An error estimator for separated representations of highly multidimensional models

A. Ammar [a,*], F. Chinesta [b], P. Diez [c], A. Huerta [c]

[a] *Laboratoire de Rhcphéologie, 13 rue de la Piscine, BP 53 Domaine Universitaire, F-38041 Grenoble cedex 9, France*
[b] *EADS Corporate Foundation International Chair, GeM: UMR CNRS-Ecole Centrale de Nantes, 1 rue de la Noe, BP 92101, F-44321 Nantes cedex 3, France*
[c] *Laboratori de Calcul Numeric, Universidad Politecnica de Cataluña, Jordi Girona, 1 Campus Nord, C2, E-08034 Barcelona, Spain*

ABSTRACT

Fine modeling of the structure and mechanics of materials from the nanometric to the micrometric scales uses descriptions ranging from quantum to statistical mechanics. Most of these models consist of a partial differential equation defined in a highly multidimensional domain (e.g. Schrodinger equation, Fokker-Planck equations among many others). The main challenge related to these models is their associated curse of dimensionality. We proposed in some of our former works a new strategy able to circumvent the curse of dimensionality based on the use of separated representations (also known as finite sums decomposition). This technique proceeds by computing at each iteration a new sum that consists of a product of functions each one defined in one of the model coordinates. The issue related to error estimation has never been addressed. This paper presents a first attempt on the accuracy evaluation of such a kind of discretization techniques.

Keywords:
Separated representations, Finite sums decomposition, Error estimation, Curse of dimensionality High dimensional models

1. Introduction

Models encountered in computational physics and engineering at the nanometric and micrometric scales usually involve too many coordinates that include the physical ones $\mathbf{x}$, time t and different conformation coordinates whose nature depends on the problem under consideration. The interested reader can refer to [8] for a survey on such models.

We proposed in some of our former works [2,3] a separated representation strategy able to circumvent the curse of dimensionality that those models involve. Its simpler version was analyzed from a numerical analysis point of view in [13]. Thus, given a model involving the unknown field $\Psi(\mathbf{x}_1,\mathbf{x}_2,\cdots,\mathbf{x}_D)$, with $(\mathbf{x}_1,\mathbf{x}_2,\cdots,\mathbf{x}_D)\in \Omega_1\times\Omega_2\times\cdots\Omega_D$, where $\Omega_i\subset\mathcal{R}^{d_i}$, $\forall i$, $1\leq i\leq D$, the solution is searched as

$$\Psi(\mathbf{x}_1,\mathbf{x}_2,\cdots,\mathbf{x}_D)\approx \sum_{i=1}^{i=N} F_1^i(\mathbf{x}_1)\cdots F_D^i(\mathbf{x}_D) \tag{1}$$

where the dimension of the model is $DIM=\sum_{j=1}^{j=D} d_j$.

Obviously, if a mesh is defined in each domain Ω_i consisting of N_n^i nodes, the whole mesh of $\Omega=\Omega_1\times\cdots\times\Omega_D$ involves $\prod_{i=1}^{i=D} N_n^i$ nodes. We can notice that the number of degrees of freedom increases exponentially with the dimension of the space. On the contrary, by assuming the separated representation (1) the number of degrees of freedom reduces to: $N\times(\sum_{i=1}^{i=D} N_n^i)$ proving that the complexity scales linearly with the dimension of the space.

In the next section we are revisiting the algorithm proposed in [2] for building-up the separated representation. For the sake of simplicity we restrict our description to a 2D model. In Section 3 the algorithm will be generalized to higher dimensions. The algorithm proceeds iteratively and at each iteration a term of the sum (1) is computed. This kind of representation involves two levels of adaptivity. The first level consists of the number of sums that must be computed to attain a certain precision; however it is clear that considering the more and more terms in Eq. (1) the solution converges to the one associated to the fully tensorial product of the bases defined on each mesh of each domain Ω_i. The second level of adaptivity lies in the refinement that must be applied to each mesh of each domain Ω_i to enhance the solution accuracy.

In this paper we are analyzing the first level of adaptivity, the second one being a work in progress. The error estimator here considered is based on the solution of the so-called dual problem, deeply analyzed in numerous works referenced in Section 5.

We illustrate in the next section the separated representation procedure for solving multidimensional models. In Section 3 we introduce a tensorial notation as a general framework to define separated representations of generic multidimensional models. In Section 4 we address some considerations on the solution error. The coupling between the error estimator based on the use of the so-called dual problem and the separated representation is then described in Section 5 and different numerical examples are analyzed in Section 6.

* Corresponding author.
E-mail addresses: Amine.Ammar@ujf-grenoble.fr (A. Ammar), Francisco.Chinesta@ec-nantes.fr (F. Chinesta), pedro.diez@upc.es (P. Diez), antonio.huerta@upc.es (A. Huerta).

2. Illustrating the separated representation solver

For the sake of simplicity we are considering a simple 2D model in which the unknown field depends on the coordinates x and y each one defined on its associated domain Ω_x and Ω_y respectively. Now, the solution is searched in the form:

$$\Psi(x,y) \approx \sum_{i=1}^{i=N} F^i(x) \cdot G^i(y) \tag{2}$$

where the different functions $F_i(x)$ and $G_i(y)$ are a priori unknown being computed during the solution process.

In order to build-up this separated representation, an iterative algorithm is proposed. If we assume that the first n functions in the sum have been already computed, the next iteration consists in finding the best functions $R(x)$ and $S(y)$ such that the updated representation given by

$$\Psi(x,y) \approx \sum_{i=1}^{i=n} F^i(x) \cdot G^i(y) + R(x) \cdot S(y) \tag{3}$$

satisfies the weak formulation.

We are assuming a generic equation

$$\mathcal{K}(\Psi(x,y)) + \mathcal{L}(\Psi(x,y)) = 0 \tag{4}$$

where $\mathcal{K}$ and $\mathcal{L}$ are two differential linear operators (the non-linear case was addressed in [4]). For the sake of simplicity we are assuming that the first one only involves derivatives with respect to the x-coordinate, the second one involving the derivatives with respect to the other coordinate. More general scenarios will be addressed in the next section.

The weak form related to Eq. (4) can be written as:

$$\int_{\Omega_x \times \Omega_y} \Psi^* \cdot (\mathcal{K}(\Psi) + \mathcal{L}(\Psi)) dx \ dy = 0. \tag{5}$$

Thus, the solution process requires at each iteration in the construction of the approximation Eq. (2) the computation of functions $R(x)$ and $S(y)$ involved in Eq. (3). As this problem is non-linear, the use of a linearization strategy is compulsory. The simplest choice consists of an alternating directions fixed point strategy that iterates, until reaching convergence, by computing $R(x)$ (with $S(y)$ given at the previous iteration of the non-linear solver) and then $S(y)$ (from the just computed $R(x)$). We summarize both steps:

• When $S(y)$ is assumed to be known, the trial function Ψ^* appearing in Eq. (5) writes $R^*(x) \cdot S(y)$, and then the weak form Eq. (5) reads:

$$\int_{\Omega_x \times \Omega_y} R^* \cdot S \cdot (\mathcal{K}(R) \cdot S + R \cdot \mathcal{L}(S)) dx \ dy$$
$$= -\int_{\Omega_x \times \Omega_y} R^* \cdot S \cdot \sum_{i=1}^{i=n} \left(\mathcal{K}\left(F^i\right) \cdot G^i + F^i \cdot \mathcal{L}\left(G^i\right) \right) dx \, dy. \tag{6}$$

Now, as all the functions depending on y are known, we can integrate Eq. (6) with respect to that coordinate. Thus, defining the following coefficients:

$$\begin{cases} \alpha^y = \int_{\Omega_y} S^2 dy \\ \alpha_i^y = \int_{\Omega_y} S \cdot G^i dy \\ \beta^y = \int_{\Omega_y} S \cdot \mathcal{L}(S) dy \\ \beta_i^y = \int_{\Omega_y} S \cdot \mathcal{L}\left(G^i\right) dy \end{cases} \tag{7}$$

Eq. (6) reduces to:

$$\int_{\Omega_x} R^* \cdot (\alpha^y \mathcal{K}(R) + \beta^y R) dx \tag{8}$$
$$= -\int_{\Omega_x} R^* \cdot \sum_{i=1}^{i=n} \left(\alpha_i^y \mathcal{K}\left(F^i\right) + \beta_i^y F^i \right) dx$$

• When $R(x)$ is assumed to be known, the trial function Ψ^* appearing in Eq. (5) writes $R(x) \cdot S^*(y)$, and then the weak form Eq. (5) reads:

$$\int_{\Omega_x \times \Omega_y} R \cdot S^* \cdot (\mathcal{K}(R) \cdot S + R \cdot \mathcal{L}(S)) dx dy \tag{9}$$
$$= -\int_{\Omega_x \times \Omega_y} R \cdot S^* \cdot \sum_{i=1}^{i=n} \left(\mathcal{K}\left(F^i\right) \cdot G^i + F^i \cdot \mathcal{L}\left(G^i\right) \right) dx \ dy.$$

Now, as all the functions depending on x are known, we can integrate Eq. (9) with respect to that coordinate. Thus, defining the following coefficients:

$$\begin{cases} \alpha^x = \int_{\Omega_x} R^2 dx \\ \alpha_i^x = \int_{\Omega_x} R \cdot F^i dx \\ \beta^x = \int_{\Omega_x} R \cdot \mathcal{K}(R) dx \\ \beta_i^x = \int_{\Omega_x} R \cdot \mathcal{K}(F^i) dx \end{cases} \tag{10}$$

Eq. (9) reduces to:

$$\int_{\Omega_y} S^* \cdot (\beta^x S + \alpha^x \mathcal{L}(S)) dy$$
$$= -\int_{\Omega_y} S^* \cdot \sum_{i=1}^{i=n} \left(\beta_i^x G^i + \alpha_i^x \mathcal{L}\left(G^i\right) \right) dy. \tag{11}$$

Eqs. (8) and (11) are solved iteratively until reaching convergence. If we denote by R^k and R^{k-1} two consecutive solutions of Eq. (8), and similarly for the solutions of Eq. (11), S^k and S^{k-1}, the convergence is reached when the condition $\|R^k \cdot S^k - R^{k-1} \cdot S^{k-1}\| < \epsilon$ is verified (ϵ being a small enough parameter). In our numerical experiments the number of iterations involved in this non-linear solution rarely exceeds some few iterations.

The computer implementation of these steps needs for a discrete representation of all the functions: $F^i(x)$, $G^i(y)$, $R(x)$ and $S(y)$ using a finite element interpolation in their associated domains Ω_x and Ω_y. For this purpose we introduce the vectors **M** and **N** containing the shape functions associated with the meshes of Ω_x and Ω_y respectively. Finally $\mathbf{F}^i$, $\mathbf{G}^i$, $\mathbf{R}$ and $\mathbf{S}$ represent the nodal description of the associated functions.

From Eqs. (8) and (11) we can notice that the following matrices should be computed to define their associated discrete forms:

$$\begin{cases} \mathbb{M} = \int_{\Omega_x} \mathbf{M} \mathbf{M}^T d\Omega_x \\ \mathbb{N} = \int_{\Omega_y} \mathbf{N} \mathbf{N}^T d\Omega_y \\ \mathbb{K} = \int_{\Omega_x} \mathbf{M} \mathcal{K} \mathbf{M}^T d\Omega_x \\ \mathbb{L} = \int_{\Omega_y} \mathbf{N} \mathcal{L} \mathbf{N}^T d\Omega_y \end{cases} \tag{12}$$

Remark. These integrals take into account the specific character of each operator. For example integration by parts is used in second order operators, upwinding for stabilizing advective terms, etc.

Using this notation the discrete form related to Eqs. (8) and (11) results:

- For a given **S** (that implies $\Psi^*(x, y) = R^*(x) \cdot S(y)$)

$$\left(\mathbb{K}\mathbf{S}^T\mathbb{N}\mathbf{S} + \mathbb{M}\mathbf{S}^T L\mathbf{S}\right)\mathbf{R}$$
$$= -\sum_{i=1}^{n}\left(\mathbb{K}\mathbf{F}^i\mathbf{S}^T\mathbb{N}\mathbf{G}^i + \mathbb{M}\mathbf{F}^i\mathbf{S}^T\mathbb{L}\mathbf{G}^i\right) \tag{13}$$

- For a given **R** (that implies $\Psi^*(x, y) = R(x) \cdot S^*(y)$)

$$\left(\mathbf{R}^T\mathbb{K}\mathbf{R}\mathbb{N} + \mathbf{R}^T\mathbb{M}\mathbf{R}\mathbb{L}\right)\mathbf{S}$$
$$= -\sum_{i=1}^{n}\left(\mathbf{R}^T\mathbb{K}\mathbf{F}^i\mathbb{N}\mathbf{G}^i + \mathbf{R}^T\mathbb{M}\mathbf{F}^i\mathbb{L}\mathbf{G}^i\right). \tag{14}$$

It must be noticed that the number of the degrees of freedom involved in such a non-linear solution is the sum of the degrees of freedom involved in each discretization (instead of the product that results in mesh based discretization strategies).

Finally, after convergence, the new approximation functions $\mathbf{F}^{n+1}$ and $\mathbf{G}^{n+1}$ are obtained directly from the converged **R** and **S**:

$$\begin{cases} \mathbf{F}^{n+1} = \mathbf{R} \\ \mathbf{G}^{n+1} = \mathbf{S} \end{cases}. \tag{15}$$

In our previous works we used as a stopping criterion of the enrichment process the norm of the residual. In those works we analyzed the convergence rate as a function of the number of terms in the finite sums decomposition [4]. In this work we are using other criteria based in a quantity of interest.

3. General framework

In this section we are introducing a tensor notation that allows generalizing the procedure described in the previous section to more general models.

Let Ω be a multidimensional domain involving the coordinates x_1, x_2, ⋯, x_D (each one not necessarily one dimensional). The solution of the weak form

$$a(\Psi(x_1, x_2, \cdots, x_D), \Psi^*(x_1, x_2, \cdots, x_D)) = b(\Psi^*(x_1, x_2, \cdots, x_D)) \tag{16}$$

can be approximated by

$$\Psi(x_1, x_2, \cdots, x_D) \approx \sum_{i=1}^{n_F} F_1^i(x_1) \cdot F_2^i(x_2) \cdots F_D^i(x_D) \tag{17}$$

whose tensor form writes:

$$\Psi = \sum_{j=1}^{n_F} \mathbf{F}_1^j \otimes \mathbf{F}_2^j \otimes \cdots \otimes \mathbf{F}_D^j. \tag{18}$$

Thus, the discrete form of Eq. (16) reads

$$\Psi^{*T}\mathcal{A}\Psi = \Psi^{*T}\mathcal{B} \tag{19}$$

where

$$\mathcal{A} = \sum_{j=1}^{n_A} \mathbb{A}_1^j \otimes \mathbb{A}_2^j \otimes \cdots \otimes \mathbb{A}_D^j \tag{20}$$

and

$$\mathcal{B} = \sum_{j=1}^{n_B} \mathbf{B}_1^j \otimes \mathbf{B}_2^j \otimes \cdots \otimes \mathbf{B}_D^j. \tag{21}$$

Remark 1. The model solved in the previous section could be written as:

$$\begin{cases} \mathcal{B} = 0 \\ \mathcal{A} = A_1^1 \otimes A_2^1 + A_1^2 \otimes A_2^2 \end{cases} \tag{22}$$

where

$$\begin{cases} \mathbb{A}_1^1 = \mathbb{K} \\ \mathbb{A}_2^1 = \mathbb{N} \\ \mathbb{A}_1^2 = \mathbb{M} \\ \mathbb{A}_2^2 = \mathbb{L} \end{cases}. \qquad \square \tag{23}$$

The enrichment step looks for the new candidates ($\mathbf{R}_1$, ⋯, $\mathbf{R}_D$) for enriching the reduced separated approximation basis:

$$\Psi = \underbrace{\sum_{j=1}^{n_F} \mathbf{F}_1^j \otimes \mathbf{F}_2^j \otimes \cdots \otimes \mathbf{F}_D^j}_{\Psi_F} + \underbrace{\mathbf{R}_1 \otimes \mathbf{R}_2 \otimes \cdots \otimes \mathbf{R}_D}_{\Psi_R}. \tag{24}$$

Within a fixed point alternating directions algorithm, we look at each iteration for the computation of a single function $\mathbf{R}_j$ assuming all the others known. Thus, when one is looking for $\mathbf{R}_j$ the test function writes

$$\Psi^* = \mathbf{R}_1 \otimes \cdots \otimes \mathbf{R}_{j-1} \otimes \mathbf{R}_j^* \otimes \mathbf{R}_{j+1} \otimes \cdots \otimes \mathbf{R}_D \tag{25}$$

where the different terms of the discrete weak form read:

$$\Psi^*\mathcal{A}\Psi_R = \sum_{k=1}^{n_A}\left(\mathbf{R}_j^{*T}\mathbb{A}_j^k\mathbf{R}_j \prod_{\substack{h=1 \\ h \neq j}}^{D} \mathbf{R}_h^T\mathbb{A}_h^k\mathbf{R}_h\right) = \mathbf{R}_j^{*T}\mathbb{K}\mathbf{R}_j \tag{26}$$

$$\Psi^*\mathcal{A}\Psi_F = \sum_{i=1}^{n_F}\sum_{k=1}^{n_A}\left(\mathbf{R}_j^{*T}\mathbb{A}_j^k\mathbf{F}_j^i \prod_{\substack{h=1 \\ h \neq j}}^{D} \mathbf{R}_h^T\mathbb{A}_h^k\mathbf{F}_h^i\right) = \mathbf{R}_j^{*T}\mathbf{V} \tag{27}$$

$$\Psi^*\mathcal{B} = \sum_{k=1}^{n_B}\left(\mathbf{R}_j^{*T}\mathbf{B}_j^k \prod_{\substack{h=1 \\ h \neq j}}^{D} \mathbf{R}_h^T\mathbf{B}_h^k\right) = \mathbf{R}_j^{*T}\mathbf{V}'. \tag{28}$$

Finally, $\mathbf{R}_j$ comes from the solution of the linear system:

$$\mathbb{K}\mathbf{R}_j + \mathbf{V} = \mathbf{V}. \tag{29}$$

We assume that the convergence is attained when the number of sums in the separated approximation of Ψ reaches the value N, i.e. when $n_F = N$.

4. Reduced versus fully separated representations

It is easy to imagine that using a fully tensorial product of the basis related to each coordinate x_1, ⋯, x_D one could solve the discrete system associated with a generic PDE:

$$\mathcal{A}\Psi - \mathcal{B} = 0. \tag{30}$$

The error of such a computed solution Ψ with respect to the exact one Ψ^{ex} depends on the mesh considered. If we assume n_d nodes distributed on each domain Ω_d, $d=1, 2, \cdots, D$, with $\Omega = \Omega_1 \times \cdots \times \Omega_D$, the number of approximation functions will be $n_1 \times \cdots \times n_D$.

When we are using the strategy just described, in the limit case when $N = n_1 \times \cdots \times n_D$, the computed solution is the same that the hypothetical one that could have been computed by using a fully tensor product approximation basis.

Thus, a first possibility to quantify the reduced separated approximation consists in computing

$$\varepsilon = \|\mathcal{A}\Psi - \mathcal{B}\|_2 \tag{31}$$

that writes:

$$\varepsilon = \sqrt{\Psi^T \mathcal{A}^T \mathcal{A}\Psi + \mathcal{B}^T \mathcal{B} - 2\mathcal{B}^T \mathcal{A}\Psi} \tag{32}$$

where each term can be computed from:

$$\Psi^T \mathcal{A}^T \mathcal{A}\Psi = \sum_{i=1}^{n_F} \sum_{j=1}^{n_F} \sum_{k=1}^{n_A} \sum_{h=1}^{n_A} \prod_{d=1}^{D} \mathbf{F}_d^{iT} \mathbb{A}_d^{kT} \mathbb{A}_d^{h} \mathbf{F}_d^{j} \tag{33}$$

$$\mathcal{B}^T \mathcal{B} = \sum_{i=1}^{n_B} \sum_{j=1}^{n_B} \prod_{d=1}^{D} \mathbf{B}_d^{iT} \mathbf{B}_d^{j} \tag{34}$$

$$\mathcal{B}^T \mathcal{A}\Psi = \sum_{i=1}^{n_B} \sum_{j=1}^{n_F} \sum_{k=1}^{n_A} \prod_{d=1}^{D} \mathbf{B}_d^{iT} \mathbb{A}_d^{k} \mathbf{F}_d^{j}. \tag{35}$$

5. Error estimator based on a dual formulation

This section aims at introducing an objective error assessment as a convergence criterion for the iterative scheme successively enriching the separated representation. The main goal is to compute the number of terms in the finite sums decomposition such that the error related to a certain quantity of interest is lower than a certain threshold value. The error assessment is performed on the basis of an arbitrary linear quantity of interest represented by a linear functional $o(\cdot)$. Thus, the enrichment scheme loops while the error in this quantity is larger than the prescribed tolerance and stops once the prescribed accuracy is reached.

Following the ideas introduced in [5,6,15] the error in the quantity of interest is represented using an auxiliary problem (denoted as dual or adjoint problem). This strategy has been widely used by many authors in the field of error estimation and adaptivity, using different error estimation techniques [1,10–12,14,16–18]. Here, in order to obtain a proper error assessment, the exact solution of the auxiliary problem (also denoted as extractor) is simply replaced by a more accurate solution. In the context of PGD, this option is very effective and costless because it only requires performing some extra enrichment iterations for the adjoint problem.

In this section we are assuming a generic multidimensional model whose weak form writes

$$a(\Psi, \Psi^*) = b(\Psi^*) \tag{36}$$

defined in $\Omega = \Omega_1 \times \cdots \Omega_D$, where each Ω_d $(d = 1, 2, \cdots, D)$ involves the coordinate x_d (not necessarily one-dimensional). From now on, the form (36) will be referred as primal form.

The discrete counterpart of Eq. (36) reads

$$\mathcal{A}\Psi = \mathcal{B} \tag{37}$$

where

$$\mathcal{A} = \sum_{j=1}^{n_A} \mathbb{A}_1^{j} \otimes \mathbb{A}_2^{j} \otimes \cdots \otimes \mathbb{A}_D^{j} \tag{38}$$

and

$$\mathcal{B} = \sum_{j=1}^{n_B} \mathbf{B}_1^{j} \otimes \mathbf{B}_2^{j} \otimes \cdots \otimes \mathbf{B}_D^{j}. \tag{39}$$

The separated representation of Ψ at iteration n_F writes:

$$\Psi^{ap} \approx \sum_{j=1}^{n_F} \mathbf{F}_1^{j} \otimes \mathbf{F}_2^{j} \otimes \cdots \otimes \mathbf{F}_D^{j} \tag{40}$$

where $\mathbf{F}_i^{j}$ is the discrete (nodal) form of $F^j(x_i)$.

Now, we are interested in a certain function of Ψ, $o(\Psi)$, of physical interest (the model output). In what follows we assume that the operator defining the output is linear. Thus, we could write

$$o(\Psi^{ap}) = \mathcal{O}\left(\sum_{j=1}^{n_F} \mathbf{F}_1^{j} \otimes \mathbf{F}_2^{j} \otimes \cdots \otimes \mathbf{F}_D^{j} \right). \tag{41}$$

If this operator accepts a separated representation, that is:

$$\mathcal{O} = \sum_{j=1}^{n_O} \mathbf{O}_1^{j} \otimes \mathbf{O}_2^{j} \otimes \cdots \otimes \mathbf{O}_D^{j} \tag{42}$$

then, the output can be evaluated from:

$$o(\Psi^{ap}) = \sum_{j=1}^{n_F} \sum_{i=1}^{n_O} \left(\mathbf{F}_1^{jT} \mathbf{O}_1^{i} \right) \cdot \left(\mathbf{F}_2^{jT} \mathbf{O}_2^{i} \right) \cdots \left(\mathbf{F}_D^{jT} \mathbf{O}_D^{i} \right) \tag{43}$$

Now, the error on the output can be evaluated by solving the so-called dual problem:

$$a(\Psi^*, \Phi) = o(\Psi^*) \tag{44}$$

whose discrete form writes:

$$\mathcal{A}^T \Phi = \mathcal{O} \tag{45}$$

where $\mathcal{O}$ was already defined and where $\mathcal{A}^T$ is given by

$$\mathcal{A}^T = \sum_{j=1}^{n_A} \mathbb{A}_1^{jT} \otimes \mathbb{A}_2^{jT} \otimes \cdots \otimes \mathbb{A}_D^{jT}. \tag{46}$$

A good error estimation needs for an accurate solution of the dual problem. Within the separated representation framework the solution of the dual problem (45) (assumed accurate enough) can be written as:

$$\Phi \approx \sum_{j=1}^{n_G} \mathbf{G}_1^{j} \otimes \mathbf{G}_2^{j} \otimes \cdots \otimes \mathbf{G}_D^{j}. \tag{47}$$

Now, the error in the output can be evaluated:

$$o(e) = b(\Phi) - a(\Psi^{ap}, \Phi) \tag{48}$$

whose discrete counterpart writes:

$$o(e) = \Phi^T \mathcal{B} - \Phi^T \mathcal{A}\Psi^{ap} \tag{49}$$

4

that is easily computed from:

$$o(e) = \sum_{i=1}^{n_B} \sum_{j=1}^{n_C} \left(\mathbf{G}_1^{jT}\mathbf{B}_1^i\right) \cdot \left(\mathbf{G}_2^{jT}\mathbf{B}_2^i\right) \cdots \left(\mathbf{G}_D^{jT}\mathbf{B}_D^i\right)$$
$$- \sum_{k=1}^{n_A} \sum_{i=1}^{n_F} \sum_{j=1}^{n_C} \left(\mathbf{G}_1^{jT}\mathbb{A}_1^k\mathbf{F}_1^i\right) \cdot \left(\mathbf{G}_2^{jT}\mathbb{A}_2^k\mathbf{F}_2^i\right) \cdots \left(\mathbf{G}_D^{jT}\mathbb{A}_D^k\mathbf{F}_D^i\right).$$

(50)

6. Numerical results

In this section we are addressing some simple numerical examples in order to illustrate the capabilities of the error estimator previously proposed. In these examples neither the model parameters not the particular form of the outputs have physical sense.

6.1. Poisson equation with exact solution of the dual problem

We consider the 2D harmonic problem defined in $\Omega = \Omega_x \times \Omega_y = (0,1) \times (0,1)$ with homogeneous boundary conditions:

$$-\Delta\Psi = \sum_{i=1}^{20} \sin(i\pi x) \cdot \sin(i\pi y)$$

(51)

whose exact solution writes:

$$\Psi^{ex}(x,y) = \sum_{i=1}^{20} \frac{1}{2i^2\pi^2} \sin(i\pi x) \cdot \sin(i\pi y).$$

(52)

In what follows we consider the following output:

$$o(\Psi) = \int \left(-2\left(x^2-x\right)-2\left(y^2-y\right)\right)\Psi d\Omega$$

(53)

that corresponds to the dual problem:

$$-\Delta\Phi = -2\left(x^2-x\right)-2\left(y^2-y\right)$$

(54)

whose exact solution is also known:

$$\Phi(x,y) = \left(x^2-x\right) \cdot \left(y^2-y\right).$$

(55)

To avoid the effect of the eventual numerical error related to a numerical solution of the dual problem, in what follows we consider the exact solution of the dual problem (55).

We are computing two errors:

• Error estimator based on the solution of the dual problem

$$\varepsilon_d = b(\Phi)-a\left(\Psi^{ap},\Phi\right) = \Phi^T\left(\mathcal{B}-\mathcal{A}\Psi^{ap}\right)$$

(56)

where in the first simulations we consider $\Phi = \Phi^{ex}$ given by Eq. (55).

• Exact error

$$\varepsilon_a = o\left(\Psi^{ex}\right)-o\left(\Psi^{ap}\right)$$

(57)

where

$$o\left(\Psi^{ex}\right) = \left|\int\left(-2\left(x^2-x\right)-2\left(y^2-y\right)\right)\Psi^{ex}d\Omega\right|$$

and

$$o\left(\Psi^{ap}\right) = \left|\int\left(-2\left(x^2-x\right)-2\left(y^2-y\right)\right)\Psi^{ap}d\Omega\right|.$$

(58)

We can notice in Fig. 1 that the error decays as expected very fast. It can be also noticed in that figure that due to the particular form of
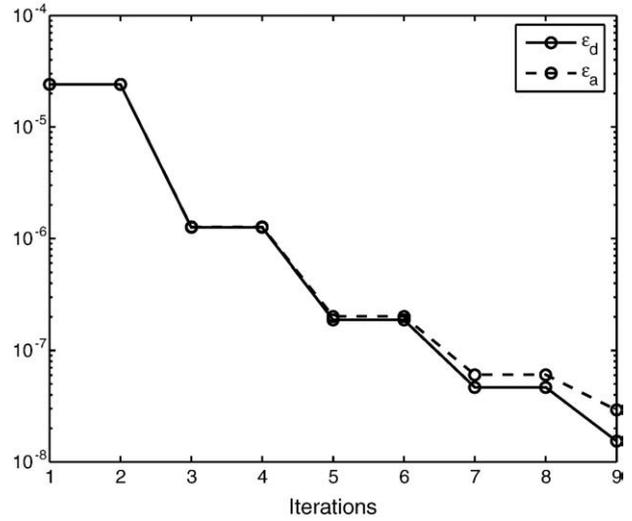

Fig. 1. Error evolution.

the error, the introduction of new terms in the finite sums decomposition does not affect in the same way the evolution of the error. Thus, the introduction of the second term in the decomposition does not affect the errors defined in Eqs. (56) and (57). However, the introduction of the third term in the decomposition reduces in more than one order of magnitude both errors. Fig. 1 also compares the evolution of both errors given by Eqs. (56) and (57). Theoretically both expressions must give the same results, but slight differences are noticed from the 7th iteration. We must recall that the error at the 7th iteration is lower than 10^{-7} and then the differences noticed have not an impact on the overall technique. On the other hand, these slight differences can be explained from the fact that both errors are subjected to the approximation, discretization and integration errors that affect differently both error expressions.

Fig. 2 shows the evolution of the residual using the L^2-norm. We can notice that the residual decreases monotonically during the first 20 iterations. This behavior was expected, because the exact solution (52) consists of a sum of 20 terms. Thus, the main part of the residual is captured in 20 iterations and further enrichments improve very slightly the residual because they try to capture the different approximation and discretization errors. In any case, when the sum reaches the 20 terms, the residual norm is of order 10^{-10}.
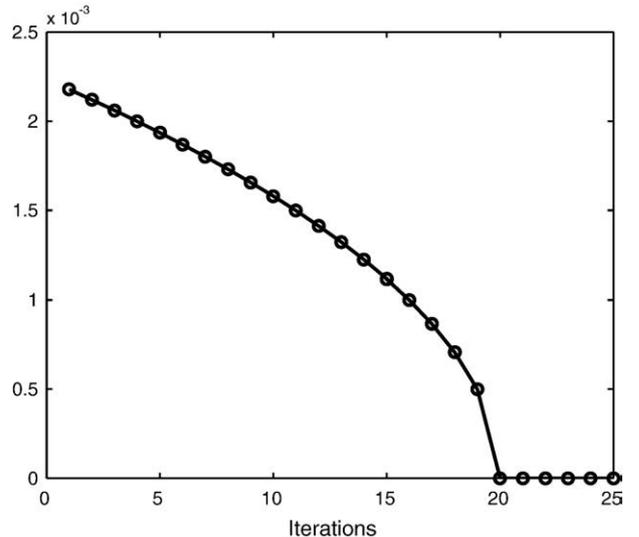

Fig. 2. Evolution of the residual L^2-norm.

Until now, the analyzed problem had a known solution of the dual problem. In the next section we consider a slightly different model.

6.2. Poisson equation involving a numerical solution of the dual problem

In this section we consider three different primal problems:

- *Problem* 1:

$$-\Delta\Psi = \sum_{i=1}^{20} sin(i\pi x)\cdot sin(i\pi y) \qquad (59)$$

- *Problem* 2:

$$-\Delta\Psi = 1 \qquad (60)$$

- *Problem* 3:

$$-\Delta\Psi = x\cdot y. \qquad (61)$$

all of them defined in $\Omega = \Omega_x \times \Omega_y = (0,1) \times (0,1)$ with homogeneous Dirichlet boundary conditions.

In the three cases the output was done by:

$$o(\Psi) = \int x\cdot y\Psi d\Omega \qquad (62)$$

whose associated dual problem writes:

$$-\Delta\Phi = x\cdot y \qquad (63)$$

whose exact solution is not given "a priori".

Remark 2. *Problem* 3 involves the same primal and dual problems.

Remark 3. *Problems* 2 and 3 imply non-separable solutions, i.e. their solutions cannot be expressed exactly from a finite sum of functional products.

Remark 4. When models involve non-separated right hand members, before performing the solution, the right hand members should be separated up to a given precision, by invoking the singular value decomposition or our separated representation algorithm itself. This practice does not involve major difficulties as noticed in [8] and [19].

It is well known [10] that the use of the error estimator described in the previous section (Section 5) needs a more accurate solution of the dual problem than of the primal one. In order to check this requirement in the separated representation framework we are considering different ways of solving the dual problem:

- The dual problem is solved until convergence. Its solution represents the finest one and it will be the reference one.
- The dual problem is solved with two times more terms in the finite sums decomposition than the primal one. Its solution will be noted as $dual \times 2$.
- The finite sums decomposition of the dual problem involves two more terms than the solution of the primal one. This solution will be noted as $dual + 2$.
- The finite sums decomposition of the dual problem involves one more term than the solution of the primal one. This solution will be noted as $dual + 1$.
- Both primal and dual solutions consist of the same number of terms in their separated representations. This solution is noted as $dual \times 1$.

Figs. 3–5 show the convergence analysis of *Problem* 1, *Problem* 2 and *Problem* 3 respectively. The results depicted in Fig. 4 prove that the computed error is quite insensible to the way in which the dual problem is solved. Moreover, we can notice that the error decreases
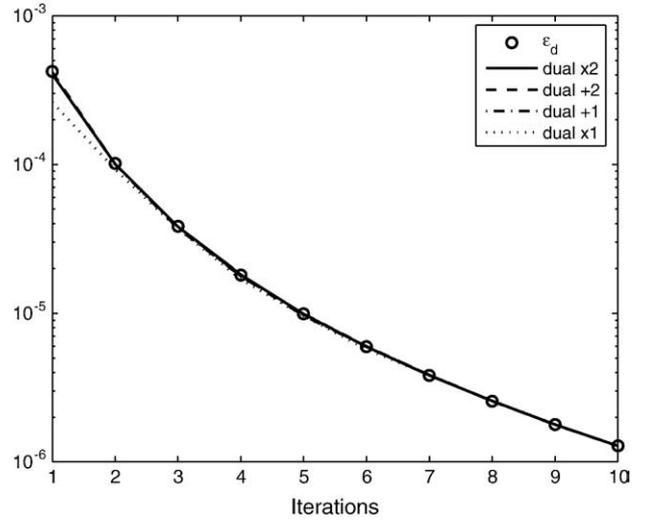

Fig. 3. Error evolution associated with *Problem* 1.

monotonically with the number of terms involved in the finite sums decomposition.

In the other two problems the reference solution of the dual problem defined the upper bound of the estimated error as noticed in Figs. 5 and 6. We can also remark that the evolution of the error is no more decreasing monotonically. Finally, it is important to note that in both cases, as expected, the worst error estimation lies in the use of the same number of terms in the decomposition for both the dual and the primal problem.

In the context of finite elements discretizations when the solutions of the primal and dual problems are computed using the same mesh the Galerkin orthogonality implies a null residual, fact that implies the necessity of solving the dual problem using a finer mesh. When the three just referred problems were solved by using a standard finite element discretization instead of the separated representation, a zero error estimation was obtained when using the same meshes for solving both the primal and the dual problems. In the separated representation framework this situation is slightly different, less severe in general, because the orthogonality is only attained when the separated representation involves all the sums of the fully tensor product of the one-dimensional bases. If we consider only a reduced number of terms in the decomposition the orthogonality is not activated, except if both dual and primal problems are the same, as was the case in *Problem* 3 in which we can notice that the solutions
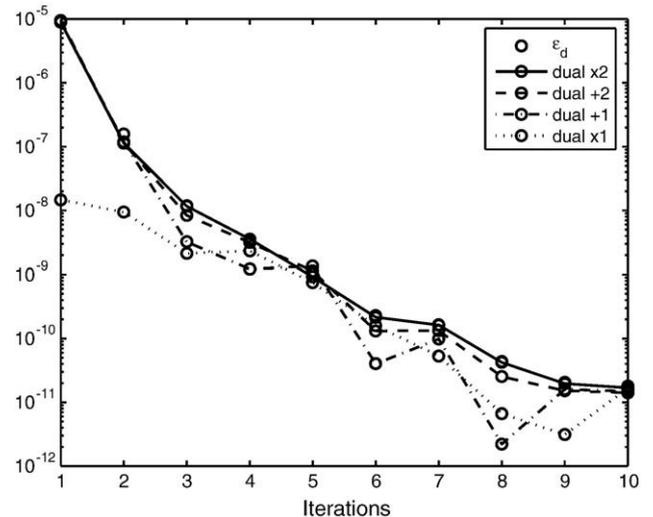

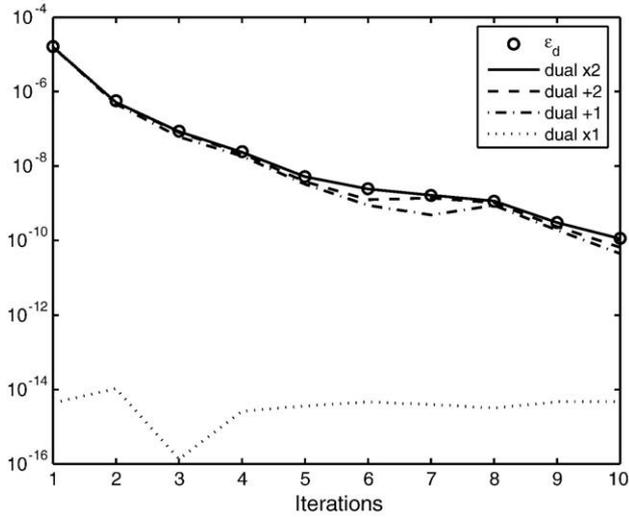**Fig. 4.** Error evolution associated with *Problem* 2.

Fig. 5. Error evolution associated with *Problem* 3.

consisting of the same sums give a null error estimation as shown in Fig. 6.

6.3. Solving a 3D model

Homogenization is a usual procedure in solid mechanics. The simplest one concerns the linear behavior of microstructured materials. For this purpose a 3*D* problem must be solved within a representative volume with quite simple boundary conditions. For accounting for all the microstructural details very fine meshes must be preferably used, however, fine meshes imply important computation time and computer resources. In this case the use of separated representations could be an appealing choice [9]. However, in some applications, the primal field is not the field of interest, and then one must define a more appropriate error estimator as a stopping criterion for the enrichment of a separated representation. This section focuses in this issue.

We are again solving the Laplace equation (as the one that usually appears in thermal homogenization) in $\Omega = \Omega_x \times \Omega_y \times \Omega_z = (0,1)^3$ using a regular mesh consisting of $2001 \times 2001 \times 2001$ nodes. Thus, the primal problem writes:

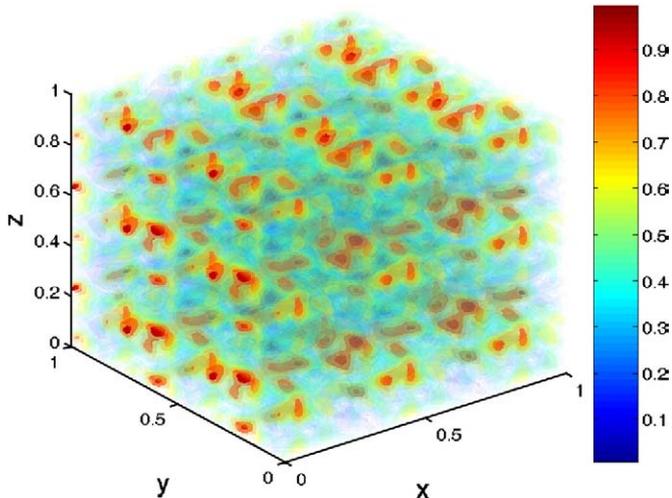$$\nabla \cdot (-\mathbf{k}(\mathbf{x}) \cdot \nabla \Psi(\mathbf{x})) = 0 \qquad (64)$$

whose solution is subjected to an affine field on the domain boundary, as for example:

$$\Psi(\mathbf{x} \in \partial\Omega) = x \qquad (65)$$

The output chosen in this numerical test was given by:

$$o(\Psi) = \int x \Psi d\Omega \qquad (66)$$

that corresponds to the dual problem

$$\nabla \cdot (-\mathbf{k}(\mathbf{x}) \cdot \nabla \Phi(\mathbf{x})) = x \qquad (67)$$

that is associated to null Dirichlet boundary conditions.

The conductivity tensor $\mathbf{k}(\mathbf{x})$ that was assumed isotropic, i.e. $\mathbf{k}(\mathbf{x}) = k(\mathbf{x})\mathbf{I}$ ($\mathbf{I}$ being the unit tensor), was built-up from a sum of 10 terms consisting of sinus functions with different frequencies and phase angles. Thus, we can define a conductivity varying in the interval $[0,1]$, where the highest values were depicted in red in Fig. 6 and the regions with lowest conductivity were represented in blue.

Fig. 7 represents the evolution of the estimated error with the number of terms involved in the separated representation. We can notice that despite the fact that the residual decreases monotonically, the estimated error follows a different tendency. Thus, numerous enrichments that contribute to decrease the residual do not contribute to decrease the error associated to the output under consideration. Table 1 depicts the second and the 17th functions of the decomposition. We can notice that the first functions are smoother that the last ones that exhibit higher "frequencies". We don't represent all the decomposition modes to avoid too cloudy images.

Remark 5. When the heat model is not isotropic there are terms involving crossed derivatives. The treatment of this kind of models does not introduce major difficulties. In fact this situation was found in the models treated in some of our former works [2,3].

6.4. Solving a highly multidimensional model

The numerical tests performed in the previous sections only involved 2*D* or 3*D* problems, which can also be solved without major difficulties by using a standard finite element discretization. However, when the dimension of the space increases, standard discretizations fail, and in our knowledge only sparse grids [7] or separated representations can be applied. The former technique can be successfully applied in spaces of moderate dimensions, whereas the last one has proved its ability for solving models defined in spaces
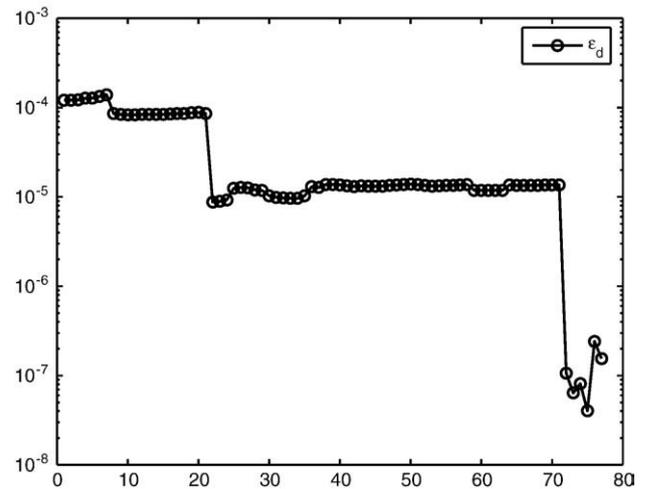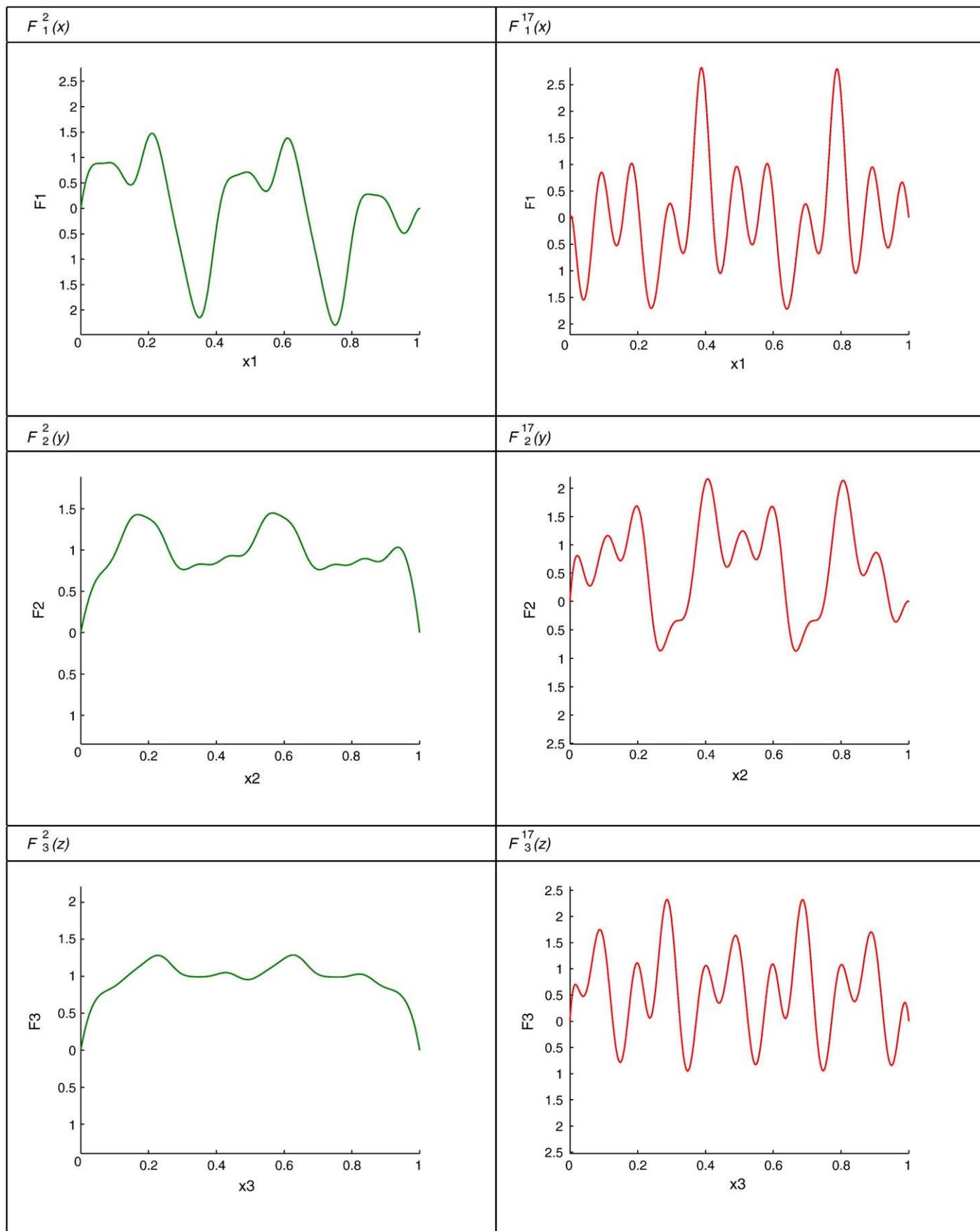

Fig. 6. Material conductivity.

Fig. 7. Evolution of the estimated error.

Table 1
Functions F_i^j, with $j=2$ and $j=17$, involved in the separated representation of $\Psi(x)$.

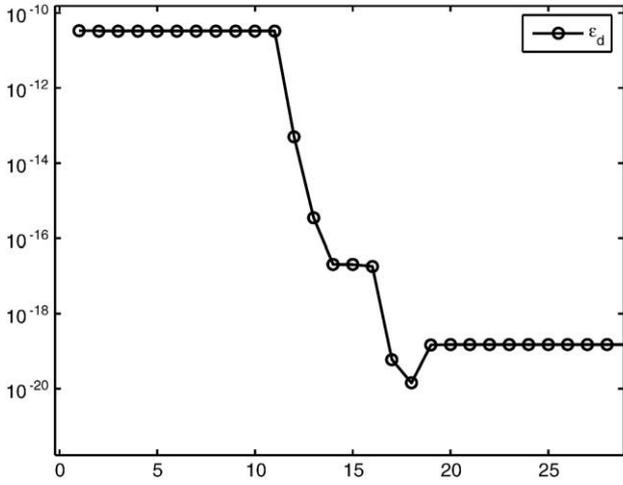| $F_1^2(x)$ | $F_1^{17}(x)$ |
| $F_2^2(y)$ | $F_2^{17}(y)$ |
| $F_3^2(z)$ | $F_3^{17}(z)$ |

Fig. 8. Evolution of the estimated error for the 10*D* problem.

involving an extremely high number of dimensions (hundreds in some cases) [2,3,8].

In this section we are applying the error estimator previously presented to a simple model defined in a space consisting of 10 dimensions.

The primal problem involves the field $\Psi(x_1,\cdots,x_{10})$ verifying

$$-\Delta\Psi = \sum_{i=1}^{i=20}\prod_{j=1}^{j=10} sin(i\pi x_j) \qquad (68)$$

that is defined in $\Omega = \Omega_1 \times \cdots \times \Omega_{10} = (0,1)^{10}$ with homogeneous Dirichlet boundary conditions.

The output was again given by:

$$o(\Psi) = \int\left(\prod_{i=1}^{i=10} x_i\right)\Psi d\Omega \qquad (69)$$

whose associated dual problem writes:

$$-\Delta\Phi = \prod_{i=1}^{i=10} x_i. \qquad (70)$$

Fig. 8 depicts the evolution of the estimated error. We can observe that the error only evolves significantly during the first 20 enrichments, because the problem admits a solution that can be represented separately by means of 20 sums. We can also notice that the error exhibits some plateaus in which the error does not evolve despite the solution enrichment. Even if the equation residual is decreasing monotonically, the error in the output quantity exhibits a different behavior, and as Figs. 7 and 8 proves, many modes contributing to the reduction of the residual do not contribute to reduce the error.

7. Conclusions

This paper presents a first attempt to define a useful error estimator for solvers operating in highly multidimensional models. Until now, this kind of models was untractable because the curse of dimensionality. Recently, we proposed a new technique based on a reduced separated representation in which the solution is approximated by a finite sums decomposition. We verified in some of our former works that for some differential operators (elliptic and symmetric) the computed decomposition is very close to the one resulting from the application of the singular value decomposition to the exact solution, proving the computed decomposition optimality. This technique was extended to non-linear and non-elliptic models in [4]. However, an important issue remains unsolved, the stopping criterion in the enrichment process. We verified numerically that the residual is not the best criterion when we are interested in some output of physical interest other than the residual itself.

Here we proposed the use of an error estimator based on the solution of the so-called dual problem as stopping criterion of the enrichment process. The proposed strategy allows accurate solutions of general models and the first numerical tests included in this paper proved its ability. However different questions remain open, needing further developments: (i) the extension of the error estimator to transient multidimensional models; (ii) the inclusion of more complex outputs (eventually non-linear); and (iii) the definition of some pertinent criteria able to adapt the mesh in each coordinate for reaching further improvements of the solution (until now the accuracy limit is dictated by the meshes considered in each direction and that remain unchanged during the whole solution process).

References

[1] M. Ainsworth, J.T. Oden, A posteriori error estimation in finite element analysis, John Wiley & Sons, 2000.

[2] A. Ammar, B. Mokdad, F. Chinesta, R. Keunings, A new family of solvers for some classes of multidimensional partial differential equations encountered in kinetic theory modeling of complex fluids, Journal of Non-Newtonian Fluid Mechanics 139 (2006) 153–176.

[3] A. Ammar, B. Mokdad, F. Chinesta, R. Keunings, A new family of solvers for some classes of multidimensional partial differential equations encountered in kinetic theory modeling of complex fluids. Part II: transient simulation using space-time separated representation, Journal of Non-Newtonian Fluid Mechanics 144 (2007) 98–121.

[4] A. Ammar, M. Normandin, F. Daim, D Gonzalez, E. Cueto, F. Chinesta, Non incremental strategies based on separated representations: application in computational rheology. Communications in Mathematical Sciences 8/3 (in press).

[5] I. Babuška, A. Miller, The post-processing approach in the finite-element method. 1. Calculation of displacements, stresses and other higher derivatives of the displacements, International Journal for Numerical Methods in Engineering 20 (1984) 1085–1109.

[6] I. Babuška, A. Miller, The post-processing approach in the finite-element method. 1. The calculation of stress intensity factors, International Journal for Numerical Methods in Engineering 20 (1984) 1111–1129.

[7] H.J. Bungartz, M. Griebel, Sparse grids, Acta Numerica 13 (2004) 1–123.

[8] F. Chinesta, A. Ammar, P. Joyot, The nanometric and micrometric scales of the structure and mechanics of materials revisited: an introduction to the challenges of fully deterministic numerical descriptions, International Journal for Multiscale Computational Engineering 6/3 (2008) 191–213.

[9] F. Chinesta, A. Ammar, F. Lemarchand, P. Beauchene, F. Boust, Alleviating mesh constraints: model reduction, parallel time integration and high resolution homogenization, Computer Methods in Applied Mechanics and Engineering 197/5 (2008) 400–413.

[10] P. Díez, G. Calderón, Goal-oriented error estimation for transient parabolic problems, Computational Mechanics 39 (2007) 631–646.

[11] P. Ladevèze, J.P. Pelle, Mastering calculations in linear and nonlinear mechanics, Mechanical Engineering Series, Springer-Verlag, New York, 2005.

[12] P. Ladevèze, Strict upper error bounds on computed outputs of interest in computational structural mechanics, Computational Mechanics 42 (2008) 271–286.

[13] C. Le Bris, T. Lelievre, Y. Maday, Results and questions on a nonlinear approximation approach for solving high-dimensional partial differential equations, Constructive Approximations (2009), doi:10.1007/s00365-009-9071-1.

[14] J.T. Oden, S. Prudhomme, Goal-oriented error estimation and adaptivity for the finite element method, computers & mathematics with applications, An International Journal 41 (2001) 735–756.

[15] M. Paraschivoiu, J. Peraire, A. Patera, A posteriori finite element bounds for linear-functional outputs of elliptic partial differential equations, Computer Methods in Applied Mechanics and Engineering 150 (1997) 289–312.

[16] M. Paraschivoiu, A. Patera, A hierarchical duality approach to bounds for the outputs of partial differential equations, Computer Methods in Applied Mechanics and Engineering 158 (1998) 389–407.

[17] N. Parés, P. Díez, A. Huerta, Bounds of functional outputs for parabolic problems. Part I: exact bounds of the discontinuous Galerkin time discretization, Computer Methods in Applied Mechanics and Engineering 197 (2008) 1641–1660.

[18] N. Parés, P. Díez, A. Huerta, Bounds of functional outputs for parabolic problems. Part II: bounds of the exact solution, Computer Methods in Applied Mechanics and Engineering 197 (2008) 1661–1679.

[19] E. Pruliere, J. Ferec, F. Chinesta, A. Ammar, An efficient reduced simulation of residual stresses in composites forming processes. International Journal of Material Forming (2010), http://dx.doi.org/10.1007/s12289-009-0675-6.