On a Time-Simultaneous Multigrid Method in Combination with Stabilization Techniques for the Convection–Diffusion Equation

Oberseminar WS 22/23

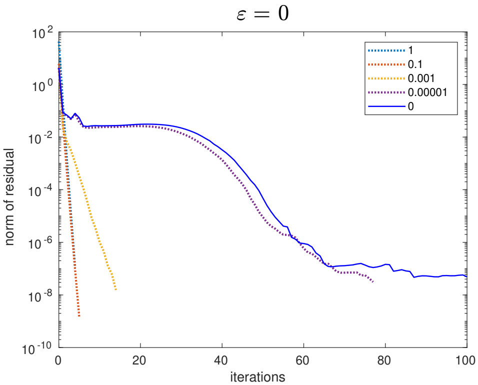
Wiebke Drews, Stefan Turek, Christoph Lohmann

Institute of Applied Mathematics (LS III)
Department of Mathematics
TU Dortmund University

January 30th, 2023

technische universität dortmund
fakultät für mathematik

Motivation



Exemplary convergence behavior for convection-dominated problems.

- solve convection-diffusion equation efficiently $\rightarrow$ parallelization
- use a **time-simultaneous** multigrid method[1]
- stabilization for convection-dominated problems

The perfect algorithm is **not** presented here!

[1] Multigrid Waveform Relaxation (by Lubich, Ostermann (1987)).

Table of contents

Preliminaries

Convection-diffusion equation in 1D

$$
\begin{aligned}
\partial_t u(x,t) - \varepsilon u_{xx}(x,t) + v(x,t)u_x(x,t) &= f(x,t) & (x,t) &\in \Omega \times (0,T] \\
u(0,t) = u(1,t) &= 0 & t &\in [0,T] \\
u(x,0) &= u_0(x) & x &\in \Omega
\end{aligned}
$$

with $\Omega = (0,1), T > 0$.

- finite difference (FD) discretization in space

$$\partial_t \boldsymbol{M}_h \boldsymbol{u}_h(t) + \varepsilon \boldsymbol{L}_h \boldsymbol{u}_h(t) = \boldsymbol{f}_h(t)$$

- Crank-Nicolson scheme for discretization in time

$$\boldsymbol{A}\boldsymbol{u}^m + \boldsymbol{B}\boldsymbol{u}^{m-1} = \boldsymbol{f}^m, \quad m = 1, \ldots, K$$

using time step size δt and

$$\boldsymbol{A} := \boldsymbol{M}_h + \tfrac{1}{2}\delta t \varepsilon \boldsymbol{L}_h, \quad \boldsymbol{B} := -\boldsymbol{M}_h + \tfrac{1}{2}\delta t \varepsilon \boldsymbol{L}_h, \quad \boldsymbol{f}^m := \tfrac{1}{2}\delta t \left(\boldsymbol{f}_h^m + \boldsymbol{f}_h^{m-1}\right).$$

Preliminaries

Convection-diffusion equation in 1D

$$
\begin{aligned}
\partial_t u(x,t) - \varepsilon u_{xx}(x,t) + \cancel{v(x,t)u_x(x,t)} &= f(x,t) && (x,t) \in \Omega \times (0,T] \\
u(0,t) = u(1,t) &= 0 && t \in [0,T] \\
u(x,0) &= u_0(x) && x \in \Omega
\end{aligned}
$$

with $\Omega = (0,1), T > 0$.

- finite difference (FD) discretization in space

$$
\partial_t \boldsymbol{M}_h \boldsymbol{u}_h(t) + \varepsilon \boldsymbol{L}_h \boldsymbol{u}_h(t) = \boldsymbol{f}_h(t)
$$

- Crank-Nicolson scheme for discretization in time

$$
\boldsymbol{A}\boldsymbol{u}^m + \boldsymbol{B}\boldsymbol{u}^{m-1} = \boldsymbol{f}^m, \quad m = 1, ..., K
$$

using time step size δt and

$$
\boldsymbol{A} := \boldsymbol{M}_h + \tfrac{1}{2}\delta t \varepsilon \boldsymbol{L}_h, \quad \boldsymbol{B} := -\boldsymbol{M}_h + \tfrac{1}{2}\delta t \varepsilon \boldsymbol{L}_h, \quad \boldsymbol{f}^m := \tfrac{1}{2}\delta t \left(\boldsymbol{f}_h^m + \boldsymbol{f}_h^{m-1}\right).
$$

Algebraic transformations

Blocking all time steps into a global linear system of equations...

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

Note: K time steps $t^1, t^2, ..., t^K$ and N spatial nodes $x_1, ..., x_N$

...and rearranging the degrees of freedom...

$$(u_1^1, u_2^1, \ldots u_N^1, u_1^2, u_2^2, \ldots, u_N^2, \ldots, u_1^K, u_2^K, \ldots, u_N^K)^\top$$

$$\downarrow$$

$$\boldsymbol{u} := (u_1^1, u_1^2 \ldots, u_1^K, u_2^1, u_2^2, \ldots, u_2^K, \ldots, u_N^1, u_N^2, \ldots, u_N^K)^\top$$

Algebraic transformations

...results in a **space-only problem** with vector-valued unknowns for each spatial node:

$$\underbrace{\begin{pmatrix} \# & \# & & \\ \# & \# & \ddots & \\ & \ddots & \ddots & \# \\ & & \# & \# \end{pmatrix}}_{=:\boldsymbol{S}\in\mathbb{R}^{NK\times NK}} \boldsymbol{u} = \boldsymbol{f}, \quad \text{with } \# := \begin{pmatrix} * & & & \\ * & * & & \\ & \ddots & \ddots & \\ & & * & * \end{pmatrix} \in \mathbb{R}^{K\times K}$$

$\rightarrow$ apply geometric multigrid method on $\boldsymbol{S}$ in space!

Aim: design of a highly parallelizable solution strategy!

Time-simultaneous multigrid algorithm

- smoothing
  - number of pre-smoothing and post-smoothing steps: ν_1 and ν_2
  - (damped) block Jacobi method: $\boldsymbol{x}^{(\nu)} = \boldsymbol{x}^{(\nu-1)} + \omega \boldsymbol{D}^{-1}(\boldsymbol{f} - \boldsymbol{S}\boldsymbol{x}^{(\nu-1)})$
  - block Jacobi preconditioning embedded into GMRES method:

$$\boldsymbol{D} := \begin{pmatrix} \# & & \\ & \ddots & \\ & & \# \end{pmatrix}, \quad \# = \underbrace{\begin{pmatrix} * & & & \\ * & * & & \\ & \ddots & \ddots & \\ & & * & * \end{pmatrix}}_{K \times K}$$

- intergrid transfer operators
  - standard coarsening in space for each time step
  - prolongation:

$$\boldsymbol{P}_{\delta t,2h}^{\delta t,h} = \boldsymbol{P}_{2h}^{h} \otimes \boldsymbol{I}_K$$

  - restriction:

$$\boldsymbol{R}_{\delta t,h}^{\delta t,2h} = \boldsymbol{R}_{h}^{2h} \otimes \boldsymbol{I}_K = \frac{1}{2}\left(\boldsymbol{P}_{\delta t,2h}^{\delta t,h}\right)^{\top}$$

Theoretical convergence results

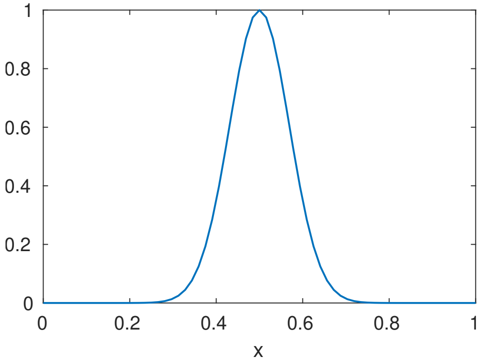
- linear multi-step methods: asymptotic convergence rate is the same as in time-stepping approach [Janssen, Vandewalle (1996)]

- Fourier analysis of time-simultaneous two-grid algorithm [Lohmann et al. (2022)]
  - 1D heat equation on uniform mesh
  - damped Jacobi (waveform relaxation) smoothing
  - spectral norm of two-grid iteration matrix $\boldsymbol{J}$ is uniformly bounded:

$$||\boldsymbol{J}||_2 < C < 1, \quad C \neq C(\delta t, h, K)$$

  for $\theta \geq \frac{1}{2}, \omega = \frac{2}{3}, \nu_1 \geq 1$.

Numerical studies on heat equation



$u(x,t)$ for different time steps.

- manufactured solution

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

  where $\eta = 100$.

- homogenous Dirichlet boundary conditions

$$u(0,t) = u(1,t) = 0.$$

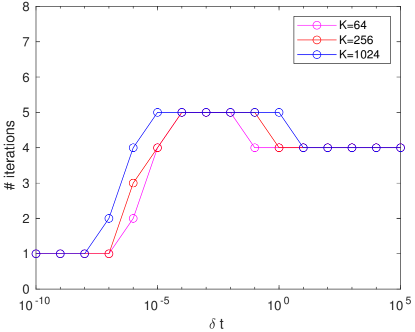
Discretization: FD in space, Crank-Nicolson scheme in time

- Level: fine level with spatial resolution $h = 2^{\text{-Level}}$
- δt: time step size
- K: number of blocked time steps

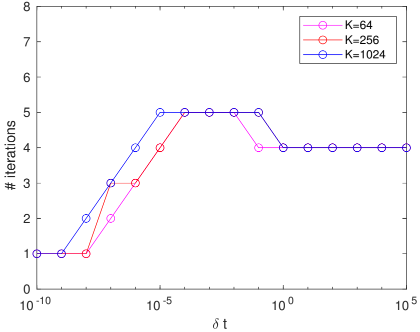
MG algorithm:

- coarse level 1
- GMRES smoother
- $\nu_1 = \nu_2 = 4$

Numerical studies on heat equation: V-cycle, $\varepsilon = 1$



Level 7.



Level 9.

Results:

- independent of K and fine mesh level
- number of iterations $\leq$ 5

Numerical studies on convection-diffusion equation

Convection-diffusion equation in 1D

$$\begin{aligned} \partial_t u(x,t) - \varepsilon u_{xx}(x,t) + \cancel{v(x,t) u_x(x,t)} &= f(x,t) && (x,t) \in \Omega \times (0,T] \\ u(0,t) = u(1,t) &= 0 && t \in [0,T] \\ u(x,0) &= u_0(x) && x \in \Omega \end{aligned}$$

with $\Omega = (0,1), T > 0$.

$$\partial_t \boldsymbol{M}_h \boldsymbol{u}_h(t) + \varepsilon \boldsymbol{L}_h \boldsymbol{u}_h(t) + \boldsymbol{K}_h \boldsymbol{u}_h(t) = \boldsymbol{f}_h(t)$$

- Convection term: discretized with first order upwind scheme
- Numerical studies with
  - $v = 1$ for convenience
  - manufactured solution as above and corresponding right hand side

Numerical studies on convection-diffusion equation

Convection-diffusion equation in 1D

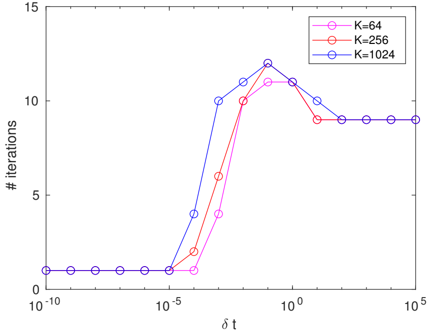
$$
\begin{aligned}
\partial_t u(x,t) - \varepsilon u_{xx}(x,t) + v(x,t)u_x(x,t) &= f(x,t) && (x,t) \in \Omega \times (0,T] \\
u(0,t) = u(1,t) &= 0 && t \in [0,T] \\
u(x,0) &= u_0(x) && x \in \Omega
\end{aligned}
$$

with $\Omega = (0,1), T > 0$.

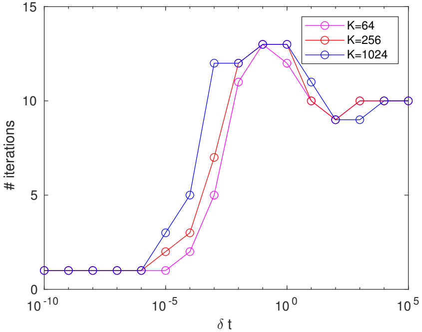
$$
\partial_t \boldsymbol{M}_h \boldsymbol{u}_h(t) + \varepsilon \boldsymbol{L}_h \boldsymbol{u}_h(t) + \boldsymbol{K}_h \boldsymbol{u}_h(t) = \boldsymbol{f}_h(t)
$$

- Convection term: discretized with first order upwind scheme
- Numerical studies with
  - $v = 1$ for convenience
  - manufactured solution as above and corresponding right hand side

Upwind discretization, V-cycle, $\varepsilon = 10^{-3}$



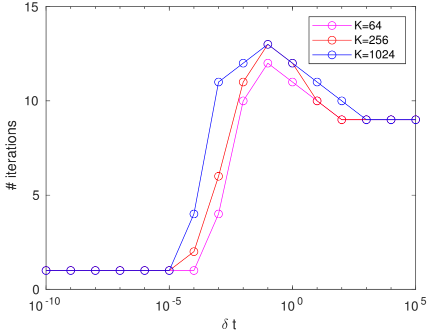
Level 7.



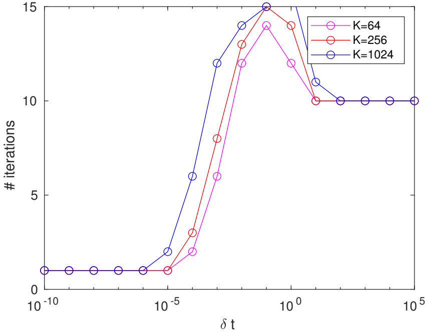
Level 9.

- Result: similar behavior as for heat equation, but only 1st order accuracy
- Goal: higher order of convergence $\rightarrow$ central difference scheme?

Upwind discretization, V-cycle, $\varepsilon = 0$



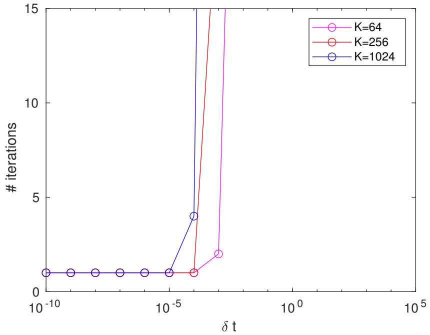
Level 7.



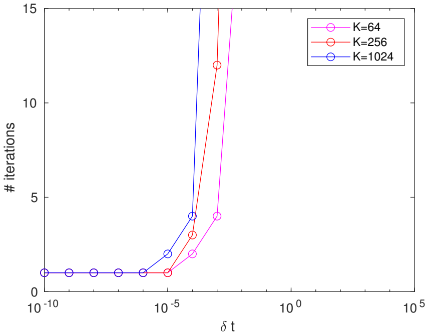
Level 9.

- Result: similar behavior as for heat equation, but only 1st order accuracy
- Goal: higher order of convergence $\rightarrow$ central difference scheme?

Central discretization, V-cycle, $\varepsilon = 10^{-3}$



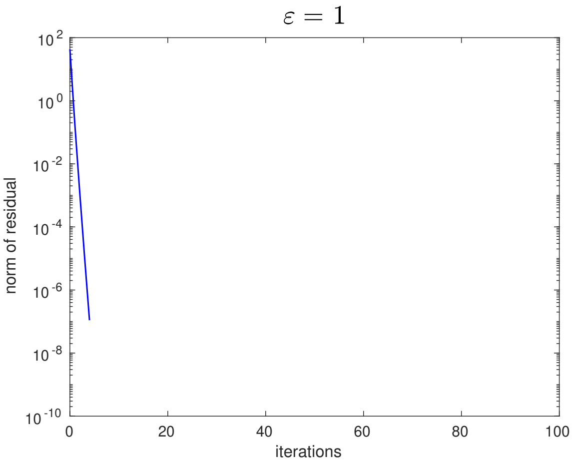
Level 7.



Level 9.

- Convection term discretized using second order central discretization
- Result: extremely increasing number of iterations!

Central discretization



Convergence behavior for different values of ε.

TG algorithm:
level $= 7$, $\delta t = \frac{1}{128}, K = 64.$

- fixed $v = 1$, varying ε

  $\rightarrow$ stability issues arise for convection-dominated problems

Higher order stabilization

- add diffusive term with stabilization parameter $\alpha_{add} \geq 0$ and compensation term[2]

$$(\partial_t u_h, \varphi_h) + \varepsilon(\nabla u_h, \nabla \varphi_h) + (v \cdot \nabla u_h, \varphi_h) + \alpha_{add}(\nabla u_h, \nabla \varphi_h) - \alpha_{add}(\mathbf{g}_h, \nabla \varphi_h) = (f, \varphi_h) \quad \forall \varphi_h \in V_h$$
$$(\mathbf{g}_h - \nabla u_h, \boldsymbol{\psi}_h) = 0 \qquad \forall \boldsymbol{\psi}_h \in (V_h)^d$$

- semi-discrete formulation in matrix form: $\boldsymbol{M}_h \sim \mathrm{id}, \boldsymbol{B}_h \sim \mathrm{grad}, \boldsymbol{B}_h^\top \sim \mathrm{div}$

$$\partial_t \boldsymbol{M}_h \boldsymbol{u}_h(t) + \varepsilon \boldsymbol{L}_h \boldsymbol{u}_h(t) + \boldsymbol{K}_h \boldsymbol{u}_h(t) + \alpha_{add} \underbrace{(\boldsymbol{L}_h - \boldsymbol{B}_h^\top \boldsymbol{M}_h^{-1} \boldsymbol{B}_h)}_{(*)} \boldsymbol{u}_h(t) = \boldsymbol{f}_h(t)$$

- $(*)$ 1D with linear FEM, uniform grid, quadrature based mass-lumping:

$$\tilde{\boldsymbol{L}}_h = \boldsymbol{M}_h^{-1} \boldsymbol{L}_h \sim -\frac{1}{h^2}[1, -2, 1], \qquad \tilde{\boldsymbol{L}}_{2h} = \boldsymbol{M}_h^{-1} \boldsymbol{B}_h^\top \boldsymbol{M}_h^{-1} \boldsymbol{B}_h \sim -\frac{1}{(2h)^2}[1, 0, -2, 0, 1]$$

$$\Longrightarrow \quad \tilde{\boldsymbol{L}}_h - \tilde{\boldsymbol{L}}_{2h} \sim \frac{1}{(2h)^2}[1, -4, 6, -4, 1]$$

[2] Fast, Mierka, Turek (2022), John, Kaya, Layton (2006).

Choice of α_{add} in multigrid algorithm

$$\alpha_{add} := \alpha \left(\frac{h_f}{h}\right)^{\gamma} \quad \Rightarrow \quad \alpha \left(\frac{h_f}{h}\right)^{\gamma} (\tilde{\boldsymbol{L}}_h - \tilde{\boldsymbol{L}}_{2h})\boldsymbol{u}_h$$

where $\alpha > 0, \gamma = 2$, $\quad h_f$: mesh size of fine level, $\quad h$: mesh size of current level.

- stabilization term corresponds to the biharmonic operator with a certain factor

  $\Rightarrow$ second order of accuracy

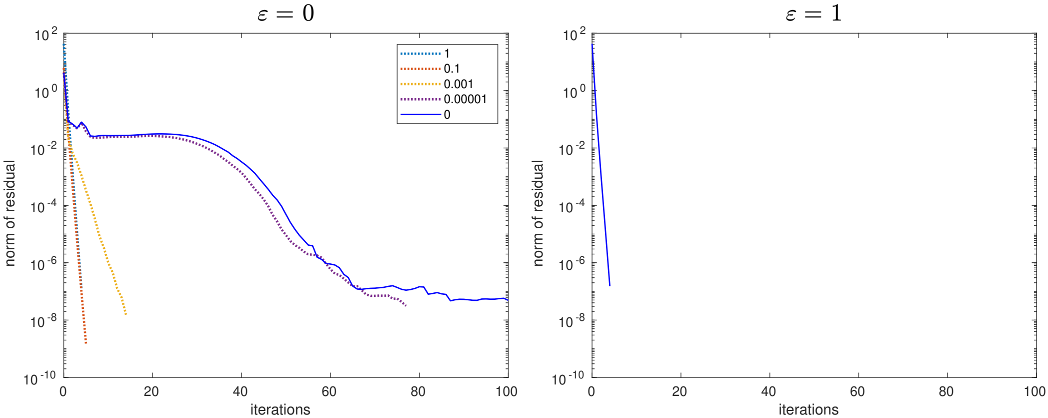
$$\begin{array}{rclllll}
\tilde{\boldsymbol{L}}_h \boldsymbol{u}_h & \sim & \frac{1}{h^2}(& h^2 u_{xx} + & C h^4 u_{xxxx} + & O(h^6)) \\
\tilde{\boldsymbol{L}}_{2h} \boldsymbol{u}_h & \sim & \frac{1}{4h^2}(& 4h^2 u_{xx} + & 2^4 C h^4 u_{xxxx} + & O(h^6)) \\
\hline
(\tilde{\boldsymbol{L}}_h - \tilde{\boldsymbol{L}}_{2h}) \boldsymbol{u}_h & \sim & & & \tilde{C} h^2 u_{xxxx} + & O(h^4)
\end{array}$$

- solve the same continuous problem on each level, **but** less stabilization on coarser levels

$$\alpha \left(\frac{h_f}{h}\right)^{\gamma} (\tilde{\boldsymbol{L}}_h - \tilde{\boldsymbol{L}}_{2h})\boldsymbol{u}_h \sim \alpha \left(\frac{h_f}{h}\right)^{\gamma} \tilde{C} h^2 u_{xxxx} = \alpha h_f^2 \tilde{C} u_{xxxx}, \qquad \text{for } \gamma = 2$$

- The stabilization term is treated fully implicit!

Numerical studies: stabilization



No stabilization. Stabilization with $\alpha = 0.1$.

- number of iterations hardly increases for small values of ε
- the plateau disappears

$\rightarrow$ stabilization can help! How to choose α quantitatively?

Choice of α

$\delta t = h$	$\alpha = 0$	$\alpha = 0.01$	$\alpha = 0.1$	$\alpha = 1$	$\alpha = 10$
1/64	2.0e-03	2.1e-03	6.9e-03	4.8e-02	1.3e-01
1/128	4.8e-04	5.1e-04	1.7e-03	1.6e-02	7.8e-02
1/256	1.2e-04	1.3e-04	4.4e-04	4.2e-03	3.4e-02

Discrete L_2-error at final time $T = 1$, $\varepsilon = 10^{-3}$.

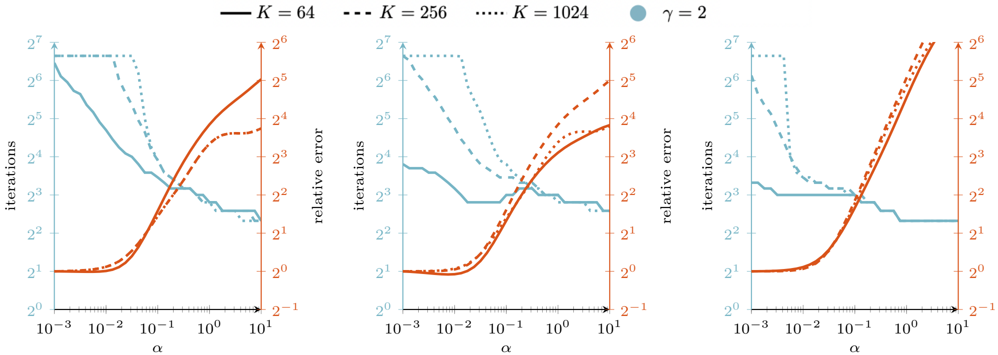
$\delta t = h$	$\alpha = 0$	$\alpha = 0.01$	$\alpha = 0.1$	$\alpha = 1$	$\alpha = 10$
1/64	5.0e-03	2.1e-03	7.0e-03	4.8e-02	1.3e-01
1/128	1.2e-03	5.1e-04	1.8e-03	1.6e-02	7.9e-02
1/256	3.0e-04	1.3e-04	4.4e-04	4.2e-03	3.4e-02

Discrete L_2-error at final time $T = 1$, $\varepsilon = 10^{-9}$.

- the error is reduced by a factor of $\approx$ 4 $\rightarrow$ 2nd order of convergence observed
- loss of accuracy for larger α

$\rightarrow$ do not choose α too large

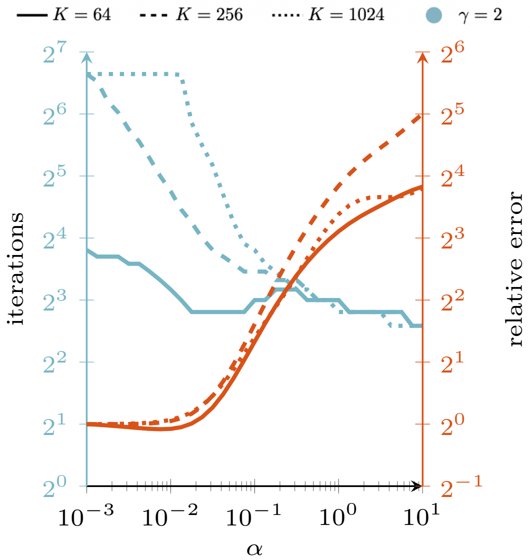
Choice of α, $\varepsilon = 10^{-3}$



Level 5, $\delta t = \frac{1}{32}$. Level 5, $\delta t = \frac{1}{128}$. Level 7, $\delta t = \frac{1}{128}$.

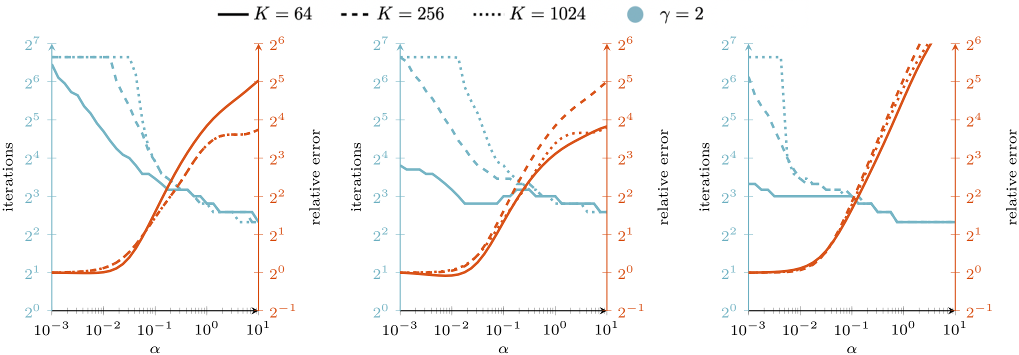
TG algorithm:
maximum number of iterations: 100.

Choice of α, $\varepsilon = 10^{-3}$



Level 5, $\delta t = \frac{1}{128}$.

Choice of α, $\varepsilon = 10^{-3}$



Level 5, $\delta t = \frac{1}{32}$. Level 5, $\delta t = \frac{1}{128}$. Level 7, $\delta t = \frac{1}{128}$.

Results:

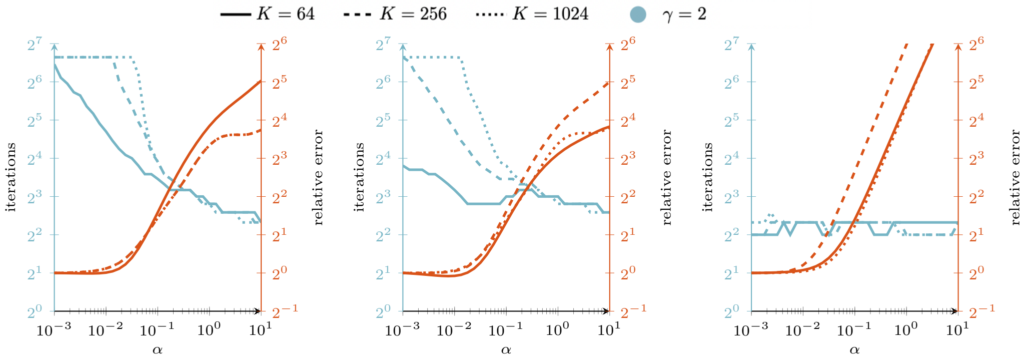
- number of iterations decreases as $\alpha \to \infty$
- independent of number of blocked time steps K and δt for sufficiently large $\alpha \approx 0.1$
- similar convergence behavior for different fine levels

$\rightarrow$ do not choose α too small

Results without stabilization:

- increasing/many iterations for large K
- δt- and level-dependency

Choice of α, $\varepsilon = 10^{-3}$



Level 5, $\delta t = \frac{1}{32}$. Level 5, $\delta t = \frac{1}{128}$. Level 9, $\delta t = \frac{1}{128}$.

Results:

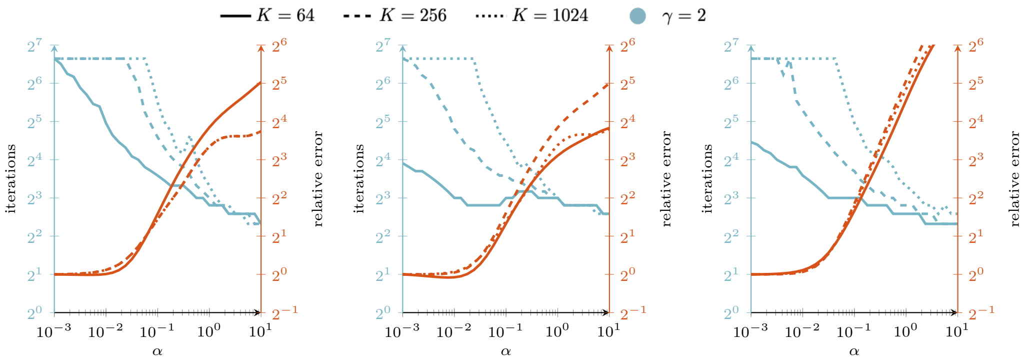
- number of iterations decreases as $\alpha \to \infty$
- independent of number of blocked time steps K and δt for sufficiently large $\alpha \approx 0.1$
- similar convergence behavior for different fine levels

$\rightarrow$ do not choose α too small

Results without stabilization:

- increasing/many iterations for large K
- δt- and level-dependency

Stabilization multigrid, F-cycle, $\varepsilon = 10^{-3}$



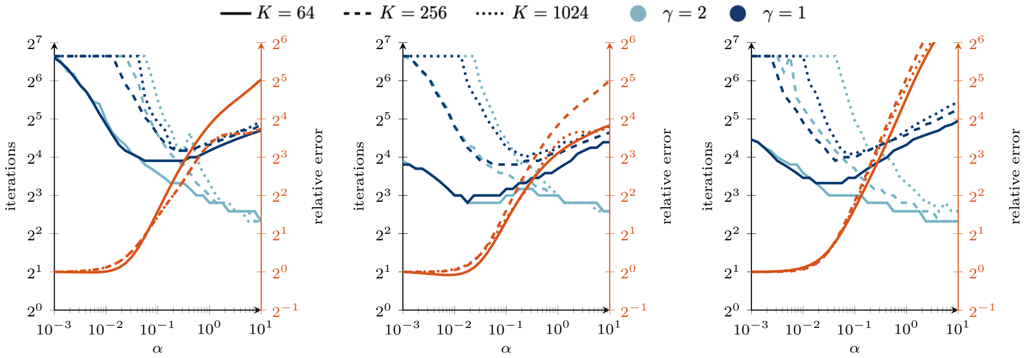
Level 5, $\delta t = \frac{1}{32}$. Level 5, $\delta t = \frac{1}{128}$. Level 7, $\delta t = \frac{1}{128}$.

Results:

- two-grid result can also be observed for multigrid
- but stabilization on coarse grid may not be enough for stable convergence rates
  $\alpha_{add} := \alpha \left(\frac{h_f}{h}\right)^\gamma \Rightarrow$ choosing $\gamma = 0$ or 1 can help!

MG algorithm:
maximum number of iterations: 100.

Stabilization multigrid, F-cycle, $\varepsilon = 10^{-3}$



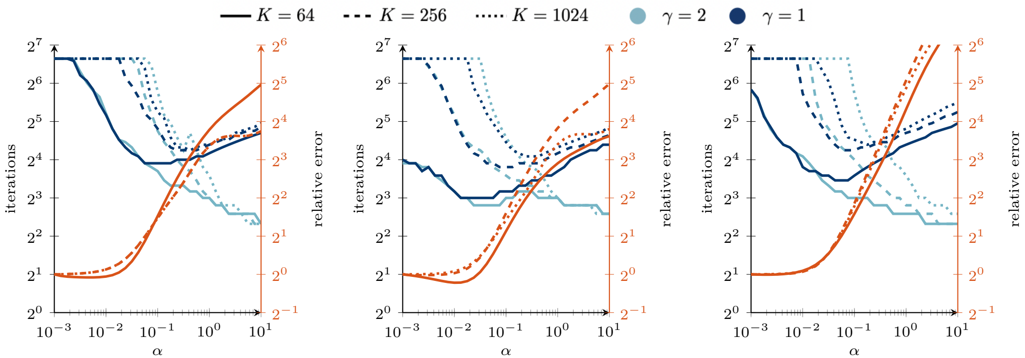
Level 5, $\delta t = \frac{1}{32}$. Level 5, $\delta t = \frac{1}{128}$. Level 7, $\delta t = \frac{1}{128}$.

Results:

- two-grid result can also be observed for multigrid
- but stabilization on coarse grid may not be enough for stable convergence rates
  $\alpha_{add} := \alpha \left(\frac{h_f}{h}\right)^{\gamma}$ $\Rightarrow$ choosing $\gamma = 0$ or 1 can help!

MG algorithm:
maximum number of iterations: 100.

Stabilization multigrid, F-cycle, $\varepsilon = 10^{-6}$



Level 5, $\delta t = \frac{1}{32}$.　　Level 5, $\delta t = \frac{1}{128}$.　　Level 7, $\delta t = \frac{1}{128}$.

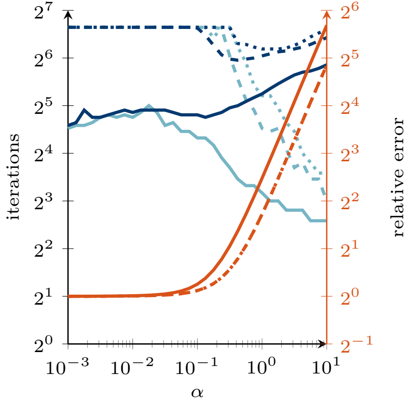
Results:

- without stabilization: ε-dependency
- with stabilization: similar results even for smallest ε

MG algorithm:
maximum number of iterations: 100.

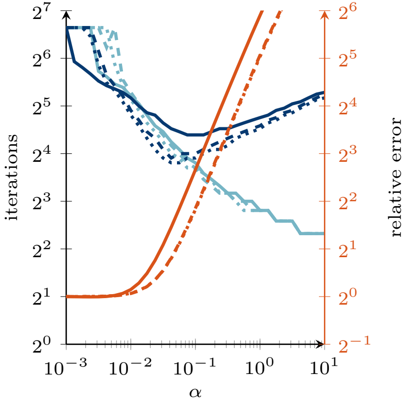
Difficulties and practical relevance

time step size $\delta t >$ fine mesh size



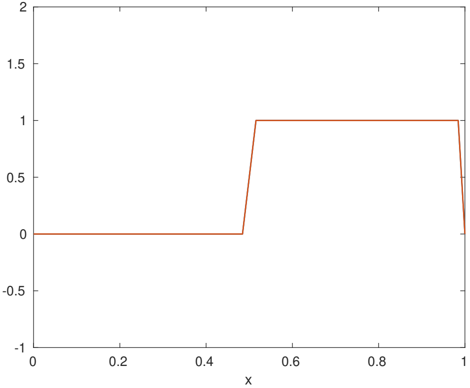
Level 9, $\delta t = \frac{1}{32}, \varepsilon = 10^{-3}$.

fixed final time T

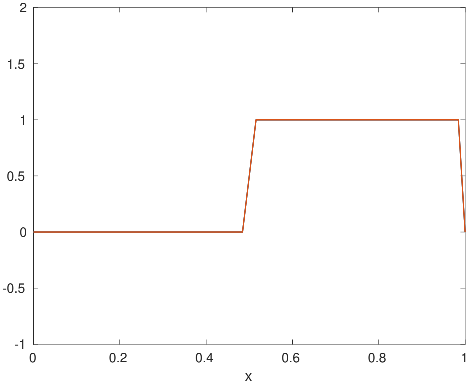


$T = 2$, Level 7, $\varepsilon = 10^{-3}$.

From another point of view: Heaviside step function



No stabilization.



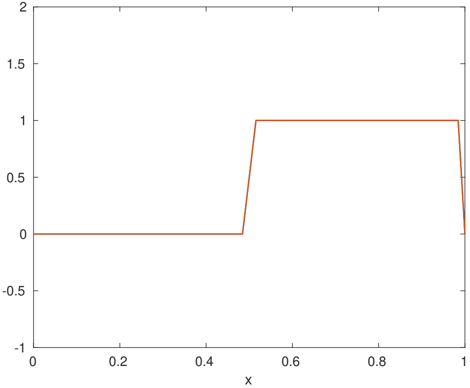
Stabilization with $\alpha = 0.01$.

- Level 6, $\delta t = \frac{1}{128}$
- oscillations in the numerical solution

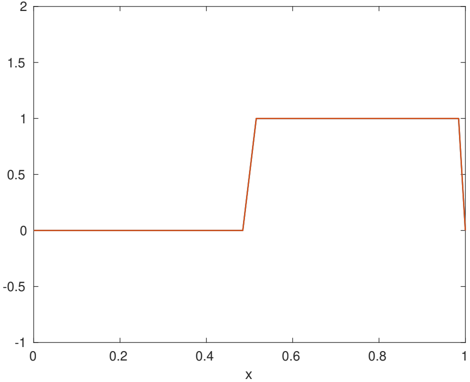
- small α can lead to a smoother numerical solution
- too large α can make it worse

$\rightarrow$ trade-off: solution vs. convergence behavior

From another point of view: Heaviside step function



No stabilization.



Stabilization with $\alpha = 0.1$.

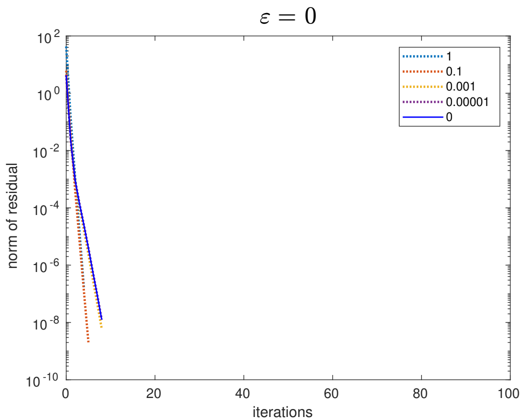
	δt	$1/32$	$1/128$	$1/512$
$\alpha = 0$	$K = 256$	100	100	100
	$K = 512$	100	100	100
	$K = 1024$	100	100	100
$\alpha = 0.1$	$K = 256$	18	14	10
	$K = 512$	24	15	12
	$K = 1024$	28	19	14

Level 7, $\varepsilon = 0$.

- small α can lead to a smoother numerical solution
- too large α can make it worse

$\rightarrow$ trade-off: solution vs. convergence behavior

Conclusion



Convergence behavior with stabilization.

Summary

- presented multigrid algorithm works fine for convection-diffusion equation if diffusion parameter is sufficiently large
- difficulties with convection-dominated problems $\rightarrow$ stabiliziation can help!
- choice of stabilization parameter is crucial

Outlook and more aspects

- extension to 2D and 3D problems
- studies on the reduction of the computational efficiency
- stabilization in time using u_{tttt}
- combination of this time-simultaneous algorithm and other parallel-in-time methods

References

J. Dünnebacke, S. Turek, C. Lohmann, A. Sokolov, P. Zajac.
Increased space-parallelism via time-simultaneous Newton-multigrid methods for nonstationary nonlinear PDE problems.
The International Journal of High Performance Computing Applications, vol. 35, no. 3, pp. 211-225, 2021.

M. Fast, O. Mierka, S. Turek.
Mathematical Modeling of Coolant Flow in Drilling Processes with Temperature Coupling.
Ergebnisberichte des Instituts für Angewandte Mathematik Nummer 651, Fakultät für Mathematik, TU Dortmund, 651, 2022.

J. Janssen, S. Vandewalle.
Multigrid Waveform Relaxation on Spatial Finite Element Meshes: The Discrete-Time Case.
SIAM J. Sci. Comput., vol. 17, no. 1, pp. 133-155, 1996.

V. John, S. Kaya, W. Layton.
A two-level variational multiscale method for convection-dominated convection–diffusion equations.
Comput. Methods Appl. Mech. Engrg., vol. 196, no. 33, pp. 4594-4603, 2006.

C. Lohmann, J. Dünnebacke, S. Turek.
Fourier analysis of a time-simultaneous two-grid algorithm using a damped Jacobi waveform relaxation smoother for the one-dimensional heat equation.
Journal of Numerical Mathematics, vol. 30, no. 3, pp. 173-207, 2022.

C. Lubich, A. Ostermann.
Multi-grid dynamic iteration for parabolic equations.
BIT., vol. 27, no. 2, pp. 216–234, 1987.