MOX–Report No. 25/2009

An anisotropic Zienkiewicz-Zhu a posteriori error estimator for 3D applications

P.E. Farrell, Stefano Micheletti, Simona Perotto

MOX, Dipartimento di Matematica "F. Brioschi"
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

mox@mate.polimi.it http://mox.polimi.it

An anisotropic Zienkiewicz-Zhu a posteriori error estimator for 3D applications*

P.E. Farrell†, S. Micheletti♯ and S. Perotto♯

August 13, 2009

† Applied Modelling and Computation Group,
Department of Earth Science and Engineering,
Imperial College London, SW7 2AZ, UK
patrick.farrell06@imperial.ac.uk

♯ MOX – Modellistica e Calcolo Scientifico
Dipartimento di Matematica "F. Brioschi"
Politecnico di Milano
via Bonardi 9, I-20133 Milano, Italy
{stefano.micheletti,simona.perotto}@polimi.it

Keywords: anisotropic mesh adaptation; Zienkiewicz-Zhu recovery procedure; a posteriori error estimator; finite elements.

AMS Subject Classification: 65N15, 65N30, 65N50, 65K10

Abstract

We extend the anisotropic Zienkiewicz-Zhu a posteriori error estimator of [1] to three dimensions. Like the standard Zienkiewicz-Zhu estimator, the proposed estimator is designed to be independent of the problem at hand, is cheap to compute and easy to implement. In contrast to the standard Zienkiewicz-Zhu estimator, the elementwise counterpart of the proposed estimator explicitly takes into account the geometrical properties of the actual tetrahedron. Thus, in a wide variety of applications, the estimator is able to detect the anisotropic features exhibited by the solution of the governing equations. A metric-based optimization procedure, rigorously addressed, drives the adaptation of the mesh. It is shown numerically to yield quasi-optimal triangulations, dictating the accuracy-vs-number of elements behaviour. Despite being heuristic to some extent, in practice the overall anisotropic adaptation procedure turns out to be effective.

*P.E. Farrell wishes to acknowledge support from the Imperial College High Performance Computing service. The multimaterial water collapse simulation was provided by C. R. Wilson.

1 Introduction and motivations

Anisotropic mesh adaptation has proved to be a powerful strategy for improving the quality and efficiency of flow simulations. Examples of anisotropic phenomena are present in several contexts in the literature: sharp fronts, e.g., shocks in compressible flows, in aerospace applications (see, e.g., [2, 3, 4, 5, 6]); multi-material flows with sharp immiscible interfaces, e.g., casting, mold filling and fluid jetting devices, in material processing applications (see, e.g., [7]); steep boundary layers in viscous flows around bodies (see, e.g., [8, 9, 10, 11, 5, 12]). Standard isotropic meshes fail to capture these directional features since they allow us to adjust only the size of the mesh elements. On the contrary, when anisotropic meshes are used we are able to control both the size, orientation and stretching of the mesh tetrahedra; thus, in principle, such meshes are able to identify the directionalities of the problem at hand in an accurate way, with a contained number of elements.

Anisotropic techniques based on heuristic approaches have been devised in the past (see, e.g., [8, 9, 2, 3, 10, 13, 11, 14, 15, 16, 17, 5, 18]. These techniques usually employ a metric-based approach where the metric is derived from a numerical approximation of the Hessian of the solution, coupled with an a priori error estimator. The main idea is to find the optimal mesh minimizing the interpolation error of a target function for a given number of mesh elements. Although the results are sometimes impressive, these techniques yet fail to link rigorously with a bound of the discretization error.

More rigorous approaches using theoretically based anisotropic adaptivity have been developed later on, both in a residual-based and in a goal-oriented framework ([19, 20, 21, 22, 23, 24]). This last class of adaptive procedures are suited to control more general quantities with respect to the interpolation error, e.g., the energy norm in the residual-based framework, or physically meaningful quantities (pointwise stresses, fluxes, vorticity, etc.) in the case of a goal-oriented analysis.

Due to an intrinsic complexity, anisotropic adaptation in 3D still represents the most challenging setting (see, e.g., [19, 20, 11, 15, 16, 17, 5, 25, 26, 27, 6]). With reference to the above different anisotropic approaches we stick to a heuristic estimator to drive a 3D mesh adaptation procedure. In particular, to limit the computational load of the adaptive algorithm, we propose a computationally straightforward error estimator, i.e., a Zienkiewicz-Zhu-like estimator. We consider piecewise linear finite elements, and devise a simple recovery technique, different from the standard one, yet enjoying the same philosophy: it is not confined to a specific problem; it is independent of the finite element formulation (except for the finite element space); it is cheap to compute and easy to implement; and, first and foremost, the method works very well in practice (see, e.g., [28, 29, 30]). Since the pioneering work [31] dealing with the linear elastic problem, and some further papers [32, 33], various efforts have been made to theoretically understand the amazingly good properties of the Zienkiewicz-Zhu

error estimator, obtained by approximating the true gradient with the recovered one. A first work in which some averaging technique is studied is [34], though the idea is nearly as old as the finite element method itself (see, e.g., [35]).

Concerning the Zienkiewicz-Zhu estimator that we propose in this paper, we follow the two-step procedure typical of the isotropic case: first a recovered gradient is constructed by performing a series of local operations over patches of tetrahedra; then such recovered gradient is employed to build the estimator itself. On the other hand, in contrast to the standard Zienkiewicz-Zhu estimator, the estimator that we propose is designed to incorporate anisotropic features of the triangulation, i.e., size, stretching and orientation of each element. The theoretical background for these anisotropic quantities, as well as for related anisotropic interpolation estimates is provided in [21, 22].

The layout of the paper is the following. After introducing the anisotropic background in Section 2, we propose and justify the anisotropic a posteriori error estimator. In Section 4, we show how to get a suitable metric tensor field out of the estimator via the introduction of local constrained minimazion problems. The solution to these problems is explicitly derived and the corresponding argument is rigorously proved. In Section 5, the actual adaptation procedure is arranged. Numerical results dealing with both academic test cases and challenging applications are collected in Section 6. Finally, some conclusions are drawn in the last section.

2 The anisotropic setting

To begin with, we extend the anisotropic theory developed in [21, 22] to the 3D case. This provides the theoretical background on which the estimator (16)-(17) is designed. Let Ω be a polyhedral domain in $\mathbb{R}^3$. We introduce a conforming partition $\mathcal{T}_h = \{K\}$ of Ω consisting of tetrahedra (e.g., [36]). The general element K is characterized geometrically by the properties of the affine map $T_K : \widehat{K} \to K$, where $\widehat{K}$ is the reference isotropic tetrahedron, centred at the origin and inscribed in the unit sphere (see Figure 1). In particular

$$\mathbf{x} = T_K(\widehat{\mathbf{x}}) = M_K\,\widehat{\mathbf{x}} + t_K,$$

with $M_K \in \mathbb{R}^{3\times 3}$ the Jacobian and $t_K \in \mathbb{R}^3$ the shift vector. The matrix M_K is factorized via the polar decomposition, yielding

$$M_K = B_K\,Z_K \tag{1}$$

where $B_K \in \mathbb{R}^{3\times 3}$ is symmetric positive definite, and $Z_K \in \mathbb{R}^{3\times 3}$ is orthogonal. Then B_K is spectrally decomposed as

$$B_K = R_K^T\,\Lambda_K\,R_K, \tag{2}$$

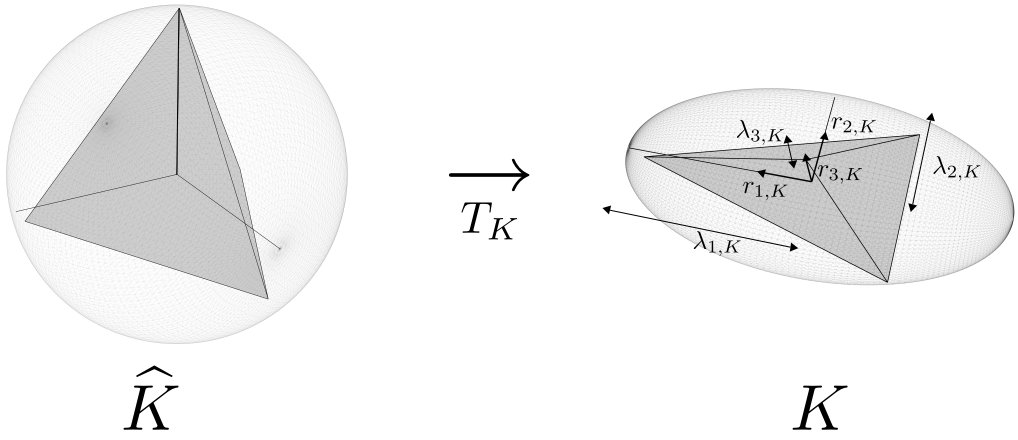

Figure 1: Sketch of the map T_K: the main anisotropic quantities are highlighted.

with $R_K^T = [\mathbf{r}_{1,K}, \mathbf{r}_{2,K}, \mathbf{r}_{3,K}]$ and $\Lambda_K = \mathrm{diag}(\lambda_{1,K}, \lambda_{2,K}, \lambda_{3,K})$ the matrices collecting the corresponding eigenvectors and eigenvalues, respectively. These quantities have a meaningful geometric interpretation. The unit sphere circumscribing $\widehat{K}$ is changed via T_K into an ellipsoid circumscribing K: the set of unit (column) vectors $\{\mathbf{r}_{i,K}\}$ defines the corresponding principal directions, whereas each $\lambda_{i,K}$ measures the length of the associated semi-axis (see Figure 1). Without loss of generality, we assume $\lambda_{1,K} \geq \lambda_{2,K} \geq \lambda_{3,K} > 0$, for any $K \in \mathcal{T}_h$.

To fix notation, let us pick the regular tetrahedron $\widehat{K}$ as the one characterized by the coordinates $V_1 = \left[-\sqrt{\frac{2}{3}}, -\frac{\sqrt{2}}{3}, -\frac{1}{3}\right]$, $V_2 = \left[\sqrt{\frac{2}{3}}, -\frac{\sqrt{2}}{3}, -\frac{1}{3}\right]$, $V_3 = \left[0, 2\,\frac{\sqrt{2}}{3}, -\frac{1}{3}\right]$, $V_4 = \left[0, 0, 1\right]$. It can be checked that $\widehat{K}$ is inscribed in the unit sphere centred at the origin, that all of its edges have length $|\widehat{e}| = 2\sqrt{2/3}$ and that the volume $|\widehat{K}|$ is equal to $8\sqrt{3}/9$.

Then given any tetrahedron K whose vertices have coordinates $\{x_1^i, x_2^i, x_3^i\}$, with $i = 1, 2, 3, 4$, the affine transformation T_K is characterized by

$$M_K = \frac{1}{4}\begin{bmatrix} \sqrt{6}\,(x_1^2 - x_1^1) & \sqrt{2}\,(2\,x_1^3 - x_1^1 - x_1^2) & 3\,x_1^4 - x_1^1 - x_1^2 - x_1^3 \\ \sqrt{6}\,(x_2^2 - x_2^1) & \sqrt{2}\,(2\,x_2^3 - x_2^1 - x_2^2) & 3\,x_2^4 - x_2^1 - x_2^2 - x_2^3 \\ \sqrt{6}\,(x_3^2 - x_3^1) & \sqrt{2}\,(2\,x_3^3 - x_3^1 - x_3^2) & 3\,x_3^4 - x_3^1 - x_3^2 - x_3^3 \end{bmatrix},$$

$$t_K = \frac{1}{4}\begin{bmatrix} x_1^1 + x_1^2 + x_1^3 + x_1^4 \\ x_2^1 + x_2^2 + x_2^3 + x_2^4 \\ x_3^1 + x_3^2 + x_3^3 + x_3^4 \end{bmatrix}.$$

The above choice for $\widehat{K}$ turns out to be particularly handy with a view to the metric construction in Section 5.

We now prove an anisotropic interpolation estimate for functions $v \in H^1(\Omega)$, standard notation being adopted to denote the Sobolev spaces of functions with Lebesgue measurable derivatives and their norms (e.g., [37]). In such a case, a more general interpolant, other than the Lagrangian one, has to be considered, i.e., the Clément-like interpolant of degree 1, I_h^1 [38, 39]. Such an estimate will provide the inspiration for the anisotropic estimator proposed in Section 3.

Proposition 2.1 *Let $v \in H^1(\Omega)$. Then, under the assumptions that* $\mathrm{card}(\Delta_K) \leq \mathcal{D}$ *and* $\mathrm{diam}(\widehat{\Delta}_K) \leq \delta$, *for any $K \in \mathcal{T}_h$, there exists a constant $C = C(\mathcal{D}, \delta)$, such that*

$$\|v - I_h^1(v)\|_{L^2(K)} \leq C \Bigg(\sum_{i=1}^{3} \lambda_{i,K}^2 \, (\mathbf{r}_{i,K}^T G_{\Delta_K}(\nabla v)\, \mathbf{r}_{i,K}) \Bigg)^{1/2}, \tag{3}$$

$G_K(\cdot)$ being the symmetric semidefinite positive matrix whose general entry is given by

$$[G_{\Delta_K}(\mathbf{v})]_{i,j} = \sum_{T \in \Delta_K} \int_T v_i \, v_j \, d\mathbf{x}, \quad \textit{with } i,j = 1,2,3, \tag{4}$$

for any vector-valued function $\mathbf{v} = (v_1, v_2, v_3)^T \in [L^2(\Omega)]^3$, and where $\Delta_K, \widehat{\Delta}_K = T_K^{-1}(\Delta_K)$ denote a suitable patch of elements associated with the tetrahedron K and its pullback, respectively.

Proof. We remark that hereafter the constant C may change its value in different expressions. First, let us recall the relation

$$\widehat{\nabla} \widehat{v} = M_K^T \nabla v \tag{5}$$

between the gradient ∇v of a function $v \in H^1(K)$ and of its pullback $\widehat{v} = v \circ T_K$, referred to the reference element $\widehat{K}$. The polar decomposition (1) combined with the spectral factorization (2) yields

$$\widehat{\nabla} \widehat{v} = M_K^T \nabla v = Z_K^T R_K^T \Lambda_K R_K \nabla v.$$

By considering the L^2-norm of $\widehat{\nabla} \widehat{v}$ on $\widehat{K}$ and thanks to the orthogonality of the matrices Z_K and R_K, we get

$$\|\widehat{\nabla} \widehat{v}\|_{L^2(\widehat{K})}^2 = \int_{\widehat{K}} |Z_K^T R_K^T \Lambda_K R_K \nabla v|^2 \, d\widehat{\mathbf{x}} = \int_{\widehat{K}} |\Lambda_K R_K \nabla v|^2 \, d\widehat{\mathbf{x}} = \int_{\widehat{K}} |\Lambda_K \boldsymbol{\phi}_K|^2 \, d\widehat{\mathbf{x}}, \tag{6}$$

$\boldsymbol{\phi}_K$ defining the $\mathbb{R}^3$-vector valued function

$$\boldsymbol{\phi}_K = R_K \nabla v = \begin{bmatrix} \phi_{1,K} \\ \phi_{2,K} \\ \phi_{3,K} \end{bmatrix} = \begin{bmatrix} \nabla v \cdot \mathbf{r}_{1,K} \\ \nabla v \cdot \mathbf{r}_{2,K} \\ \nabla v \cdot \mathbf{r}_{3,K} \end{bmatrix}. \tag{7}$$

By substituting (7) into (6), we deduce

$$\|\widehat{\nabla}\widehat{v}\|^2_{L^2(\widehat{K})} = \int_{\widehat{K}} \left(\lambda^2_{1,K}\,\phi^2_{1,K} + \lambda^2_{2,K}\,\phi^2_{2,K} + \lambda^2_{3,K}\,\phi^2_{3,K}\right)\,d\widehat{\mathbf{x}}$$

$$= \frac{\lambda_{1,K}}{\lambda_{2,K}\,\lambda_{3,K}} \int_K \phi^2_{1,K}\,d\mathbf{x} + \frac{\lambda_{2,K}}{\lambda_{1,K}\,\lambda_{3,K}} \int_K \phi^2_{2,K}\,d\mathbf{x} + \frac{\lambda_{3,K}}{\lambda_{1,K}\,\lambda_{2,K}} \int_K \phi^2_{3,K}\,d\mathbf{x}$$

$$= \frac{\lambda_{1,K}}{\lambda_{2,K}\,\lambda_{3,K}} \|\nabla v\cdot\mathbf{r}_{1,K}\|^2_{L^2(K)} + \frac{\lambda_{2,K}}{\lambda_{1,K}\,\lambda_{3,K}} \|\nabla v\cdot\mathbf{r}_{2,K}\|^2_{L^2(K)} + \frac{\lambda_{3,K}}{\lambda_{1,K}\,\lambda_{2,K}} \|\nabla v\cdot\mathbf{r}_{3,K}\|^2_{L^2(K)}$$

$$= \left(\prod_{j=1}^{3}\lambda_{j,K}\right)^{-1} \sum_{i=1}^{3} \lambda^2_{i,K} \|\nabla v\cdot\mathbf{r}_{i,K}\|^2_{L^2(K)}, \tag{8}$$

where the relation $|K| = \lambda_{1,K}\,\lambda_{2,K}\,\lambda_{3,K}|\widehat{K}|$ is also exploited.

Let us focus now on the interpolation error on K. We first recall that both the standard Clément and Scott-Zhang interpolants satisfy the properties

$$\|w - I^1_h(w)\|_{L^2(\mathcal{K})} \le C\,h_{\mathcal{K}}\,|w|_{H^1(\Delta_{\mathcal{K}})}, \tag{9}$$

$$[I^1_h(w)]\widehat{\ } = \widehat{I}^1_h(\widehat{w}) \tag{10}$$

for any $w \in H^1(\Omega)$, $\widehat{\ }$ denoting the corresponding pullback quantities as above, and where the constant C depends on the regularity of the patch $\Delta_{\mathcal{K}}$ ([40, 21]). Notice that $\mathcal{K}$ is a certain tetrahedron, $h_{\mathcal{K}} = \mathrm{diam}(\mathcal{K})$, while $\Delta_{\mathcal{K}}$ represents a suitable patch of elements surrounding $\mathcal{K}$.

Using relation (10), we get

$$\|v - I^1_h(v)\|^2_{L^2(K)} = \lambda_{1,K}\,\lambda_{2,K}\,\lambda_{3,K}\,\|[v - I^1_h(v)]\widehat{\ }\|^2_{L^2(\widehat{K})} = \lambda_{1,K}\,\lambda_{2,K}\,\lambda_{3,K}\,\|\widehat{v} - \widehat{I}^1_h(\widehat{v})\|^2_{L^2(\widehat{K})}\,. \tag{11}$$

Estimate (9) applied to the term $\|\widehat{v} - \widehat{I}^1_h(\widehat{v})\|^2_{L^2(\widehat{K})}$, on letting $\mathcal{K} = \widehat{K}$, $\Delta_{\mathcal{K}} = \widehat{\Delta}_K$ and $w = \widehat{v}$, yields

$$\|v - I^1_h(v)\|^2_{L^2(K)} \le C\,\lambda_{1,K}\,\lambda_{2,K}\,\lambda_{3,K}\,h^2_{\widehat{K}}\,|\widehat{v}|^2_{H^1(\widehat{\Delta}_K)} = C\,\lambda_{1,K}\,\lambda_{2,K}\,\lambda_{3,K} \sum_{\widehat{T}\in\widehat{\Delta}_K} |\widehat{v}|^2_{H^1(\widehat{T})}. \tag{12}$$

In the last equality, the quantity $h^2_{\widehat{K}}$ is included into the constant C, being a known $\mathcal{O}(1)$ value. By applying relation (8) to each seminorm $|\widehat{v}|^2_{H^1(\widehat{T})}$ (after identifying $\widehat{T}$ with $\widehat{K}$), we have

$$\|v - I^1_h(v)\|_{L^2(K)} \le C \left\{ \sum_{T\in\Delta_K} \left[\sum_{i=1}^{3} \lambda^2_{1,K}\,\|\nabla v\cdot\mathbf{r}_{i,K}\|^2_{L^2(T)} \right] \right\}^{1/2}. \tag{13}$$

Result (3) follows by straightforward algebraic manipulations of the right-hand side in (13). □

Remark 2.1 *The hypotheses of Proposition 2.1 can be considered smoothness requirements: they do not limit the anisotropy of each element K, rather they constrain the variation of $\{\mathbf{r}_{i,K}\}$ and $\{\lambda_{i,K}\}$ over the patch Δ_K.*

3 An anisotropic Zienkiewicz-Zhu error estimator

We now show how to employ Proposition 2.1 to devise an a posteriori error estimator. For simplicity, we refer to the standard Poisson problem completed with homogeneous Dirichlet boundary conditions. More general problems can be tackled as well, as shown in Section 6.
In harmony with a Zienkiewicz-Zhu approach, we distinguish between two steps: first we furnish a procedure for obtaining an approximate recovered gradient, more accurate than the actual gradient of the Galerkin solution; second, we employ this recovered gradient for a posteriori error purposes.

Let u be the (weak) solution to the model problem: find $u \in V$ such that

$$\int_\Omega \nabla u \cdot \nabla v \, d\mathbf{x} = \int_\Omega f v \, d\mathbf{x} \quad \forall v \in V, \tag{14}$$

where $V = H_0^1(\Omega)$. Let $u_h \in V_h^0$ be its Galerkin approximation satisfying

$$\int_\Omega \nabla u_h \cdot \nabla v_h \, d\mathbf{x} = \int_\Omega f v_h \, d\mathbf{x} \quad \forall v_h \in V_h^0, \tag{15}$$

with $V_h^0 = V_h \cap V$ and V_h the space of the piecewise linear finite elements.
The objective of the recovery procedure is to build an approximation to ∇u, only using information related to u_h. Several approaches can be undertaken for this purpose (see, e.g., [41, 32, 42, 40, 43, 44, 45, 46]). For example, in the case of affine finite elements, the Zienkiewicz-Zhu recovery procedure in [32] consists of the following steps:

- suitable L^2 -projections $P_N(\nabla u_h)$ of ∇u_h onto the space $[\mathbb{P}_1]^2$ are computed, namely

  $$\int_{\Delta_N} (\nabla u_h - P_N(\nabla u_h)) \cdot \mathbf{v} \, d\mathbf{x} = 0 \quad \forall \mathbf{v} \in [\mathbb{P}_1]^2,$$

  with $\Delta_N = \{T \ni \mathbf{x}_N\}$ the patch associated with the general node $\mathbf{x}_N$ of the partition $\mathcal{T}_h$, and $\mathbb{P}_1 = \text{span}\{1, x_1, x_2\}$;

- the recovered gradient $G^{ZZ} u_h \in [V_h]^2$ is then formed as

  $$G^{ZZ} u_h(\mathbf{x}) = \sum_N P_N(\nabla u_h)(\mathbf{x}_N) \, \varphi_N(\mathbf{x}) \quad \forall \mathbf{x} \in \Omega,$$

  φ_N being the standard hat function associated with the node N.

We propose here a different and simpler approach, whereby the recovered gradient, denoted by $P_{\Delta_K}(\nabla u_h)(\cdot)$ is assumed constant over the patch $\Delta_K = \{T \in \mathcal{T}_h : T \cap K \neq \emptyset\}$ comprising all of the tetrahedra that share a node, an edge, or a face with K. We let

$$P_{\Delta_K}(\nabla u_h)(\mathbf{x}) = \frac{1}{|\Delta_K|} \sum_{T \in \Delta_K} |T| \, \nabla u_h|_T, \quad \text{with } \mathbf{x} \in \Delta_K,$$

namely, we compute the volume-weighted average over the patch Δ_K of the gradients of the discrete solution. Notice that the recovered gradient $P_{\Delta_K}(\nabla u_h)(\cdot)$ is strictly attached to K, and not to the elements comprising Δ_K, i.e., for an element $K' \neq K$, even when $K' \in \Delta_K$, $P_{\Delta_{K'}}(\nabla u_h)(\cdot)$ is in general different from $P_{\Delta_K}(\nabla u_h)(\cdot)$, and it is constant over $\Delta_{K'}$.

Hereafter we denote by $\mathbf{E}_{\Delta_K} = P_{\Delta_K}(\nabla u_h) - \nabla u_h|_{\Delta_K}$, dropping the dependence on $\mathbf{x}$, the recovered error on the gradient over Δ_K. Then we define the anisotropic Zienkiewicz-Zhu local estimator for the H^1-seminorm of the discretization error as

$$\eta^2_{K,\,\mathrm{aniso}} = \frac{1}{(\lambda_{1,K}\lambda_{2,K}\lambda_{3,K})^{2/3}} \sum_{i=1}^{3} \lambda^2_{i,K}\,(\mathbf{r}^T_{i,K} G_{\Delta_K}(\mathbf{E}_{\Delta_K})\mathbf{r}_{i,K}), \tag{16}$$

where the matrix $G_{\Delta_K}(\cdot)$ is defined as in (4). Then the corresponding global error estimator is given by

$$\eta^2_{\mathrm{aniso}} = \sum_{K\in\mathcal{T}_h} \eta^2_{K,\,\mathrm{aniso}}. \tag{17}$$

The estimator (16)-(17) is essentially heuristic. Nevertheless the rationale is twofold: on the one hand the terms summed on the right-hand side of (16) are suggested by Proposition 2.1, after substituting the partial derivatives with the corresponding components of the recovered gradient (notice that, in general, the recovered gradient is not computed as the gradient of a scalar function); on the other hand, the volumetric scaling factor $(\lambda_{1,K}\lambda_{2,K}\lambda_{3,K})^{2/3}$ guarantees that when $\lambda_{1,K} = \lambda_{2,K} = \lambda_{3,K}$, that is when K is an isotropic tetrahedron, the estimator becomes

$$\eta^2_{K,\,\mathrm{iso}} = \int_{\Delta_K} |\mathbf{E}_{\Delta_K}|^2\, d\mathbf{x},$$

i.e., an isotropic Zienkiewicz-Zhu like estimator for the H^1-seminorm of the discretization error, based on a patchwise constant recovered gradient [1]. Further theoretical background is discussed in [1] in the 2D case.
Estimator (16)-(17) can be applied to more general problems, such as the elasticity or Navier-Stokes equations. One possibility is to replace the gradient by the stress (rate) tensor (see, e.g., [31]). Alternately, a suitable scalar variable, representative of the problem, can be identified as the "dummy" variable u, and (16)-(17) can be extended in a straightforward way, as shown in Section 6. More sophisticated recovery procedures can easily find room in variants of the present estimator.

4 The estimator-to-metric procedure

We describe the procedure used for adapting the mesh moving from estimator (16)-(17).

We employ a *metric-based* adaptive procedure as a predictive tool. Two reciprocal approaches can be pursued: either given a constraint on the maximum number of elements, find the mesh providing the most accurate numerical solution; or given a constraint on the accuracy of the numerical solution, find the mesh with the least number of elements. We here focus on the latter approach.

Let us emphasize that metric and mesh are intertwined. The objective is to build a mesh that is optimal with respect to a matching condition involving a suitable metric, induced by a symmetric positive-definite tensor field $\widetilde{M} : \Omega \to \mathbb{R}^{3\times 3}$ (see, e.g., [47]). Typically, this mesh is obtained via an iterative procedure that processes intermediate tentative meshes.
Thus, with any given tentative mesh $\mathcal{T}_h$ we associate a piecewise constant metric $\widetilde{M}_{\mathcal{T}_h}$, such that $\widetilde{M}_{\mathcal{T}_h}|_K = \widetilde{M}_K = B_K^{-2} = R_K^T \Lambda_K^{-2} R_K$, for any $K \in \mathcal{T}_h$, the matrices R_K and Λ_K being defined as in Section 2. With respect to this metric, any tetrahedron K is unit equilateral, i.e., $\left(\mathbf{e}^T \widetilde{M}_K\, \mathbf{e}\right)^{1/2} = 1$, with $\mathbf{e}$ the (arbitrarily oriented) vector identifying any edge e of K.

Vice versa let now $\widetilde{M}$ be a given metric. We first diagonalize the tensor field $\widetilde{M}$ as $\widetilde{M} = \widetilde{R}^T \widetilde{\Lambda}^{-2} \widetilde{R}$, with $\widetilde{\Lambda} = \mathrm{diag}(\widetilde{\lambda}_1, \widetilde{\lambda}_2, \widetilde{\lambda}_3)$ and $\widetilde{R}^T = [\widetilde{\mathbf{r}}_1, \widetilde{\mathbf{r}}_2, \widetilde{\mathbf{r}}_3]$ a positive diagonal and an orthogonal matrix, respectively. Notice that this can be done, ideally, for every $\mathbf{x} \in \Omega$. We then approximate the quantities $\left\{\widetilde{\lambda}_i\right\}$, $\{\widetilde{\mathbf{r}}_i\}$ via piecewise constants over a tentative mesh $\mathcal{T}_h$, and denote these quantities by $\overline{\mathbf{r}}_{i,K} \in \mathbb{R}^3$, $\overline{\lambda}_{i,K} \in \mathbb{R}$, for any $K \in \mathcal{T}_h$ and with $i = 1, 2, 3$. For example, this can be carried out by averaging the pointwise functions $\widetilde{\mathbf{r}}_i, \widetilde{\lambda}_i$ over K. The averaged quantities define a piecewise constant metric, say $\overline{M}_{\mathcal{T}_h}$.

Thus we state

Definition 4.1 *The mesh $\mathcal{T}_h$ matches $\widetilde{M}$ if, for any $K \in \mathcal{T}_h$, $\widetilde{M}_{\mathcal{T}_h}|_K = \overline{M}_{\mathcal{T}_h}|_K$.*

In the spirit of a predictive procedure the tensor field $\widetilde{M}$ represents the actual unknown. At each iteration of the adaptive process, say j, we deal with three quantities:

i) the actual mesh $\mathcal{T}_h^{(j)}$;

ii) the new metric $\widetilde{M}^{(j+1)}$ computed on $\mathcal{T}_h^{(j)}$ and piecewise constant;

iii) the updated mesh $\mathcal{T}_h^{(j+1)}$ matching $\widetilde{M}^{(j+1)}$.

In more detail, at each step j, first problem (15) is solved on $\mathcal{T}_h^{(j)}$ and $\widetilde{M}^{(j+1)}$ is built elementwise through a suitable local optimization procedure (one for each $K \in \mathcal{T}_h^{(j)}$). Then the new mesh $\mathcal{T}_h^{(j+1)}$ is built via the matching condition. This last task is accomplished by the mesh optimisation procedure described in [11]. A quality function is defined for each element, measuring its conformity to the ideal element described by the tensor field (both size and shape). The quality of the mesh is defined to be the quality of the worst element within the mesh. Iterations of optimisation procedures such as edge collapsing, edge splitting, edge

and face swapping and node movement are applied and the operation accepted if the quality of the mesh improves. This procedure is applied until the mesh quality function satisfies a user-specified threshold. For background reading on mesh optimisation procedures, see [47].

For practical reasons, the metric $\widetilde{M}^{(j+1)}$ is averaged nodewise before being passed to the mesh optimisation procedure, as it takes as input a piecewise-linear representation of the tensor field. This nodewise averaging can change the desired number of elements of the mesh encoded by the tensor field. Empirically, it is found that rescaling the averaged piecewise-linear representation to match the expected number of elements of the piecewise-constant representation is important for the convergence of the adaptive procedure.

The local optimization procedure involved in point ii) consists first in minimizing the estimator $\eta^2_{K,\,\mathrm{aniso}}$ in (16) with respect to stretching and orientation, and then, by an equidistribution criterion, in computing the actual value of $\lambda_{1,K}, \lambda_{2,K}$ and $\lambda_{3,K}$. For the purpose of minimization, we introduce the stretching factors

$$s_{1,K} = \left(\frac{\lambda^2_{1,K}}{\lambda_{2,K}\lambda_{3,K}}\right)^{2/3},\ s_{2,K} = \left(\frac{\lambda^2_{2,K}}{\lambda_{1,K}\lambda_{3,K}}\right)^{2/3},\ \text{and}\ s_{3,K} = \left(\frac{\lambda^2_{3,K}}{\lambda_{1,K}\lambda_{2,K}}\right)^{2/3}. \tag{18}$$

The estimator can thus be rewritten as

$$\begin{aligned}\eta^2_{K,\,\mathrm{aniso}} &= \sum_{i=1}^{3} s_{i,K}\,(\mathbf{r}^T_{i,K} G_{\Delta_K}(\mathbf{E}_{\Delta_K})\mathbf{r}_{i,K}) \\ &= \Big(\prod_{i=1}^{3}\lambda_{i,K}\Big)|\widehat{\Delta}_K|\sum_{i=1}^{3} s_{i,K}\,(\mathbf{r}^T_{i,K}\widehat{G}_{\Delta_K}(\mathbf{E}_{\Delta_K})\mathbf{r}_{i,K}),\end{aligned} \tag{19}$$

where $\widehat{G}_{\Delta_K}(\cdot)$ is the scaled matrix $G_{\Delta_K}(\cdot)/|\Delta_K|$, and $\widehat{\Delta}_K$ is the patch defined as in Proposition 2.1 Thus it holds $|\Delta_K| = \lambda_{1,K}\lambda_{2,K}\lambda_{3,K}\,|\widehat{\Delta}_K|$.
Notice also that $s_{1,K} \geq s_{2,K} \geq s_{3,K}$ and that these stretching factors are not independent as

$$\prod_{i=1}^{3} s_{i,K} = 1. \tag{20}$$

The idea behind expression (19) is that we have singled out the volume information (the term before the summation) from quantities that just depend on orientation and stretching. We now state a result about the minimization of the terms involved in the summation.

Proposition 4.1 *Let*

$$J(\{s_{i,K}, \mathbf{r}_{i,K}\}_{i=1,2,3}) = \sum_{i=1}^{3} s_{i,K}\,(\mathbf{r}^T_{i,K}\widehat{G}_{\Delta_K}(\mathbf{E}_{\Delta_K})\mathbf{r}_{i,K}), \tag{21}$$

and let $\{\mathbf{g}_i, g_i\}$, $i = 1, 2, 3$, *represent the eigen-pairs associated with* $\widehat{G}_{\Delta_K}(\mathbf{E}_{\Delta_K})$, *where it is understood that* $g_1 \geq g_2 \geq g_3 > 0$ *as well as that* $\{\mathbf{g}_i\}$ *defines an orthonormal frame. Under the constraint* (20) *and assuming that* $\{\mathbf{r}_{i,K}\}$ *defines a right-handed orthonormal basis, then* $J(\cdot)$ *is minimized when*

$$s_{1,K} = \frac{\sqrt[3]{g_1\, g_2\, g_3}}{g_3}, \quad s_{2,K} = \frac{\sqrt[3]{g_1\, g_2\, g_3}}{g_2}, \quad s_{3,K} = \frac{\sqrt[3]{g_1\, g_2\, g_3}}{g_1} \tag{22}$$

and

$$\mathbf{r}_{1,K} = \mathbf{g}_3, \quad \mathbf{r}_{2,K} = \mathbf{g}_2, \quad \mathbf{r}_{3,K} = \mathbf{g}_1. \tag{23}$$

Proof. The proof consists of three parts: in *1)* we minimize $J(\cdot)$ with respect to $\{\mathbf{r}_{i,K}\}$, thus proving that the optimal unit vectors coincide with $\{\mathbf{g}_i\}$, up to a permutation; in *2)* we exploit the ordering $s_{1,K} \geq s_{2,K} \geq s_{3,K}$ to unequivocally determine the optimal $\{\mathbf{r}_{i,K}\}$; in *3)* we compute the optimal values of the stretching factors by minimizing $J(\cdot)$ with respect to $\{s_{i,K}\}$. As intermediate steps, we study in *1a)* the case when all elements of $\{s_{i,K}\}$ are distinct; in *1b)* when any and only two are equal; in *1c)* when all elements of $\{s_{i,K}\}$ are equal.

1) Let us first minimize $J(\cdot)$ with respect to $\{\mathbf{r}_{i,K}\}$. In order to enforce orthonormality, we require that the Gateaux derivative of $J(\cdot)$ be zero for rotations about the arbitrary unit vector $\boldsymbol{\omega} \in \mathbb{R}^3$, i.e.,

$$\delta J = \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \Big(J(\{s_{i,K}, \mathbf{r}_{i,K} + \varepsilon\, \boldsymbol{\omega} \wedge \mathbf{r}_{i,K}\}) - J(\{s_{i,K}, \mathbf{r}_{i,K}\}) \Big) = 0. \tag{24}$$

Notice that this way allows us to avoid using the Lagrange multiplier approach to enforce the constraint. We obtain

$$\begin{aligned} J(\{s_{i,K}, \mathbf{r}_{i,K} + \varepsilon\, \boldsymbol{\omega} \wedge \mathbf{r}_{i,K}\}) &= J(\{s_{i,K}, \mathbf{r}_{i,K}\}) \\ &+ 2\,\varepsilon \sum_{i=1}^{3} s_{i,K} \left(\mathbf{r}_{i,K}^T \widehat{G}_{\Delta_K}(\mathbf{E}_{\Delta_K})(\boldsymbol{\omega} \wedge \mathbf{r}_{i,K}) \right) \\ &+ \varepsilon^2 \sum_{i=1}^{3} s_{i,K} \left((\boldsymbol{\omega} \wedge \mathbf{r}_{i,K})^T \widehat{G}_{\Delta_K}(\mathbf{E}_{\Delta_K})(\boldsymbol{\omega} \wedge \mathbf{r}_{i,K}) \right). \end{aligned} \tag{25}$$

This shows that, provided that the first order condition (24) is satisfied, that is

$$\sum_{i=1}^{3} s_{i,K} \left(\mathbf{r}_{i,K}^T \widehat{G}_{\Delta_K}(\mathbf{E}_{\Delta_K})(\boldsymbol{\omega} \wedge \mathbf{r}_{i,K}) \right) = 0 \quad \forall \text{ unit vector } \boldsymbol{\omega} \in \mathbb{R}^3, \tag{26}$$

the stationary point of $J(\cdot)$ is an actual minimum, since the $\mathcal{O}(\varepsilon^2)$ quantity in (25) is nonnegative, because $s_{i,K} > 0$ and due to the positive semidefiniteness of $\widehat{G}_{\Delta_K}(\mathbf{E}_{\Delta_K})$. From (26), first choosing $\boldsymbol{\omega} = \mathbf{r}_{1,K}$, we have $\boldsymbol{\omega} \wedge \mathbf{r}_{1,K} = \mathbf{0}$, $\boldsymbol{\omega} \wedge \mathbf{r}_{2,K} = \mathbf{r}_{3,K}$ and $\boldsymbol{\omega} \wedge \mathbf{r}_{3,K} = -\mathbf{r}_{2,K}$. Thus $\delta J = (s_{2,K} - s_{3,K})\, \mathbf{r}_{2,K}^T \widehat{G}_{\Delta_K}(\mathbf{E}_{\Delta_K})\, \mathbf{r}_{3,K}$. Taking successively $\boldsymbol{\omega} = \mathbf{r}_{2,K}$ and $\boldsymbol{\omega} = \mathbf{r}_{3,K}$, the first order conditions (26) for the minimum become

$$\begin{cases} (s_{2,K} - s_{3,K})\, \mathbf{r}_{2,K}^T \widehat{G}_{\Delta_K}(\mathbf{E}_{\Delta_K})\, \mathbf{r}_{3,K} = 0 \\ (s_{3,K} - s_{1,K})\, \mathbf{r}_{3,K}^T \widehat{G}_{\Delta_K}(\mathbf{E}_{\Delta_K})\, \mathbf{r}_{1,K} = 0 \\ (s_{1,K} - s_{2,K})\, \mathbf{r}_{1,K}^T \widehat{G}_{\Delta_K}(\mathbf{E}_{\Delta_K})\, \mathbf{r}_{2,K} = 0. \end{cases} \tag{27}$$

1a) To deal with (27), suppose first that all the elements of $\{s_{i,K}\}$ are distinct. Then it must hold $\mathbf{r}_{2,K}^T \widehat{G}_{\Delta_K}(\mathbf{E}_{\Delta_K})\,\mathbf{r}_{3,K} = \mathbf{r}_{3,K}^T \widehat{G}_{\Delta_K}(\mathbf{E}_{\Delta_K})\,\mathbf{r}_{1,K} = \mathbf{r}_{1,K}^T \widehat{G}_{\Delta_K}(\mathbf{E}_{\Delta_K})\,\mathbf{r}_{2,K} = 0$. From these relations it follows that $\widehat{G}_{\Delta_K}(\mathbf{E}_{\Delta_K})\,\mathbf{r}_{1,K} \perp \mathrm{span}\{\mathbf{r}_{2,K}, \mathbf{r}_{3,K}\}$. Similar relations can be obtained by permuting cyclically the indices $\{1,2,3\}$. Thus we must have that $G_{\Delta_K}(\mathbf{E}_{\Delta_K})\,\mathbf{r}_{1,K}$ is parallel to $\mathbf{r}_{1,K}$, i.e., $\mathbf{r}_{1,K}$ is an eigenvector of $G_{\Delta_K}(\mathbf{E}_{\Delta_K})$. This shows that, when all the elements of $\{s_{i,K}\}$ are distinct, $\{\mathbf{r}_{i,K}\}$ must coincide with $\{\mathbf{g}_i\}$, up to a permutation.

1b) Suppose now that any and only two of $\{s_{i,K}\}$ are equal. Since the argument applies whatever pair is picked, to fix ideas, let $s_{2,K} = s_{3,K}$ and $s_{1,K} \neq s_{2,K}$. Moreover, let $\{l,m,n\}$ denote a permutation of $\{1,2,3\}$. From $(27)_{2,3}$ we can only infer that $\widehat{G}_{\Delta_K}(\mathbf{E}_{\Delta_K})\,\mathbf{r}_{1,K} \perp \mathrm{span}\{\mathbf{r}_{2,K}, \mathbf{r}_{3,K}\}$, that is, $\mathbf{r}_{1,K}$ is an eigenvector of $\widehat{G}_{\Delta_K}(\mathbf{E}_{\Delta_K})$, say $\mathbf{g}_l$. Relation $(27)_1$ is trivially satisfied. As for $\mathbf{r}_{2,K}, \mathbf{r}_{3,K}$, we can only state that $\mathrm{span}\{\mathbf{r}_{2,K}, \mathbf{r}_{3,K}\} = \mathrm{span}\{\mathbf{g}_m, \mathbf{g}_n\}$. Enforcing $\mathbf{r}_{1,K} = \mathbf{g}_l$ and $s_{2,K} = s_{3,K}$, the functional $J(\cdot)$ in (21) becomes

$$J(\{s_{i,K}, \mathbf{r}_{i,K}\}_{i=1,2,3}) = s_{1,K}\, g_l + s_{2,K} \left(\mathbf{r}_{2,K}^T \widehat{G}_{\Delta_K}(\mathbf{E}_{\Delta_K})\,\mathbf{r}_{2,K} + \mathbf{r}_{3,K}^T \widehat{G}_{\Delta_K}(\mathbf{E}_{\Delta_K})\,\mathbf{r}_{3,K}\right). \tag{28}$$

We show now that the term multiplying $s_{2,K}$ takes a value independent of $\mathbf{r}_{2,K}, \mathbf{r}_{3,K}$. This allows us to pick, $\mathbf{r}_{2,K} = \mathbf{g}_m$ and $\mathbf{r}_{3,K} = \mathbf{g}_n$, say. To prove this, we recall that the symmetric matrix $\widehat{G}_{\Delta_K}(\mathbf{E}_{\Delta_K})$ can be decomposed as

$$\widehat{G}_{\Delta_K}(\mathbf{E}_{\Delta_K}) = \sum_{i=1}^{3} g_i\, \mathbf{g}_i\, \mathbf{g}_i^T.$$

Substituting this into (28), we get

$$\begin{aligned}
&J(\{s_{i,K}, \mathbf{r}_{i,K}\}_{i=1,2,3}) = s_{1,K}\, g_l + s_{2,K} \sum_{i=1}^{3} \left\{ g_i\, (\mathbf{r}_{2,K}^T \mathbf{g}_i)(\mathbf{g}_i^T \mathbf{r}_{2,K}) + g_i\, (\mathbf{r}_{3,K}^T \mathbf{g}_i)(\mathbf{g}_i^T \mathbf{r}_{3,K}) \right\} \\
&= s_{1,K}\, g_l + s_{2,K} \sum_{i=1}^{3} \left\{ g_i \left[(\mathbf{g}_i^T \mathbf{r}_{2,K})^2 + (\mathbf{g}_i^T \mathbf{r}_{3,K})^2 \right] \right\} \\
&= s_{1,K}\, g_l + s_{2,K} \left\{ g_m \left[(\mathbf{g}_m^T \mathbf{r}_{2,K})^2 + (\mathbf{g}_m^T \mathbf{r}_{3,K})^2 \right] + g_n \left[(\mathbf{g}_n^T \mathbf{r}_{2,K})^2 + (\mathbf{g}_n^T \mathbf{r}_{3,K})^2 \right] \right\},
\end{aligned}$$

where the contribution due to $i = l$ is zero since $\mathbf{g}_l^T \mathbf{r}_{2,K} = \mathbf{g}_l^T \mathbf{r}_{3,K} = 0$. Introducing the angle $\theta \in [0, \pi)$ such that $\mathbf{g}_m^T \mathbf{r}_{2,K} = \cos\theta$, and using the orthonormality and coplanarity properties of $\{\mathbf{r}_{2,K}, \mathbf{r}_{3,K}\}$ and $\{\mathbf{g}_m, \mathbf{g}_n\}$, it holds $(\mathbf{g}_m^T \mathbf{r}_{2,K})^2 = (\mathbf{g}_n^T \mathbf{r}_{3,K})^2 = \cos^2\theta$ and $(\mathbf{g}_m^T \mathbf{r}_{3,K})^2 = (\mathbf{g}_n^T \mathbf{r}_{2,K})^2 = \sin^2\theta$. Thus we obtain that $J(\{s_{i,K}, \mathbf{r}_{i,K}\}_{i=1,2,3}) = s_{1,K}\, g_l + s_{2,K}\, (g_m + g_n)$, that is, the independence of $J(\{s_{i,K}, \mathbf{r}_{i,K}\}_{i=1,2,3})$ of $\mathbf{r}_{2,K}, \mathbf{r}_{3,K}$.

1c) The case when $s_{1,K} = s_{2,K} = s_{3,K}$ can be dealt with using algebraic arguments to show, again, that the functional $J(\{s_{i,K}, \mathbf{r}_{i,K}\}_{i=1,2,3})$ does not depend on the choices of $\mathbf{r}_{1,K}, \mathbf{r}_{2,K}, \mathbf{r}_{3,K}$. In such a case relations (27) are identically satisfied. However, we can still match $\mathbf{r}_{1,K}, \mathbf{r}_{2,K}, \mathbf{r}_{3,K}$ to the eigenvectors of $\widehat{G}_{\Delta_K}(\mathbf{E}_{\Delta_K})$. In fact, in this case, we

have

$$J(\{s_{i,K}, \mathbf{r}_{i,K}\}_{i=1,2,3}) = s_{1,K} \sum_{i=1}^{3} \mathbf{r}_{i,K}^T \, \widehat{G}_{\Delta_K}(\mathbf{E}_{\Delta_K}) \, \mathbf{r}_{i,K} = s_{1,K} \, \mathrm{Tr}\big(R_K \widehat{G}_{\Delta_K}(\mathbf{E}_{\Delta_K}) R_K^T\big), \tag{29}$$

where the orthogonal matrix R_K^T is defined as in Section 2 and $\mathrm{Tr}(\cdot)$ denotes the trace of a matrix. Actually it can be checked that $\big[R_K \widehat{G}_{\Delta_K}(\mathbf{E}_{\Delta_K}) R_K^T\big]_{ij} = \mathbf{r}_{i,K}^T \widehat{G}_{\Delta_K}(\mathbf{E}_{\Delta_K}) \, \mathbf{r}_{j,K}$, with $i, j = 1, 2, 3$. On noting that $R_K \widehat{G}_{\Delta_K}(\mathbf{E}_{\Delta_K}) R_K^T$ is similar to $\widehat{G}_{\Delta_K}(\mathbf{E}_{\Delta_K})$, it clearly holds that $\mathrm{Tr}\big(R_K \widehat{G}_{\Delta_K}(\mathbf{E}_{\Delta_K}) R_K^T\big) = \mathrm{Tr}\big(\widehat{G}_{\Delta_K}(\mathbf{E}_{\Delta_K})\big) = \sum_{i=1}^{3} g_i$, showing that $J(\{s_{i,K}, \mathbf{r}_{i,K}\}_{i=1,2,3})$ in (29) does not depend on the choice of $\mathbf{r}_{1,K}, \mathbf{r}_{2,K}, \mathbf{r}_{3,K}$.

Summarising what we have proved so far, we can state that, whatever the values of the stretching factors $\{s_{i,K}\}$, we can always pick the orthonormal vectors $\{\mathbf{r}_{i,K}\}$ such that $\mathbf{r}_{1,K} = \mathbf{g}_l$, $\mathbf{r}_{2,K} = \mathbf{g}_m$, and $\mathbf{r}_{3,K} = \mathbf{g}_n$, for a suitable permutation $\{l, m, n\}$ of the indices $\{1, 2, 3\}$.

2) In the following, we exploit the ordering $s_{1,K} \geq s_{2,K} \geq s_{3,K}$ to unequivocally determine l, m and n. Let then

$$J(\{s_{i,K}, \mathbf{r}_{i,K}\}_{i=1,2,3}) = J_{lmn}(\{s_{i,K}, \mathbf{r}_{i,K}\}_{i=1,2,3}) = s_{1,K} \, g_l + s_{2,K} \, g_m + s_{3,K} \, g_n,$$

where we have used the property of the Rayleigh quotients, i.e., $\mathbf{g}_i^T \, \widehat{G}_{\Delta_K}(\mathbf{E}_{\Delta_K}) \, \mathbf{g}_i = g_i$, for $i = 1, 2, 3$. To shorten the notation we neglect in this point the dependence of $J(\{s_{i,K}, \mathbf{r}_{i,K}\}_{i=1,2,3})$ on its arguments. It can be proved that J is minimized with the identification $\{l, m, n\} = \{3, 2, 1\}$. In fact, letting $J^* = J_{321} = s_{1,K} \, g_3 + s_{2,K} \, g_2 + s_{3,K} \, g_1$, we show that any other permutation yields a non-smaller value of J. Any of the five other permutations of $\{3, 2, 1\}$ is at distance d (1 or 2) to $\{3, 2, 1\}$, i.e., d pairwise operations of swapping are applied to $\{3, 2, 1\}$. Let us first consider any of the three permutations of $\{3, 2, 1\}$ at distance 1, e.g., $\{3, 2, 1\} \mapsto \{2, 3, 1\}$. For this choice, we have

$$J_{231} - J^* = s_{1,K} \, (g_2 - g_3) + s_{2,K} \, (g_3 - g_2) = (s_{1,K} - s_{2,K})(g_2 - g_3) \geq 0. \tag{30}$$

The other two permutations at distance 1, i.e., $\{3, 2, 1\} \mapsto \{1, 2, 3\}$ and $\{3, 2, 1\} \mapsto \{3, 1, 2\}$, can be dealt with in the same fashion and yield the same result. Consider now one of the two permutations at distance 2, e.g., $\{3, 2, 1\} \mapsto \{1, 3, 2\}$ (the other is $\{3, 2, 1\} \mapsto \{2, 1, 3\}$). This can be equivalently obtained by composing two distance-1 permutations, i.e., $\{3, 2, 1\} \mapsto \{2, 3, 1\} \mapsto \{1, 3, 2\}$. We thus have

$$J_{132} - J^* = (J_{132} - J_{231}) + (J_{231} - J^*).$$

The second term is nonnegative thanks to (30). Let us check the first one.

$$J_{132} - J_{231} = s_{1,K} \, (g_1 - g_2) + s_{3,K} \, (g_2 - g_1) = (s_{1,K} - s_{3,K})(g_1 - g_2) \geq 0.$$

Thus it holds $J_{132} - J^* \geq 0$. The optimality of J_{321} can be checked analogously for the second permutation at distance 2. This proves the desired result that the minimum value of J is attained at $J = J^* = J_{321}$.

3) We are left with the task of minimizing $J(\{s_{i,K}, \mathbf{r}_{i,K}\}_{i=1,2,3})$ with respect to $\{s_{i,K}\}$. To take into account the constraint (20), we introduce the Lagrangian

$$\mathcal{L}(\{s_{i,K}\}_{i=1,2,3}, \mu) = s_{1,K}\, g_3 + s_{2,K}\, g_2 + s_{3,K}\, g_1 + \mu \Big(\prod_{i=1}^{3} s_{i,K} - 1 \Big),$$

where $\mu \in \mathbb{R}$ is the Lagrange multiplier. The optimality conditions yield

$$\begin{cases} \dfrac{\partial \mathcal{L}}{\partial s_{1,K}} = g_3 + \mu\, s_{2,K}\, s_{3,K} = 0 \\ \dfrac{\partial \mathcal{L}}{\partial s_{2,K}} = g_2 + \mu\, s_{3,K}\, s_{1,K} = 0 \\ \dfrac{\partial \mathcal{L}}{\partial s_{3,K}} = g_1 + \mu\, s_{1,K}\, s_{2,K} = 0 \\ \dfrac{\partial \mathcal{L}}{\partial \mu} = s_{1,K}\, s_{2,K}\, s_{3,K} - 1 = 0. \end{cases} \tag{31}$$

These equations provide the value of the Lagrange multiplier $\mu = -\sqrt[3]{g_1\, g_2\, g_3}$ as well as the optimal stretching factors (22). □

Remark 4.1 *The optimal* $\{\mathbf{r}_{i,K}\}$ *and* $\{s_{i,K}\}$ *in* (23)-(22) *equalize the three terms* $s_{i,K}\, (\mathbf{r}_{i,K}^T \widehat{G}_{\Delta_K}(\mathbf{E}_{\Delta_K}) \mathbf{r}_{i,K})$ *in* (21)*, i.e.,*

$$s_{1,K}\, g_3 = s_{2,K}\, g_2 = s_{3,K}\, g_1. \tag{32}$$

This equalization in turn yields

$$\sum_{i=1}^{3} s_{i,K}\, (\mathbf{r}_{i,K}^T \widehat{G}_{\Delta_K}(\mathbf{E}_{\Delta_K}) \mathbf{r}_{i,K}) = 3\, \sqrt[3]{g_1\, g_2\, g_3},$$

that is, on the optimized mesh, the functional $J(\cdot)$ *in* (21) *does not depend any longer on the stretching factors of the elements. Although we do not have a rigorous proof yet, we expect this property as well as the global equidistribution principle (invoked shortly below) to be the reason why the estimator (17), built on the local contributions (19), performs well with respect to the effectivity index, as shown in the numerical test cases in Section 6.*

To construct the optimal metric, we need the optimal $\{\mathbf{r}_{i,K}\}$ and $\{\lambda_{i,K}\}$. Thus we just have to compute $\{\lambda_{i,K}\}$. For this purpose we move from the optimal $\{s_{i,K}\}$ in (22), combined with (18) and with the equidistribution criterion

$$\eta_{K,\,\mathrm{aniso}}^2 = \frac{\tau^2}{\#\mathcal{T}_h}, \tag{33}$$

with τ a user-defined global tolerance on the H^1-seminorm of the discretization error and $\#\mathcal{T}_h$ being the cardinality (number of elements) of the actual mesh. In more detail, from (19), (33), and employing (32) three times, we have

$$\prod_{i=1}^{3} \lambda_{i,K} = \frac{\tau^2}{\#\mathcal{T}_h |\widehat{\Delta}_K| (3\, s_{1,K}\, g_3)} = \frac{\tau^2}{\#\mathcal{T}_h |\widehat{\Delta}_K| (3\, s_{2,K}\, g_2)} = \frac{\tau^2}{\#\mathcal{T}_h |\widehat{\Delta}_K| (3\, s_{3,K}\, g_1)}.$$

Thus

$$\left(\prod_{i=1}^{3}\lambda_{i,K}\right)^3 = \left(\frac{\tau^2}{3\#\mathcal{T}_h|\widehat{\Delta}_K|}\right)^3 \frac{1}{g_1\, g_2\, g_3},$$

where also (20) has been used. From this we obtain

$$\prod_{i=1}^{3}\lambda_{i,K} = \frac{\tau^2}{3\#\mathcal{T}_h|\widehat{\Delta}_K|}\left(\prod_{i=1}^{3} g_i\right)^{-1/3} \equiv V_K. \tag{34}$$

Now, using (18) and (34), we have

$$\begin{aligned}
\lambda_{1,K} &= V_K^{1/3}\left(\frac{\prod_{i=1}^3 g_i}{g_3^3}\right)^{1/6} = \left(\frac{\tau^2}{3\#\mathcal{T}_h|\widehat{\Delta}_K|}\right)^{1/3}\left(\prod_{i=1}^{3} g_i\right)^{1/18} g_3^{-1/2} \\
\lambda_{2,K} &= V_K^{1/3}\left(\frac{\prod_{i=1}^3 g_i}{g_2^3}\right)^{1/6} = \left(\frac{\tau^2}{3\#\mathcal{T}_h|\widehat{\Delta}_K|}\right)^{1/3}\left(\prod_{i=1}^{3} g_i\right)^{1/18} g_2^{-1/2} \\
\lambda_{3,K} &= V_K^{1/3}\left(\frac{\prod_{i=1}^3 g_i}{g_1^3}\right)^{1/6} = \left(\frac{\tau^2}{3\#\mathcal{T}_h|\widehat{\Delta}_K|}\right)^{1/3}\left(\prod_{i=1}^{3} g_i\right)^{1/18} g_1^{-1/2}.
\end{aligned} \tag{35}$$

We are in a position to identify the optimal elementwise metric, $\widetilde{M}_K$, given by $\widetilde{M}_K = R_K^T \Lambda_K^{-2} R_K$, with $R_K^T = [\mathbf{r}_{1,K}, \mathbf{r}_{2,K}, \mathbf{r}_{3,K}]$ and $\Lambda_K = \mathrm{diag}(\lambda_{1,K}, \lambda_{2,K}, \lambda_{3,K})$, where $\{\mathbf{r}_{i,K}\}$ and $\{\lambda_{i,K}\}$ are provided by (23) and (35), respectively. The corresponding nodewise metric is thus defined by

$$\widetilde{M}_N = \frac{1}{|\Delta_N|}\sum_{K\in\Delta_N} |K|\, \widetilde{M}_K,$$

where N is any node of the mesh, Δ_N is the patch of the elements that share N and $|\Delta_N|$ is the corresponding volume.

Finally the metric is obtained dividing by $(2\sqrt{2/3})^2 = 8/3$ every $\widetilde{M}_N$. This scaling amounts to shrinking the reference tetrahedron to a unit edge one.

Remark 4.2 *The hypothesis $g_1 \geq g_2 \geq g_3 > 0$ in Proposition 4.1 can be relaxed by assuming $g_1 \geq g_2 \geq g_3 \geq 0$ which happens when $\widehat{G}_{\Delta_K}(\mathbf{E}_{\Delta_K})$ is actually positive semidefinite. However this degenerate case can be tackled by defining a minimum value*

$$g_{min} = h_\Omega^{-12}\left(\frac{\tau^2}{3\#\mathcal{T}_h|\widehat{\Delta}_K|}\right)^4, \tag{36}$$

for g_i where h_Ω is the diameter of the domain. The eigenvalues are modified by taking g_i to be $\max(g_i, g_{min})$. When this expression is used for g_{min}, if all the eigenvalues $\{g_i\}$ are degenerate, then, from (35), $\lambda_{i,K}$ is equal to h_Ω for $i = 1, 2, 3$.

5 The adaptive procedure

We summarize here the main steps of the actual adaptive procedure. The objective is to approximate the solution to the constrained minimization problem:

$$\min_{\mathcal{T}_h \in \mathcal{T}(\overline{\Omega})} \#\mathcal{T}_h \qquad \text{s.t.} \sum_{K \in \mathcal{T}_h} \eta^2_{K,\,\text{aniso}} = \tau^2,$$

where $\mathcal{T}(\overline{\Omega})$ is the set of all possible conformal meshes of $\overline{\Omega}$, with τ the global tolerance enforced on the H^1-seminorm.

We detail in the following the resulting algorithm.

1. Let $\mathcal{T}_h^{(0)}$ be a background mesh;

2. Set the iteration counter $j = 0$;

3. While ''a convergence criterion is not satisfied''
   Compute u_h from (15);
   Loop on $K \in \mathcal{T}_h^{(j)}$:

   (a) Compute
   $$P_{\Delta_K}(\nabla u_h) = \frac{1}{|\Delta_K|} \sum_{T \in \Delta_K} |T|\, \nabla u_h|_T$$
   and $\mathbf{E}_{\Delta_K} = P_{\Delta_K}(\nabla u_h) - \nabla u_h|_{\Delta_K}$;

   (b) Compute $\widehat{G}_{\Delta_K}(\mathbf{E}_{\Delta_K}) = G_{\Delta_K}(\mathbf{E}_{\Delta_K})/|\Delta_K|$ with
   $$[G_{\Delta_K}(\mathbf{E}_{\Delta_K})]_{ij} = \int_{\Delta_K} (\mathbf{E}_{\Delta_K})_i\, (\mathbf{E}_{\Delta_K})_j \, d\mathbf{x}, \quad \text{with } i,j = 1,2,3;$$

   (c) Compute $M_K = R_K^T\, \Lambda_K\, R_K\, Z_K$ by the SVD algorithm:
   $[R_K^T, \Lambda_K, Q_K] = \text{svd}(M_K)$, such that $M_K = R_K^T\, \Lambda_K\, Q_K$,
   with $Q_K = R_K\, Z_K$;

   (d) Let $\lambda_{1,K}^{(j)} \geq \lambda_{2,K}^{(j)} \geq \lambda_{3,K}^{(j)}$ be the entries of Λ_K;

   (e) Compute $|\widehat{\Delta}_K| = |\Delta_K|/(\prod_{i=1}^3 \lambda_{i,K}^{(j)})$;

   (f) Compute the eigen-pairs $\{g_i, \mathbf{g}_i\}$ of $\widehat{G}_{\Delta_K}(\mathbf{E}_{\Delta_K})$:
   $[g_i, \mathbf{g}_i] = \text{eig}(\widehat{G}_{\Delta_K}(\mathbf{E}_{\Delta_K}))$, with $g_1 \geq g_2 \geq g_3$;
   set $g_i = \max(g_i, g_{\min})$ for $i = 1,2,3$, with $g_{\min}$ as in (36);

(g) Set $\mathbf{r}_{1,K}^{(j+1)} = \mathbf{g}_3$, $\mathbf{r}_{2,K}^{(j+1)} = \mathbf{g}_2$, and $\mathbf{r}_{3,K}^{(j+1)} = \mathbf{g}_1$,

$$\lambda_{1,K}^{(j+1)} = \left(\frac{\tau^2}{3\#\mathcal{T}_h^{(j)}|\widehat{\Delta}_K|}\right)^{1/3}\left(\prod_{i=1}^{3} g_i\right)^{1/18} g_3^{-1/2}$$
$$\lambda_{2,K}^{(j+1)} = \left(\frac{\tau^2}{3\#\mathcal{T}_h^{(j)}|\widehat{\Delta}_K|}\right)^{1/3}\left(\prod_{i=1}^{3} g_i\right)^{1/18} g_2^{-1/2}$$
$$\lambda_{3,K}^{(j+1)} = \left(\frac{\tau^2}{3\#\mathcal{T}_h^{(j)}|\widehat{\Delta}_K|}\right)^{1/3}\left(\prod_{i=1}^{3} g_i\right)^{1/18} g_1^{-1/2},$$

and form the matrices

$$R_K^T = [\mathbf{r}_{1,K}^{(j+1)}, \mathbf{r}_{2,K}^{(j+1)}, \mathbf{r}_{3,K}^{(j+1)}], \quad \Lambda_K = \mathrm{diag}(\lambda_{1,K}^{(j+1)}, \lambda_{2,K}^{(j+1)}, \lambda_{3,K}^{(j+1)});$$

(h) Build the metric $\widetilde{M}_K = R_K^T \Lambda_K^{-2} R_K$;

4. Loop on the nodes and build the nodewise metric

$$\widetilde{M}_N = \frac{3}{8}\frac{1}{|\Delta_N|}\sum_{K\in\Delta_N} |K|\,\widetilde{M}_K;$$

5. Compute the mesh $\mathcal{T}_h^{(j+1)}$;

6. Set $j = j + 1$;

7. End While

Step 5. assumes the availability of an adaptive remeshing procedure, as already anticipated in Section 4. This could consist of global remeshing (e.g., [8, 48]), local remeshing (e.g., [49, 50]), or mesh optimisation (e.g., [11]). In this work, the mesh optimisation algorithm of [11] is used.
Concerning the convergence criterion at point 3., an example is provided in the next section.

6 Examples

In this section we assess the performance of both the error estimator (17),(19) and the adaptive procedure in Section 5 on some numerical simulations dealing with academic test cases as well as fluid dynamic and multi-material applications.

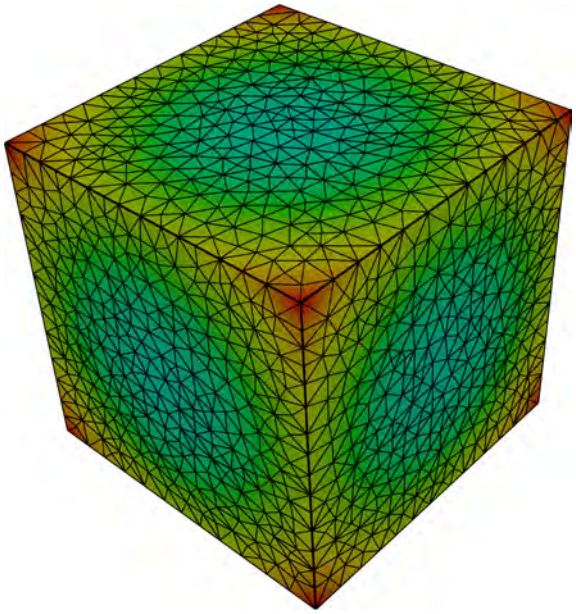

Figure 2: Adapted mesh for the academic test case with isotropic solution, for $\tau = 1$.

6.1 Academic test case with isotropic solution

The convergence of η_{aniso} and of the adaptive procedure is investigated for an analytic solution, for different values of the tolerance τ. An initial mesh of 17000 elements is generated for the domain $\Omega = (0,1)^3$. On this mesh, the field u is applied by nodal evaluation, with

$$u(\mathbf{x}) = ||\mathbf{x} - \mathbf{c}||^2, \tag{37}$$

to give u_h, where $\mathbf{c}$ is the centre of the cube and $\|\cdot\|$ denotes the Euclidean norm on $\mathbb{R}^3$. In practice u_h is the Lagrange interpolant of u.
The adaptive procedure is then applied to generate a new mesh, upon which u is re-evaluated. The adopted convergence criterion is

$$\frac{3}{4}\tau \leq \eta_{\text{aniso}} \leq \frac{5}{4}\tau. \tag{38}$$

Results are shown in Table 1, where $\theta = \eta_{\text{aniso}}/|u - u_h|_{H^1(\Omega)}$ denotes the effectivity index. It appears that, despite its large values, θ is independent of the chosen tolerance.
As can be seen from Figure 2, the resulting mesh is isotropic as expected. The interpolation error behaviour in Figure 3 (left) shows that the error scales like $\mathcal{O}(\#\mathcal{T}_h^{-1/3})$. This is the expected trend in a 3D framework.

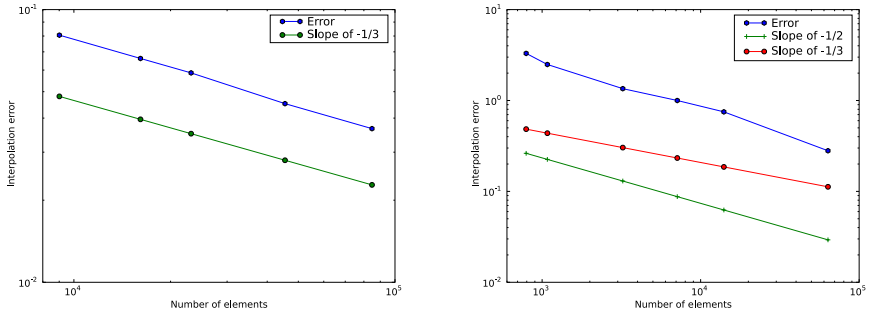

Figure 3: Plots of interpolation error in the H^1-seminorm for the academic test case with isotropic (left) and anisotropic (right) solution.

τ	$\lvert u - u_h \rvert_{H^1(\Omega)}$	$\#\mathcal{T}_h$	η_{aniso}	θ	iterations
2.0	8.06×10^{-2}	9011	1.90	23.6	1
1.5	6.62×10^{-2}	16118	1.59	24.0	1
1.25	5.86×10^{-2}	23172	1.43	24.4	1
1.0	4.52×10^{-2}	45439	1.16	25.7	2
0.8	3.66×10^{-2}	84799	0.97	26.5	2

Table 1: Convergence data for the academic test case with isotropic solution: tolerance; H^1-seminorm of the error; number of elements; estimator; effectivity index; number of iterations.

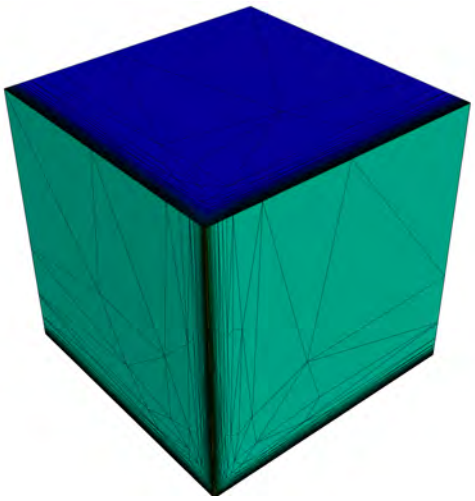

Figure 4: Adapted mesh for the academic test case with anisotropic solution, for $\tau = 1$.

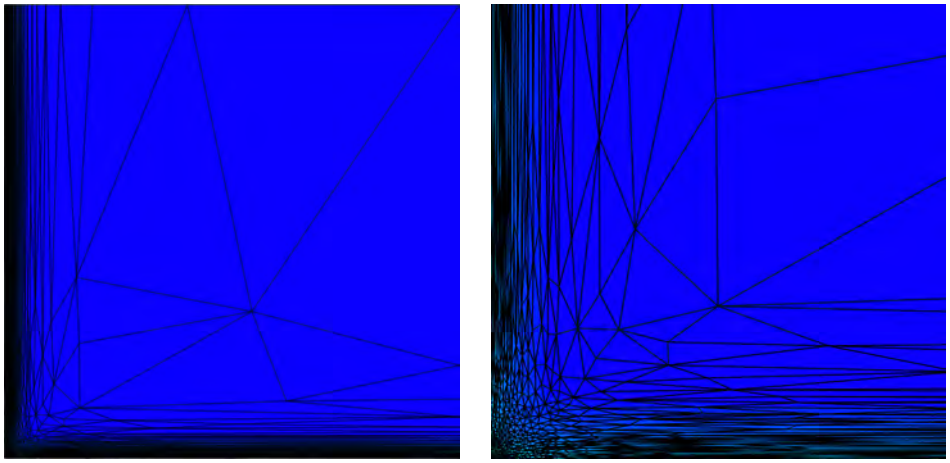

Figure 5: Top-down view (left) and zoom (right) of the bottom-left corner of the plane $x_3 = 1$ for the academic test case with anisotropic solution, for $\tau = 1$.

6.2 Academic test case with anisotropic solution

To assess the performance of η_{aniso} and of the adaptive procedure on a test case exhibiting anisotropic features, we consider the field

$$u(x_1, x_2, x_3) = e^{-x_1/\epsilon} + e^{-x_2/\epsilon} + e^{-x_3/\epsilon}, \tag{39}$$

with $\epsilon = 0.01$. This function emulates the presence of three boundary layers along $x_1 = 0$, $x_2 = 0$ and $x_3 = 0$. This field is applied to the same initial mesh as in the previous section. The same convergence criterion is used.
Results are shown in Table 2. Notice the reduced values of the effectivity index, indicating a higher reliability of the estimator, as well as a larger number of adaptive iterations since now more involved adaptation is required due to the steep boundary layers. As the field is poorly represented on the initial isotropic mesh, the algorithm takes more iterations to converge to a highly anisotropic mesh. In general, we can observe that the meshes yielded by the adaptive procedure are quasi-optimal. Actually, the effectivity index is not unitary, yet being independent of the accuracy required on the true error. In fact, for a given problem, the family of (optimally) adapted triangulations characterized by a different level of accuracy, τ, in the H^1-seminorm, shows an effectivity index which is independent of τ, as well as of the stretching factors (related to the shape regularity) of the family of triangulations. Thus the quasi-optimality has essentially an impact only on the constant in the error-vs-number of elements relation.

As can be seen from Figures 4 and 5, the resulting mesh is highly anisotropic. The interpolation error behaviour is shown in Figure 3 (right). The slope of the trend is now about $-1/2$. This substantial improvement is related to the essentially 2D nature of the problem.

τ	$\lvert u - u_h \rvert_{H^1(\Omega)}$	$\#\mathcal{T}_h$	η_{aniso}	θ	iterations
10.0	3.30	793	10.5	3.2	6
8.0	2.49	1079	7.9	3.2	6
5.0	1.35	3228	4.2	3.1	5
3.0	1.00	7122	2.9	2.9	5
2.0	0.75	14016	2.1	2.8	5
1.0	0.28	63588	0.9	3.2	6

Table 2: Convergence data for the academic test case with anisotropic solution: tolerance; H^1-seminorm of the error; number of elements; estimator; effectivity index; number of iterations.

τ	$\lvert u - u_h \rvert_{H^1(\Omega)}$	$\#\mathcal{T}_h$	η_{aniso}	θ	iterations
34.0	2.37	1747944	35.9	15.1	3
35.0	3.07	788518	42.2	13.7	3
36.0	2.67	1169779	39.4	14.7	3

Table 3: Convergence data for the academic test case with anisotropic solution with isotropic adaptivity: tolerance; H^1-seminorm of the error; number of elements; estimator; effectivity index; number of iterations.

Table 3 shows results where the adaptive algorithm is prevented from exploiting anisotropy by forcing the metric to be isotropic. This is achieved by setting each $\lambda_{i,K}$ to be the minimum of $\{\lambda_{i,K}\}$. As can be seen, anisotropic adaptivity is vastly more efficient for representing anisotropic functions; similar levels of error are achieved for approximately 10^3 fewer elements. This strongly motivates the development of error estimators which are able to exploit anisotropy.

6.3 Laminar flow past a cylinder

The cylinder flow benchmark of [51] is considered. In this problem, laminar flow through a channel is interrupted by a cylindrical body. The benchmark specifies that the incompressible Navier-Stokes equations are to be solved in the domain $\Omega = (0, H) \times (0, H) \times (0, 2.5) \setminus C$, where $H = 0.41$ m is the height of the domain and C is a cylinder of radius 0.1 m and length H positioned 0.45 m from the inflow boundary in the channel. The inflow condition of case 3D-1Z is used:

$$\mathbf{u}(0, x_2, x_3) = (16\, U_m\, x_2\, x_3\, (H - x_2)(H - x_3)/H^4,\, 0,\, 0)^T, \tag{40}$$

where $\mathbf{u}$ is the velocity field and U_m is chosen to be 0.45 m/s such that the Reynolds number is 20. Homogenous Dirichlet boundary conditions are imposed on velocity on the horizontal sides of the channel and the cylindrical body. This motivates placing extra resolution along the sides and in the wake of the

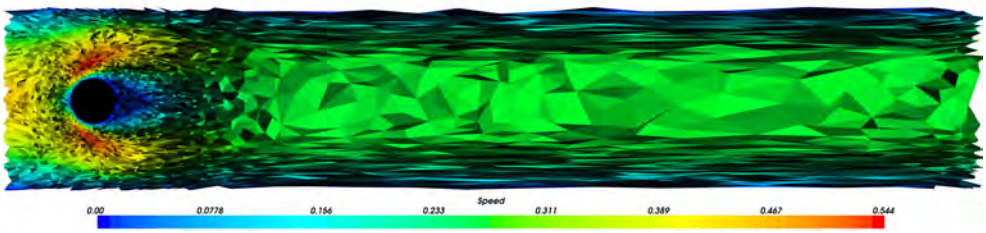

Figure 6: Overlay of the norm of the velocity field and mesh for laminar flow past a cylinder.

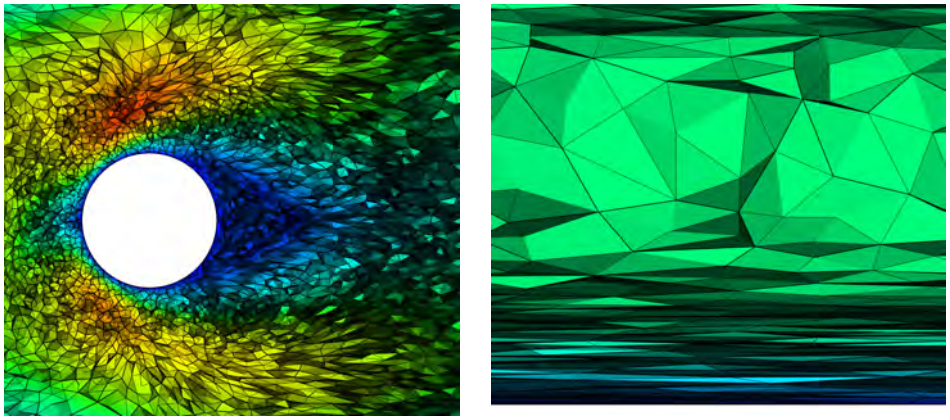

Figure 7: Zoom of the mesh and norm of the velocity field in the wake of the cylinder (left) and along the sides of the channel (right).

cylinder to resolve the flow features. At the outflow, a homogenous Neumann boundary condition is imposed on velocity and a homogenous Dirichlet boundary condition is imposed on pressure. The incompressible Navier-Stokes equations are discretised with the stabilised $\mathbb{P}1$-$\mathbb{P}1$ element pair [52]. The speed is chosen as the field to guide adaptivity, with $\tau = 1.5$. The maximum number of nodes allowed is set to 500000, and the minimum element size set to 4.1×10^{-3}, i.e., one hundredth of the height of the domain. These settings are chosen so as not to interfere with the adaptive procedure. Moreover we additionally employ the gradation algorithm of [53] with $\gamma = 2$, to smooth the variation in mesh spacing requested by the metric.

The results can be seen in Figures 6 and 7. The adaptive procedure converges in 5 iterations, with $\eta_{\text{aniso}} = 1.64$. The resulting mesh has approximately 220000 nodes, or 1.2 million elements. Note the concentration of the resolution in the wake past the cylinder and along the boundary layers along the side of the channel. As can be seen in Figure 7, the elements along the sides of the channel are extremely anisotropic, reflecting the anisotropy of the solution caused by the boundary conditions there. The elements in the wake of the cylinder are mostly isotropic, again reflecting the nature of the dynamics in that region. This result

gives confidence that the anisotropic resolution of boundary layers analysed in the previous section also holds for flows of practical interest.

If the objective of the simulation is to compute the lift or drag coefficients, the use of goal-based adaptivity is extremely efficient [54]. Here, no dual-based information is incorporated into the adaptive procedure. Future work will combine this approach with the dual-based approach described in [24].

6.4 Multimaterial application

Simulations of multimaterial flows are numerically challenging. The interfaces of the material volume fractions recording the materials are sharp and anisotropic, and this must be reflected in the mesh upon which these flows are discretised. Furthermore, the discretisation must be conservative and bounded; otherwise nonphysical phenomena such as mass exchange may occur.

The adaptive algorithm is applied to an unsteady multimaterial simulation of a water column collapsing under gravity. The models solved comprise the incompressible Navier-Stokes equation and the advection equation for the evolution of the material volume fraction. The simulation is conducted in the domain

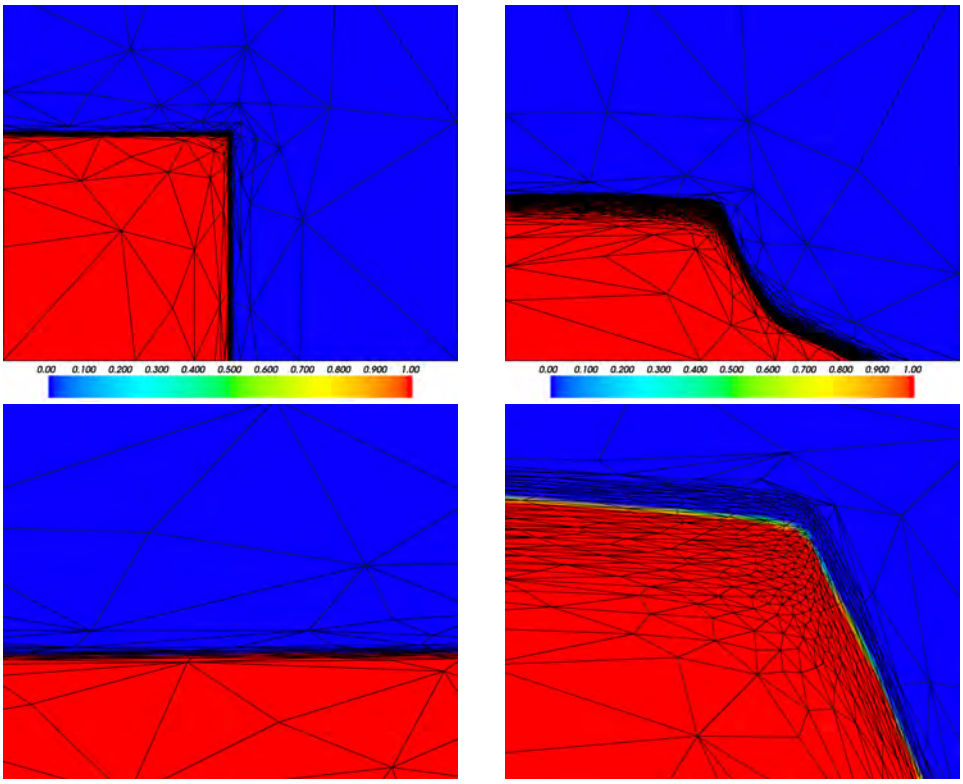

Figure 8: The multimaterial application: material volume fraction (top) and zoom (bottom) of the mesh at $t = 0$ (left) and $t = 2$ (right) time units.

$\Omega = (-0.5, 0.5) \times (-0.5, 2) \times (-0.5, 0.5)$. A material volume fraction representing water is initialised to be 1 in the region $[-0.5, -0.25] \times [-0.5, 0] \times [-0.5, 0]$ and zero elsewhere. No-normal flow is imposed on velocity on all boundaries except for the top. At the top, a homogenous Neumann boundary condition is imposed on velocity, and a homogenous Dirichlet boundary condition is imposed on pressure. The $\mathbb{P}0$-$\mathbb{P}1$cv element pair is used for the velocity-pressure discretisation; the HyperC control volume face value algorithm is used for the advection equation [55]. The discretisation is described in [56]. The material volume fraction representing water is chosen as the field to guide adaptivity, with $\tau = 25$. A minimum edge length of 0.001 is enforced to constrain the adaptive algorithm. The mesh is adapted every 10 timesteps. As conservation and boundedness are crucial, the interpolation algorithm presented in [57] is used to transfer data between meshes. To spread resolution ahead of the dynamics, the metric tensor formed is advected forward for one adaptivity period and superimposed with itself. This has the effect of extending resolution to where the interface will be over the course of the adaptivity period.

The results can be seen in Figures 8 and 9. The error estimator clearly detects

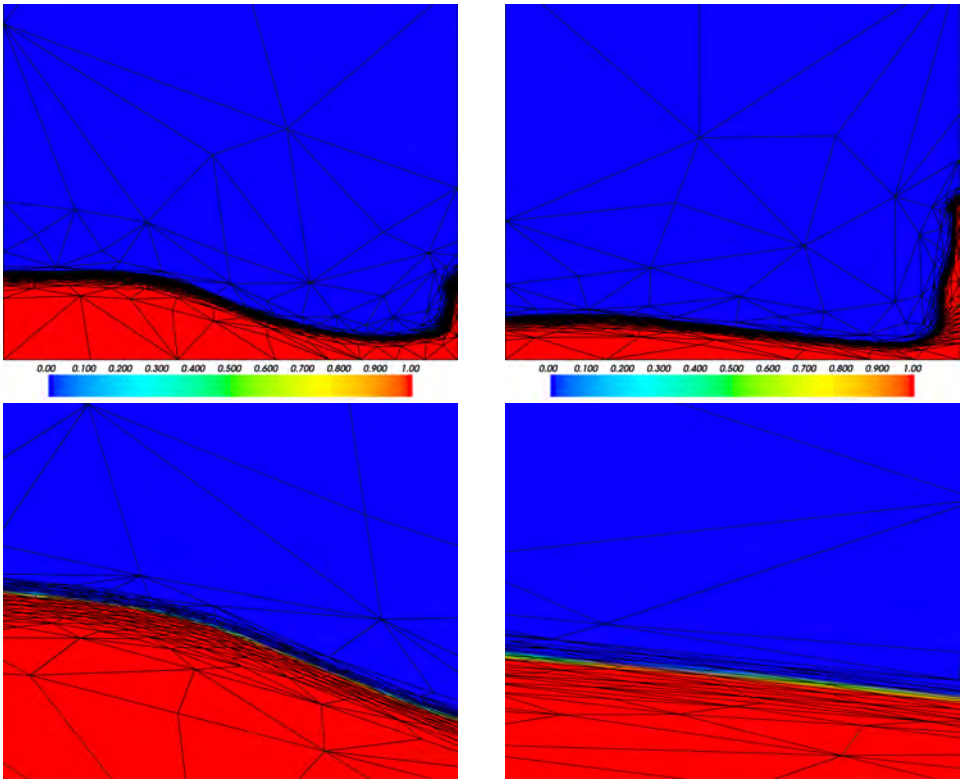

Figure 9: The multimaterial application: material volume fraction (top) and zoom (bottom) of the mesh at $t = 4$ (left) and $t = 6$ (right) time units.

the anisotropic nature of the material interface and places anisotropic elements

there to resolve it. The estimate η_{aniso} stays within the range $[20, 30]$ for most of the simulation, except at the end when the resolution requested is finer than the minimum element size imposed. The computation of this simulation to the same accuracy on any fixed grid would be prohibitively expensive.

7 Conclusions

A suitable 3D adaptation algorithm based on an optimization procedure is devised and proved, numerically, to yield quasi-optimal meshes. In fact, academic test cases with both smooth (isotropic) and nonsmooth (anisotropic) solutions, show that the error measured in the H^1-seminorm depends on the number of mesh elements, $\#\mathcal{T}_h$, as $\mathcal{O}(\#\mathcal{T}_h^{-1/3})$, the optimal scaling in 3D. Moreover, when the 3D problem at hand exhibits 2D features, such as boundary layers, the dependence of the error on the cardinality of the mesh becomes $\mathcal{O}(\#\mathcal{T}_h^{-1/2})$, still optimal in 2D. Despite its heuristic nature, the proposed anisotropic Zienkiewicz-Zhu a posteriori estimator goes to prove effectiveness. Indeed, the error estimator is able to detect the main features of the solution even in complex applications. Results in Section 6.4 are promising with a view to unsteady problems as well.

References

[1] Micheletti S, Perotto S. 2009; In preparation.

[2] Formaggia L, Selmin V. Simulation of hypersonic flows on unstructured grids. *Int. J. Numer. Meth. Engng* 1992; **34**:569–606.

[3] Simpson RB. Anisotropic mesh transformations and optimal error control. *Appl. Numer. Math.* 1994; **14**:183–198.

[4] Venditti DA. Grid adaptation for functional outputs of compressible flow simulations. PhD Thesis, Massachusetts Institute of Technology 2002.

[5] Frey PJ, Alauzet F. Anisotropic mesh adaptation for CFD computations. *Comput. Methods Appl. Mech. Engrg.* 2005; **194**:5068–5082.

[6] Bourgault Y, Picasso M, Alauzet F, Loseille A. On the use of anisotropic a posteriori error estimators for the adaptative solution of 3D inviscid compressible flows. *Int. J. Numer. Meth. Fluids* 2009; **59**(1):47–74.

[7] Boussetta R, Coupez T, Fourment L. Adaptive remeshing based on a posteriori error estimation for forging simulation. *Comput. Methods Appl. Mech. Engrg.* 2006; **195**(48-49):6626–6645.

[8] Peraire J, Vahdati M, Morgan K, Zienkiewicz O. Adaptive remeshing for compressible flow computations. *J. Comput. Phys.* 1987; **72**(2):449–466.

[9] D'Azevedo E. Optimal triangular mesh generation by coordinate transformation. *SIAM J. Sci. Stat. Comput.* 1991; **12**(4):755–786.

[10] Castro-Diaz MJ, Hecht F, Mohammadi B, Pironneau O. Anisotropic unstructured mesh adaptation for flow simulations. *Int. J. Numer. Meth. Fluids* 1997; **25**(4):475–491.

[11] Pain C, Umpleby A, de Oliveira C, Goddard A. Tetrahedral mesh optimisation and adaptivity for steady-state and transient finite element calculations. *Comput. Methods Appl. Mech. Engrg.* 2001; **190**(29–30):3771–3796.

[12] Micheletti S, Perotto S. Output functional control for nonlinear equations driven by anisotropic mesh adaption: the Navier-Stokes equations. *SIAM J. Sci. Comput.* 2008; **30**(6):2817–2854.

[13] Almeida R, Feijóo R, Galeão A, Padra C, Silva R. Adaptive finite element computational fluid dynamics using an anisotropic error estimator. *Comput. Methods Appl. Mech. Engrg.* 2000; **182**(3-4):379–400.

[14] Dompierre J, Vallet MG, Bourgault Y, Fortin M, Habashi W. Anisotropic mesh adaptation: towards user-independent, mesh-independent and solver-independent CFD. III. Unstructured meshes. *Int. J. Numer. Meth. Fluids* 2002; **39**(8):675–702.

[15] Belhamadia Y, Fortin A, Chamberland E. Three-dimensional anisotropic mesh adaptation for phase change problems. *J. Comput. Phys.* 2004; **201**(2):753–770.

[16] Piggott MD, Pain CC, Gorman GJ, Power PW, Goddard AJH. h, r, and hr adaptivity with applications in numerical ocean modelling. *Ocean Model.* 2005; **10**(1-2):95–113.

[17] Gruau C, Coupez T. 3D tetrahedral, unstructured and anisotropic mesh generation with adaptation to natural and multidomain metric. *Comput. Methods Appl. Mech. Engrg.* 2005; **194**(48-49):4951–4976.

[18] Remacle JF, Frazao SS, Li X, Shephard MS. An adaptive discretization of shallow-water equations based on discontinuous Galerkin methods. *Int. J. Numer. Meth. Fluids* 2006; **52**(8):903–923.

[19] Apel T. *Anisotropic Finite Elements: Local Estimates and Applications.* Advances in Numerical Mathematics, Teubner: Stuttgart, 1999.

[20] Kunert G. An a posteriori residual error estimator for the finite element method on anisotropic tetrahedral meshes. *Numer. Math.* 2000; **86**(3):471–490.

[21] Formaggia L, Perotto S. New anisotropic a priori error estimates. *Numer. Math.* 2001; **89**(4):641–667.

[22] Formaggia L, Perotto S. Anisotropic error estimates for elliptic problems. *Numer. Math.* 2003; **94**(1):67–92.

[23] Venditti DA, Darmofal DL. Anisotropic grid adaptation for functional outputs: application to two-dimensional viscous flows. *J. Comput. Phys.* 2003; **187**:22–46.

[24] Formaggia L, Micheletti S, Perotto S. Anisotropic mesh adaptation in computational fluid dynamics: application to the advection–diffusion–reaction and the Stokes problems. *Appl. Numer. Math.* 2004; **51**(4):511–533.

[25] Remaki L, Habashi W. 3-d mesh adaptation on multiple weak discontinuities and boundary layers. *SIAM J. Sci. Comput.* 2006; **28**(4):1379–1397.

[26] Alauzet F, Frey P, George P, Mohammadi B. 3D transient fixed point mesh adaptation for time dependent problems. *J. Comput. Phys.* 2007; **222**(2):592–623.

[27] Narski J, Picasso M. Adaptive 3d finite elements with high aspect ratio for dentritic growth of a binary alloy including flow induced by shrinkage. *Fluid Dynam. Mater. Process.* 2007; **3**(1):49–64.

[28] Bruch J. Free surface seepage problems solved using: (a) the Zienkiewicz-Zhu error estimation procedure; and (b) a parallel computer. *Pitman Res. Notes Math. Ser.* 1993; **282**:98–104.

[29] Lawrence K, Nambiar R. The Zienkiewicz-Zhu error estimator for multiple material problems. *Commun. Appl. Numer. Methods* 1992; **8**:273–277.

[30] Pawlak T, Wheeler M, Yunus S. Application of the Zienkiewicz-Zhu error estimator for plate and shell analysis. *Int. J. Numer. Meth. Engng* 1990; **29**:1281–1298.

[31] Zienkiewicz O, Zhu J. A simple error estimator and adaptive procedure for practical engineering analysis. *Int. J. Numer. Meth. Engng* 1987; **24**:337–357.

[32] Zienkiewicz O, Zhu J. The superconvergent patch recovery and a posteriori error estimates. I: The recovery technique. *Int. J. Numer. Meth. Engng* 1992; **33**:1331–1364.

[33] Zienkiewicz O, Zhu J. The superconvergent patch recovery and a posteriori error estimates. II: Error estimates and adaptivity. *Int. J. Numer. Meth. Engng* 1992; **33**:1365–1382.

[34] Krízek M, Neittaanmäki P. Superconvergence phenomenon in the finite element method arising from averaging gradients. *Numer. Math* 1984; **45**:105–116.

[35] Zienkiewicz O, Cheung Y. *The Finite Element Method in Structural and Continuum Mechanics. Numerical Solution of Problems in Structural and Continuum Mechanics*. McGraw-Hill, London, 1967.

[36] Ciarlet P. *The Finite Element Method for Elliptic Problems*. North-Holland Publishing Company: Amsterdam, 1978.

[37] Lions JL, Magenes E. *Non-Homogeneous Boundary Value Problems and Applications, vol. I*. Springer-Verlag, Berlin, 1972.

[38] Clément P. Approximation by finite element functions using local regularization. *RAIRO Anal. Numér.* 1975; **2**:77–84.

[39] Scott LR, Zhang S. Finite element interpolation of non-smooth functions satisfying boundary conditions. *Math. Comp.* 1990; **54**:483–493.

[40] Ainsworth MT, Oden JT. *A Posteriori Error Estimation in Finite Element Analysis*. Wiley: New York, 2000.

[41] Zienkiewicz O, Zhu J. The superconvergent patch recovery (SPR) and adaptive finite element refinement. *Comput. Methods Appl. Mech. Engrg.* 1992; **101**:207–224.

[42] Rodríguez R. Some remarks on Zienkiewicz-Zhu estimator. *Numer. Methods Partial Differential Equations* 1994; **10**:625–635.

[43] Carstensen C. All first-order averaging techniques for a posteriori finite element error control on unstructured grids are efficient and reliable. *Math. Comp.* 2004; **73**:1153–1165.

[44] Naga A, Zhang Z. A posteriori error estimates based on the polynomial preserving recovery. *SIAM J. Numer. Anal.* 2004; **42**:1780–1800.

[45] Yan N, Zhou A. Gradient recovery type a posteriori error estimation for finite element approximations on irregular meshes. *Comput. Methods Appl. Mech. Engrg.* 2001; **190**:4289–4299.

[46] Bottasso C, Maisano G, Micheletti S, Perotto S. On some new recovery based a posteriori error estimators. *Comput. Methods Appl. Mech. Engrg.* 2006; **195**:4794–4815.

[47] George PL, Borouchaki H. *Delaunay Triangulation and Meshing: Application to Finite Elements*. Hermes: Paris, 1998.

[48] Peraire J, Peiró J, Formaggia L, Morgan K, Zienkiewicz OC. Finite element Euler computations in three dimensions. *Int. J. Numer. Meth. Engng* 1988; **26**(10):2135–2159.

[49] Hassan O, Probert EJ, Morgan K. Unstructured mesh procedures for the simulation of three-dimensional transient compressible inviscid flows with moving boundary components. *Int. J. Numer. Meth. Fluids* 1998; **27**(1-4):41–55.

[50] Hassan O, Probert EJ, Morgan K, Weatherill NP. Unsteady flow simulation using unstructured meshes. *Comput. Methods Appl. Mech. Engrg.* 2000; **189**(4):1247–1275.

[51] Schäfer M, Turek S, Durst F, Krause E, Rannacher R. Benchmark computations of laminar flow around a cylinder. *Notes Numer. Fluid Mech.* 1996; **52**:547–566.

[52] Gresho PM, Sani RL. *Incompressible Flow and the Finite Element Method, Volume 2, Isothermal Laminar Flow*. John Wiley & Sons, 2000.

[53] Li X, Remacle J, Chevaugeon N, Shephard MS. Anisotropic mesh gradation control. *Proceedings of the 13th International Meshing Roundtable*, Williamsburg, Virginia, 2004.

[54] Bangerth W, Rannacher R. *Adaptive finite element methods for differential equations*. ETH Zürich Lectures in Mathematics, Birkhäuser Verlag: Basel, 2003.

[55] Leonard BP. The ultimate conservative difference scheme applied to unsteady one-dimensional advection. *Comput. Methods Appl. Mech. Engrg.* 1991; **88**(1):17–74.

[56] Wilson CR. Modelling multiple-material flows on adaptive unstructured meshes. PhD Thesis, Imperial College London, London, UK 2009.

[57] Farrell PE, Piggott MD, Pain CC, Gorman GJ. Conservative interpolation between unstructured meshes via supermesh construction. *Comput. Methods Appl. Mech. Engrg.* 2009; **198**(33-36):2632–2642.

MOX Technical Reports, last issues

Dipartimento di Matematica "F. Brioschi",
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

25/2009 P.E. Farrell, S. Micheletti, S. Perotto:
An anisotropic Zienkiewicz-Zhu a posteriori error estimator for 3D applications

24/2009 F. Di Maio, P. Secchi, S. Vantini, E. Zio:
Optimized Fuzzy C-Means Clustering and Functional Principal Components for Post-Processing Dynamic Scenarios in the Reliability Analysis of a Nuclear System

23/2009 L. Gerardo Giorda, F. Nobile, C. Vergara:
Analysis and optimization of Robin-Robin partitioned procedures in Fluid-Structure Interaction Problems

22/2009 L. Formaggia, S. Minisini, P. Zunino:
Modeling erosion controlled drug release and transport phenomena in the arterial tissue

21/2009 L. Bonaventura, S. Castruccio, L.M. Sangalli:
A Bayesian approach to geostatistical interpolation with flexible variogram models

20/2009 N.Accoto, T. Rydén, P. Secchi:
Bayesian hidden Markov models for performance-based regulation of continuity of electricity supply

19/2009 L. Formaggia, A.Mola, N. Parolini, M. Pischiutta:
A three-dimensional model for the dynamics and hydrodynamics of rowing boats

18/2009 J. Silva Soares, P. Zunino:
A mathematical model for water uptake, degradation, erosion, and drug release from degradable polydisperse polymeric networks

17/2009 M. Longoni, C. Magistroni, P. Ruffo, G. Scrofani:
3D Inverse and Direct Structural Modeling Workflow

16/2009 G. Aletti, C. May, P. Secchi:
A functional equation whose unknown is P([0; 1]) valued