,

Weierstraß-Institut
für Angewandte Analysis und Stochastik
Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 2198-5855

Existence of solutions of a finite element flux-corrected-transport scheme

Volker John[1], Petr Knobloch[2]

submitted: September 21, 2020

[1] Weierstrass Institute
Mohrenstr. 39
10117 Berlin, Germany
and
Freie Universität Berlin
Department of Mathematics and Computer Science
Arnimallee 6
14195 Berlin, Germany
E-Mail: john@wias-berlin.de

[2] Department of Numerical Mathematics
Faculty of Mathematics and Physics
Charles University
Sokolovská 83
18675 Praha 8
Czech Republic
E-Mail: knobloch@karlin.mff.cuni.cz

No. 2761
Berlin 2020

2010 *Mathematics Subject Classification.* 65M60.

Key words and phrases. evolutionary convection-diffusion-reaction equations, transport equations, finite element method, nonlinear flux-corrected-transport scheme, existence of a solution, Brouwer's fixed-point theorem.

The work of P. Knobloch has been supported through the grant No. 19-04243S of the Czech Science Foundation.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Existence of solutions of a finite element flux-corrected-transport scheme

Volker John, Petr Knobloch

Abstract

The existence of a solution is proved for a nonlinear finite element flux-corrected-transport (FEM-FCT) scheme with arbitrary time steps for evolutionary convection-diffusion-reaction equations and transport equations.

1 Introduction

This note considers the transient convection–diffusion–reaction equation

$$\begin{aligned} u_t - \varepsilon\,\Delta u + \boldsymbol{b}\cdot\nabla u + c\,u &= f \qquad \text{in } (0,T]\times\Omega\,, && (1)\\ u &= u_b \qquad \text{on } [0,T]\times\Gamma_D\,, && (2)\\ \varepsilon\,\frac{\partial u}{\partial \boldsymbol{n}} &= g \qquad \text{on } [0,T]\times\Gamma_N\,, && (3)\\ u(0,\cdot) &= u_0 \qquad \text{in } \Omega, && (4)\end{aligned}$$

where $\Omega\subset\mathbb{R}^d$, $d=2,3$, is a bounded polygonal or polyhedral domain with a Lipschitz-continuous boundary $\partial\Omega$ that is composed of disjoint subsets Γ_D and Γ_N, $\boldsymbol{n}$ is the outer unit normal vector to $\partial\Omega$, $[0,T]$ is a time interval, $\varepsilon>0$ is a constant diffusivity, $\boldsymbol{b}:[0,T]\to W^{1,\infty}(\Omega)^d$ is a convection field, $c:[0,T]\to L^\infty(\Omega)$ is a reaction coefficient, $f:[0,T]\to L^2(\Omega)$ is an outer source of the unknown quantity u, $u_b:[0,T]\to H^{1/2}(\Gamma_D)$ and $g:[0,T]\to L^2(\Gamma_N)$ are the boundary conditions, and $u_0\in H^1_0(\Omega)$ is the initial condition. Without loss of generality, it can be assumed that

$$c-\frac{1}{2}\,\mathrm{div}\,\boldsymbol{b}\geq 0 \qquad \text{in } [0,T]\times\Omega, \tag{5}$$

which can be always achieved by a transform of variables $u\mapsto u\exp(-\kappa t)$ with $\kappa>0$ sufficiently large. In addition, it is assumed that

$$\Gamma_D\supset\partial\Omega^- := \{x\in\partial\Omega\,;\ \boldsymbol{b}(x)\cdot\boldsymbol{n}(x)<0\}\,. \tag{6}$$

The analysis of this paper also covers the case $\varepsilon=0$. Since it is a first order partial differential equation, a boundary condition is prescribed only on $\partial\Omega^-$. We shall again consider a

DOI 10.20347/WIAS.PREPRINT.2761 Berlin 2020

Dirichlet boundary condition so that the initial–boundary value problem consists of (1), (2), (4) with $\varepsilon = 0$ and $\Gamma_D = \partial\Omega^-$. An important particular case for $\varepsilon = 0$ is the transport equation, i.e., also c vanishes identically.

The considered classes of problems obey, under certain conditions on their data, maximum principles. Often, these conditions are satisfied in applications. For a numerical method to be physically consistent and to be accepted by practitioners, it is of importance that it satisfies a discrete maximum principle (DMP). There are only very few finite element methods that possess this property and are not excessively diffusive, among them FEM-FCT (flux-corrected-transport) schemes, e.g., proposed in [1, 2, 3]. In particular, a nonlinear FEM-FCT scheme has been proven to compute very accurate solutions, see [4]. However, concerning the solvability of the nonlinear problem, there is only one very recent result in [5]. It shows the existence and local uniqueness of a solution for sufficiently small time steps. The analysis in [5] is based on the implicit function theorem for Lipschitz functions and utilizes tools from non-smooth optimization. In the present note, a new result will be shown: the existence of a solution for arbitrary time steps. The proof is based on a consequence of Brouwer's fixed-point theorem.

2 FEM-FCT schemes

For the discretization of (1)–(4), the time interval is decomposed by $0 = t_0 < t_1 < \cdots < t_K = T$ with $\Delta t_k = t_k - t_{k-1}$. We consider conforming finite element spaces, where it is assumed that the basis functions $\varphi_1, \ldots, \varphi_N$ are nonnegative, as it is the case for standard piecewise linear or multilinear basis functions or for bases constructed using Bernstein polynomials. Let the basis functions be numbered such that $\varphi_1, \ldots, \varphi_M$, $M \leq N$, are associated with degrees of freedom that are not on the Dirichlet boundary so that they vanish on Γ_D.

Using a one-step θ-scheme and the usual approach for deriving a Galerkin finite element discretization leads for the time instant t_k to a discrete problem of the form

$$\mathbb{M}\,\frac{\mathrm{U}^k - \mathrm{U}^{k-1}}{\Delta t_k} + \theta\,\mathbb{A}^k\,\mathrm{U}^k + (1-\theta)\,\mathbb{A}^{k-1}\,\mathrm{U}^{k-1} = \theta\,\mathrm{F}^k + (1-\theta)\,\mathrm{F}^{k-1}\,, \tag{7}$$

$$u_i^k = u_i^b(t_k)\,, \qquad i = M+1, \ldots, N\,, \tag{8}$$

with $\mathrm{U}^0 = \mathrm{U}_0$ and $\theta \in [0,1]$. In (7), $\mathrm{U}^k = (u_1^k, \ldots, u_N^k)^T$ denotes the vector of unknowns at t_k, U_0 and $u_i^b(t_k)$ are the coefficients of finite element representations of the initial condition and the boundary condition at t^k, respectively. Further, $\mathbb{M} = (m_{ij})_{j=1,\ldots,N}^{i=1,\ldots,M}$ is the mass matrix, $\mathbb{A}^k = (a_{ij}(t_k))_{j=1,\ldots,N}^{i=1,\ldots,M}$ the stiffness matrix, and $\mathrm{F}^k = (f_1(t_k), \ldots, f_M(t_k))^T$ the right-hand side vector defined by

$$m_{ij} = (\varphi_j, \varphi_i)_\Omega\,, \quad a_{ij}(t) = a(t)(\varphi_j, \varphi_i)\,, \quad f_i(t) = (f(t), \varphi_i)_\Omega + (g(t), \varphi_i)_{\Gamma_N}\,.$$

Here, $(\cdot,\cdot)_\Omega$ denotes the inner product in $L^2(\Omega)$ or $L^2(\Omega)^d$, $(\cdot,\cdot)_{\Gamma_N}$ is the inner product in $L^2(\Gamma_N)$, and

$$a(t)(u,v) = \varepsilon\,(\nabla u, \nabla v)_\Omega + (\boldsymbol{b}(t)\cdot\nabla u, v)_\Omega + (c(t)\,u, v)_\Omega\,.$$

It is well known that, if convection dominates diffusion, a stabilization has to be introduced, e.g., see [6]. One possibility is to apply a FCT approach, e.g., see [1, 2, 3]. To this end, one extends the matrices $\mathbb{A}^k$ to $(a_{ij}^k)_{i,j=1,\dots,N}$ by setting $a_{ij}^k = a(t_k)(\varphi_j,\varphi_i)$, $i,j=1,\dots,N$. Then, one introduces artificial diffusion matrices $\mathbb{D}^k = (d_{ij}^k)_{j=1,\dots,N}^{i=1,\dots,M}$ possessing the entries $d_{ij}^k = -\max\{a_{ij}^k, 0, a_{ji}^k\}$ for all $i\neq j$ and $d_{ii}^k = -\sum_{j\neq i} d_{ij}^k$. In addition, one defines the lumped mass matrix $\mathbb{M}_L = (m_{ij}^L)_{j=1,\dots,N}^{i=1,\dots,M}$ with the entries $m_{ij}^L = 0$ for all $i\neq j$ and $m_{ii}^L = \sum_{j=1}^N m_{ij}$. Denoting $\mathbb{L}^k := \mathbb{A}^k + \mathbb{D}^k$, (7) can be written in the form

$$\mathbb{M}_L\,\frac{\mathrm{U}^k - \mathrm{U}^{k-1}}{\Delta t_k} + \theta\,\mathbb{L}^k\,\mathrm{U}^k + (1-\theta)\,\mathbb{L}^{k-1}\,\mathrm{U}^{k-1} = \theta\,\mathrm{F}^k + (1-\theta)\,\mathrm{F}^{k-1} + \mathrm{R}^k(\mathrm{U}^k,\mathrm{U}^{k-1})$$

with

$$\mathrm{R}^k(\mathrm{U}^k,\mathrm{U}^{k-1}) = -(\mathbb{M}-\mathbb{M}_L)\,\frac{\mathrm{U}^k - \mathrm{U}^{k-1}}{\Delta t_k} + \theta\,\mathbb{D}^k\,\mathrm{U}^k + (1-\theta)\,\mathbb{D}^{k-1}\,\mathrm{U}^{k-1}\,.$$

Note that $\mathbb{L}^k$ has non-positive off-diagonal entries. The matrix $\mathbb{D}^k$ has zero row sums and hence $(\mathbb{D}^k\,\mathrm{U})_i = \sum_{j=1}^N d_{ij}^k\,(u_j - u_i)$, $i=1,\dots,M$, for any $\mathrm{U} = (u_1,\dots,u_N)^T$. Since also the matrix $\mathbb{M}-\mathbb{M}_L$ has zero row sums, one deduces that

$$(\mathrm{R}^k(\mathrm{U}^k,\mathrm{U}^{k-1}))_i = \sum_{j=1}^N r_{ij}^k\,,\qquad i=1,\dots,M\,,$$

with so-called algebraic fluxes

$$\begin{aligned} r_{ij}^k = &-\frac{1}{\Delta t_k}\,m_{ij}\,(u_j^k - u_i^k) + \frac{1}{\Delta t_k}\,m_{ij}\,(u_j^{k-1} - u_i^{k-1}) \\ &+ \theta\,d_{ij}^k\,(u_j^k - u_i^k) + (1-\theta)\,d_{ij}^{k-1}\,(u_j^{k-1} - u_i^{k-1})\,. \end{aligned}$$

Now the idea of flux correction is to limit those fluxes r_{ij}^k that would cause spurious oscillations. To this end, $(\mathrm{R}^k(\mathrm{U}^k,\mathrm{U}^{k-1}))_i$ is replaced by

$$(\tilde{\mathrm{R}}^k(\mathrm{U}^k,\mathrm{U}^{k-1}))_i = \sum_{j=1}^N \alpha_{ij}^k\,r_{ij}^k,\quad \alpha_{ij}^k\in[0,1],\quad \alpha_{ij}^k = \alpha_{ji}^k\,,\quad i,j=1,\dots,N\,, \tag{9}$$

where the limiters α_{ij}^k depend on the solution. Then, the discrete solution at the time instant

t_k satisfies the following system of (nonlinear) algebraic equations

$$\begin{aligned}
&\sum_{j=1}^{N} m_{ij}\,(u_j^k - u_j^{k-1}) + \Delta t_k\,\theta \sum_{j=1}^{N} a_{ij}^k\, u_j^k + \Delta t_k\,(1-\theta) \sum_{j=1}^{N} a_{ij}^{k-1}\, u_j^{k-1} \\
&- \sum_{j=1}^{N} (1-\alpha_{ij}^k)\, m_{ij}\,(u_j^k - u_i^k) + \sum_{j=1}^{N} (1-\alpha_{ij}^k)\, m_{ij}\,(u_j^{k-1} - u_i^{k-1}) \\
&+ \Delta t_k\,\theta \sum_{j=1}^{N} (1-\alpha_{ij}^k)\, d_{ij}^k\,(u_j^k - u_i^k) + \Delta t_k\,(1-\theta) \sum_{j=1}^{N} (1-\alpha_{ij}^k)\, d_{ij}^{k-1}\,(u_j^{k-1} - u_i^{k-1}) \\
&= \Delta t_k\,\theta\, f_i^k + \Delta t_k\,(1-\theta)\, f_i^{k-1}\,, \qquad i = 1,\dots,M\,, \qquad (10)
\end{aligned}$$

$$u_i^k = u_i^b(t_k)\,, \qquad i = M+1,\dots,N\,, \qquad (11)$$

where $f_i^k = f_i(t_k)$.

3 Solvability of the nonlinear FEM-FCT scheme

For proving the solvability of the nonlinear problem, we shall use a consequence of Brouwer's fixed-point theorem, see [7, p. 164, Lemma 1.4].

Lemma 1 *Let X be a finite-dimensional Hilbert space with inner product $(\cdot,\cdot)_X$ and norm $\|\cdot\|_X$. Let $\Pi : X \to X$ be a continuous mapping and $B > 0$ a real number such that $(\Pi x, x)_X > 0$ for any $x \in X$ with $\|x\|_X = B$. Then there exists $x \in X$ such that $\|x\|_X < B$ and $\Pi x = 0$.*

Theorem 2 *For any $i, j \in \{1,\dots,N\}$, let $\alpha_{ij}^k : \mathbb{R}^N \to [0,1]$ be such that $\alpha_{ij}^k\, r_{ij}^k$ is a continuous function of $u_1^k,\dots,u_N^k$. Let* (5) *and* (6) *be satisfied and let the functions α_{ij}^k satisfy* (9)*. Then there exists a solution of the nonlinear problem* (10)–(11)*.*

Proof. For a vector $\mathrm{U} = (u_1,\dots,u_N)^T$, we set $\tilde{\mathrm{U}} = (u_1,\dots,u_M)^T$. On the other hand, for a vector $\tilde{\mathrm{U}} = (u_1,\dots,u_M)^T$, we set $\mathrm{U} = (u_1,\dots,u_M,u_{M+1}^b(t_k),\dots,u_N^b(t_k))^T$. With this notation, we define an operator $\Pi : \mathbb{R}^M \to \mathbb{R}^M$ by

$$\begin{aligned}
(\Pi\,\tilde{\mathrm{U}})_i &= \sum_{j=1}^{M} m_{ij}\, u_j + \Delta t_k\,\theta \sum_{j=1}^{M} a_{ij}^k\, u_j \\
&\quad + \sum_{j=1}^{N} (1-\alpha_{ij}^k(\mathrm{U}))\,[\Delta t_k\,\theta\, d_{ij}^k - m_{ij}]\,(u_j - u_i) + g_i(\mathrm{U})\,, \qquad i = 1,\dots,M\,,
\end{aligned}$$

where for $i = 1, \ldots, M$

$$
\begin{aligned}
g_i(\mathrm{U}) = & \sum_{j=M+1}^{N} m_{ij}\, u_j^b(t_k) + \Delta t_k\, \theta \sum_{j=M+1}^{N} a_{ij}^k\, u_j^b(t_k) - \sum_{j=1}^{N} m_{ij}\, u_j^{k-1} \\
& + \Delta t_k\, (1-\theta) \sum_{j=1}^{N} a_{ij}^{k-1}\, u_j^{k-1} + \sum_{j=1}^{N} (1 - \alpha_{ij}^k(\mathrm{U}))\, m_{ij}\, (u_j^{k-1} - u_i^{k-1}) \\
& + \Delta t_k\, (1-\theta) \sum_{j=1}^{N} (1 - \alpha_{ij}^k(\mathrm{U}))\, d_{ij}^{k-1}\, (u_j^{k-1} - u_i^{k-1}) - \Delta t_k\, \theta\, f_i^k - \Delta t_k\, (1-\theta)\, f_i^{k-1}\,.
\end{aligned}
$$

Then, $\mathrm{U}^k \in \mathbb{R}^N$ solves the algebraic problem (10)–(11) if an only if $\Pi\, \tilde{\mathrm{U}}^k = 0$ and $u_i^k = u_i^b(t_k)$ for $i = M+1, \ldots, N$. Thus, it suffices to show that the operator Π satisfies the assumptions of Lemma 1.

Let $(\cdot,\cdot)$ denote the Euclidean inner product in $\mathbb{R}^M$ and $\|\cdot\|$ the corresponding norm. Then, for any $\tilde{\mathrm{U}} \in \mathbb{R}^M$, one has

$$
\begin{aligned}
(\Pi\, \tilde{\mathrm{U}}, \tilde{\mathrm{U}}) = & \sum_{i,j=1}^{M} u_i\, m_{ij}\, u_j + \Delta t_k\, \theta \sum_{i,j=1}^{M} u_i\, a_{ij}^k\, u_j \\
& + \sum_{i,j=1}^{N} (1 - \alpha_{ij}^k(\mathrm{U}))\, [\Delta t_k\, \theta\, d_{ij}^k - m_{ij}]\, u_i\, (u_j - u_i) \\
& - \sum_{i=M+1}^{N} \sum_{j=1}^{N} (1 - \alpha_{ij}^k(\mathrm{U}))\, [\Delta t_k\, \theta\, d_{ij}^k - m_{ij}]\, u_i^b(t_k)\, (u_j - u_i) + \sum_{i=1}^{M} u_i\, g_i(\mathrm{U})\,,
\end{aligned}
$$

where we extended the matrix $\mathbb{M}$ to a symmetric $N \times N$ matrix. In view of the symmetry of $\mathbb{M}$, $\mathbb{D}^k$, and the limiters, one has

$$
\begin{aligned}
& \sum_{i,j=1}^{N} (1 - \alpha_{ij}^k(\mathrm{U}))\, [\Delta t_k\, \theta\, d_{ij}^k - m_{ij}]\, u_i\, (u_j - u_i) \\
& \qquad = -\frac{1}{2} \sum_{i,j=1}^{N} (1 - \alpha_{ij}^k(\mathrm{U}))\, [\Delta t_k\, \theta\, d_{ij}^k - m_{ij}]\, (u_j - u_i)^2 \ge 0\,,
\end{aligned}
$$

where we used that $m_{ij} \ge 0$, $d_{ij}^k \le 0$ for $i \neq j$, and $\alpha_{ij}^k \in [0,1]$. The last property also implies that the values $g_i(\mathrm{U})$ are bounded independently of U. Consequently,

$$
(\Pi\, \tilde{\mathrm{U}}, \tilde{\mathrm{U}}) \ge \sum_{i,j=1}^{M} u_i\, m_{ij}\, u_j + \Delta t_k\, \theta \sum_{i,j=1}^{M} u_i\, a_{ij}^k\, u_j - C_1\, \|\tilde{\mathrm{U}}\| - C_2
$$

with some $C_1, C_2 > 0$. Obviously, the matrix $(m_{ij})_{i,j=1,\ldots,M}$ is positive definite. Moreover, in view of (5) and (6), one has $a(t_k)(v,v) \ge \varepsilon\, |v|_{1,\Omega}^2$ for any $v \in H^1(\Omega)$ with $v = 0$ on Γ_D so that the matrix $(a_{ij}^k)_{i,j=1,\ldots,M}$ is positive semi-definite. This gives $(\Pi\, \tilde{\mathrm{U}}, \tilde{\mathrm{U}}) \ge C_3\, \|\tilde{\mathrm{U}}\|^2 - C_4$ with some C_3, $C_4 > 0$, which implies that $(\Pi\, \tilde{\mathrm{U}}, \tilde{\mathrm{U}}) > 0$ if $\|\tilde{\mathrm{U}}\| \ge \sqrt{2\, C_4/C_3}$. Since Π is continuous, the statement of the theorem follows from Lemma 1. $\square$

Remark 3 *If the problem* (10)–(11) *is defined using data that do not satisfy some of the assumptions* (5) *and* (6)*, then the term* $\Delta t_k\, \theta\, \sum_{i,j=1}^{M} u_i\, a_{ij}^k\, u_j$ *in* $(\Pi\, \tilde{\mathrm{U}}, \tilde{\mathrm{U}})$ *may be negative and has to be estimated from below by* $-C_5\, \Delta t_k\, \|\tilde{\mathrm{U}}\|^2$*. This allows to prove the solvability only for a sufficiently small time step* Δt_k*.*

4 Example of limiters α_{ij}^k satisfying the assumptions of Theorem 2

Let $\mathscr{T}_h$ be a simplicial triangulation of Ω possessing the usual compatibility properties and let the considered finite element space consist of continuous piecewise linear functions with respect to $\mathscr{T}_h$. Then, the basis functions $\varphi_1, \dots, \varphi_N$ are assigned to vertices $x_1, \dots, x_N$ of $\mathscr{T}_h$ and satisfy $\varphi_i(x_j) = \delta_{ij}$, $i, j = 1, \dots, N$.

We shall present a limiting strategy described in [2] which is motivated by [8] and utilizes an explicit solution $\hat{\mathrm{U}}^k$ of the low order scheme

$$\mathbb{M}_L \hat{\mathrm{U}}^k = \left(\mathbb{M}_L - (1-\theta)\, \Delta t_k \mathbb{L}^{k-1}\right) \mathrm{U}^{k-1} + (1-\theta)\, \Delta t_k\, \mathrm{F}^{k-1}\,.$$

To assure that $\hat{\mathrm{U}}^k$ satisfies the DMP, if the continuous solution satisfies a weak maximum principle, the time step has to obey a CFL-like condition. Then, for $i = 1, \dots, N$, one defines the local quantities

$$P_i^+ := \sum_{j\in S(i)} (r_{ij}^k)^+, \qquad P_i^- := \sum_{j\in S(i)} (r_{ij}^k)^-,$$

$$Q_i^+ := \max_{j\in S(i)\cup\{i\}} \hat{u}_j^k - \hat{u}_i^k, \qquad Q_i^- := \min_{j\in S(i)\cup\{i\}} \hat{u}_j^k - \hat{u}_i^k,$$

$$R_i^+ := \begin{cases} \min\left(1, \frac{m_{ii}^L\, Q_i^+}{\Delta t_k\, P_i^+}\right) & \text{if } P_i^+ > 0, \\ 1 & \text{if } P_i^+ = 0, \end{cases} \qquad R_i^- := \begin{cases} \min\left(1, \frac{m_{ii}^L\, Q_i^-}{\Delta t_k\, P_i^-}\right) & \text{if } P_i^- < 0, \\ 1 & \text{if } P_i^- = 0, \end{cases}$$

where $(r_{ij}^k)^+ = \max\{r_{ij}^k, 0\}$ and $(r_{ij}^k)^- = \min\{r_{ij}^k, 0\}$ are the positive and negative parts of r_{ij}^k, respectively, and $S(i) = \{j \in \{1, \dots, N\} \setminus \{i\}\,;\ \exists\, T \in \mathscr{T}_h : x_i, x_j \in T\}$. Finally, the correction factors α_{ij}^k are defined by

$$\alpha_{ij}^k := \begin{cases} \min(R_i^+, R_j^-) & \text{for } r_{ij}^k \geq 0, \\ \min(R_i^-, R_j^+) & \text{for } r_{ij}^k < 0. \end{cases} \tag{12}$$

This choice of α_{ij}^k guarantees that the scheme (10)–(11) satisfies the DMP.

These limiters α_{ij}^k are clearly symmetric (if $r_{ij}^k \neq 0$) with values in $[0, 1]$ and the following lemma shows that they also satisfy the continuity assumption from Theorem 2.

Lemma 4 *For any* $i, j \in \{1, \dots, N\}$*, the function* α_{ij}^k *defined in* (12) *is such that* $\alpha_{ij}^k\, r_{ij}^k$ *is a continuous function of* $u_1^k, \dots, u_N^k$*.*

Proof. For $\mathrm{U} \in \mathbb{R}^N$, denote $\Phi(\mathrm{U}) = (\alpha_{ij}^k\, r_{ij}^k)(\mathrm{U})$, i.e., we dropped the index k in U^k. Consider any $\overline{\mathrm{U}} \in \mathbb{R}^N$. If $r_{ij}^k(\overline{\mathrm{U}}) \neq 0$, then the denominators in the formulas defining α_{ij}^k do not vanish in a neighborhood of $\overline{\mathrm{U}}$ and hence α_{ij}^k is continuous at $\overline{\mathrm{U}}$. Consequently, also $\alpha_{ij}^k\, r_{ij}^k$ is continuous at $\overline{\mathrm{U}}$. If $r_{ij}^k(\overline{\mathrm{U}}) = 0$, then

$$\begin{aligned} &|(\alpha_{ij}^k\, r_{ij}^k)(\mathrm{U}) - (\alpha_{ij}^k\, r_{ij}^k)(\overline{\mathrm{U}})| \\ &\quad = \ |(\alpha_{ij}^k\, r_{ij}^k)(\mathrm{U})| \le |r_{ij}^k(\mathrm{U})| = |r_{ij}^k(\mathrm{U}) - r_{ij}^k(\overline{\mathrm{U}})| \le C\, \|\mathrm{U} - \overline{\mathrm{U}}\| \end{aligned}$$

and hence again $\alpha_{ij}^k\, r_{ij}^k$ is continuous at $\overline{\mathrm{U}}$. □

Acknowledgment

The work of P. Knobloch has been supported through the grant No. 19-04243S of the Czech Science Foundation.

References

[1] D. Kuzmin, S. Turek, Flux correction tools for finite elements, J. Comput. Phys. 175 (2002) 525–558.

[2] D. Kuzmin, M. Möller, Algebraic flux correction. I. Scalar conservation laws, in: Flux-corrected transport, Sci. Comput., Springer, Berlin, 2005, pp. 155–206.

[3] D. Kuzmin, Explicit and implicit FEM-FCT algorithms with flux linearization, J. Comput. Phys. 228 (2009) 2517–2534.

[4] V. John, E. Schmeyer, Finite element methods for time-dependent convection–diffusion–reaction equations with small diffusion, Comput. Methods Appl. Mech. Engrg. 198 (2008) 475–494.

[5] V. John, P. Knobloch, P. Korsmeier, On the solvability of the nonlinear problems in an algebraically stabilized finite element method for evolutionary transport-dominated equations, Math. Comp. (2020). In press.

[6] H.G. Roos, M. Stynes, L. Tobiska, Robust Numerical Methods for Singularly Perturbed Differential Equations. Convection–Diffusion–Reaction and Flow Problems. 2nd ed., Springer-Verlag, Berlin, 2008.

[7] R. Temam, Navier-Stokes equations. Theory and numerical analysis, North-Holland, Amsterdam, 1977.

[8] S.T. Zalesak, Fully multidimensional flux-corrected transport algorithms for fluids, J. Comput. Phys. 31 (1979) 335–362.