No. 670 January 2024

Improving Convergence of Time-Simultaneous Multigrid Methods for Convection-Dominated Problems using VMS Stabilization Techniques

W. Drews, S. Turek, C. Lohmann

ISSN: 2190-1767

Improving Convergence of Time-Simultaneous Multigrid Methods for Convection-Dominated Problems using VMS Stabilization Techniques

Wiebke Drews[0009−0006−7269−6838],
Stefan Turek[0000−0002−9740−6087] and
Christoph Lohmann[0000−0002−9681−9314]

Abstract We present the application of a time-simultaneous multigrid algorithm closely related to multigrid waveform relaxation for stabilized convection-diffusion equations in the regime of small diffusion coefficients. We use Galerkin finite elements and the Crank-Nicolson scheme for discretization in space and time. The multigrid method blocks all time steps for each spatial unknown, enhancing parallelization in space. While the number of iterations of the solver is bounded above for the 1D heat equation, convergence issues arise in convection-dominated cases. In singularly perturbed advection-diffusion scenarios, Galerkin FE discretizations are known to show instabilities in the numerical solution. We explore a higher-order variational multiscale stabilization, aiming to enhance solution smoothness and improve convergence without compromising accuracy.

1 Introduction

In the scope of convection-dominated transport problems, Galerkin finite element solutions are known to encounter spurious artifacts. Amid various stabilization techniques like GLS and SUPG, this study focuses on a higher-order, fully implicit variational multiscale (VMS) method [1, 2]. This method enhances solution accuracy by altering the problem's variational form, adding a diffusive term, and eliminat-

Wiebke Drews
Institute of Applied Mathematics (LSIII), TU Dortmund University, Vogelpothsweg 87, D-44227 Dortmund, Germany, e-mail: wiebke.drews@math.tu-dortmund.de

Stefan Turek
Institute of Applied Mathematics (LSIII), TU Dortmund University, Vogelpothsweg 87, D-44227 Dortmund, Germany, e-mail: stefan.turek@math.tu-dortmund.de

Christoph Lohmann
Institute of Applied Mathematics (LSIII), TU Dortmund University, Vogelpothsweg 87, D-44227 Dortmund, Germany, e-mail: christoph.lohmann@math.tu-dortmund.de

ing low-frequency diffusion. Beyond accuracy enhancement, we aim to improve the convergence behavior of the time-simultaneous multigrid solver closely related to multigrid waveform relaxation (WRMG) [3]. This solver belongs to time-parallel integration methods, is highly parallelizable in space, and characterized by treating all time steps simultaneously and consequently solving a space-only problem. Unlike conventional methods handling initial value problems sequentially, these newer algorithms enable simultaneous solutions across all time steps, augmenting parallelization limited by spatial resolution. [4, 5] provide a comprehensive introduction to parallel-in-time methods, that aim to reduce the run times of applications by fully exploiting the potential of massively parallel computing. Although parallel-in-time methods date back over 50 years, challenges persist in achieving optimal efficiency for higher-order discretizations of convection-dominated problems. Recent investigations of Multigrid Reduction-In-Time (MGRIT) algorithms [6] and Schwarz waveform relaxation algorithms [7] aim to tackle these challenges.

We enhance the convection-diffusion equation by a VMS stabilization technique and numerically solve the system by the time-simultaneous multigrid method. The concept involves perturbing the system using higher-order diffusion to maintain high accuracy while improving convergence behavior in convection-dominated scenarios, which is investigated in the numerical studies. This work is highly related to [8].

2 Convection-dominated problems

We consider the convection-diffusion equation for the solution $u : \Omega \times (0,T) \to \mathbb{R}$ such that

$$\partial_t u(x,t) - \varepsilon \Delta u(x,t) + \boldsymbol{v}(x,t) \cdot \nabla u(x,t) = f(x,t), \qquad (x,t) \in \Omega \times (0,T), \tag{1}$$

where the spatial domain is given by $\Omega \subset \mathbb{R}^d, d \in \{1,2,3\}$, the time interval is limited by the final time $T > 0$ and $f : \Omega \times (0,T) \to \mathbb{R}$ denotes the right hand side. Using the subspace $V_h \subset V = H^1(\Omega)$ of linear finite elements (FE), then $u_h : (0,T) \to V_h$ satisfies the variational formulation

$$(\partial_t u_h, \varphi_h) + \varepsilon(\nabla u_h, \nabla \varphi_h) + (\boldsymbol{v} \cdot \nabla u_h, \varphi_h) = (f, \varphi_h) \quad \forall \varphi_h \in V_h. \tag{2}$$

In convection-dominated cases, where the diffusion coefficient $\varepsilon \geq 0$ is much smaller than the velocity field $\boldsymbol{v} \in \mathbb{R}^d$ and the grid size h, it is known that oscillations occur in the Galerkin FE solution of the problem. Stabilization techniques are commonly used to reduce those instabilities [9, Sec. 8.3]. While oscillations only occur to a limited extent in the convection-dominated cases under consideration, the slow convergence of the solver causes particular problems. We therefore aim to improve the solution behavior and introduce the chosen stabilization technique in what follows.

2.1 Stabilization technique

We consider a projection-based variational multiscale (VMS) stabilization method as published in [1, 2]. Using the definition $\tilde{u}_h := [u_h + \frac{\delta t}{2}\partial_t u_h]$, which is specifically investigated in [8] for the application of the Crank-Nicolson scheme with time step size δt, problem (2) is enhanced by a diffusive and a compensation part as follows

$$
\begin{aligned}
(\partial_t u_h, \varphi_h) + \varepsilon(\nabla u_h, \nabla\varphi_h) + (\boldsymbol{v}\cdot\nabla u_h, \varphi_h) \\
+ \alpha_{add}[(\nabla\tilde{u}_h, \nabla\varphi_h) - (\mathbf{g}_h, \nabla\varphi_h)] = (f, \varphi_h) \quad & \forall\varphi_h \in V_h, \quad (3)\\
(\mathbf{g}_h - \nabla\tilde{u}_h, \boldsymbol{\psi}_h) = 0 \quad & \forall\boldsymbol{\psi}_h \in (V_h)^d,
\end{aligned}
$$

where $\alpha_{add} \geq 0$ is the stabilization parameter and will be defined in more detail within the multigrid context below. The projection-based gradient $\mathbf{g}_h : (0.T) \to (V_h)^d$ enforces the stabilization term to vanish in the continuous problem while $(V_h)^d$ is a d-dimensional FE subspace of $(L^2(\Omega))^d$. The use of the same FE subspace for the approximation of the gradient was also considered in [10, Sec. 5]. By substituting $\mathbf{g}_h$ into the first equation of (3), the problem in matrix form reads

$$
\boldsymbol{M}_h\partial_t\boldsymbol{u}_h(t) + \varepsilon\boldsymbol{L}_h\boldsymbol{u}_h(t) + \boldsymbol{K}_h\boldsymbol{u}_h(t) + \alpha_{add}\boldsymbol{W}_h[\boldsymbol{u}_h(t) + \tfrac{\delta t}{2}\partial_t\boldsymbol{u}_h(t)] = \boldsymbol{f}_h(t), \quad (4)
$$

where $\boldsymbol{M}_h, \boldsymbol{L}_h, \boldsymbol{K}_h$ are the mass, diffusion, and convection matrices, and $\boldsymbol{f}_h$ is the discretized right hand side. The substitution results in the stabilization matrix $\boldsymbol{W}_h := \boldsymbol{L}_h - \boldsymbol{B}_h^\top\boldsymbol{N}_h^{-1}\boldsymbol{B}_h$, where $\boldsymbol{B}_h$ is the discrete counterpart of the gradient and $\boldsymbol{N}_h$ is the mass matrix corresponding to the vector-valued FE space $(V_h)^d$. Since all integrals are approximated using the Trapezoidal rule, both mass matrices are diagonal and $\boldsymbol{W}_h$ can be determined explicitly. All occurring matrices are defined in $\mathbb{R}^{N\times N}$ for $N \in \mathbb{N}$ spatial unknowns. Using the Crank-Nicolson scheme as the time-integrator with the discrete initial condition $\boldsymbol{u}_h^0$ results in the sequential solution procedure

$$
\boldsymbol{A}_I\boldsymbol{u}_h^m + \boldsymbol{A}_E\boldsymbol{u}_h^{m-1} = \boldsymbol{f}^m, \quad m = 1, ..., K, \quad (5)
$$

$$
\begin{aligned}
\text{where } \boldsymbol{A}_I &:= \quad (\boldsymbol{M}_h + \alpha_{add}\tfrac{\delta t}{2}\boldsymbol{W}_h) + \tfrac{\delta t}{2}(\varepsilon\boldsymbol{L}_h + \boldsymbol{K}_h + \alpha_{add}\boldsymbol{W}_h),\\
\boldsymbol{A}_E &:= -(\boldsymbol{M}_h + \alpha_{add}\tfrac{\delta t}{2}\boldsymbol{W}_h) + \tfrac{\delta t}{2}(\varepsilon\boldsymbol{L}_h + \boldsymbol{K}_h + \alpha_{add}\boldsymbol{W}_h)
\end{aligned}
$$

for time step size δt, K time steps and the right hand side $\boldsymbol{f}^m := \frac{1}{2}\delta t\left(\boldsymbol{f}_h^m + \boldsymbol{f}_h^{m-1}\right)$. Due to the definition of $\tilde{u}_h$ used in (3), the stabilization matrix in $\boldsymbol{A}_E$ cancels out and is therefore treated fully implicitly within this time stepping scheme.

2.2 Solution technique

Based on the sequential solution method (5) using the Crank-Nicolson scheme, we block all equations into a single linear system of equations

$$\begin{pmatrix} \boldsymbol{A}_I & & & \\ \boldsymbol{A}_E & \boldsymbol{A}_I & & \\ & \ddots & \ddots & \\ & & \boldsymbol{A}_E & \boldsymbol{A}_I \end{pmatrix} \begin{pmatrix} \boldsymbol{u}^1 \\ \boldsymbol{u}^2 \\ \vdots \\ \boldsymbol{u}^K \end{pmatrix} = \begin{pmatrix} \boldsymbol{f}^1 - \boldsymbol{A}_E \boldsymbol{u}^0 \\ \boldsymbol{f}^2 \\ \vdots \\ \boldsymbol{f}^K \end{pmatrix} \in \mathbb{R}^{NK}$$

and perform the following rearrangement to switch from a time-major ordering to a space-major ordering:

$$(u_1^1, u_2^1, ..., u_N^1, u_1^2, ..., u_1^K, ...) \quad \rightsquigarrow \quad (u_1^1, u_1^2, ..., u_1^K, u_2^1, ..., u_N^1, ...) =: \boldsymbol{u}$$

In the one-dimensional case of (1), this results in a block tridiagonal structured global linear system

$$\underbrace{\begin{pmatrix} \square & \square & & \\ \square & \square & \ddots & \\ & \ddots & \ddots & \square \\ & & \square & \square \end{pmatrix}}_{=:\boldsymbol{S} \in \mathbb{R}^{NK \times NK}} \boldsymbol{u} = \boldsymbol{f} \quad \text{with block matrix entries} \quad \boxed{\begin{matrix} * & & & \\ * & * & & \\ & \ddots & \ddots & \\ & & * & * \end{matrix}} \in \mathbb{R}^{K \times K}, \tag{6}$$

where no stabilization is incorporated into the system using $\alpha_{add} = 0$. While the matrix block entries are lower bidiagonal matrices due to the time integrator, the overlying block tridiagonal structure results from the FE discretization in space. Therefore, (6) can be interpreted as a space-only system with vector-valued unknowns, i.e., all time steps are treated simultaneously for each spatial node.

A time-simultaneous multigrid algorithm, which uses spatial coarsening and common multigrid components such as smoothing and intergrid transfer operators is applied to (6). For smoothing purposes, we consider the GMRES method with block Jacobi preconditioning. The preconditioner $\boldsymbol{D}$ then corresponds to the block diagonal of $\boldsymbol{S}$ and provides a high degree of parallelization, since each block can be considered independently and only couples the (temporal) degrees of freedom associated with a single spatial node. Its lower bidiagonal structure makes it easy to solve the resulting linear systems of equations by other appropriate approaches. To transfer between space-time grids, we define the global restriction operator by the Kronecker product $\boldsymbol{R} := \boldsymbol{R}_h^H \otimes \boldsymbol{I}_K$, where $\boldsymbol{I}_K$ is an identity matrix of size $K \times K$, and the global prolongation operator $\boldsymbol{P} = \boldsymbol{R}^\top$, accordingly. While $\boldsymbol{R}_h^H$ denotes the restriction operator in space induced by the FE space, where $H = 2h$ is the coarser mesh size, the time step size stays fixed on all grid levels. In the coarse grid correction of the special case of the two-grid algorithm, the system matrix $\bar{\boldsymbol{S}} \in \mathbb{R}^{\bar{N}K \times \bar{N}K}$ is designed in the same way as (6), but with grid size H. The coarse grid problem is then solved directly. A detailed description of this algorithm can be found in [8, 11]. The multigrid solver is closely related to WRMG [3] when using the block Jacobi smoother and a corresponding time stepping method. This was already discussed in [11] and literature on the convergence analysis of WRMG can be found in [12, 13].

For problem (1) in 1D, an equidistant triangulation and a constant velocity field, the FE formulation (4) is equivalent to the second order finite difference (FD) discretization, ensuring second-order accuracy in space. In this context, the stabilization matrix can be interpreted as a scaled FD discretization of the biharmonic operator. This was shown in more detail in [8], where the following stabilization parameter for a coarser level with grid size H was derived:

$$\alpha_{add} \coloneqq \alpha \left(\frac{h}{H}\right)^{\gamma}$$

We study three different choices of γ: The parameter $\gamma = 2$ preserves solving the same continuous problem on each level of the multigrid algorithm (cf. [8]). In this case, the stabilization parameter is reduced by a factor of 4 when the mesh size H is doubled due to a mesh refinement within the multigrid algorithm. The second choice $\gamma = 0$, which results from the definition of the VMS stabilization (4), is an intuitive level-independent option. As a compromise, also $\gamma = 1$ is taken into account, keeping α_{add} larger on the coarser levels.

3 Numerical studies

The described VMS stabilization technique using the Crank-Nicolson time integrator in combination with the time-simultaneous multigrid algorithm will be examined for convection-dominated problems in this section. First, we shortly investigate a non-smooth exact solution to illustrate the solution behavior for the chosen discretization in unstabilized and stabilized cases in Sec. 3.1. Afterwards the influence of the stabilization on the multigrid solver is studied in Sec. 3.2.

3.1 Non-smooth solution

We consider the Heaviside step function with periodic boundary values, $v = 1$ and $\varepsilon = 0$ on $\Omega = (0, 1)$ at the final time $T = 1$ for different discretization techniques without any stabilization and one stabilized case with $\alpha_{add} > 0$ in Fig. 1. Spurious oscillations occur in the Galerkin approximation of the exact solution. These can be smoothed out by using the upwind scheme to discretize the convective part of the problem at hand. However, this result is very diffusive and the method only leads to first-order accuracy [9, Sec. 8.3]. Therefore, we prefer the stabilized case, since the spurious artifacts of the Galerkin solution can be smoothed and the higher order is preserved by this VMS stabilization.

In what follows, the influence of the stabilization on the solver is investigated. The improvement of the convergence behavior is the main reason for the application of the stabilization in this work.

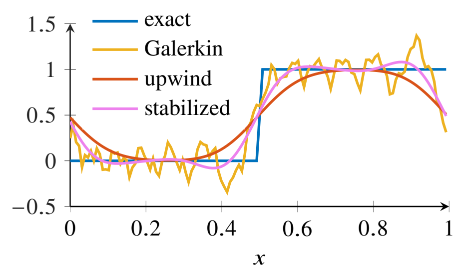

Fig. 1 Heaviside step function and corresponding numerical solutions at final time $T = 1$ in the case of $h = \delta t = \frac{1}{128}$. The parameter $\alpha = 0.1$ is used for the stabilized case.

3.2 Smooth solution

This study explores the numerical solution of the stabilized convection-diffusion problem (3) in 1D for convection-dominated cases and the fixed velocity field $v = 1$ using the time-simultaneous multigrid algorithm. We consider

$$u(x,t) = \exp\left(-\eta\left(\tfrac{1}{2} - x + \tfrac{1}{4}\sin\left(\tfrac{\pi}{2}t\right)\right)^2\right)\sin(\pi x), \quad (x,t) \in (0,1)\times(0,T),$$

as the manufactured solution of the inital value problem satisfying homogeneous Dirichlet boundary conditions and the corresponding initial guess $u(x,0) = \exp(-\eta(\frac{1}{2} - x)^2)\sin(\pi x)$. The parameter $\eta = 100$ is chosen to keep the temporal and spatial error in balance and characterizes the steepness of u. We differ between the multgrid case, where the coarse level is chosen to be level 1, and the two-grid case with coarse level $l - 1$, while $h = 2^{-l}$ is the mesh size of fine level l. In the numerical experiments, we compute the numerical solution up to a relative tolerance of 10^{-8} and the smoother performs 4 pre- and post-smoothing steps. For further algorithmic aspects, we refer to Sec. 2.2.

In Fig. 2, the number of multigrid iterations for the unstabilized Galerkin discretization is illustrated for different numbers of blocked time steps K. While the number of iterations is bounded above for moderate diffusion coefficients, the solver requires more than 100 iterations as ε decreases and more time steps are blocked. This convergence behavior is to be improved for a large number of simultaneously

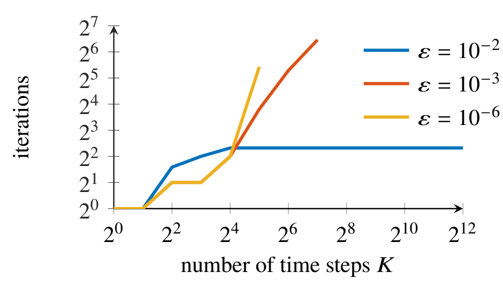

Fig. 2 Number of iterations for multigrid solver using V-cycle, $h = \delta t = \frac{1}{128}$, $v = 1$, $\alpha_{add} = 0$ and Galerkin discretization of convective term.

treated time steps in order to obtain a fast multigrid solver. The convection-dominated case $\varepsilon = 10^{-3}$ is therefore investigated in Fig. 3 to illustrate the influence of the stabilization on the two-grid solver and to justify the choice of the stabilization parameter α. While the method does not converge within 100 iterations when only a small

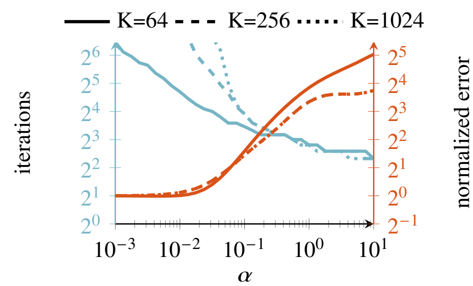

Fig. 3 Number of iterations and normalized error for two-grid solver with $h = \delta t = \frac{1}{32}$, $v = 1$, $\varepsilon = 10^{-3}$, and stabilization parameter $\gamma = 2$.

amount of stabilization is incorporated into the system, the number of iterations can be reduced and is even bounded above for different numbers of blocked time steps for $\alpha \geq 10^{-1}$. At the same time, the error, which is normalized with respect to the error for $\alpha = 10^{-3}$, becomes larger if α increases. From the point to view of solution quality, α should not be chosen too large. Comprising both, the error and the convection behavior, we suggest to choose $\alpha = 10^{-1}$. Then the solver shows the desired effects and the error of the solution grows only by a factor of 4, which corresponds to one level of mesh refinement due to the second order of convergence of the FE discretization at hand.

In Table 1a we examine the different choices of γ for the preferred value $\alpha = 10^{-1}$ in the case of $\varepsilon = 10^{-3}$. In the two-grid approach, the number of iterations is bounded above for all configurations. However, the solver converges fastest here for $\gamma = 2$. In the multigrid case, however, the number of iterations for this value of γ increases significantly if a large number of time steps are blocked. As already mentioned above, $\gamma = 2$ results in a smaller stabilization parameter α_{add} on the coarser levels. This seems to affect the convergence of the multigrid solver, so that in particular the level-independent stabilization with $\gamma = 0$ provides better convergence rates, especially for a large number of blocked time steps. In Table 1b, similar effects can be observed for an even smaller diffusion coefficient $\varepsilon = 10^{-6}$. The preferred choice of $\gamma = 0$ for the multigrid case and large K is emphasized here once again since the solver reaches the maximum number of iterations using the other values of γ.

4 Conclusion

The study presented the time-simultaneous multigrid algorithm applied to the convection-diffusion equation in 1D, revealing convergence issues in convection-dominated problems, while the convergence rate is uniformly bounded regardless

Table 1: Number of iterations in case of $v = 1$, and stabilization parameter $\alpha = 10^{-1}$. A dash "-" indicates that the solver did not converge within 100 iterations.

(a) $\varepsilon = 10^{-3}$

	Two-grid						Multigrid (F-cycle)					
	$\gamma = 0$		$\gamma = 1$		$\gamma = 2$		$\gamma = 0$		$\gamma = 1$		$\gamma = 2$	
$K \setminus h = \delta t$	1/128	1/512	1/128	1/512	1/128	1/512	1/128	1/512	1/128	1/512	1/128	1/512
16	10	5	7	4	6	4	10	13	7	9	6	4
64	19	8	11	6	8	4	19	17	11	9	8	4
256	26	17	13	10	8	4	31	26	16	12	13	6
1024	25	19	13	11	8	5	34	32	17	16	31	16
4096	24	18	12	11	7	5	33	33	20	21	60	67
16384	23	17	12	10	7	4	32	32	64	23	-	43

(b) $\varepsilon = 10^{-6}$

	Two-grid						Multigrid (F-cycle)					
	$\gamma = 0$		$\gamma = 1$		$\gamma = 2$		$\gamma = 0$		$\gamma = 1$		$\gamma = 2$	
$K \setminus h = \delta t$	1/128	1/512	1/128	1/512	1/128	1/512	1/128	1/512	1/128	1/512	1/128	1/512
16	10	5	8	4	6	4	11	14	8	10	6	4
64	21	8	12	6	8	4	21	19	12	10	8	4
256	32	21	15	11	9	4	37	32	19	13	16	8
1024	31	24	15	12	9	5	46	50	23	22	53	42
4096	30	23	14	12	8	5	44	61	26	28	98	-
16384	28	21	13	11	8	4	42	59	-	34	-	96

of the number of blocked time steps if the diffusion parameter is sufficiently large. The use of a higher-order VMS stabilization technique improves convergence to a bounded number of iterations and maintains second-order accuracy in space and time. The choice of the stabilization parameter is a crucial aspect. A level-dependent parameter significantly reduces the number of iterations in the two-grid studies. Multigrid scenarios emphasize the trade-off between the added stabilization and the consistency of the coarse grid problem, which affects convergence behavior. Therefore, future investigations aim to extend this method by adapting stabilization parameters or varying smoothing steps based on levels, and may explore time-based stabilization.

Acknowledgements This work was supported by the Federal Ministry of Education and Research (BMBF) through the project "StroemungsRaum" 16ME0706K, which is part of the initiative "Neue Methoden und Technologien für das Exascale-Höchstleistungsrechnen" (SCALEXA).

References

1. John, V., Kaya, S., Layton, W.: A two-level variational multiscale method for convection-dominated convection-diffusion equations. In: Computer Methods in Applied Mechanics and Engineering, 195.33, pp. 4594–4603 (2006)
2. Layton, W.: A connection between subgrid scale eddy viscosity and mixed methods. In: Applied Mathematics and Computation, 133.1, pp. 147–157 (2002)
3. Lubich, C., Ostermann, A.: Multi–grid dynamic iteration for parabolic equations. In: BIT Numerical Mathematics, 27.2, pp. 216–234 (1987)
4. Gander, M.J.: 50 Years of Time Parallel Time Integration. In: Multiple Shooting and Time Domain Decomposition Methods, Springer International Publishing, pp. 69–113 (2015)
5. Ong, B.W., Schroder, J.B.: Applications of time parallelization. In: Computing and Visualization in Science, 23.11 (2020)
6. De Sterck, H., Falgout, R. D., Krzysik, O. A., Schroder, J. B.: Efficient Multigrid Reduction-in-Time for Method-of-Lines Discretizations of Linear Advection. In: Journal of Scientific Computing, 96.1 (2023)
7. Dong, W.B., Tang, H.S.: Convergence analysis of Schwarz waveform relaxation method to compute coupled advection–diffusion–reaction equations. In: Mathematics and Computers in Simulation, 218, pp. 462–481 (2024)
8. Drews, W., Turek, S., Lohmann, C.: Numerical Analysis of a Time–Simultaneous Multigrid Solver for Stabilized Convection–Dominated Transport Problems in 1D. In: Ergebnisberichte des Instituts für Angewandte Mathematik, Nummer 668, Fakultät für Mathematik, TU Dortmund (2023)
9. Quarteroni, A., Valli, A.: Numerical Approximation of Partial Differential Equations. Springer Berlin, Heidelberg (1994)
10. Lohmann, C., Kuzmin, D., Shadid, J. N., Mabuza, S.: Flux-corrected transport algorithms for continuous Galerkin methods based on high order Bernstein finite elements. In: Journal of Computational Physics, 344, pp. 151–186 (2017)
11. Dünnebacke, J., Turek, S., Lohmann, C., Sokolov, A., Zajac, P.: Increased space-parallelism via time-simultaneous Newton-multigrid methods for nonstationary nonlinear PDE problems. In: The International Journal of High Performance Computing Applications, 35.3, pp. 211–225 (2021)
12. Lohmann, C., Dünnebacke, J., Turek, S.: Fourier analysis of a time–simultaneous two–grid algorithm using a damped Jacobi waveform relaxation smoother for the one–dimensional heat equation. In: Journal of Numerical Mathematics, 30.3, pp. 173–207 (2022)
13. Notay, Y.: Rigorous convergence proof of space-time multigrid with coarsening in space. In: Numerical Algorithms, 89.2, pp. 675–699 (2022)