INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN FLUIDS
Int. J. Numer. Meth. Fluids 2008; **58**:1009–1015
Published online 13 August 2008 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/fld.1597

Anisotropic mesh adaption for time-dependent problems[‡]

S. Micheletti and S. Perotto*, †

MOX, Modeling and Scientific Computing, Dipartimento di Matematica 'F. Brioschi', Politecnico di Milano, via Bonardi 9, I-20133 Milano, Italy

SUMMARY

We propose a space–time adaptive procedure for a model parabolic problem based on a theoretically sound anisotropic *a posteriori* error analysis. A space–time finite element scheme (continuous in space but discontinuous in time) is employed to discretize this problem, thus allowing for non-matching meshes at different time levels. Copyright © 2008 John Wiley & Sons, Ltd.

Received 22 March 2007; Revised 12 July 2007; Accepted 15 July 2007

KEY WORDS: space–time adaptivity; anisotropic meshes; time-dependent problems

1. THE GOAL

We aim at providing an adaptive algorithm for an efficient approximation of a pure diffusive parabolic problem. This is pursued *via* both an adaptive choice of the temporal step and an anisotropic mesh adaption strategy, suitably combined to minimize the computational cost. The key point is the derivation of an *a posteriori* error estimator where the spatial contribution is kept distinct from the temporal one, to balance the different sources of error (see e.g. [1, 2]). In more detail, we control a proper global norm of the discretization error in compliance, e.g. with [3, 4]. An approximation scheme based on space–time finite elements, continuous in space but discontinuous in time (a so-called cGdG scheme), provides us with a natural framework where an estimator, with the space and time contributions split, can be ensued. The cGdG approach is widely used in the literature (we refer, e.g. to [4, 5]), though essentially confined to the isotropic setting. The anisotropic focus represents the first innovative contribution of this paper.

*Correspondence to: S. Perotto, MOX, Modeling and Scientific Computing, Dipartimento di Matematica 'F. Brioschi', Politecnico di Milano, via Bonardi 9, I-20133 Milano, Italy.
†E-mail: simona.perotto@polimi.it, http://www1.mate.polimi.it/∼simona
‡This article was initially submitted for publication in the special issue 'Institute for Computational Fluid Dynamics', FLD 56:8.

Contract/grant sponsor: COFIN 2006 'Numerical Approximation of Multiscale and Multiphysics Problems with Adaptive Methods'

Indeed, as far as we know, the only paper dealing with anisotropic grids in the parabolic framework is [6], in a standard finite element-backward Euler framework. In this work the author skipped the time adaption issue, simply by proposing the successive halving of the time step to make the time contribution with respect to the spatial one negligible. However, in our opinion, one of the crucial weak points of [6] consists in neglecting the unavoidable exchange of information among the grids associated with successive time intervals. In reference to these last two issues, our work is an improvement over [6].

Let us focus on the model problem at hand

$$\begin{gathered}\frac{\partial u}{\partial t}-\nabla\cdot(D\nabla u)=f \quad \text{in } \Omega\times J \\ u(\mathbf{x},t)=0 \text{ on } \Gamma_D\times J, \quad D\nabla u(\mathbf{x},t)\cdot\mathbf{n}=g \text{ on } \Gamma_N\times J, \quad u(\mathbf{x},0)=u_0(\mathbf{x}) \text{ on } \Omega\end{gathered} \tag{1}$$

where $J=(0,T)$, with $T>0$, Ω is a bounded polygonal domain in $\mathbb{R}^2$ with boundary $\partial\Omega$, Γ_D and Γ_N are measurable non-overlapping partitions of $\partial\Omega$ such that $\partial\Omega=\overline{\Gamma}_D\cup\overline{\Gamma}_N$, and $\mathbf{n}$ is the unit outward normal vector to $\partial\Omega$. Concerning the data, the source $f\in L^2(0,T;(H^1_{\Gamma_D}(\Omega))')$; the diffusivity tensor $D\in[L^\infty(\Omega)]^{2\times 2}$ and satisfies the standard ellipticity condition; the Neumann datum $g\in L^2(0,T;H^{-1/2}(\Gamma_N))$ while the initial one $u_0\in L^2(\Omega)$. Observe that the notations adopted for the function spaces are standard (see e.g. [7]). The anisotropic nature of problem (1) can be emphasized, e.g. by an anisotropic diffusivity tensor or a scalar one varying anisotropically in Ω (see Section 4).

The weak formulation of (1) reads: find $u\in U=L^2(0,T;H^1_{\Gamma_D}(\Omega))\cap H^1(0,T;(H^1_{\Gamma_D}(\Omega))')$ s.t.

$$\int_0^T\int_\Omega\left\{\frac{\partial u}{\partial t}v+D\nabla u\cdot\nabla v\right\}\mathrm{d}\mathbf{x}\,\mathrm{d}t=\int_0^T\int_\Omega f\,v\,\mathrm{d}\mathbf{x}\,\mathrm{d}t+\int_0^T\int_{\Gamma_N}g\,v\,\mathrm{d}s\,\mathrm{d}t\quad\forall v\in U \tag{2}$$

with $u(\mathbf{x},0)=u_0(\mathbf{x})$. It is known that U is continuously embedded in $C^0([0,T];L^2(\Omega))$ [7].

1.1. Managing the space–time

Among the possible cGdG schemes, we adopt the cG(1)dG(0) method, i.e. continuous piecewise affine in space and discontinuous piecewise constant in time finite elements. The choice of linear elements is essentially dictated by the anisotropic framework [8] we are interested in. The employment of piecewise constants in time aims at containing the computational cost.

Concerning the time discretization we partition the t-axis *via* the time levels $0=t_0<t_1<\cdots<t_{N-1}<t_N=T$, and set $J_n=(t_{n-1},t_n]$, $k_n=t_n-t_{n-1}$ and $S_n=\Omega\times J_n$, with $n=1,\ldots,N$. The cG(1)dG(0) approximate solution to (2) reduces, on each sub-interval J_n, to a function belonging to the space $\mathscr{S}_k=\{v:(0,T]\to H^1_{\Gamma_D}(\Omega):v(\cdot,t)|_{J_n}\equiv\widetilde{v}(\cdot)\in H^1_{\Gamma_D}(\Omega)\}$. These functions may be discontinuous at each time level, being continuous to the left. This leads us to distinguish between the two values $v_m^\pm=\lim_{t\to 0^+}v(\mathbf{x},t_m\pm t)$, and to define the temporal jump $[v]_m=v_m^+-v_m^-$, with $m=1,\ldots,N-1$. Note that the jump $[v]_m$ vanishes when $v\in U$. Moreover, since $0\notin J_1$, the value $v(\mathbf{x},0)$ has to be specified separately, $\forall v\in\mathscr{S}_k$. Finally, while $\mathscr{S}_k\not\subset U$, $\mathscr{S}_k|_{S_n}\subset U|_{S_n}$, where $\mathscr{S}_k|_{S_n}=\{v:J_n\to H^1_{\Gamma_D}(\Omega):v(\cdot,t)\equiv\widetilde{v}(\cdot)\in H^1_{\Gamma_D}(\Omega),\forall t\in J_n\}$ and $U|_{S_n}=L^2(J_n;H^1_{\Gamma_D}(\Omega))\cap H^1(J_n;(H^1_{\Gamma_D}(\Omega))')$. Thus, we can rewrite the weak form (2) as: find $u\in U$ s.t.

$$B_{\mathrm{DG}}(u,v)=\sum_{n=1}^N\int_{J_n}\int_\Omega\left\{\frac{\partial u}{\partial t}v+D\nabla u\cdot\nabla v\right\}\mathrm{d}\mathbf{x}\,\mathrm{d}t+\sum_{m=1}^{N-1}\int_\Omega[u]_m v_m^+\,\mathrm{d}\mathbf{x}+\int_\Omega u_0^+v_0^+\,\mathrm{d}\mathbf{x}$$

DOI: 10.1002/fld

$$= \sum_{n=1}^{N} \int_{J_n} \int_{\Omega} f v \, \mathrm{d}\mathbf{x} \, \mathrm{d}t + \sum_{n=1}^{N} \int_{J_n} \int_{\Gamma_N} g v \, \mathrm{d}s \, \mathrm{d}t + \int_{\Omega} u_0^- v_0^+ \, \mathrm{d}\mathbf{x} = F(v) \quad \forall v \in U \tag{3}$$

Trivially $[u]_m = 0, \forall m = 1, \ldots, N-1$, while $u_0^+ = u_0^- = u_0(\mathbf{x})$. Throughout this paper we denote by $|||w|||_{\mathrm{DG}} = (B_{\mathrm{DG}}(w, w))^{1/2}$ the norm induced by the bilinear form $B_{\mathrm{DG}}(\cdot, \cdot)$ on the space $W = U \cup \mathscr{S}_k$. We refer to [9] for details.

To discretize the space we resort to a family of conformal triangulations of $\overline{\Omega}$. The temporal discontinuity allows us to employ non-matching families $\{\mathscr{T}_{h_n}\}_{h_n}$ of meshes between successive slabs. In particular we set $\mathscr{T}_{h_n} = \{K_n\}$, with K_n triangle of diameter h_{K_n} and $h_n = \max_{K_n} h_{K_n}$, $S_{K_n} = K_n \times J_n$ and $L_{K_n} = \partial K_n \times J_n$.

Thus we define the cG(1)dG(0) space $\mathscr{S}_{hk} = \{v_{hk} \in \mathscr{S}_k : v_{hk}(\cdot, t)|_{J_n} \equiv \psi(\cdot) \in X_{h_n}^1 \cap H_{\Gamma_D}^1(\Omega)\}$, $X_{h_n}^1$ being the space of the linear finite elements associated with the mesh $\mathscr{T}_{h_n}$. The spatial continuity of the functions in $\mathscr{S}_{hk}$ is guaranteed, while the discontinuity in time of $\mathscr{S}_k$ is maintained.

Inspired by (3), we state the cG(1)dG(0) formulation of (1), looking for $u_{hk} \in \mathscr{S}_{hk}$ s.t.

$$B_{\mathrm{DG}}(u_{hk}, v_{hk}) = F(v_{hk}) \quad \forall v_{hk} \in \mathscr{S}_{hk} \tag{4}$$

with $u_h^0 \in X_{h_1}^1 \cap H_{\Gamma_D}^1(\Omega)$ a proper finite element approximation of u_0. We let the space–time discretization error associated with (4) be $e_{hk} = u - u_{hk}$. Finally, we remark that (4) can be actually referred to as a generic cG(r)dG(q) formulation by properly redefining the spaces $\mathscr{S}_k$ and $\mathscr{S}_{hk}$.

2. THE ANISOTROPIC SOURCE

We introduce the essential concepts of the anisotropic setting in [8] used for enriching the *a posteriori* analysis with an anisotropic counterpart. Let us focus on the generic slab S_n, $\mathscr{T}_{h_n}$ being the associated mesh. The source of the anisotropic information resides in the invertible affine map $T_{K_n} : \widehat{K} \to K_n$ from the unitary equilateral reference triangle $\widehat{K}$ to the general element $K_n \in \mathscr{T}_{h_n}$, identified by the relation $\mathbf{x} = (x_1, x_2)^T = T_{K_n}(\widehat{\mathbf{x}}) = M_{K_n}\widehat{\mathbf{x}} + \mathbf{t}_{K_n}$, $\forall \mathbf{x} \in K_n$, with $\widehat{\mathbf{x}} \in \widehat{K}$, and where $M_{K_n} \in \mathbb{R}^{2\times 2}$ and $\mathbf{t}_{K_n} \in \mathbb{R}^2$. In more detail, we exploit the spectral properties of the (non-singular) Jacobian M_{K_n} *via* two factorizations: the polar decomposition $M_{K_n} = B_{K_n} Z_{K_n}$ of M_{K_n} into a symmetric positive-definite matrix $B_{K_n} \in \mathbb{R}^{2\times 2}$ and an orthogonal matrix $Z_{K_n} \in \mathbb{R}^{2\times 2}$; the spectral factorization $B_{K_n} = R_{K_n}^T \Lambda_{K_n} R_{K_n}$ of B_{K_n} in terms of its eigenvectors $\mathbf{r}_{i,K_n}$ and eigenvalues λ_{i,K_n}, for $i = 1, 2$, with $\Lambda_{K_n} = \mathrm{diag}(\lambda_{1,K_n}, \lambda_{2,K_n})$ and $R_{K_n}^T = [\mathbf{r}_{1,K_n}, \mathbf{r}_{2,K_n}]$. The shape and the orientation of each element K_n are completely described through the quantities $\mathbf{r}_{i,K_n}$ and λ_{i,K_n}. Indeed, the unit circle circumscribed to $\widehat{K}$ is mapped into an ellipse circumscribing K_n: the eigenvectors $\mathbf{r}_{i,K_n}$ and the eigenvalues λ_{i,K_n} provide us with the directions and the lengths of the semi-axes of such an ellipse, respectively. In particular, we introduce the so-called stretching factor $s_{K_n} = \lambda_{1,K_n}/\lambda_{2,K_n}$ to quantify the deformation of the element K_n, with the agreement $s_{K_n} \geqslant 1$, $s_{\widehat{K}}$ being identically equal to 1.

3. THE WAY

We provide the theoretical tool basis for the proposed space–time adaption procedure. It consists of a real anisotropic *a posteriori* error estimator for the DG-norm $|||e_{hk}|||_{\mathrm{DG}}$, namely not driven by

DOI: 10.1002/fld

heuristic considerations but rather by a founded anisotropic theory. In particular, at this stage, we are essentially interested in the reliability of this error estimator, i.e. in an upper bound for e_{hk} *via* the estimator itself.

Let us anticipate some useful notation. We define the local residuals by distinguishing between the spatial and the temporal ones. For a fixed time interval J_n, let

$$\rho_{K_n}=\left[f-\frac{\partial u_{hk}}{\partial t}+\nabla\cdot(D\nabla u_{hk})\right]\Bigg|_{S_{K_n}},\quad j_{K_n}=\begin{cases}0 & \text{on } (\partial K_n\cap\Gamma_D)\times J_n\\ 2(g-D\nabla u_{hk}\cdot\mathbf{n})|_{S_{K_n}} & \text{on } (\partial K_n\cap\Gamma_N)\times J_n\\ -[D\nabla u_{hk}\cdot\mathbf{n}] & \text{on } (\partial K_n\cap\mathscr{E}_h^n)\times J_n\end{cases}$$

be the interior and the boundary residual, respectively, associated with u_{hk} and with the prism S_{K_n}, $\forall K_n\in\mathscr{T}_{h_n}$, for $n=1,\ldots,N$, where $\mathscr{E}_h^n$ is the set of the internal edges of $\mathscr{T}_{h_n}$ while $[D\nabla u_{hk}\cdot\mathbf{n}]=D\nabla u_{hk}\cdot\mathbf{n}_{K_n}+D\nabla u_{hk}\cdot\mathbf{n}_{K_n'}$ denotes the jump of the diffusive flux across the internal interfaces of K_n, for $K_n'\cap K_n\neq\emptyset$.

Then we introduce the temporal residual $\mathscr{J}_n=[-u_{hk}]_n$ associated with u_{hk} and with the time level t_n, together with the initial residual $e_0^-=u_0-u_h^0$. While ρ_{K_n} and j_{K_n} are related to the space discretization, the quantities $\mathscr{J}_n$ and e_0^- are due to the discontinuity of the temporal scheme. Moreover, the residual $\mathscr{J}_n$ has to be carefully computed by merging the information coming from the two different meshes, $\mathscr{T}_{h_n}$ and $\mathscr{T}_{h_{n+1}}$ (see [9] for details).

Thus, the *a posteriori* result can be stated as follows:

Proposition 1

Let $u\in U$ be the solution to (2) and let $u_{hk}\in\mathscr{S}_{hk}$ be the corresponding cG(1)dG(0)-approximation, solution to (4). Let Δ_{K_n} be the patch of the elements sharing at least a vertex with K_n. Then, if $\#(\Delta_{K_n})\leqslant N_\Delta$ and $\operatorname{diam}(\Delta_{\widehat{K}})\leqslant C_\Delta\simeq O(1)$, with $C_\Delta\geqslant h_{\widehat{K}}$ and $\Delta_{\widehat{K}}=T_{K_n}^{-1}(\Delta_{K_n})$ the reference patch, there exists a constant $C=C(N_\Delta,C_\Delta)$ such that

$$|||e_{hk}|||_{\mathrm{DG}}^2\lesssim\eta^2=C\sum_{n=1}^{N}\sum_{K_n\in\mathscr{T}_{h_n}}\left(\underbrace{\alpha_{K_n}^S R_{K_n}^S\omega_{K_n}^S}_{\eta_{K_n}^S}+\underbrace{\alpha_{K_n}^{T1}R_{K_n}^{T1}\omega_{K_n}^{T1}+\alpha_{K_n}^{T2}R_{K_n}^{T2}\omega_{K_n}^{T2}}_{\eta_{K_n}^T}\right)\tag{5}$$

where $\alpha_{K_n}^S=|\widehat{K}|\lambda_{1,K_n}^{3/2}\lambda_{2,K_n}^{3/2}$, $\alpha_{K_n}^{T1}=\alpha_{K_n}^{T2}=k_n^2$,

$$R_{K_n}^S=\frac{1}{|K_n|^{1/2}}\left\{\frac{h_{K_n}^{1/2}\|\overline{j}_{K_n}\|_{L^2(L_{K_n})}}{2\lambda_{1,K_n}^{1/2}\lambda_{2,K_n}^{1/2}}+\left[\|\overline{\rho}_{K_n}\|_{L^2(S_{K_n})}+\frac{1}{k_n^{1/2}}(\|\mathscr{J}_{n-1}\|_{L^2(K_n)}+\delta_{1n}\|e_0^-\|_{L^2(K_n)})\right]\right\}$$

$$R_{K_n}^{T1}=k_n^{-1/2}[\|\rho_{K_n}-\overline{\rho}_{K_n}\|_{L^2(S_{K_n})}+k_n^{-1/2}(\|\mathscr{J}_{n-1}\|_{L^2(K_n)}+\delta_{1n}\|e_0^-\|_{L^2(K_n)})]$$

$$R_{K_n}^{T2}=(4k_n)^{-1/2}\|j_{K_n}-\overline{j}_{K_n}\|_{L^2(L_{K_n})},\quad \omega_{K_n}^{T1}=k_n^{-1/2}\left\|\frac{\partial e_{hk}^*}{\partial t}\right\|_{L^2(S_{K_n})}$$

DOI: 10.1002/fld

$$\omega_{K_n}^{T2} = k_n^{-1/2} \left\| \frac{\partial e_{hk}^*}{\partial t} \right\|_{L^2(L_{K_n})}$$

$$\omega_{K_n}^{S} = [s_{K_n}(\mathbf{r}_{1,K_n}^T \widetilde{G}_{K_n}^n(e_{hk}^*)\mathbf{r}_{1,K_n}) + s_{K_n}^{-1}(\mathbf{r}_{2,K_n}^T \widetilde{G}_{K_n}^n(e_{hk}^*)\mathbf{r}_{2,K_n})]^{1/2}$$

the standard notation $\|\cdot\|_{L^2(\cdot)}$ being adopted for the L^2-norm, $\overline{\rho}_{K_n} = k_n^{-1} \int_{J_n} \rho_{K_n}(\mathbf{x}, t)\, \mathrm{d}t$, $\overline{j}_{K_n} = k_n^{-1} \int_{J_n} j_{K_n}(\mathbf{x}, t)\, \mathrm{d}t$, δ_{1n} is the Kronecker symbol, while $\widetilde{G}_{K_n}^n(e_{hk}^*) = |K_n|^{-1} G_{K_n}^n(e_{hk}^*)$ is the scaled counterpart of the symmetric positive semi-definite matrix $(G_{K_n}^n(e_{hk}))_{ij} = \int_{\Delta_{K_n} \times J_n} (\partial e_{hk}/\partial x_i)(\partial e_{hk}/\partial x_j)\, \mathrm{d}\mathbf{x}\, \mathrm{d}t$, for $i, j = 1, 2$, where the spatial derivatives of the (unknown) discretization error are computed via a Zienkiewicz–Zhu gradient recovery procedure combined with a proper linear reconstruction of the time dependence.

Proof of result (5) is provided in [9]. Notice that the quantities $R_{K_n}^S$ and $R_{K_n}^{Ti}$ are scaled (with respect to the space and time, respectively), so that all the spatial and temporal dimensional information is collected into the coefficients $\alpha_{K_n}^S$ and $\alpha_{K_n}^{Ti}$. The anisotropic information is lumped into the weights $\omega_{K_n}^S$, while the $\omega_{K_n}^{Ti}$'s drive the temporal adaption. In more detail, both these weights will be involved in a suitable optimization procedure leading to the predicted mesh spacing and time step [9].

4. THE CROSSCHECK

We pursue a slab-wise error equidistribution criterion. We aim at guaranteeing on each slab S_n the same global tolerance τ. For this purpose we split τ into a space (τ^S) and a time (τ^T) contribution. The time step and the spatial mesh are successively adapted till both the time and space estimators $\eta_n^T = \sum_{K_n \in \mathcal{T}_{h_n}} \eta_{K_n}^T$ and $\eta_n^S = \sum_{K_n \in \mathcal{T}_{h_n}} \eta_{K_n}^S$ are within the corresponding tolerances. We sketch the overall procedure in Figure 1. Notice that, at the end, if the time tolerance is widely satisfied (focus on the parameter θ in the flowchart), a new (larger) time step is guessed for the next slab.

We consider the groundwater flow equation describing the flow of groundwater through an aquifer [10]. The transient equation has the form (1) where u is the hydraulic head, D represents the hydraulic conductivity scaled by the specific storativity S_0, the scaled specific discharge $\mathbf{q} = -D\nabla u$ is given by the Darcy law, while the velocity of the water is $\mathbf{q}/p$, where p is the volume porosity. We consider (1) on the domain $\Omega = (0, 100)^2$(m), modeling a vertical cross-section of a pervious aquifer interlaced with the two impervious layers, $\Omega_1 = (0, 80) \times (56, 60)$ and $\Omega_2 = (20, 100) \times (40, 44)$. Moreover we assume $T = 10^5$(s), $f = u_0 = 0$, $S_0 = 0.02\,(\mathrm{m}^{-1})$, $p = 0.01$, while D takes on the value $1.16 \times 10^{-3}(\mathrm{m\,s}^{-1})$ in the pervious zone and $1.16 \times 10^{-10}\,(\mathrm{m\,s}^{-1})$ in the barriers $\Omega_1 \cup \Omega_2$. We modify (1) to support non-homogeneous Dirichlet conditions. In particular, we choose $u = 1 - \exp(-t/1000)$ (m) on the top horizontal side and $u = 0$ on the bottom one, while a homogeneous Neumann condition holds elsewhere. The tolerances for the adaptive algorithm are $\tau^S = \tau^T = 10$. This is a multiscale problem in both space and time. There exist three main phases: a very fast transient where the velocity field sets up near the top boundary; an intermediate period when the flow curves around Ω_1 and is pushed inside the channel between the two impervious layers; and the final phase when the velocity becomes steady inside the channel (where most of the kinetic energy gathers up) while some water flows out of the bottom boundary. We show in Figure 2 two

DOI: 10.1002/fld

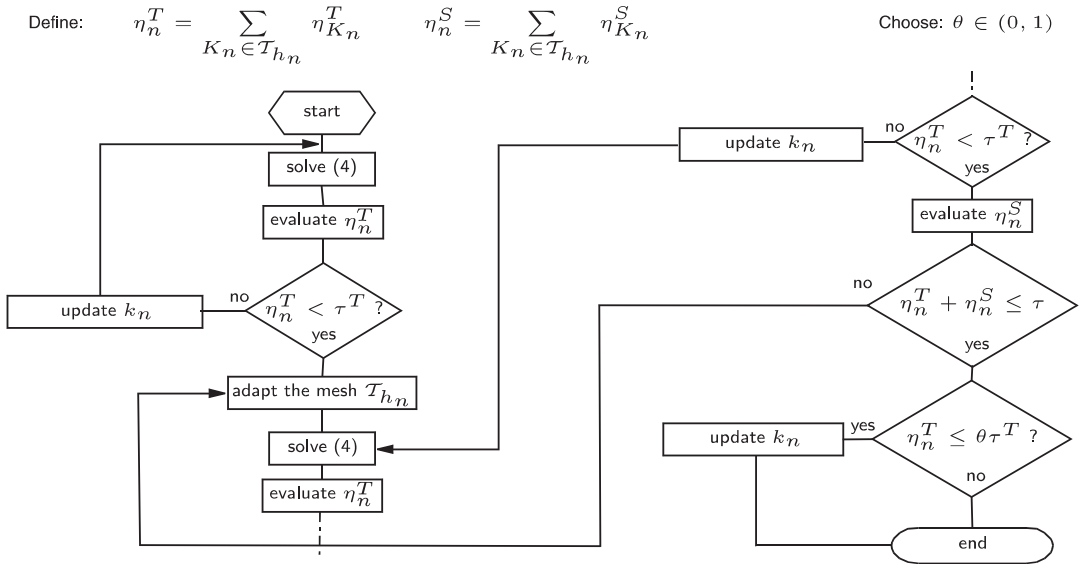

Figure 1. Flowchart of the space–time adaptive algorithm for each slab S_n.

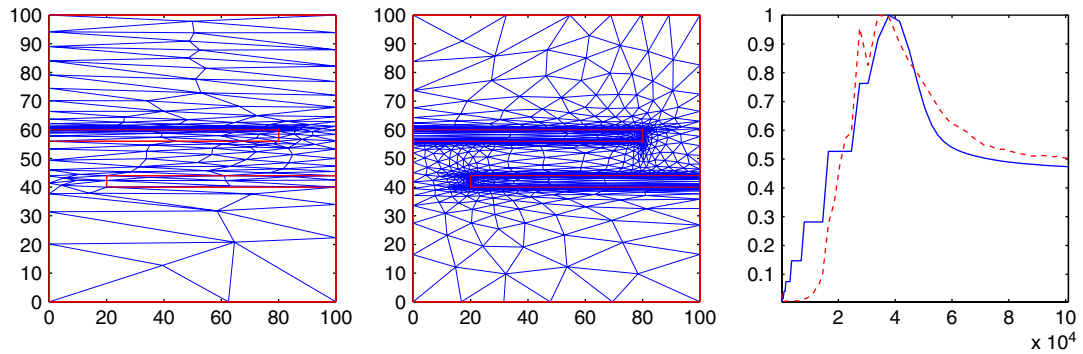

Figure 2. Adapted grids: $t \simeq 10^4$ (left), $t \simeq T$ (center); time histories of k_n and $\#(\mathcal{T}_{h_n})$ (right).

adapted grids, on the left when the fluid starts entering the channel, and in the middle at the final time. In both cases the meshes are extremely anisotropic, especially around Ω_1 and Ω_2. Figure 2 (right) displays the time histories of k_n (solid line) and of the number of mesh elements (dashed line), both scaled to their maximum value, 3800 (s) and 14 700, respectively. In particular, at the very initial stage, the time step is really small (about 30–60 (s)), while the meshes are very coarse (only 60–80 elements). Both curves reach their maximum when the water leaves the channel, and eventually decrease smoothly.

For a more exhaustive numerical validation, we refer to [9]. We anticipate that the estimator is always reliable and the effectivity index is independent of s_{K_n} on the adapted grids.

DOI: 10.1002/fld

REFERENCES

1. Bergam A, Bernardi C, Mghazli Z. A posteriori analysis of the finite element discretization of some parabolic equations. *Mathematics of Computation* 2004; **74**(251):1117–1138.
2. Verfürth R. A posteriori error estimates for nonlinear problems. $L^r(0, T; L^\rho(\Omega))$-error estimates for finite element discretizations of parabolic equations. *Mathematics of Computation* 1998; **67**(224):1335–1360.
3. Picasso M. Adaptive finite elements for a linear parabolic problem. *Computer Methods in Applied Mechanics and Engineering* 1998; **167**:223–237.
4. Eriksson K, Johnson C. Adaptive finite element methods for parabolic problems. I: a linear model problem. *SIAM Journal on Numerical Analysis* 1991; **28**:43–77.
5. Eriksson K, Johnson C, Thomée V. Time discretization of parabolic problems by the discontinuous Galerkin method. *RAIRO Modelisation Mathématique et Analyse Numérique* 1985; **19**:611–643.
6. Picasso M. An anisotropic error indicator based on Zienkiewicz–Zhu error estimator: application to elliptic and parabolic problems. *SIAM Journal on Scientific Computing* 2003; **24**(4):1328–1355.
7. Dautray R, Lions J-L. *Mathematical Analysis and Numerical Methods for Science and Technology*: *Evolution Problems I*, vol. 5. Springer: Berlin, 1992.
8. Formaggia L, Perotto S. New anisotropic a priori error estimates. *Numerische Mathematik* 2001; **89**:641–667.
9. Micheletti S, Perotto S. Space–time adaption for parabolic problems in an anisotropic framework, 2008, in preparation.
10. Wang HF, Anderson MP. *Introduction to Groundwater Modeling*: *Finite Difference and Finite Element Methods*. Academic Press: San Diego, 1995.

DOI: 10.1002/fld