement to the existing tools of uncertainty quantification will have both mathematical and industrial benefits.

This is particularly true in the area of nuclear reactor design and control, where greater engineering precision results in significant preservation of resources. The usual difficulties in modeling the work of the nuclear reactor include the large size of the associated systems of equations, the nonlinearity, and the implicit dependence of equations on parameters. As a result, one can afford to run the computational model of a nuclear reactor only for a small number of scenarios involving the values of its physical parameters. In addition, although the information on the behavior of parameters is available in formats convenient for experimental physics and engineering purposes, such formats are not necessarily appropriate for uncertainty analysis.

As a test case for our method, we consider a mathematical model of heat transport in the nuclear reactor core. We create a computationally efficient method to describe the dependence of a merit function, the maximum centerline temperature, on the model uncertainties.

A. Problem Definition

We first present our mathematical problem in the most general form. Consider an arbitrary system of (discretized, algebraic-differential) equations with main variables $T = (T_1, T_2, \dots, T_n)$ and *intermediate parameters* $\hat{R} = (R_1, R_2, \dots, R_N)$:

$$F(T, \hat{R}) = 0 \quad (\text{I.1})$$

Here and in the sequel, we denote by R a generic intermediate parameter, one of R_1, R_2, \dots, R_N , and by \hat{R} the vector of all such parameters.

The parameter set \hat{R} is not independent. It is related to the variables by a set of expressions

$$R = R(T) + \Delta R(T; \alpha) \quad (\text{I.2})$$

with the experimental error $\Delta R(T; \alpha)$, which is also dependent on temperature and on a set of parameters α that quantifies the uncertainty. The parameters α become the *primary* uncertainty parameters. Then the structural equation of the nonlinear system (I.1) becomes

$$F(T, \hat{R}(T; \alpha)) = 0 \quad (\text{I.3})$$

Strictly speaking, equation (I.3) now results in the primary variable T being a function of (α) and not of \hat{R} (which is itself a function of temperature). To abide by the physical meaning of the respective parameters R , we may still write $T = T(\hat{R})$.

For a given *merit function*

$$\mathcal{J} = \mathcal{J}(T) : R^n \rightarrow R \quad (\text{I.4})$$

we need to find the influence of uncertainties in the parameters on the uncertainty of the output. To find the effects of the uncertainty on the merit function \mathcal{J} ,

$$\Delta \mathcal{J} = \mathcal{J}(T(R)) - \mathcal{J}(T(R + \Delta R)) \quad (\text{I.5})$$

^{*}Electronic address: roderick@mcs.anl.gov; Also at: Department of Mathematics, Portland State University

[†]Address all correspondence to this author. anitescu@mcs.anl.gov

[‡]Electronic address: fischer@mcs.anl.gov

we express an output as a function of uncertainties of the inputs, represented by the parameters α :

$$\mathcal{J} \approx \hat{\mathcal{J}} = \hat{\mathcal{J}}(T(R(T) + \Delta R(T; \alpha))) = \hat{\mathcal{J}}(\alpha) \quad (\text{I.6})$$

This representation, or *surrogate model*, is created by using a *stochastic finite-element approximation* [12, 13, 25]. That is, we create a set of polynomials in α , $\{\Psi_q(\alpha)\}_{q \in Q}$, and we define $\hat{\mathcal{J}}(\alpha) = \sum_{q \in Q} x_q \Psi_q(\alpha)$. The coefficients x_q are obtained by requiring that the function and derivative values of the surrogate model $\hat{\mathcal{J}}(\alpha)$ match the ones of the real model $\mathcal{J}(\alpha)$, in a least-squares sense.

The main motivation for our approach is the observation that, for most computational multiphysics applications, there exist methods for computing the full gradient of a merit function \mathcal{J} that require an effort comparable to or even smaller than the effort of computing \mathcal{J} itself. We have in mind primarily the adjoint sensitivity method [5] and reverse automatic differentiation approach [14]. Therefore, as the number of parameters increases, using \mathcal{J} derivative information is appealing because it provides more information than does the evaluation of \mathcal{J} by a factor equal to the number of parameters, while having a similar overall comparable cost. We thus expect to have to compute the solution of (I.3) far fewer times.

B. Comparison with Previous Approaches

Among the most commonly encountered approaches for uncertainty quantification in engineering systems are Monte Carlo methods [8] and first-order sensitivity approaches [5]. The first approach is flexible, but slow, whereas the second approach is reasonably efficient, but inexact. Our approach is a hybrid between a sampling approach and a first-order sensitivity approach. We thus expect it to inherit the flexibility of the first and the efficiency of the second. Such methods have been recently successfully used with some environmental and scientific applications [17, 19], but our work includes some key differences, as we explain below.

The only other method, that we are aware of, for uncertainty quantification of engineering systems that is a hybrid between a probabilistic-type method and a sensitivity approach is based on Gaussian processes [24]. The difficulty with that approach is that one needs a good Gaussian process prior description of the uncertainty. Such priors have been produced in the case of essentially stationary spatial uncertainty in low dimensions, but they have not been generated for arbitrary nonlinear maps like the ones described here. For such random variables, any form of approximate stationarity is highly unlikely to hold.

Key parts of our method are related to polynomial approximations of complex systems with uncertain parameters, stochastic finite element methods (SFEM) [1, 2, 12, 13, 25]. A method that has recently attracted

major interest is the one of SFEM-Galerkin methods [1, 13]. Such methods, however, must approximate the state variables before they can approximate a merit function. That would require substantially more storage compared to our method, as well as a specialized solver for the resulting nonlinear problem. SFEM collocation methods [3, 4, 27] are nonintrusive and do not need specialized solvers, but they still use a state variable approximation, with the same memory requirements for representation as SFEM-Galerkin. Surface response approximation [6, 17, 19, 23] use only sampled information about a merit function to construct the response surface, though rarely using derivative information as well, as we do here, [17, 19].

Hermite interpolation approaches, to which our work is also related, have rarely been used for spaces of dimension comparable to the one described here (12 dimensions) [21]. Some reasons for this are described in Section II C.

Nonetheless, we have yet to develop theory for the stability for this approach. We therefore describe a control variate approach [20] where our SFEM/derivative approximation method can be used as an accelerator for a sampling approach that depends on efficiency but not on correctness on the quality of our approximation.

Our method shares with the references [17, 19] the feature that it is a hybrid between sampling methods and sensitivity methods. Nonetheless, our method is applied to a very different application, thermo-hydraulic behavior of sodium-cooled nuclear reactor [7]. A representation of the uncertainty that would be usable in our context is not provided in prior work [10, 11]; in Section III C we derive one that is consistent with all available information. The mechanism for reducing the size of the basis to make the representation tractable is quite different from [19] (reference [17] uses a full basis of a given degree). In addition, in [19] a linear model turned out to be sufficient, which is not the case here, as we demonstrate in Section IV. Moreover, the issue of eliminating the bias in the polynomial approximation was not proposed in any of these references; we describe a method to do so by using a control variate.

The rest of the paper is organized as follows: In Section II, we describe a stochastic finite-element collocation procedure with derivative information. In Section III, we introduce a simplified model of the nuclear reactor core, describe its uncertainties, and obtain the derivative information. In Section IV, we present the numerical results. In Section V, we summarize the study and discuss research plans.

II. STOCHASTIC FINITE-ELEMENT METHOD WITH INCOMPLETE BASIS

Choose a set Ψ of multivariable orthogonal polynomials of the parameter set $S = \{s_i\}$. A subset $\{\Psi_q\}$ is used

to approximate the merit function:

$$\mathcal{J} \approx \hat{\mathcal{J}} = \sum_q x_q \Psi_q \quad (\text{II.1})$$

$$\Psi_q = p_0^q + \sum_i p_i^q s_i + \sum_{i,j} p_{ij}^q s_i s_j + \sum_{i,j,k} p_{ijk}^q s_i s_j s_k + \dots \quad (\text{II.2})$$

The coefficients $\{p^q\} = \{p^q, p_{ij}^q, p_{ijk}^q, \dots\}$ are chosen so that $\{\Psi_q\}$ satisfy the orthogonality condition in some probability measure π :

$$\int \Psi_q \Psi_w d\pi = 0, \quad w \neq q \quad (\text{II.3})$$

The basis Ψ_q is the *stochastic finite-element basis*, and the approximation $\hat{\mathcal{J}} = \sum x_q \Psi_q$ is called the SFEM approximation. The term stochastic refers to the fact that, if the variables s are random, then the random variable approximation (II.1) can be used approximate the statistical properties of \mathcal{J} [13]. Our paper is concerned mainly with the procedure by which the approximation $\hat{\mathcal{J}}$ is constructed. This approximation procedure is deterministic in nature.

This problem of basis (or *polynomial chaos*) construction has standard solutions [12, 25]. For the Gaussian measure $\pi = \exp(-\frac{1}{2}|S|^2)$, the formal definition of the basis functions reads as follows:

$$\Psi_q(s_1, s_2, s_3, \dots) = \frac{\partial^{(\cdot)}}{\partial s_1 \partial s_2 \partial s_3 \partial \dots} \exp(-\frac{1}{2}|S|^2) \quad (\text{II.4})$$

which can be shown to be equivalent to a family of Hermite polynomials. We will refer to the Hermite polynomial basis in the following form:

$$\begin{aligned} \Psi(s_1, s_2, s_3, \dots) &= H(s_1)H(s_2)H(s_3)\dots = \prod_j H^{(k_j)}(s_j) \\ H^{(0)}(s) &= 1 \\ H^{(1)}(s) &= 2s \\ H^{(2)}(s) &= 4s^2 - 2 \\ H^{(3)}(s) &= 8s^3 - 12s \\ H^{(4)}(s) &= 16s^4 - 48s^2 + 12 \\ &\dots \end{aligned} \quad (\text{II.5})$$

Collocation is the most convenient way to determine the coefficients $x = (x_1, x_2, \dots, x_q, \dots)^T$ for the expression (II.1) [3, 27]. In this context, collocation is identical to the surface response approach [17]. The collocation conditions are enforced at the nodes $\{S_i\}$, $i = 1, 2, \dots, m$, resulting in a system of linear equations

$$\begin{pmatrix} \Psi(S_1) \\ \Psi(S_2) \\ \vdots \\ \Psi(S_m) \end{pmatrix} x = \begin{pmatrix} \mathcal{J}(S_1) \\ \mathcal{J}(S_2) \\ \vdots \\ \mathcal{J}(S_m) \end{pmatrix} \quad (\text{II.6})$$

where $\{S_i\}$, $S_i = (s_{1i}, s_{2i}, \dots, s_{Ni})$ is a sample of points from the collocation parameter space and $\{\mathcal{J}(S_i)\}$ is the

set of the corresponding exact values of the output. The rows of the collocation matrix are defined by

$$\Psi(S_i) = (\Psi_1(S_i), \Psi_2(S_i), \dots) \quad (\text{II.7})$$

It is also possible to obtain x by using orthogonality of the basis functions, by Galerkin [1, 2, 13]. This approach would involve creating a weak form of (I.3) and solving it coupled with (II.1). That approach is more robust for surfaces with a relatively low degree of smoothness, but it is harder to implement because it requires a major intrusion in the application code and the development of different nonlinear equations solvers. Therefore, in this work we address only collocation approaches.

A. Ranking the Parameters and Selecting the Polynomial Degrees

The nonlinear dependence of the output on the parameters implies the use of high-degree polynomials. This could result in a large number of polynomials in basis even for a moderate number of variables. For example, we need 66 polynomials for a basis on 10 parameters and maximal degree 2, 286 polynomials for degree 3, and 10,286 polynomials for degree 4. For 15 parameters and degree 4, we already need 51,441 polynomials.

For computational efficiency, we use a small number of variables. When physically justified, various uncertainty effects should be lumped together and attributed to a small list of independent random inputs [6]. Moreover, we use a smaller basis, which includes high-order polynomials only in some variables (or, more generally, only in some ‘‘important’’ directions in the parameter space). This basis can even be constructed adaptively, as long as the adaptive procedure does not require too many additional evaluations of $\mathcal{J}(S)$.

In the presense of information on statistical distribution of uncertain parameters, it is possible to explicitly estimate their influence on the distribution of the output. For an illustration, assume normal, zero-mean distribution with diagonal covariance matrix $\Sigma = (\sigma_{ij})$: $S \sim N[0, \Sigma]$. Create the linear approximation of the output function around $S = (0, 0, \dots)$:

$$\mathcal{J}(S) \approx \Psi_0(0) + \sum_i \Psi'_i(0)s_i + \frac{1}{2} \sum_{i,j} \Psi''_{ij}(0)s_i s_j \quad (\text{II.8})$$

Let's compute the variance:

$$\text{var}(\mathcal{J}) = E(\mathcal{J}(S)^2) - E(\mathcal{J}(S))^2 \quad (\text{II.9})$$

From the definition of variance σ_{ii}^2 , and by zero-mean assumption

$$E(s_i^2) = \sigma_i^2, \quad (\text{II.10})$$

Then, the variance of \mathcal{J} can be approximated by

$$\text{var}[\mathcal{J}(S)] = \sum_i \Psi'_i(0)^2 \sigma_i^2. \quad (\text{II.11})$$

A simple (first-order) test may look as follows. If we observe

$$\left| \frac{\partial \mathcal{J}}{\partial s_i} \sigma_i^2 \right| > \left| \frac{\partial \mathcal{J}}{\partial s_j} \sigma_j^2 \right| \quad (\text{II.12})$$

at the neutral setup of the system, $S = (0, 0..)$, we conclude that the parameter s_i is relatively more important for correct representation of uncertainty of the output. In the construction of the basis, we will use a higher degree in s_i than in s_j .

In the case where the statistical information is difficult to determine or manipulate, the importance test (II.12) may be reduced to

$$\left| \frac{\partial \mathcal{J}}{\partial s_i} \right| > \left| \frac{\partial \mathcal{J}}{\partial s_j} \right| \quad (\text{II.13})$$

We point out that our approximation $\hat{\mathcal{J}}$ will ultimately be validated by a control variate. Therefore the approximation $\hat{\mathcal{J}}$ does not need to be of a prescribed quality for correctness of the method. Hence, we will use (II.13) for our ranking, since it is easier to implement.

In our numerical experiments, we assume an existence of a modest computational budget limiting the number of direct evaluations of $\mathcal{J}(S)$. We start with a full basis of degree 3 and use (II.13) to put the variables, depending on their importance, into groups I, II, and III of sizes n_I, n_{II}, n_{III} . We allow the polynomials in the variables from group III to have maximal degree 3, and only keep the polynomials of degree 1 in the variables from group I. This (arbitrary) procedure appears to be sufficient to decide on the practicality of the polynomial basis truncation. We take $n_I > n_{II} \gg n_{III}$.

The sizes of the groups, their numbers, and values of maximal degrees may be adjusted to suit the modeling needs as well as the computational resource limits.

B. Using Derivative Information to Compute the Collocation

We compute derivatives of the output function and add derivative information to the collocation matrix. Then a single right-side entry $\mathcal{J}(S_i)$, will generate a subcolumn of entries, $(\frac{\partial \mathcal{J}(S_i)}{\partial s_j})$, providing right-side information for several rows at once.

The computational time saved by using fewer sample points may be partially spent on oversampling, or sensitivity analysis, to improve the performance of the model.

Matching the values of the derivatives with the corresponding polynomial derivatives on the right, we build

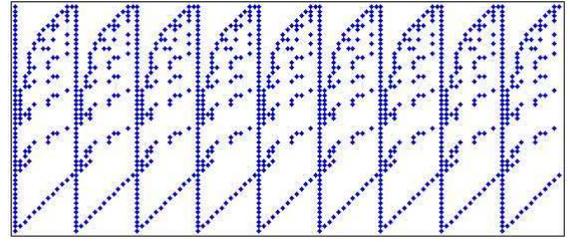


FIG. II.1: SFEM with derivative information: a typical collocation matrix sparsity pattern (transposed)

an extended system of collocation equations:

$$\begin{pmatrix} \Psi_1(S_1) & \Psi_2(S_1) & \cdots & \Psi_k(S_1) \\ \vdots & \vdots & \vdots & \vdots \\ \Psi_1(S_m) & \Psi_2(S_m) & \cdots & \Psi_k(S_m) \\ \frac{\partial \Psi_1(S_1)}{\partial s_1} & \frac{\partial \Psi_2(S_1)}{\partial s_1} & \cdots & \frac{\partial \Psi_k(S_1)}{\partial s_1} \\ \frac{\partial \Psi_1(S_1)}{\partial s_2} & \frac{\partial \Psi_2(S_1)}{\partial s_2} & \cdots & \frac{\partial \Psi_k(S_1)}{\partial s_2} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \Psi_1(S_1)}{\partial s_n} & \frac{\partial \Psi_2(S_1)}{\partial s_n} & \cdots & \frac{\partial \Psi_k(S_1)}{\partial s_n} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \Psi_1(S_2)}{\partial s_1} & \frac{\partial \Psi_2(S_2)}{\partial s_1} & \cdots & \frac{\partial \Psi_k(S_2)}{\partial s_1} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial \Psi_1(S_m)}{\partial s_n} & \frac{\partial \Psi_2(S_m)}{\partial s_n} & \cdots & \frac{\partial \Psi_k(S_m)}{\partial s_n} \end{pmatrix} X = \begin{pmatrix} \mathcal{J}(S_1) \\ \vdots \\ \mathcal{J}(S_m) \\ \frac{\partial \mathcal{J}(S_1)}{\partial s_1} \\ \frac{\partial \mathcal{J}(S_1)}{\partial s_2} \\ \vdots \\ \frac{\partial \mathcal{J}(S_1)}{\partial s_n} \\ \frac{\partial \mathcal{J}(S_2)}{\partial s_1} \\ \vdots \\ \frac{\partial \mathcal{J}(S_m)}{\partial s_n} \end{pmatrix} \quad (\text{II.14})$$

This system includes $m + mn$ equations based on m sample points. In comparison with an approach without derivative information, the minimal required number of sample points drops by a factor of $\frac{1}{1+n}$. An additional advantage of the extended collocation system is its sparse structure (see Figure II.1), allowing fast linear algebra operations, and having a relatively lower chance of accidental rank deficiency. It turns out, as explained and demonstrated in Section IV, that finding all first-order partial derivatives for a sample point is computationally more expensive only by a small factor compared to adding another point.

C. Connection with Multivariate Hermite Interpolation

The solution of the problem (II.14) is closely related to the Hermite interpolation problem [21]. Given the coordinate maximum degrees n_1, n_2, \dots, n_N , and the homogeneous maximum degree n_T , we define the following set of polynomials:

$$\Pi_{n_1, n_2, \dots, n_N}^{N, n_T} = \left\{ \Psi(S) = \sum_{0 \leq \beta_i \leq n_i} a_{\beta_1, \beta_2, \dots, \beta_N} \prod_{k=1}^N s_k^{\beta_k} \right\}$$

We call $\Pi_{n_1, n_2, \dots, n_N}^{N, n_T}$ the set of polynomials of N variables of total degree no more than n_T and of coordinate degrees

no more than n_1, n_2, \dots, n_N .

The collocation problem becomes the following. Given the points S_1, S_2, \dots, S_m , determine a polynomial $\Psi(S)$ such that

$$\Psi(S_i) = \mathcal{J}(S_i); \frac{\partial \Psi}{\partial s_j}(S_i) = \frac{\partial \mathcal{J}}{\partial s_j}(S_i) \quad \begin{array}{l} i = 1, 2, \dots, m \\ j = 1, 2, \dots, N. \end{array} \quad (\text{II.15})$$

This Hermite interpolation problem can have a unique solution in the general case only when the number of monomials in $\Pi_{n_1, n_2, \dots, n_N}^{N, n_T}$ equals $m(N+1)$, the number of interpolation conditions, generally a rare occurrence. Partial first-derivative information can be included or dropped to make the system square. Even in that case, however, the issue of existence and uniqueness of a polynomial satisfying those conditions is not settled. All the results we are aware of concerning existence and uniqueness for almost all choices of the collocation set $\{S_i\}$ in the multidimensional case involve either maximum coordinated degree or maximum total degree, but not both [21].

We use both total degree and coordinate degree limitations to prune the number of polynomials needed to generate our approximation, which would otherwise rapidly grow with increasing N . If the resulting system is underdetermined, we have the option of adding additional collocation points (though this does not guarantee its well-posedness). If the system is overdetermined, we can solve it in a least squares sense. To account for either type of ill-posedness, we solve the system (II.15) using a generalized pseudoinverse approach based on singular value decomposition [3]. The generalized pseudoinverse uses singular value decomposition where exceedingly small singular values are replaced with $+\infty$ before carrying out the inversion.

The choice of sampling points and coordinate degrees in our method has a major heuristic component in the absence of theory for Hermite interpolation of the type considered here. Nonetheless, as we show in Section IV, the approach can be effective in situations where direct Monte Carlo sampling [20] would require orders of magnitude more effort.

D. The Use of a Control Variate for Variance Reduction

While the polynomial approximation $\hat{\mathcal{J}}$ can have excellent accuracy, as we show in Section IV, it introduces a bias. To correct for that bias, we use a control variate approach [20]. Recall that our setup is such that the function evaluation involves a run with a computationally expensive software to determine the values of \mathcal{J} at the collocation nodes, but once the SFEM model is computed, evaluations of it are cheap.

In that case, we assume that the object is to estimate the average of a functional $E[\Phi(\mathcal{J})]$ of the merit function \mathcal{J} with respect to some probability density over the uncertain parameter space. We use $\hat{\mathcal{J}}$ as a control variate

using the following identity.

$$E[\Phi(\mathcal{J})] = E\left[\Phi(\mathcal{J}) - \rho\left(\Phi(\hat{\mathcal{J}}) - E\left[\Phi(\hat{\mathcal{J}})\right]\right)\right] \quad (\text{II.16})$$

which is valid for any value of ρ . The functional Φ can be a power function, a maximum operator, or the characteristic function of an interval. In the last case, in nuclear reactor applications this would result in the estimation of the probability that a merit function will be in a certain range (for example, leading to the widely used 95% probability margins, after some extra processing). Since evaluations involving $\hat{\mathcal{J}}$ are inexpensive, the average quantity $E[\Phi(\hat{\mathcal{J}})]$ can be evaluated very inexpensively.

The optimal choice of ρ is the correlation coefficient between $\Phi(\mathcal{J})$ and $\Phi(\hat{\mathcal{J}})$ [20]. This coefficient is not known a priori, but can be inferred from data. If the approximation $\hat{\mathcal{J}}$ is of high quality, then $\rho \approx 1$. We assume that to be the case, and we use the right-hand side of the identity

$$E[\Phi(\mathcal{J})] = E\left[\Phi(\mathcal{J}) - \left(\Phi(\hat{\mathcal{J}}) - E\left[\Phi(\hat{\mathcal{J}})\right]\right)\right] \quad (\text{II.17})$$

to estimate $E[\Phi(\mathcal{J})]$. Note, however, that the random variable in the left has variance $\text{var}(\Phi(\mathcal{J}))$, whereas the one on the left-hand side has the variance $\text{Var}\left(\Phi(\mathcal{J}) - \Phi(\hat{\mathcal{J}})\right)$, which is potentially much smaller than $\text{var}(\Phi(\mathcal{J}))$, while introducing no additional bias. Should that occur, we will be able to obtain a confidence interval on $E[\Phi(\mathcal{J})]$ while using far fewer samples.

III. MODEL OF THE REACTOR CORE

For the convenience of description and numerical experimentation, we use a steady-state model of the reactor core, with uniform fuel elements, simple heat transport model (including convection and diffusion), and no control mechanisms. The idea is to preserve the general behavior of the physical system and to avoid the model-specific complexities of nuclear reactor analysis. The models describing the homogenized heat distribution in the core, and specific heat distribution inside the fuel pin follow the ones given in [7].

A. The Computational Model

A basic unit of the core is a cylindrical fuel pin, surrounded by flowing coolant (of constant velocity pattern). The chemical and physical properties of the coolant and the pin are spatially homogeneous, dependent only on temperature. There are two sources of heat: nuclear fission inside the fuel pin (generating thermal source q'''), and thermal energy carried by the incoming coolant (of constant temperature T_0). Heat is transported inside the

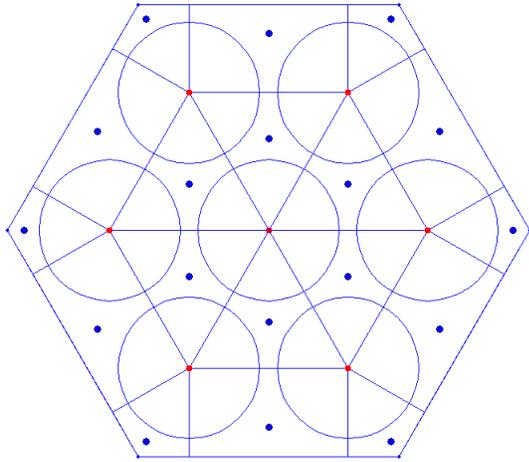


FIG. III.1: Finite volumes model of the reactor core

pin (from the centerline to the surface, and along the centerline) by conduction, in the coolant by convection and diffusion. A complete reactor core includes a hexagonal assembly of the fuel pins inside a large, insulated cylinder filled with coolant. We introduce a finite-volume grid of identical horizontal layers, as shown in Figure III.1. The heat is exchanged through the coolant volume elements and vertically inside the pins.

In this work we are concerned only with assessment of the thermohydraulics of a sodium-cooled reactor [9]. Our goal is to evaluate the benefit of derivative-based uncertainty assessment for realistic configurations involving highly nonlinear and important state variables. In the case of thermohydraulics, such a quantity is the maximum temperature in the fuel pin [7].

A code that computes the maximum fuel pin temperature coupled with heat transport in the channel and that includes derivative information was not available to us. We decided to develop a simple approximation technique that captures the essential physical phenomena and provides the afferent derivative information. On the other hand, direct calculation of the fuel pin temperature involves a complex, multiphysics, three dimensional model. Developing such models would result in an exceedingly large development cost to evaluate a nonlinear uncertainty quantification model. We therefore decided to implement a hierarchical approximation procedure. First, thermal flux balance equations are developed on a very coarse mesh (one node per channel or per pin at multiple height levels, as displayed in Figure III.1). Then, a radially symmetric model of the temperature in the center pin is developed, with boundary conditions generated by the surface temperature computed by the coarse thermal flux balance calculation. The latter model is solved on a fine mesh, resulting in an accurate description of the centerline temperature, where the maximum fuel pin temperature can be found in our model.

1. Three-Dimensional Thermohydraulics Calculations

Our computational approach is based on a classical finite volume method in which we compute distinct convective and diffusive fluxes [9].

We assume a prescribed form of the nuclear heat source:

$$q'''(x, y, z) = C_s \cdot \sin\left(\pi \frac{z}{H}\right) \quad (\text{III.1})$$

where H is the length of the fuel pin. The parameter C_s is chosen so that the temperature excursion of the reactor is 770 degrees Kelvin at the input, and 970 degrees Kelvin at the output. Here and in sequel we use degrees Kelvin (K) to measure the temperature.

Given a velocity field \vec{u} , we write the main heat transport equation as

$$0 = -\nabla \cdot K \nabla T - \rho c_p \vec{u} \cdot \nabla T + q''' \quad (\text{III.2})$$

where K is the thermal conductivity, ρ is the density and c_p is the heat capacity of the medium (fuel or coolant, depending on the location).

Integrate (III.2) over a volume cell Ω and apply the divergence theorem:

$$0 = \int_{\partial\Omega} K \nabla T \cdot \vec{n} dS + \int_{\partial\Omega} \rho c_p T \vec{u} \cdot \vec{n} dS - \int_{\Omega} q''' dV \quad (\text{III.3})$$

Note that the explicit expressions for the outward normal vector \vec{n} , variables of integration dS , dV , depend on the shape of the cell.

In the finite-volume formulation, equation (III.3) can be written as a conservation law for thermal energy flux Φ over all interfaces $\partial\Omega$ between the neighboring volume cells:

$$0 = \sum_{\partial\Omega} \Phi - \int_{\Omega} q''' dV \quad (\text{III.4})$$

The flux conservation law (III.4) is discretized on a coarse mesh with one node per pin and one node per channel at a given height level, as displayed in Figure III.1, for multiple height levels. The details of the calculations can be found in Appendix A. This approach results in a system of linear equations with matrix Λ^{th} :

$$\Lambda_{th}(c_p, K_c, K_f, h)T = B_{th} \quad (\text{III.5})$$

Here we denote by the subscript c quantities associated with the coolant and by subscript f quantities associated with the fuel pin. Hence, we denote by K_c the thermal conductivity of the coolant and by K_f the thermal conductivity of the fuel pin. The convective heat transfer coefficient h measures the rate of heat exchange from the coolant to the surface of the pin. It is a unit-free construct, related to other constructed variables such as

Nusselt number Nu , dependent on the hydraulic characteristics of the coolant flow, not discussed in detail here. It is provided to us in the form

$$h = \text{const} \cdot c_{p,c} \cdot Nu = \text{const} \cdot c_{p,c} \cdot Nu(K_c, c_{p,c}) \quad (\text{III.6})$$

In equation (III.5) we use c_p to denote $c_{p,c}$, the heat capacity of the coolant, since $c_{p,f}$ does not enter it.

We impose in-flow conditions in the form $T = T_0$ for volume cells in the first horizontal layer, and $\nabla T = 0$ everywhere else. At the outflow, this means that the entire heat output is due to transport.

In addition, we enforce the mass conservation for the coolant flow in the form

$$\rho \vec{u} = \text{const} \quad (\text{III.7})$$

The described setup produces physically reasonable distributions of temperature in the reactor channels.

The values of the material properties (ρ, c_p, K, h) depend on temperature. The assumption (III.7) excludes ρ from the analysis. For the rest of the parameters, $R = (c_p, K, h)$, we use the prescribed material properties obtained from experimental data.

$$c_p = c_p(T), \quad K = K(T) \quad h = h(Nu) = h(Nu(T)) \quad (\text{III.8})$$

At this point, the expressions are uncertainty-free.

The dependencies (III.5), (III.8) are coupled by a fixed-point iteration procedure started from a temperature-independent guess R_0 and repeated until convergence within numerical tolerance:

$$T_n := T(\hat{R}_{n-1}), \quad \hat{R}_n := \hat{R}(T_n), \quad |T_n - T_{n-1}| > \varepsilon \quad (\text{III.9})$$

Note that the individual components of temperature T and temperature-dependent parameters $R(T)$ are indexed differently. Equations (III.5) provide one value of temperature for each volume element.

The relationships (III.8) produce one value of thermodynamical parameter for each interface between two volume elements and require an estimate of the temperature on that interface:

$$R(T) = R^{(I,J)}(T_{I \rightarrow J}) = R^{(I,J)}\left(\frac{T_I + T_J}{2}\right) \quad (\text{III.10})$$

For technical convenience, for each thermodynamical parameter R , we store a sparse pattern matrix Π_R :

$$\Pi_R = (\Pi_{IJ}) \quad (\text{III.11})$$

where a component $\Pi_{IJ} = 1$ if cells I, J are neighbors, and the heat flux $\Phi_{I \rightarrow J}$ explicitly depends on R ; $\Pi_{IJ} = 0$ otherwise.

2. Centerline Temperature Model

Any temperature-dependent characteristic of the reactor core can be chosen as a merit function. For a demonstration of the capabilities of the method, we define \mathcal{J} as

a function of a derived quantity T_{center} , the temperature along the central axis of the reactor core (coincides with the central axis of a fuel pin), since we expect it to display substantial nonlinear variation in the uncertainty range. In the following, we ignore the gap and the cladding, and we assume that the entire fuel pin is made of uranium oxide.

After the convergence of the fixed-point iteration (III.9) is achieved, we find the distribution of temperature inside the central pin by solving an additional steady-state equation [7]

$$-\nabla \cdot K \nabla T_{pin} + q''' = 0 \quad (\text{III.12})$$

In cylindrical coordinates (θ, r, z) , assuming radial symmetry, we have

$$\frac{1}{r} \frac{\partial r K \frac{\partial T_{pin}(r,z)}{\partial r}}{\partial r} + \frac{\partial^2 K T_{pin}(r,z)}{\partial z^2} + q''' = 0 \quad (\text{III.13})$$

with boundary conditions

$$\nabla T_{pin}(r, z) = 0 \quad (\text{III.14})$$

at $r = 0$, and

$$T_{pin}(r, z) = T_{surface} \quad (\text{III.15})$$

at the surface of the pin, $r = D/2$, where $T_{surface}$ can be estimated from the finite volume model as an average of the (homogenized) temperatures in the pin and in the coolant (a more precise estimation involves the heat transfer coefficient h).

We apply a two-dimensional discretization grid to (III.13), coinciding with the one used for a finite-volume model in the horizontal direction, more refined in the radial direction. The obtained system of linear equations is similar to (III.5), with fewer parametric dependencies:

$$\Lambda_{pin} T_{pin} = B_{pin}(h, K) \quad (\text{III.16})$$

For each horizontal level, the value $K_{average}$ can be obtained as an average of the values $K^{(I,J)}$ over the interfaces (I, J) between the fuel pin and the neighboring coolant cells. This value of heat conductivity corresponds to the average temperature in the volume element I . To account for the much higher temperatures near the centerline of the pin, we use a simplified estimate of the form $K = K(T_{pin}) \approx \frac{1}{C_K} \cdot K_{average}$. The parameter C_K is chosen by a homogenization procedure: the flux average over a pin cross-section matches the one with a temperature dependent K for a reference temperature distribution that approximates the one at the center of the uncertainty region.

For the boundary conditions, $T_{surface}$ is obtained from $T_I, T_J, h^{(I,J)}$, as shown in (A.7). The information on the distribution of temperature in the centerline of the pin is extracted from the solution of (III.16):

$$T_{center}(z) = T_{pin}(r = 0, z) \quad (\text{III.17})$$

and substituted into the expression for the output function

$$\mathcal{J} = \mathcal{J}(T_{center}) \quad (\text{III.18})$$

More details on discretizing and solving (III.2), (III.12) can be found in the appendix. In the next section we discuss the introduction of uncertainty into relationships (III.9).

B. Uncertainty of the Physical Parameters

In the available literature on material properties [10, 11, 26] the dependencies of thermo-dynamical parameters on temperature are estimated from the experimental data and provided in the Laurent sum form

$$R = \sum_i r^{(i)} T^i \left(1 + \frac{\Delta R}{R^0} \right) \quad (\text{III.19})$$

where we denote by $\frac{\Delta R}{R^0}$ the relative error uncertainty term and by $R^0 = \sum_i r^{(i)} T^i$ the reference value.

For material properties of the coolant (liquid sodium) and fuel (uranium dioxide) we use the following expressions [11],[10]:

The reference value of the heat capacity of the coolant:

$$c_p^0 \approx 1.6582 - 8.470 \cdot 10^{-4} T + 4.4541 \cdot 10^{-7} T^2 - 2992.6 T^{-2} \quad [\text{J/kg} \cdot \text{K}] \quad (\text{III.20})$$

with relative uncertainty $\left| \frac{\Delta c_p}{c_p^0} \right|$ estimated as 0.1% at 300 K, 3% at 1000 K and 8% at 2000 K.

The reference value of the thermal conductivity of the coolant:

$$K_c^0 \approx 124.67 - 0.11381 T + 5.5226 \cdot 10^{-5} T^2 - 1.1842 \cdot 10^{-8} T^3 \quad [\text{W/m} \cdot \text{K}] \quad (\text{III.21})$$

with uncertainty $\left| \frac{\Delta K_c}{K_c^0} \right|$ estimated as 5% at 700 K, 12% at 1100 K and 15% at 1500 K.

The reference value of the thermal conductivity of the fuel:

$$K_f^0 \approx 12.57829 - 2.311 \cdot 10^{-2} T + 2.366675 \cdot 10^{-5} T^2 - 1.30812 \cdot 10^{-8} T^3 + 3.6373 \cdot 10^{-12} T^4 - 3.90508 \cdot 10^{-16} T^5 \quad [\text{W/m} \cdot \text{K}] \quad (\text{III.22})$$

with uncertainty $\left| \frac{\Delta K_f}{K_f^0} \right|$ taken to be 10% for all temperatures below 2000 K.

A different type of uncertainty description is available for the dimensionless heat transfer coefficient h [26]. We describe it briefly and note that in practice its influence on our outputs of the model is very small. It is available indirectly, through the relationship between dimensionless *Nusselt* and *Peclet* numbers, given in the form

$$Nu = \text{const} \cdot (Pe)^{0.3} \quad (\text{III.23})$$

with a constant coefficient dependent on geometry of the flow. The dependence of h on temperature is given by the expressions

$$Pe = \frac{\rho dh \bar{u} c_{p,c}}{K_c} \quad (\text{III.24})$$

$$h = \text{const} \cdot c_{p,c} \cdot Nu \quad (\text{III.25})$$

Specifically, we use the expression [26]

$$h(Pe) = \frac{4}{DH} (-16.15 + 24.96(P/D) - 8.55(P/D)^2) (Pe)^{0.3}. \quad (\text{III.26})$$

The uncertainty in the convective heat transfer parameter appears through the one in the Peclet number. The reference expression for the Peclet number is, using (III.23), with the $\rho \bar{u}$ computed with $\bar{u} = 5 \text{ m/s}$ and ρ computed at 800K,

$$Pe^0 = 10.1752 \frac{c_{p,c}}{K_c}. \quad (\text{III.27})$$

The other parameters used in the expression are $dh = 0.01$, the hydraulic diameter, in meters, $D = 0.01$, the fuel pin, in meters, $H = 3.70$, the length of the fuel pin in meters, and $P/D = 1.1394$. The uncertainty $\left| \frac{\Delta Pe}{Pe} \right|$ we have estimated from the data plots in [26] to be 70-80%.

C. Constructing the Parametric Representation of the Uncertainty

The multiplicative expression (III.19) contains no structure for dependence $\Delta R(T)$ of uncertainty on temperature. The specific values of $\left| \frac{\Delta R}{R^0}(T) \right|$ are given only for very few values of T . In addition, we have no uncertainty structure about the correlations between the values of the uncertainty at various levels. We are left with two choices. In a conservative approach, we could consider worst-case uncertainty calculations, in which only extreme values of the merit function \mathcal{J} would be sought, subject to the constraints that the uncertainty parameters satisfy the bounds developed in Section III B. Such problems are optimization problems that have been amply studied [15] but nonetheless lead to very conservative estimates.

We take a different point of view, in which we create an uncertainty structure that is perhaps imperfect but is consistent with the information available to us. We expect that future activity in the characterization of the uncertainty structure for advanced burner reactors will provide the information that is missing here. Once provided, this uncertainty structure can be accommodated by the computational framework we have developed in Section II.

In addition, the worst-case approach ignores the fact that every realization of the dependence of the physical parameters in Section III B, from the available literature

[10, 11, 26], is smooth with respect to its independent parameter (T for all parameters except h , for which the independent parameter is Pe). We use this observation to develop an approximate polynomial uncertainty structure. We choose a well-conditioned set of basis functions $\{C(T)\}$ and expand the *multiplicative error*:

$$\begin{aligned} R &= \left(\sum_i r^{(i)} T^i \right) \left(1 + \sum_{j=0} \alpha_j C_j(T) \right) \\ h &= \left(\sum_i r^{(i)} T^i \right) \left(1 + \sum_{j=0} \alpha_j C_j(Pe(T)) \right) \end{aligned} \quad (\text{III.28})$$

For a numerically stable process, and unit-free error we chose the multiplicative model (III.28). To expand the dependencies on temperature, we use a set of Chebyshev polynomials

$$\begin{aligned} C_0(T) &= 1 \\ C_1(T) &= T + 1 \\ C_2(T) &= 2T^2 - 1 \\ C_3(T) &= 4T^3 - 3T^2 \\ &\dots \end{aligned} \quad (\text{III.29})$$

We choose to stop the expansions at the second-order term $\alpha_2 C_2$ in all parameters, since every instance of experimental data is monotonous in temperature, without noticeable oscillations.

The list of collocation parameters now reads:

$$S = \{(\alpha_0, \alpha_1, \alpha_2)_{c_{p,c}}, (\alpha_0, \alpha_1, \alpha_2)_{K_c}, (\alpha_0, \alpha_1, \alpha_2)_{K_f}, (\alpha_0, \alpha_1, \alpha_2)_h\} \quad (\text{III.30})$$

We acknowledge that specific statistical information on the distribution of parameters (III.30) is missing. We assume uniform distribution and no correlations between the parameter triplets. To describe the possible of values for $(\alpha_0, \alpha_1, \alpha_2)$ for each material property, we perform large-scale random sampling inside a sufficiently large cubic region

$$\alpha_{i,min} \leq \alpha_i \leq \alpha_{i,max}, i = 1, 2, 3 \quad (\text{III.31})$$

around $(0, 0, 0)$. For each point of the random sample, we check the uncertainty conditions in the form

$$-\xi \leq \alpha_0 C_0(T_\xi) + \alpha_1 C_1(T_\xi) + \alpha_2 C_2(T_\xi) \leq \xi \quad (\text{III.32})$$

for a small set of values of T_ξ in the region of interest: $600 \text{ K} \leq T_\xi \leq 2600 \text{ K}$. If that condition is not satisfied, the triplet $(\alpha_0, \alpha_1, \alpha_2)$ is removed from the valid parameters list. The value ξ is obtained from the multiplicative uncertainty structure described in (III.19) and from the expressions of $\frac{\Delta R}{R^0}$ and R^0 described in Section III B. For values T_ξ that are not among the ones enumerated in Section III B, and for which uncertainty information is not available to us, we take ξ corresponding to the values of $\frac{\Delta R}{R^0}$ from the adjacent points for which this information is provided. This results in a slightly larger uncertainty set. The approach results in an empirical distribution for the parameters S , which is uniform on the set that is feasible for the uncertainty constraints.

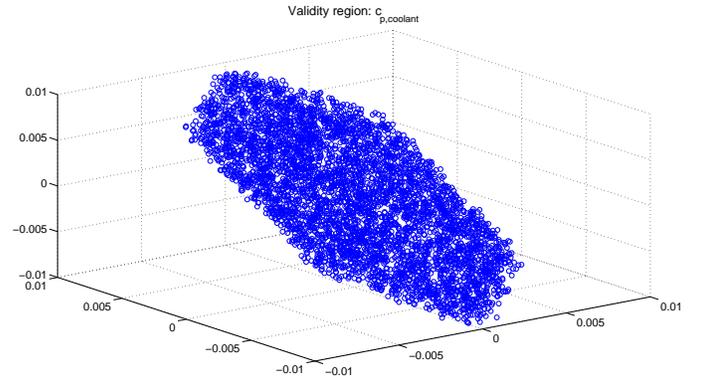


FIG. III.2: Validity region for $(\alpha_0, \alpha_1, \alpha_2)_{c_{p,c}}$

We store the *cloud of points* satisfying (III.32) (see III.2). We use a randomly (and uniformly) selected subset from this validity region to create the collocation matrix and construct the collocation model as described in (II.14). Oversampling helped to achieve a wider spread of sample points and improve the condition number of the collocation matrix.

We also used Hammersley and Smolyak sparse sampling [22] during the described two-step sampling procedure. We found that it results in an approximately equivalent performance of the resulting surrogate model.

The main critique of the approach is perhaps that we have no information to assume that the uniform distribution, compatible with the uncertainty constraints. On the other hand, in the absence of any additional information, uniform prior distribution is among the most reasonable (“non informative”) assumptions [18]. In addition, the extreme values of the merit function \mathcal{J} approach those of the worst-case assessment for increasing sample size. Therefore the information we provide would be useful in that case as well.

In addition, our approach creates an uncertainty structure that is fully consistent with the prior information. Our main goal is to investigate the scope of SFEM based uncertainty quantification approaches for realistic uncertainty structures for sodium-based reactor calculations. Therefore the performance of the SFEM approach for different uncertainty structure that is also compatible with the prior information is likely to be comparable to the one we will observe in our numerical experiments.

D. Computing Derivative Information

There are a number of possible approaches to computing derivative information (directly or through the use of the adjoint variable) [14, 17, 24]. The factors taken into account when choosing the approach are computational efficiency, and the ease at which the procedure can be adapted for a class of examples.

We use a simple chain rule approach, with elements

of automatic differentiation. To keep the description of the procedure general, we first suppose that the output function $\mathcal{J}(T)$ is not hierarchical (i.e., there is an explicit dependence of \mathcal{J} on T). We therefore consider \hat{R} related by a quasilinear system with a constant source term B

$$\begin{aligned} F(T_n, \hat{R}(T_{n-1})) &= \Lambda(\hat{R}) \cdot T_n - B = 0 \\ T_n &\approx T_{n-1} = T \end{aligned} \quad (\text{III.33})$$

and a set of expressions dependent on the collocation parameters α :

$$R(T, \alpha) = R(T) \cdot (1 + \alpha_1 C_1(T) + \alpha_2 C_2(T) + \dots) \quad (\text{III.34})$$

as specified by (III.26) – (III.32) and (III.28).

Taking into account the implicit dependencies present in (III.33), (III.34), we obtain the derivative of \mathcal{J} by the following sequence of operations:

By chain rule,

$$\frac{dJ}{d\alpha} = \frac{\partial \mathcal{J}}{\partial T_n} \cdot \frac{dT_n}{d\alpha} \quad (\text{III.35})$$

Consider (III.33) in the form

$$F(T_n(\alpha), \hat{R}(T_n, \alpha)) = 0 \quad (\text{III.36})$$

Note that we formally treat the two instances of T_n in (III.36) as separate variables. Differentiate (III.36) with respect to α :

$$\left(\frac{\partial F}{\partial T_n} + \frac{\partial F}{\partial \hat{R}} \cdot \frac{\partial \hat{R}}{\partial T_n} \right) \cdot \frac{dT_n}{d\alpha} + \frac{\partial F}{\partial \hat{R}} \cdot \frac{\partial \hat{R}}{\partial \alpha} = 0 \quad (\text{III.37})$$

Two partial derivatives are required for this expression:

$$\frac{\partial F}{\partial T_n} = \Lambda \quad (\text{III.38})$$

$$\frac{\partial F}{\partial \hat{R}} = \frac{\partial \Lambda}{\partial \hat{R}} \cdot T_n \quad (\text{III.39})$$

From (III.37) we have

$$\frac{dT_n}{d\alpha} = - \left(\frac{\partial F}{\partial T_n} + \frac{\partial F}{\partial \hat{R}} \cdot \frac{\partial \hat{R}}{\partial T_n} \right)^{-1} \cdot \frac{\partial F}{\partial \hat{R}} \cdot \frac{\partial \hat{R}}{\partial \alpha} \quad (\text{III.40})$$

Finally, the derivative is expressed as

$$\frac{dJ}{d\alpha} = - \frac{\partial \mathcal{J}}{\partial T} \cdot \left(\Lambda + \frac{\partial \Lambda}{\partial \hat{R}} \cdot T_n \cdot \frac{\partial \hat{R}}{\partial T_n} \right)^{-1} \cdot \frac{\partial \Lambda}{\partial \hat{R}} \cdot T_n \cdot \frac{\partial \hat{R}}{\partial \alpha} \Big|_{T_n=T} \quad (\text{III.41})$$

This generic approach is applied to the model described in III A in the following fashion, using the structural equations (III.5) and (III.16), where

$$\Lambda = \begin{bmatrix} \Lambda_{th} & 0 \\ \Lambda_c & \Lambda_{pin} \end{bmatrix} \quad B = \begin{bmatrix} B_{th} \\ B_{pin} \end{bmatrix}, \quad T_n = \begin{bmatrix} T \\ T_{pin} \end{bmatrix},$$

We have introduced the linear operator Λ_c to quantify the effect of the boundary conditions on the pin problem in Section III A 2, which is linear.

Details about how the derivative information is obtained, including the hierarchical case, are given in Appendix B.

IV. NUMERICAL RESULTS

In our numerical tests we use a model of the reactor core with 7 pins and 20 horizontal layers. The output function is a measure of temperature on the centerline of the central pin:

$$\mathcal{J}(T) = \|T_{center}\|_p \quad (\text{IV.1})$$

with $p = 1000$, resulting in a differentiable expression estimating the maximal temperature in the reactor core. We compared the performance of different approaches to uncertainty quantification: random sampling, linear approximation, SFEM with full basis and SFEM with truncated basis.

A. General Algorithm

The algorithm used to create the models is described below. Note the variations required for the use of *derivative information* and *basis truncation*.

• Create Valid Sample Points Sets

1. Choose a list of parameters $S = (s_1, s_2, s_3, \dots) = ((\alpha_1, \alpha_2, \alpha_3), \dots)$, as in (III.30).
2. Generate a set of admissible values of S according to conditions on maximal uncertainty: (III.31), (III.32).

• Initialize and Run the Physics Model

1. For each value S_i , evaluate $J(S_i)$ by running a fixed point iteration (III.9): (III.5) to solve a finite-volume model with given material properties; (III.28) to update the values of material properties according to given temperature. Twenty iterations are sufficient for convergence.
2. *Derivative information*: at the last step of the fixed point iteration, store the required intermediate components and evaluate the partial derivatives $\frac{\partial \mathcal{J}}{\partial s_i}$ by the method described in Section B.

• Create SFEM Polynomial Basis

1. *Evaluate parametric sensitivity*: define a neutral state of the system with no uncertainty: $S = (0, 0, 0, \dots)$; $\mathcal{J}_0 = \mathcal{J}(S_0)$. Evaluate the derivatives $\frac{\partial \mathcal{J}_0}{\partial s_i}$.
2. *Basis truncation*: using the derivative magnitude test (II.13), at $S = S_0$, rank components of S by importance. Choose importance groups I, II, III . We take $n_{III} = n_{II} = 2$,

$n_I = 8$. The construction of the corresponding truncated SFEM model requires 12 full runs of the model.

3. Assemble complete basis Ψ (II.5), of multivariate polynomials in variables S . We choose the maximal degree to be 3.
4. *Basis construction*: use importance groups *I, II, III*.
Remove basis polynomials of degree 3 *that do not include* variables from group *I*.
Remove basis polynomials of degree 2 *that do not include* variables from group *II*.
Use the remaining set of polynomials as a new basis Ψ .

• SFEM Collocation Model Creation and Validation

1. Create collocation equations (II.6) for the chosen sample points and chosen polynomial basis. *Derivative information*: use a collocation matrix with derivative information (II.14). The minimal number of collocation matrix rows is equal to the size of the basis Ψ . We use a tall matrix (oversample by a factor of 2).
2. Solve collocation equations to obtain the polynomial surrogate model $\hat{\mathcal{J}} = \hat{\mathcal{J}}(S)$ (II.1).
3. *Additional analysis*: use derivative information from Section 3.2 to create a linear approximation $\hat{\mathcal{J}}_{linear} = \mathcal{J}_0 + \sum_i \frac{\partial \mathcal{J}}{\partial s_i} s_i$.
4. *Validation* Create a new valid sample set \hat{S}_i , $i = 1, 2, \dots, N_V$ by a procedure similar to **Create Valid Sample Point Set** step, and evaluate the error of the SFEM and the linear model. Alternatively, use the control variate approach described in Section IID.

We note that the validation procedure is essentially equivalent to the control variate procedure for computing the expected value of \mathcal{J} at the number of samples N_V . Indeed, the root mean square error of the model error, evaluated on the validation sample, is the sample variance of the control variate. When the control variate is used by itself, it is not generally used with a fixed sample; rather, the number of samples is increased until a certain confidence range is met. To emulate this behavior, we assume a $\frac{\sigma}{\sqrt{n}}$ behavior of the error estimate, using the estimated sample standard deviation σ at N_V . These results are plotted in Figure IV.2.

B. Performance of the Model

Our model has 12 uncertain parameters, 3 per physical variable: $c_{p,c}$, K_f , K_c , and h .

On a validating set of $N_V = 40$ points, we observe *range* (lowest and highest observed outputs), *variance*

(variance of the observed outputs) and *error variance* (variance of the difference between the surrogate and the exact outputs).

We used the procedures described in Section IV A to create two versions of SFEM surrogate model. The first version rigorously follows the procedure described in Section IV A; we call it the version with basis truncation. The second version skips the *Basis Truncation* step in the **Create SFEM Polynomial Basis** stage of the algorithm presented in the Section IV A; we call it the version with full basis. We will use the comparison between these two versions as a way to assess the benefit of our basis truncation method.

We present the summary of our numerical findings in Table IV.1. The absolute error produced, on the $N_V = 40$ validation samples by linear approximation, SFEM, and truncated SFEM approaches is shown in Figure IV.1.

To evaluate the relative benefit of the approaches discussed here and the ones closely related to them, we evaluate their potential benefit versus computational cost when used as a control variate. The methods compared are SFEM with truncated basis, SFEM with full basis, the linear model evaluated at the center of the uncertainty area, and SFEM with truncated basis evaluated without a derivative information. The last two models are included for discussion because they are the ones the most commonly used in nonintrusive uncertainty quantification calculations.

The benefit measure used in our experiment is the theoretical variance of a Monte Carlo approach applied to the control variate (II.17), based on the error estimate produced by the validation, as described at the end of Section IV A. The cost measure we have used is the number of model function evaluations to reach that variance level. In evaluating the cost measure, we assume that one evaluation of the full gradient \mathcal{J} has, as our implementation has, twice the cost of a new evaluation of \mathcal{J} at a new point. Therefore, one model and gradient calculation has the computational cost of three separate model evaluations, that is, three units on the x axis of Figure IV.2. The results are displayed in Figure IV.2. In that figure, not all methods can be used below a certain threshold of number of samples used, so they produce no result for that x value.

Insofar the likelihood of the validity of generalizing our findings in Subsection IV C a key assumption, also used in the generation of Figure IV.2, is that the calculation of the gradient of an output functional has a computation cost that is at most a small multiple of the cost of the functional itself at a new uncertainty scenario point. That assumption is, of course, arguable. The ratio of the CPU time taken by the gradient calculation to the one of the CPU time taken by the function evaluation is dependent on the algorithms used for computing either, on the hardware used, and on the size of the problem. We posit that our assumption is reasonable based on the following evidence:

- One full run for a 7-pin model takes 1-2 minutes of

computational time (higher estimate corresponds to an average desktop computer). The calculation of the full gradient of \mathcal{J} by the derivative calculation procedure described in Appendix B takes between 150% and 200% of the time of the function evaluation procedure described in Appendix A.

- When other implementations and more parameters are considered, adjoint method approaches are known to provide derivative information at a cost that is essentially at most the one of solving the full nonlinear model [5]. This comes from the fact that one adjoint variable calculation solves a linear system, whereas the full nonlinear model needs multiple similar linear solves while it iterates to a solution.
- For an arbitrary function, the gradient of the function is surely computable in a time that is a small multiple of the time cost of the function evaluation; that multiple never exceeds 5 [14]. We use a cost ratio of 2 but we point out that even if we use 5, in the limit of increasing number of parameters our method is bound to overtake methods that do not use derivative information.

C. Discussion of the Numerical Results

The numerical results are presented in Table IV.1, in Figure IV.1, and in Figure IV.2.

We see from Figure IV.1 that the 12-parameter SFEM full model and SFEM truncated basis model consistently approximate the merit function better than does the linear model. This is reinforced by the data in Table IV.1, which shows that we estimate the range of the merit function \mathcal{J} and the error variance far better with the SFEM methods than with the linear model. One should bear in mind that we use the same code to construct the linear model as we use to compute the derivatives used in the SFEM model. Also, the fact that the tangent approximation underestimates everywhere the output functional always occurs for convex functionals \mathcal{J} [16]. We cannot prove the convexity of our \mathcal{J} choice (IV.1), but we wish to point out that the linear model behavior from Figure IV.1 is definitely possible. Of course, it is likely that changing the centerpoint of the linear approximation would result in better error behavior for the linear model. But there is no inexpensive guideline on how to do that. In addition, we use the same derivative information to rank our variables and generate the truncated polynomial basis, so the comparison is done under comparable conditions.

Insofar as our truncation mechanism is used to reduce the number of function evaluation needed to fit the model, we see that the SFEM truncated model is better than the linear model, both in terms of approximation error as seen in Table IV.1 and in terms of error per amount of effort, as seen in Figure IV.2. Therefore, the truncation captures a substantial portion of the nonlinearity,

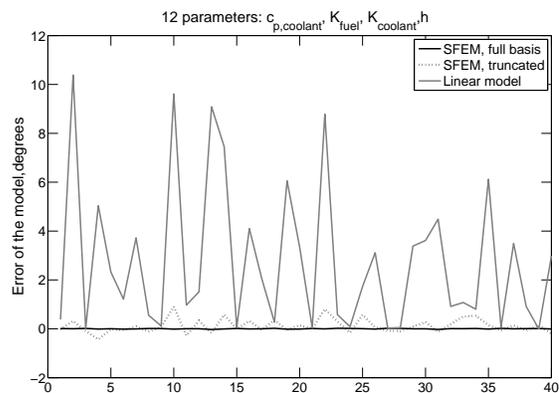


FIG. IV.1: SFEM model performance: 12 parameters

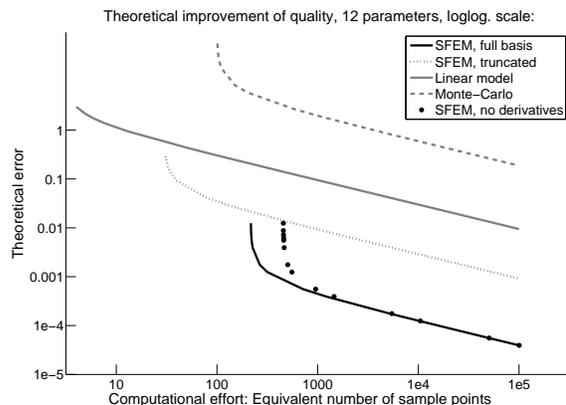


FIG. IV.2: Theoretical improvement of quality with the increase of calibration set, log-log scale

without needing the complexity of the SFEM full model. As the available computational resources increase, the SFEM full model eventually outperforms both the SFEM truncated and the SFEM full no-derivative-used model.

Finally, insofar as the use of derivative information, examination of the control variate results, Figure IV.2 is revealing. It shows that SFEM methods that use derivatives need far fewer samples to reach a certain control variate theoretical sample variance level when compared to direct Monte Carlo, linear approximation, or SFEM approaches that use collocation only.

We also point out that, for complex multiphysics models, the region of interest is precisely the region of up to 100 total samples (total model evaluations). For expensive models, it is unlikely that we will be allowed a large budget of model evaluations, so the asymptotic regime is truly not interesting for practical applications. In the below-100-samples region, the SFEM approach that uses truncation and that is calibrated with derivative information does particularly well.

Finally, an interesting conclusion can be reached by comparing the performance of each method when predicting the range of the output functional relative to the

TABLE IV.1: Performance of the 12-parameter SFEM model, uncertainty in $C_{p,coolant}, K_{fuel}, K_{coolant}, h$

	Sampling	Linear Model	SFEM Full	SFEM Truncated
Sample size	40	1	72	12
Range	2237.82 - 2460.54	2227.43 - 2450.09	2237.82 - 2460.55	2237.51 - 2459.63
Model σ^2	3487.08	3495.38	3486.87	3476.01
Model σ	59.05	59.12	59.05	58.96
Error σ^2	*	8.94	0.00015	0.084
Error σ	*	2.99	0.01	0.29

total variation of the output functional that is due to parametric uncertainty. As can be seen in Table IV.1 the range size is about 223 K. From Table IV.1 we see that, when predicting the range, the first-order sensitivity method errs by about 4.5%, the SFEM truncated by 0.5% and the SFEM full by essentially 0%. We see that, if the goal is to predict correctly the range to within 10 % error, the linear model based on first-order sensitivity calculation is adequate. If this error level were not sufficient, and the goal is to predict the range to within 1% error, then only the SFEM approaches are adequate. Since the main development cost of any of these approaches is the computation of derivative information for the multiphysics model, our method achieves the extra precision level with virtually no additional development cost (though it may need about 10 times more computational cost). Another option would be to use a higher-order sensitivity approach [5]. In our experience, such an approach has a far higher development cost than the first-order sensitivity approach. In addition, it needs a much larger increase in computational cost compared to the increase required by adding derivative information to a functional evaluation [14]. We thus conclude that “sampling-with-derivatives” SFEM approaches of the type presented here are convenient approaches to extend the capabilities of multiphysics codes which provide first-order derivative information of an output functional. We point out that, in the example discussed here, such a conclusion is largely unaffected by the assumptions concerning the probability distribution of the uncertainty. We also point out that, as the number of parameters increases and as additional physics details are included, we expect the percentages above to change, and many times, in the favor of SFEM approaches of the type presented here.

V. CONCLUSION AND RESEARCH PLANS

In our work, we found that effects of a moderate number of uncertain parameters on a complex system can be efficiently modeled through a synthetic method, combining stochastic finite-element interpolation with the use of derivative information and basis truncation. We have

applied this method for a model of the thermohydraulics of a nuclear reactor, where the uncertainty originates in the physical parameters of the system. From the available information we have created a probabilistic model, consistent with the experimental information available to us.

We observed the substantial advantage of SFEM methods computed using derivative information over classical methods of uncertainty quantification: pure random sampling, linear approximation, and SFEM methods that do not use derivative information. In addition, we observed that our basis truncation heuristic efficiently produces a far better approximation of the output function than the linear model while using far less computational resources than the full SFEM model.

Naturally, further research is needed to determine whether the trends observed in this work hold for other applications and for larger instances of uncertainty assessment in nuclear reactors. But we have shown that SFEM approaches using derivative information have great potential for real applications.

From the engineering point of view, application of the method to models with greater range of possible uncertainty and with more parameters is of interest. We plan to extend the model of the nuclear reactor core, with additional uncertainty coming from the description of neutron interaction (an eigenvalue problem), nonuniform flow of the coolant (Navier-Stokes equations), and structural deformations of the reactor elements. Additional parts of the model will be coupled with each other through temperature. We expect the most notable (and highly unstructured) contribution of uncertainty to come from the equations for nuclear fission.

Another important engineering issue needing further consideration is the one of creating appropriate uncertainty models from available data about the physical properties of the reactor. At the moment, we used information from the best available processed assessments for nuclear reactors [10, 11, 26]. But these assessments provide uncertainty information in a way that, while very valuable, is not optimally usable in a probabilistic approach. We have reduced the representation of the uncertainty by observing that all observed measurement realizations are smooth, and we have introduced a probabilistic structure based on a uniform prior. Since the real distribution of the data is likely to be more peaked, the estimates we produce should be worse (they should show larger ranges of the centerline temperature) than the calculation with the actual distribution. We expect that direct processing of the experimental data, which is a nontrivial endeavor, will show that our spectral representation of the uncertainty approach is reasonable and that the uniform prior is indeed a conservative approach, though not as conservative as worst-case approaches.

From a mathematical point of view, this research poses several interesting questions. The first is what is an appropriate basis truncation procedure. A simple heuristic for basis truncation, based on first-order sensitivity anal-

ysis, produced reasonable performance of SFEM model. We would like to find a more sophisticated procedure from basis selection, allowing us to keep the basis small without significant loss of quality, and simultaneously extend interpolation order in some variables to high degrees (5 or more). The list of stochastic variables should not necessarily be kept intact: there is a possibility of applying factor lumping and other model reduction techniques to construct optimal combinations of variables and rank them by importance.

The second important question is how to select the sample points for collocation with derivative information. This problem was addressed in the past for the case of full SFEM basis of a given degree involving collocation without derivative information [27]. To our knowledge, no result exists for the case described here, the one using a truncated basis and derivative information. We used a simple procedure to select sample points for collocation where we compensated for the possibly unfavorable condition number and other quality issues with oversampling. This was sufficient for our test case, but we foresee problems for models with more uncertain parameters (“curse of dimensionality”: random sampling may miss important locations) or with more oscillatory behavior of the model (surrogate model may behave incorrectly at points of high sensitivity). In the selection of sample points, there is a greater variety of choices, while the selection criteria is mostly unclear. We expect optimal sampling to be more difficult than optimal selection of basis.

Moreover, we hope to provide additional theoretical justification for the suggested uncertainty quantification techniques. If possible, we would like to derive error estimates for the performance of the reduced model, depending on the choice of the polynomial basis and on the number of sample points used to calibrate the model.

Acknowledgment

We are grateful to Won-Sik Yang, Andrew Siegel, Tom Fanning, and our other SHARP team colleagues for valuable comments. Mihai Anitescu is grateful to Vince Mousseau and Paul Turinsky for comments on the presentation of parts of this work at the Verification and Validation for Nuclear Systems Analysis Workshop, 2008. The authors were supported by Contract No. DE-AC02-06CH11357 of the U.S. Department of Energy.

Appendix A: Details of Finite-Volume Reactor Core Modeling

For the interface $\partial\Omega_L$ between the cells Ω_I, Ω_J , denote the temperature gradient by $\nabla T^{(L)} = \nabla T_{I \rightarrow J}^{(L)}$; the corresponding flux by $\Phi^{(L)} = \Phi_{I \rightarrow J}^{(L)} = -\Phi_{J \rightarrow I}^{(L)}$. We use a

simple scheme for estimating the gradient:

$$\nabla T_{I \rightarrow J} = \frac{T_J - T_I}{D_x} \quad (\text{A.1})$$

where D_x is the distance between the volume element center. We estimate the temperature on an interface as

$$T_{I \rightarrow J} = T_{J \rightarrow I} = \frac{T_J + T_I}{2} \quad (\text{A.2})$$

where T_I, T_J are temperatures in the center of the corresponding cells.

Let the temperature on the interface between the cells be T_{IJ} and the distance between the centers of the cells be $H_I^{(L)} = H_{I \rightarrow J}$. Denoting the coordinates of the geometric center of cell I as (x_I, y_I, z_I) , we express outward unit normal as

$$n = \vec{n}_{I \rightarrow J} = \frac{1}{D_x} \cdot (x_J - x_I, y_J - y_I, z_J - z_I) \quad (\text{A.3})$$

The finite-volume model (III.4) can now be assembled from the information about the heat fluxes across every type of boundary.

Pin to pin, diffusion:

$$\Phi = \int_{\partial\Omega} K_f \nabla T \cdot \vec{n} dS \quad (\text{A.4})$$

Coolant to coolant, diffusion and convection:

$$\Phi = \int_{\partial\Omega} K_c \nabla T \cdot \vec{n} dS + \int_{\partial\Omega} \rho_{c,p,c} T \vec{u} \cdot \vec{n} dS \quad (\text{A.5})$$

Coolant to outflow, convection only:

$$\Phi = \int_{\partial\Omega} \rho_{c,p,c} T \vec{u} \cdot \vec{n} dS \quad (\text{A.6})$$

The flux from pin to coolant can be derived from two descriptions of temperature exchange:

$$\Phi = h(T_J - T_{surface}) = K_f \frac{T_{surface} - T_I}{D/2}. \quad (\text{A.7})$$

where D is the diameter of the pin, the rightmost expression is a diffusive heat flux from the centerline of the pin to the surface, and h is the convective heat transfer coefficient measuring the rate of heat exchange from the coolant to the surface of the pin. Exclude the unknown $T_{surface}$:

$$\Phi = 2 \frac{hK_f(T_J - T_I)}{hD + 2K_f} \quad (\text{A.8})$$

We now write the detailed flux conservation equations at each of the cells created by our discretization. The flux equations include integrals over cell boundaries. We denote such integration boundaries as $\partial\Omega$, and the integration is taken over the one between the cell I and

the cell whose other index is mentioned in the expression. Also note that each temperature appearing in the following expression is considered a constant for that integration.

Every pin cell I has coolant cell neighbors (record their index as J) and pin neighbors directly above and below (recorded as $I+$, $I-$). It contributes the following:

$$\int_{\Omega} q''' dV = \sum_J 2 \frac{hK}{hD+2K} T_J - \sum_J 2 \frac{hK}{hD+2K} T_I + \sum_{I-,I+} K \frac{1}{H_{I \rightarrow I-,+}} \int_{\partial\Omega} ndST_{I-,+} - \sum_{I-,I+} K \frac{1}{H_{I \rightarrow I-,+}} \int_{\partial\Omega} ndST_I. \quad (\text{A.9})$$

$$0 = \sum_J K \frac{1}{H_{I \rightarrow J}} \int_{\partial\Omega} ndST_J + \sum_{J+,J-} K \frac{1}{H_{I \rightarrow J}} \int_{\partial\Omega} ndST_{J+,J-} - \sum_J K \frac{1}{H_{I \rightarrow J}} \int_{\partial\Omega} ndST_I + \dots - \sum_{J+,J-} K \frac{1}{H_{I \rightarrow J}} \int_{\partial\Omega} ndST_I + \frac{1}{2} \sum_J \rho c_p \int_{\partial\Omega} \vec{u} ndST_J + \frac{1}{2} \sum_{J+,J-} \rho c_p \int_{\partial\Omega} \vec{u} ndST_{J+,J-} + \dots + \frac{1}{2} \sum_J \rho c_p \int_{\partial\Omega} \vec{u} ndST_I + \frac{1}{2} \sum_{J+,J-} \rho c_p \int_{\partial\Omega} \vec{u} ndST_I + \sum_{J^*} 2 \frac{hK}{hD+2K} T_{J^*} - \sum_{J^*} 2 \frac{hK}{hD+2K} T_I \quad (\text{A.10})$$

Equation (III.12) for the temperature in the interior of the central pin is written directly in terms of temperature T_{pin} rather than in terms of fluxes. We assume cylindrical symmetry of the problem, and we write the equation in cylindrical coordinates. The primary variable is discretized on a $R \times Z$ grid and thus becomes $T_{pin}^{n,m}$, for $m = 1, 2, \dots, R$, and $n = 1, 2, \dots, Z$.

Then the discretized heat equation in the pin reads as follows:

$$K_n \frac{(m+1)T_{pin}^{(m+1,n)} - 2mT_{pin}^{(m,n)} + (m-1)T_{pin}^{(m-1,n)}}{m\Delta r^2} + \dots$$

$$\frac{K_{n+1}T_{pin}^{(m,n+1)} - 2K_nT_{pin}^{(m,n)} + K_{n-1}T_{pin}^{(m,n-1)}}{\Delta z^2} = q_n''' \quad (\text{A.11})$$

with $m = 2..R - 1$, $n = 2..Z - 1$, $\Delta r = \frac{D}{2R}$, $\Delta z = \frac{H}{Z}$, $K_n = K(z = n\Delta z)$, $q_n''' = q'''(z = n\Delta z)$.

The system is augmented with boundary conditions. At $m=1$, $\frac{\partial T_{pin}}{\partial r} = 0$, and the equation (A.11) reads as

$$\frac{K_{n+1}T_{pin}^{(1,n+1)} - 2K_nT_{pin}^{(1,n)} + K_{n-1}T_{pin}^{(1,n-1)}}{\Delta z^2} = q_n''' \quad (\text{A.12})$$

At $n = 1, Z$

$$T_{pin}^{(n,R)} = T_{surface} \quad (\text{A.13})$$

with the surface temperature estimate $T_{surface}$ computed from (A.7) with value $h_n = h(z = n\Delta z)$ and averaged over the neighbors of the fuel pin at the same height.

Every coolant element I has coolant cell neighbors in the same horizontal layer (record their index as J), cell neighbors directly above and below (recorded as $J+$, $J-$), and pin neighbors in the same horizontal layer (recorded as J^*). It contributes the following:

Appendix B: Details of Derivative Calculations

The temperature is stored in a vector T with n components, corresponding to the number of volume cells in the model. Each material property R is stored in a $n \times n$ matrix, with nonzero entries indicated by an indexing system (III.11). The components of sparse array operations (III.40), (III.41) are generated as follows:

A sparse matrix

$$\frac{\partial F}{\partial R} = \frac{\partial \Lambda}{\partial R} \cdot T = \frac{\partial \Lambda}{\partial \Phi} \cdot \frac{\partial \Phi}{\partial R} \cdot T \quad (\text{B.1})$$

has n^2 columns. The column " IJ ", $I = 1 \dots n, J = 1 \dots n$ has nonzero entries only if $\Pi_{IJ} = 1$. It is assembled as

$$\frac{\partial F^{I,J}}{\partial R} = \left(\frac{\partial \Lambda}{\partial \Phi} \right)^{(I,J)} \cdot \frac{\partial \Phi^{(I,J)}}{\partial R} \cdot T \quad (\text{B.2})$$

Since heat flux $\Phi(IJ)$ is influencing only its origin cell I and destination cell J , the matrix $\left(\frac{\partial \Lambda}{\partial \Phi} \right)^{(I,J)}$ has two nonzero entries: -1 in position (I, J) and $+1$ in position (I, I) . The expression $\frac{\partial \Phi}{\partial R}(IJ)$ is a scalar, explicit derivative of one of the formulas (A.4) - (A.6), (A.8), depending on the type of interface $I \rightarrow J$.

The entry (I, J) of the $n \times n$ matrix $\frac{\partial R}{\partial \alpha}$ is nonzero only if $\Pi_{IJ} = 1$. It is computed as the derivative of (III.28):

$$\frac{\partial R}{\partial \alpha_i}(I, J) = \left(\sum_i r^{(i)} T^i \right) \cdot C_i(T) \quad (\text{B.3})$$

evaluated at $R^{(I,J)}$, $T^{(I,J)}$.

The $n \times n$ matrix $\frac{\partial R}{\partial T}$ has two nonzero entries in each column, in positions I and J , both equal to the derivative

of (III.28) with respect to T :

$$\begin{aligned} \frac{\partial R}{\partial T_I} &= \frac{1}{2} \left(\sum_i i r^{(i)} T^{i-1} \right) \cdot (1 + \alpha_0 C_0(T) + \alpha_1 C_1(T) + \alpha_2 C_2(T)) \\ &+ \frac{1}{2} \left(\sum_i r^{(i)} T^i \right) \cdot (\alpha_0 C'_0(T) + \alpha_1 C'_1(T) + \alpha_2 C'_2(T)) + \dots \end{aligned} \quad (\text{B.4})$$

evaluated at $R^{(I,J)}$, $T^{(I,J)}$, where C'_i is a derivative of a Chebyshev polynomial (III.29); the coefficient $\frac{1}{2} = \frac{d}{dT_I} \left(\frac{T_I + T_J}{2} \right)$ comes from (III.10).

The components $\frac{\partial \mathcal{J}(T)}{\partial T^{(I)}}$ are derivatives of the expression for $\mathcal{J}(T)$ with respect to temperature.

Now let us consider the case (III.18) where the explicit expression $\mathcal{J}(T)$ is not available. When an output is defined hierarchically, we can compute the derivative by chain rule:

$$\frac{d\mathcal{J}(T_{center})}{d\alpha} = \frac{dJ}{dT_{center}} \cdot \frac{dT_{center}}{dT} \cdot \frac{dT}{d\alpha} \quad (\text{B.5})$$

where $\frac{\partial J}{\partial T_{center}}$ is available explicitly and $\frac{dT}{d\alpha}$ is available from (III.40). To find $\frac{dT_{center}}{dT}$ directly, we differentiate the expression

$$F = \Lambda_{pin} T_{pin} - B_{pin}(K, h) = 0 \quad (\text{B.6})$$

with respect to T on both sides, noting the dependencies of K and h on T . The expression simplifies to

$$\frac{dT_{pin}}{dT} = \Lambda_{pin}^{-1} \cdot \left(\frac{\partial \Lambda_{pin}}{\partial T} - \frac{\partial B_{pin}}{\partial T} - \frac{\partial B_{pin}}{\partial K} \frac{\partial K}{\partial T} - \frac{\partial B_{pin}}{\partial h} \frac{\partial h}{\partial T} \right) \quad (\text{B.7})$$

The components $\frac{\partial K}{\partial T}$, $\frac{\partial h}{\partial T}$ are computed above.

The direct dependence of (III.16) appears only in the initial conditions and influences only B_{pin} . $\frac{\partial \Lambda_{pin}}{\partial T} = 0$. In the $m \times n$ cylindrical grid, $\frac{\partial B_{pin}}{\partial T}$ is an $mZ \times nZ$ matrix with entries of 1 in positions (I, J) where I corresponds to the inflow, or the outflow horizontal layer; and J is an index of the lowest or the highest volume element in the central pin, 0 elsewhere.

An $mZ \times n^2$ matrix $\frac{\partial \Lambda_{pin}}{\partial K} \cdot \frac{\partial K}{\partial T}$, has nonzero components of the form $\frac{1}{12} \cdot \frac{\partial \Lambda_{pin}^{(I)}}{\partial K_I} \cdot (\frac{\partial K}{\partial T})^{(J_1, J_2)}$ in positions (I, J_1) , (I, J_2) . The numerical coefficient comes from the fact that K used in the calculation of the center line is an average of six interface values $K^{(J_1, J_2)}$, each depending on two temperature values T_{J_1} , T_{J_2} . Similarly, an $m \times n^2$ matrix $\frac{\partial B_{pin}}{\partial h} \cdot \frac{\partial h}{\partial T}$ has nonzero entries $\frac{1}{12} \cdot \frac{\partial B_{pin}^{(I)}}{\partial h_I} \cdot (\frac{\partial h}{\partial T})^{(J_1, J_2)}$ in positions (I, J_1) , (I, J_2) .

The derivative of the centerline distribution $\frac{\partial T_{center}}{\partial T}$ is extracted from $\frac{\partial T_{pin}}{\partial T}(r=0)$ and used in the expression

$$\begin{aligned} \frac{d\mathcal{J}(T_{center})}{d\alpha} &= \frac{dJ}{dT_{center}} \cdot \left(\frac{dT_{center}}{dT} \cdot \frac{dT}{d\alpha} + \frac{\partial T_{center}}{\partial K} \cdot \frac{dK}{d\alpha} \right. \\ &\quad \left. + \frac{\partial T_{center}}{\partial h} \cdot \frac{dh}{d\alpha} \right) \end{aligned} \quad (\text{B.8})$$

where $\frac{\partial T_{center}}{\partial K} \cdot \frac{dK}{d\alpha} = \Lambda_{pin}^{-1} \cdot \frac{\partial B_{pin}}{\partial K} \cdot \frac{dK}{d\alpha}$, $\frac{\partial T_{center}}{\partial h} \cdot \frac{dh}{d\alpha} = \Lambda_{pin}^{-1} \cdot \frac{\partial B_{pin}}{\partial h} \cdot \frac{dh}{d\alpha}$. The components $\frac{\partial B_{pin}}{\partial K}$, $\frac{dK}{d\alpha}$, $\frac{\partial B_{pin}}{\partial h}$, $\frac{dh}{d\alpha}$ are computed above.

The overall process of finding the derivative (III.41) is reduced to loading precomputed components into sparse arrays, finding an inverse of one matrix (two in the hierarchical example), and performing some sparse array multiplications.

-
- [1] M. ANITESCU, *Constrained optimization formulation of stochastic finite element methods for minimum eigenvalue problems*, tech. rep., Argonne National Laboratory, 2006.
 - [2] M. ANITESCU, G. PALMIOTTI, W. YANG, AND M. NEDA, *Stochastic Finite-Element Approximation of the Parametric Dependence of Eigenvalue Problem Solution*, in Proc. of Mathematics, Computation and Supercomputing in Nuclear Applications Conference, American Nuclear Society, 2007.
 - [3] I. BABUSKA, F. NOBILE, AND R. TEMPONE, *Reliability of computational science*, Numerical Methods for Partial Differential Equations, 23 (2007), p. 753.
 - [4] I. BABUSKA, R. TEMPONE, AND G. ZOURARIS, *Solving elliptic boundary value problems with uncertain coefficients by the finite element method: the stochastic formulation*, Computer Methods in Applied Mechanics and Engineering, 194 (2005), pp. 1251–1294.
 - [5] D. CACUCI, *Sensitivity and Uncertainty Analysis*, Chapman & Hall/CRC, 2003.
 - [6] H. CHENG AND A. SANDU, *Numerical study of uncertainty quantification techniques for implicit stiff systems*, in ACM-SE 45: Proceedings of the 45th annual southeast regional conference, New York, NY, USA, 2007, ACM, pp. 367–372.
 - [7] J. DUDERSTADT AND L. HAMILTON, *Nuclear reactor analysis*, Wiley, New York, 1976.
 - [8] M. S. ELDRED, A. A. GIUNTA, B. G. VAN BLOEMEN WAANDERS, S. F. WOJTKIEWICZ, W. E. HART, AND M. P. ALLEVA, *Dakota, a multilevel parallel object-oriented framework for design optimization, parameter estimation, uncertainty quantification, and sensitivity analysis*, SAND Report SAND2001-3796, Sandia National Laboratories, April 2002.
 - [9] J. FERGINZER AND M. PERIC, *Computational Methods for Fluid Dynamics*, Springer, 1997.
 - [10] J. FINK, *Thermophysical properties of uranium dioxide*, Journal of Nuclear Materials, 279 (2000), pp. 1–18.
 - [11] J. FINK AND L. LEIBOWITZ, *Thermodynamic and transport properties of sodium liquid and vapor*, tech. rep., ANL/RE-95/2 (Argonne, IL: Argonne National Laboratory), 1995.
 - [12] R. GHANEM AND P. SPANOS, *Polynomial chaos in stochastic finite elements*, Journal of Applied Mechanics, 57 (1990), p. 197.
 - [13] R. GHANEM AND P. SPANOS, *The Stochastic Finite Ele-*

- ment Method: A Spectral Approach*, Springer, New York, 1991.
- [14] A. GRIEWANK, *Evaluating Derivatives: Principles and Techniques of Algorithmic Differentiation*, no. 19 in Frontiers in Applied Mathematics, SIAM, Philadelphia, 2000.
- [15] M. D. GUNZBURGER, *Perspectives in Flow Control and Optimization*, Society for Industrial and Applied Mathematics, Philadelphia, 2003.
- [16] J. HIRIART-URRUTY AND C. LEMARÉCHAL, *Convex Analysis and Minimization Algorithms*, Springer-Verlag, Berlin, 1993.
- [17] S. ISUKAPALLI, A. ROY, AND P. GEORGOPOULOS, *Efficient sensitivity/uncertainty analysis using the combined stochastic response surface method and automated differentiation: Application to environmental and biological systems*, Risk Analysis, 20 (2000), pp. 591–602.
- [18] E. JAYNES AND G. BRETTHORST, *Probability Theory: The Logic of Science*, Cambridge University Press Cambridge, Cambridge, UK, 2003.
- [19] N. KIM, H. WANG, AND N. QUEIPO, *Adaptive reduction of random variables using global sensitivity in reliability-based optimisation*, International Journal of Reliability and Safety, 1 (2006), pp. 102–119.
- [20] J. LIU, *Monte Carlo Strategies in Scientific Computing*, Springer, Berlin, 2001.
- [21] R. LORENTZ, *Multivariate hermite interpolation by algebraic polynomials: A survey*, Journal of Computational and Applied Mathematics, 122 (2000), pp. 167–201.
- [22] H. NIEDERREITER, *Random Number Generation and Quasi-Monte Carlo Methods*, Society for Industrial and Applied Mathematics, Philadelphia, 1992.
- [23] W. ROUX, N. STANDER, AND R. HAFTKA, *Response surface approximations for structural optimization*, International Journal for Numerical Methods in Engineering, 42 (1998), pp. 517–534.
- [24] E. SOLAK, R. MURRAY-SMITH, W. LEITHEAD, D. LEITH, AND C. RASMUSSEN, *Derivative Observations in Gaussian Process Models of Dynamic Systems*, Advances in Neural Information Processing Systems, (2003), pp. 1057–1064.
- [25] P. SPANOS AND R. GHANEM, *Stochastic finite element expansion for random media*, Journal of Engineering Mechanics, 115 (1989), pp. 1035–1053.
- [26] Y. S. TANG, J. R. D. COFFIELD, AND J. R. A. MARKLEY, *Thermal Analysis of Liquid-Metal Fast Breeder Reactors*, American Nuclear Society, LaGrange Park, USA, 1978.
- [27] D. XIU AND J. HESTHAVEN, *High-order collocation methods for differential equations with random inputs*, SIAM Journal on Scientific Computing, 27 (2005), pp. 1118–1139.

<p>The submitted manuscript has been created by the University of Chicago as Operator of Argonne National Laboratory ("Argonne") under Contract No. DE-AC02-06CH11357 with the U.S. Department of Energy. The U.S. Government retains for itself, and others acting on its behalf, a paid-up, nonexclusive, irrevocable worldwide license in said article to reproduce, prepare derivative works, distribute copies to the public, and perform publicly and display publicly, by or on behalf of the Government.</p>
--