Mathematical Modelling and Numerical Analysis Will be set by the publisher
Modélisation Mathématique et Analyse Numérique

PDE-CONVERGENCE IN EUCLIDEAN NORM OF AMF-W METHODS FOR MULTIDIMENSIONAL LINEAR PARABOLIC PROBLEMS [*, **]

Severiano González-Pinto [1], Ernst Hairer [2] and Domingo Hernández-Abreu [1]

Abstract. This work considers space-discretised parabolic problems on a rectangular domain subject to Dirichlet boundary conditions. For the time integration s-stage AMF-W-methods, which are ADI (alternating direction implicit) type integrators, are considered. They are particularly efficient when the space dimension m of the problem is large. Optimal results on PDE-convergence have recently been obtained in [J. Comput. Appl. Math., 417:114642, 2023] for the case $m = 2$. The aim of the present work is to extend these results to arbitrary space dimension $m \geq 3$. It is explained which order statements carry over from the case $m = 2$ to $m \geq 3$, and which do not.

Résumé. Ce travail considère des problèmes paraboliques discrétisés en espace sur un domaine rectangulaire soumis à conditions aux limites de Dirichlet. Pour l'intégration temporelle, on considère les méthodes AMF-W à s étages qui sont des intégrateurs de type ADI (implicites dans des directions alternantes). Ces méthodes sont particulièrement efficaces lorsque la dimension spatiale m du problème est grande. Des résultats optimaux sur la convergence indépendamment de la résolution spatiale ont récemment été obtenus dans [J. Calcul. Appl. Math., 417:114642, 2023] pour le cas $m = 2$. L'objectif du présent travail est d'étendre ces résultats à la dimension arbitraire $m \geq 3$. Nous expliquons quels résultats sur l'ordre de convergence persistent ou non en dimension $m \geq 3$ en comparaison avec la dimension $m = 2$.

2020 Mathematics Subject Classification. 65M12, 65M15, 65M20.

August 15, 2023.

1. Introduction

This article is concerned with the time integration of linear diffusion problems

$$\partial_t u(t,\vec{x}) = \sum_{j=1}^{m} d_j\, \partial^2_{x_j x_j} u(t,\vec{x}) + c(t,\vec{x}), \qquad t \geq 0, \tag{1}$$

Keywords and phrases: multidimensional parabolic problem, ADI-type AMF-W method, PDE-convergence, order conditions, fractional order

[*] *Dedicated to the memory of Assyr Abdulle*

[**] *This work was partially supported by the Swiss National Science Foundation, grant No. 200020_192129, and by the Spanish Project PID2022-141385NB-I00 of Ministerio de Ciencia e Innovación.*

[1] Departamento de Análisis Matemático. Universidad de La Laguna. 38071, La Laguna, Spain. (spinto@ull.edu.es, dhabreu@ull.edu.es)

[2] Section de Mathématiques, Université de Genève, CH-1205 Genève, Switzerland. (Ernst.Hairer@unige.ch)

for $\vec{x} = (x_1, \dots, x_m)^\top \in [0,1]^m$, with constants $d_j > 0$, $1 \le j \le m$, and Dirichlet boundary conditions. An arbitrarily large space dimension m is admitted, and time-dependent boundary conditions are considered. The inhomogeneity $c(t, \vec{x})$ is assumed to be sufficiently smooth. A standard second order central finite difference discretisation on a uniform grid

$$x_j^{(i)} = i \cdot \Delta x_j, \qquad 0 \le i \le N_j + 1, \quad 1 \le j \le m, \tag{2}$$

with $\Delta x_j = 1/(N_j + 1)$, yields the ordinary differential equation

$$\dot{U} = D\,U + g(t), \qquad D = D_1 + \cdots + D_m, \tag{3}$$

where $D_j = d_j\,(I_{N_m} \otimes \cdots \otimes D_{x_j x_j} \otimes \cdots \otimes I_{N_1})$. Here, the differentiation matrices $D_{x_j x_j}$ are the tridiagonal Toeplitz matrices $tridiag(1, -2, 1)/\Delta x_j^2$ of dimension N_j, respectively, and $\otimes$ stands for the Kronecker product of matrices. Note that the dimension of the system (3) is $N := N_1 \cdots N_m$. Here, $U(t) = (U_{i_1,\dots,i_m}(t))_{i_1,\dots,i_m}$, where $U_{i_1,\dots,i_m}(t) \approx u(t, x_1^{(i_1)}, \dots, x_m^{(i_m)})$ approximates the solution of (1) on the grid (2). The vector $g(t) = (g_{i_1,\dots,i_m}(t))_{i_1,\dots,i_m}$ contains contributions from the inhomogeneity and from non homogeneous boundary conditions.

The numerical integration of the differential equation (3) typically requires the approximate solution of a large linear system with matrix $(I - \tau\theta D)$. An interesting approach is to replace this matrix by the product $(I - \tau\theta D_1) \cdots (I - \tau\theta D_m)$, so that only the solution of linear systems with tridiagonal matrices is needed. This idea goes back to the pioneering work by Douglas [2] and Douglas & Rachford [3]; see also the excellent review by van der Houwen & Sommeijer [13]. For such integrators PDE-convergence in the weighted euclidean norm up to order two has been addressed for 2D problems in [9] for the Peaceman–Rachford method, in [1] for a modified Douglas splitting and in [11] for problems with mixed derivatives. PDE-convergence for 1-stage AMF-W methods is considered in [4] for arbitrary space dimension m. Recently, PDE-convergence of s-stage AMF-W methods (with arbitrary $s \ge 1$) is treated in [5] for PDE problems (1) with $m = 2$.

Outline of the paper. The aim of the present work is to investigate which convergence results of [5] are valid in arbitrary space dimension $m \ge 3$, and which are valid only for $m = 2$. Section 2 introduces the class of AMF-W methods (AMF-W stands for "approximate matrix factorization with inexact Jacobian") and presents assumptions on the stability function that are essential for the study of PDE-convergence. The main convergence results – up to order 3.25 for time-independent boundary conditions and up to order 2 for time-dependent boundary conditions – are given in Section 3. Section 4 presents criteria on the method coefficients that imply the stability assumptions of the convergence theorems, and Section 5 gives a numerical confirmation of the results.

The proofs for convergence are based on three techniques: Lady Windermere's Fan, use of partial summation, and use of Fourier analysis. They are explained in Section 6 with technical details postponed to Section 9. The proofs for the main results are in Section 7 for the case of time-independent boundary conditions, and in Section 8 for time-dependent boundary conditions. Parts that extend straight-forwardly from the proofs of [5] are kept to a minimum, and parts that do not carry over from $m = 2$ to higher space dimension are emphasised.

2. AMF-W methods and their stability function

ADI-type time integrators are based on a splitting (3) of the discretised differential operator. For an application of the method also the inhomogeneity in (3) has to be split,

$$g(t) = g_1(t) + \cdots + g_m(t). \tag{4}$$

It is natural to do this in such a way that, in the case of inhomogeneous Dirichlet boundary conditions, $g_j(t)$ contains the contribution of the discretisation of $d_j \partial^2_{x_j x_j} u$, $1 \le j \le m$. The contribution of the inhomogeneity $c(t, \vec{x})$ is included in $g_1(t)$.

For the numerical solution of (3) we consider so-called AMF-W methods [5]. They are one-step methods, which permit to compute, for a given numerical approximation $U_n \approx U(t_n)$ at time t_n, the approximation $U_{n+1} \approx U(t_{n+1})$ at time $t_{n+1} = t_n + \tau$ as follows

$$\begin{aligned} K_i^{(0)} &= \tau D\Big(U_n + \sum_{j=1}^{i-1} a_{ij} K_j\Big) + \tau\, g(t_n + c_i\tau) + \sum_{j=1}^{i-1} \ell_{ij} K_j, \\ (I - \theta\tau D_k) K_i^{(k)} &= K_i^{(k-1)} + \theta\rho_i \tau^2 \dot{g}_k(t_n + \eta\tau), \quad k = 1, 2, \dots, m, \\ K_i &= K_i^{(m)}, \quad i = 1, 2, \dots, s, \\ U_{n+1} &= U_n + \sum_{i=1}^{s} b_i K_i. \end{aligned} \tag{5}$$

The method is characterised by the coefficients (A, L, b, θ, η), where $A = (a_{i,j})_{j<i}$ and $L = (\ell_{i,j})_{j<i}$ are lower triangular matrices, $b = (b_i)_i$ is a vector of dimension s, and $\theta > 0$, $\eta \geq 0$ are two constants. The coefficients ρ_i and c_i are defined by $\rho_i = 1 + \sum_{j=1}^{i-1} \ell_{ij}\rho_j$ and $c_i = \sum_{j=1}^{i-1} a_{ij}\rho_j$. In vector form they are

$$\rho = (I - L)^{-1}\mathbf{1} \qquad \text{and} \qquad c = A\rho, \tag{6}$$

where $\mathbf{1}$ denotes the vector of dimension s with all entries equal to 1. We also use the notation $c^r = (c_i^r)_i$.

Stability function. The propagation matrix of the one-step method is obtained by removing the inhomogeneities $g_j(t), j = 1, \dots, m$, from the formulas (5). This yields

$$U_{n+1} = R(\tau D_1, \dots, \tau D_m) U_n$$

where, using the Kronecker product notation,

$$\begin{aligned} R(\tau D_1, \dots, \tau D_m) &= I + (b^\top \otimes I) P(\tau D_1, \dots, \tau D_m)^{-1} (\mathbf{1} \otimes \tau D), \\ P(\tau D_1, \dots, \tau D_m) &= I \otimes \pi(\tau D_1, \dots, \tau D_m) - A \otimes \tau D - L \otimes I, \\ \pi(\tau D_1, \dots, \tau D_m) &= (I - \theta\tau D_1) \cdots (I - \theta\tau D_m). \end{aligned} \tag{7}$$

Note that the matrices D_j commute pairwise. Substituting τD_j by a scalar (real or complex) variable z_j, we obtain the so-called stability function

$$\begin{gathered} R(z_1, \dots, z_m) = 1 + b^\top P(z_1, \dots, z_m)^{-1}\mathbf{1} \cdot (z_1 + \cdots + z_m), \qquad \text{where} \\ P(z_1, \dots, z_m) = \pi(z_1, \dots, z_m) I - (z_1 + \cdots + z_m) A - L, \qquad \pi(z_1, \dots, z_m) = (1 - \theta z_1) \cdots (1 - \theta z_m). \end{gathered} \tag{8}$$

Here, $P(z_1, \dots, z_m)$ is a triangular matrix of dimension s. Since we consider only purely diffusion problems in the present work, we assume z_j to be real variables.

A necessary condition for unconditional convergence is that the stability function $R(z_1, \dots, z_m)$ satisfies

$$-1 \leq R(z_1, \dots, z_m) \leq 1, \qquad \text{for} \quad z_1, \dots, z_m \leq 0. \tag{9}$$

This property is usually called A_0-stability of the method. To get optimal convergence orders we also consider slightly stronger properties

$$-1 \leq R(z_1, \dots, z_m) \leq 1 - C_1 \frac{|z_1 + \ldots + z_m|}{\pi(z_1, \dots, z_m)}, \qquad \text{for} \quad z_1, \dots, z_m \leq 0, \tag{10}$$

or

$$-1 \le R(z_1,\dots,z_m) \le 1 - C_1 \frac{|z_1+\ldots+z_m|}{\pi(z_1,\dots,z_m)^2}, \qquad \text{for} \quad z_1,\dots,z_m \le 0, \tag{11}$$

where $C_1 > 0$ is a positive constant[1]. Criteria for these stability requirements in terms of the coefficients of the AMF-W method are given in Section 4.

Special case $s = 1$. The most simple integrator among the methods (5) is obtained for $s = 1$ and $\eta = 0$. It is given by

$$\begin{aligned} &K^{(0)} = \tau D U_n + \tau\, g(t_n), \\ &(I - \theta\tau D_k)K^{(k)} = K^{(k-1)} + \theta\tau^2 \dot g_k(t_n), \quad k = 1,2,\dots,m, \\ &U_{n+1} = U_n + K^{(m)}, \end{aligned} \tag{12}$$

and it is related to the θ-method. This method is of order $p = 1$ for all values of the parameter $\theta \ge 0$. It is of order $p = 2$ for $\theta = 1/2$. The stability function of this method is

$$R(z_1,\dots,z_m) = 1 + \frac{z_1+\cdots+z_m}{(1-\theta z_1)\cdots(1-\theta z_m)}.$$

For $\theta \ge 1/2$, it not only satisfies the A_0-condition (9), but it also satisfies the condition (10), and consequently also (11).

3. Main PDE-convergence results

For vectors $U = (U_{i_1,\dots,i_m})$ and $V = (V_{i_1,\dots,i_m})$ in $\mathbf{R}^N$ (where $i_j = 1,\dots,N_j$) we consider the weighted inner product and the induced ℓ_2 norm

$$\|U\| = \sqrt{\langle U,U\rangle}, \qquad \langle U,V\rangle = \Delta x_1 \cdots \Delta x_m \sum_{i_1=1}^{N_1} \cdots \sum_{i_m=1}^{N_m} U_{i_1,\dots,i_m} V_{i_1,\dots,i_m}. \tag{13}$$

We call a one-step method PDE-convergent of order p, if the global error satisfies

$$\|U_n - U(t_n)\| \le C\tau^p \qquad \text{for} \quad 0 \le t_n = n\tau \le T, \tag{14}$$

where the constant C is independent of the space discretisation. In the following we also call a method PDE-convergent of order p^*, if it is PDE-convergent of order $p - \epsilon$ for all fixed $\epsilon > 0$. The constant C in (14) now depends on ϵ and tends to infinity for $\epsilon \to 0$.

We are interested in to which extent the classical order conditions are relevant for PDE-convergence. Using the notation

$$\tilde A = A(I-L)^{-1}, \quad \tilde b^\top = b^\top (I-L)^{-1}, \quad \tilde\Gamma = \theta(I-L)^{-1} \tag{15}$$

together with ρ and c from (6), we obtain the following classical order conditions (see [5] or [8, p. 114-117])

- order $p = 1 \iff \tilde b^\top \mathbf{1} = 1.$
- order $p = 2 \iff \tilde b^\top \mathbf{1} = 1$ and $\tilde b^\top(\tilde A + \tilde\Gamma)\mathbf{1} = 1/2.$
- order $p = 3 \iff \tilde b^\top \mathbf{1} = 1,\ \ \tilde b^\top \tilde A \mathbf{1} = 1/2,\ \ \tilde b^\top \tilde\Gamma \mathbf{1} = 0$ and $\tilde b^\top c^2 = 1/3,\ \ \tilde b^\top(\tilde A + \tilde\Gamma)^2 \mathbf{1} = 1/6.$

In the following convergence statements we assume that the time step size satisfies $\tau \ge c_0 \max(\Delta x_1^2,\dots,\Delta x_m^2)$ for some $c_0 > 0$, which is a natural assumption when applying linearly implicit integration methods. This is motivated by the fact that, if the product of the time step size τ with the Lipschitz constant of the vector field

[1] Throughout the paper, $C, C_0, C_1, \dots$ will stand for positive constants independent of τ and Δx_j, $1 \le j \le m$, which may take different values at each appearance.

(3) (which is proportional to the inverse of $\min(\Delta x_1^2, \ldots, \Delta x_m^2)$) is small or of moderate size, an explicit time integrator would be more efficient.

Theorem 3.1 (time-independent boundary conditions)**.** *Consider the s-stage AMF-W method* (5) *applied to* (3) *in space dimension $m \geq 2$ with time-independent Dirichlet boundary conditions. Assume that*

- *the classical conditions for order p ($p \leq 3$) hold,*
- *the stability condition* (9) *holds if $p = 1$ or $p = 2$, condition* (11) *holds if $p = 3$,*

then the AMF-W method is PDE-convergent of order p.

The statement of this theorem does not hold for order $p = 4$, even in the case that the classical conditions for order four are satisfied. However, we can improve the order by 0.25, if the following subset of 4th order conditions are imposed:

$$\begin{aligned} &\tilde{b}^\top (\tilde{A} + \tilde{\Gamma})\, \tilde{\Gamma} \mathbf{1} = 0, \\ &\tilde{b}^\top (\tilde{A} + \tilde{\Gamma})^3 \mathbf{1} = 1/24, \\ &\tilde{b}^\top \tilde{\Gamma}\, (\tilde{A} + \tilde{\Gamma}) \mathbf{1} = 0, \\ &\tilde{b}^\top (\tilde{A} + \tilde{\Gamma})\, c^2 = 1/12, \\ &\tilde{b}^\top c^3 = 1/4. \end{aligned} \tag{16}$$

Theorem 3.2 (time-independent boundary conditions)**.** *Consider the s-stage AMF-W method* (5) *applied to* (3) *in space dimension $m \geq 2$ with time-independent Dirichlet boundary conditions. Assume that*

- *the classical conditions for order $p = 3$ and the five order conditions* (16) *hold,*
- *the stability condition* (11) *holds,*

then the AMF-W method is PDE-convergent of order $p^* = 3.25$. Here, PDE-convergence of order p^* means PDE-convergence of order $p - \epsilon$, for all fixed $\epsilon > 0$.

For the general case of time-dependent Dirichlet boundary conditions, the derivatives $\dot{g}_j(t)$ typically contain terms with negative powers of Δx_j, $1 \leq j \leq m$, which lead to an order reduction. In Theorem 3.3 below, we replace the assumption (11) by the slightly stronger assumption (10).

Theorem 3.3 (time-dependent boundary conditions)**.** *Consider the s-stage AMF-W method* (5) *applied to* (3) *in space dimension $m \geq 2$ with time-dependent Dirichlet boundary conditions. Assume that*

- *the classical conditions for order p ($p \leq 2$) hold,*
- *the stability condition* (9) *holds if $p = 1$, condition* (10) *holds if $p = 2$,*

then the AMF-W method is PDE-convergent of order p for $p = 1$, and of order p^* for $p = 2$. Again, PDE-convergence of order p^* means PDE-convergence of order $p - \epsilon$, for all fixed $\epsilon > 0$.

The previous three theorems are an extension of Theorems 1, 3, and 6 of [5] from the case $m = 2$ to arbitrary space dimension $m \geq 2$. Note that the improvement from order $p^* = 2$ to order $p = 2$ in the situation of Theorem 3.3 extends to $m \geq 2$ if the step size satisfies $\tau \leq c_1 \min(\Delta x_1, \ldots, \Delta x_m)$ with some $c_1 > 0$ (Remark 1 of [5]). If the condition (10) is relaxed to (11) in Theorem 3.3 the convergence order drops from $p^* = 2$ to $p^* = 1.5$ (Remark 2 of [5]).

However, the statement of Theorem 5 of [5] (convergence of order $p^* = 2.25$ for time-dependent boundary conditions) does not carry over to space dimension $m \geq 3$ (see Section 5 for a numerical confirmation).

4. Stability criteria for AMF-W methods

AMF-W methods are in general constructed in such a way that not only (9) holds, but that the stability function satisfies $|R(z_1, \ldots, z_m)| \leq 1$ for z_j in a complex neighbourhood of the negative real line. For methods of order $p \geq 1$ the maximum principle for analytic functions then implies

$$-1 \leq R(z_1, \ldots, z_m) < 1, \qquad \text{for} \quad z_1, \ldots, z_m \leq 0, \quad z_1 + \ldots + z_m < 0. \tag{17}$$

Additional assumptions are necessary for satisfying (10) or (11). To obtain criteria in terms of the coefficients of an s-stage AMF-W method we write the stability function as

$$R(z_1,\ldots,z_m) = 1 + \frac{z_1+\ldots+z_m}{\pi(z_1,\ldots,z_m)}\, b^\top (I-X)^{-1}\mathbf{1}, \qquad X = \frac{z_1+\ldots+z_m}{\pi(z_1,\ldots,z_m)}\, A + \frac{1}{\pi(z_1,\ldots,z_m)}\, L \tag{18}$$

and we consider the real-valued function

$$q(X) = \sum_{k=0}^{s-1} b^\top X^k\, \mathbf{1}, \tag{19}$$

defined for real $s\times s$ matrices X. Note that $q(\zeta A)$ is a scalar polynomial of degree $s-1$ of the real variable ζ.

Lemma 4.1. *Consider an AMF-W method* (5) *with coefficients* (A, L, b, θ, η) *of classical order* $p \geq 1$, *and assume that the stability function* $R(z_1,\ldots,z_m)$ *satisfies* (17). *We then have:*

(A1) *if* $q(\zeta A) > 0$ *for* $\zeta \in [-\theta^{-1}, 0]$, *then the stability function satisfies* (10);

(A2) *if* $q(\zeta A) > 0$ *for* $\zeta \in (-\theta^{-1}, 0]$, *and* $q'(-\theta^{-1}A)(\theta^{-1}A + L) > 0$ *if* $q(-\theta^{-1}A) = 0$, *then the stability function satisfies* (11).

Proof. (A1) The left inequality of (10) is satisfied by (17). For the right inequality we have to prove that

$$G_1(z_1,\ldots,z_m) := \Big(R(z_1,\ldots,z_m) - 1\Big)\frac{\pi(z_1,\ldots,z_m)}{z_1+\ldots+z_m} = b^\top\big(I + X + \ldots + X^{s-1}\big)\mathbf{1} = q(X)$$

satisfies $G_1(z_1,\ldots,z_m) \geq C_1 > 0$ for all $z_1,\ldots,z_m \leq 0$. At the origin we have $X = L$ which implies $G_1(0,\ldots,0) = b^\top(I-L)^{-1}\mathbf{1} = \tilde b^\top \mathbf{1} = 1$, because $p \geq 1$. In the limit $z_m \to -\infty$ we obtain

$$G_1(z_1,\ldots,z_{m-1},-\infty) = b^\top\big(I + \zeta A + \ldots + \zeta^{s-1}A^{s-1}\big)\mathbf{1} = q(\zeta A), \qquad \zeta = \frac{1}{(-\theta)(1-\theta z_1)\cdot\ldots\cdot(1-\theta z_{m-1})}.$$

Since $\zeta \in [-\theta^{-1}, 0]$ for non positive z_j, the assumption on the polynomial $q(\zeta A)$ implies the existence of a constant $c > 0$, such that $G_1(z_1,\ldots,z_{m-1},-\infty) \geq c$. By symmetry, $G_1(z_1,\ldots,z_m) \geq c$ holds if at least one element among $z_1,\ldots,z_m$ is equal to $-\infty$. Therefore, there exists $K > 0$ (typically very large), such that $G_1(z_1,\ldots,z_m) > c/2$ if at least one element among $z_1,\ldots,z_m$ satisfies $z_j \leq -K$. The positivity of $G_1(z_1,\ldots,z_m)$ on the compact set $[-K,0]^m$ proves the existence of a constant $C_1 > 0$ such that $G_1(z_1,\ldots,z_m) \geq C_1$ for all $z_j \leq 0$. This proves the inequality (10).

(A2) If $q(\zeta A)$ vanishes for $\zeta = -\theta^{-1}$ we see that $G_1(z_1,\ldots,z_m)$ vanishes only if one argument is $-\infty$ and all other arguments are zero. We therefore consider the function

$$G_2(z_1,\ldots,z_m) := \Big(R(z_1,\ldots,z_m) - 1\Big)\frac{\pi(z_1,\ldots,z_m)^2}{z_1+\ldots+z_m} = \pi(z_1,\ldots,z_m)\, q(X),$$

which satisfies $G_2(z_1,\ldots,z_m) \geq G_1(z_1,\ldots,z_m)$ for all $z_j \leq 0$. We have to prove that $G_2(z_1,\ldots,z_m) > 0$ whenever $G_1(z_1,\ldots,z_m) = 0$. Without loss of generality we assume $z_1 = \ldots = z_{m-1} = 0$ and $z_m \to -\infty$. Since $X = \zeta A + z_m^{-1}\zeta\,(\theta^{-1}A + L) + \mathcal{O}(z_m^{-2})$, and $q(-\theta^{-1}A) = 0$, we have by Taylor series expansion that

$$G_2(0,\ldots,0,z_m) = q'(-\theta^{-1}A)(\theta^{-1}A + L) + \mathcal{O}(z_m^{-1}) \qquad \text{for} \quad z_m \to -\infty,$$

which, in the limit, is strictly positive by our assumption. Consequently, we conclude the existence of $C_1 > 0$, such that $G_2(z_1,\ldots,z_m) \geq C_1$ for all $z_j \leq 0$. This proves the estimate (11). □

The stability criteria of Lemma 4.1 are independent of the space dimension m and identical to those of [5]. In [5] it has been verified that the 2-stage W-methods collected in [10, p. 400-405] for $\theta > 1/4$, the 3-stage AMF-W methods in [7, Theorem 1] for $\theta > 1/3$, the 4-stage AMF-W methods in [6, Corollary 1 and Theorem

3] for $\theta > \theta_1 = 0.36367\ldots$, the 3- and 4-stage AMF-W methods in [12] all satisfy (A1). The 2-stage W-method with $\theta = 1/4$, the methods of [7, Theorem 1] with $s = p = 3$ and $\theta = 1/3$, and the methods in [6, Corollary 1 and Theorem 3] with $s = 4$, $p = 3$, and $\theta = \theta_1$ satisfy (A2).

5. Numerical experiments

For a numerical confirmation of the convergence statements in Section 3 we consider the linear diffusion partial differential equation (1) with $d_j = 1$ for all j, where $c(t, \vec{x})$ is selected in such way that

$$u(t, \vec{x}) = u_e(t, \vec{x}) := \mathrm{e}^t \Big(4^m \prod_{j=1}^{m} x_j (1 - x_j) + \kappa \sum_{j=1}^{m} \Big(x_j + \frac{1}{j+2} \Big)^2 \Big) \tag{20}$$

is the exact solution of (1). We impose the initial condition $u(0, \vec{x}) = u_e(0, \vec{x})$ and Dirichlet boundary conditions. If $\kappa = 0$ we have homogeneous boundary conditions, but when $\kappa = 1$ we get non-homogeneous time-dependent Dirichlet conditions.

For a linear parabolic problem (1) the space discretization (3) introduces an error of size $\mathcal{O}(\Delta x_1^2 + \ldots + \Delta x_m^2)$. The error of the full discretization is a combination of the spatial error and that of the time integrator. When using a time step size τ that is proportional to $\min(\Delta x_1, \ldots, \Delta x_m)$, results on PDE-convergence of order up to 2 are important, but also higher order convergence is of interest for the case when the error of the space discretization is small.

The present work is devoted to an analysis of the time integration error. Therefore we have chosen the exact solution (20) for our test equation as a polynomial of degree 2 in each spatial variable so that the global errors come only from the time discretisation. We apply the MOL approach on a uniform grid with meshwidth $\Delta x_j = 1/(N+1)$, $1 \le j \le m$. This yields the ordinary differential equation (3) in dimension N^m. AMF-W-methods (5) with $\eta = 1/2$ will be applied to (3) with fixed step size $\tau = \Delta x_j = 2^{-j}$, $2 \le j \le 8$.

$N+1$	PDE-GE2	PDE-ORD2
4	0.4298E-01	—
8	0.4334E-02	3.310
16	0.5555E-03	2.964
32	0.7369E-04	2.914
64	0.7807E-05	3.239
128	0.7399E-06	3.399
256	0.6919E-07	3.419

$N+1$	PDE-GE2	PDE-ORD2
4	0.5217E-01	—
8	0.7818E-02	2.738
16	0.1277E-02	2.614
32	0.2341E-03	2.448
64	0.4838E-04	2.274
128	0.1076E-04	2.169
256	0.2470E-05	2.123

Table 1. Statistics for $m = 3$ and the AMF-W-method ($s = 4$, $p = 4$, $\theta = 1/2$) based on Kutta's 3/8-rule [6, p. 154], with $\kappa = 0$ (left table) and with $\kappa = 1$ (right table).

All our numerical experiments are for space dimension $m = 3$. In Table 1 we apply the 4-stage AMF-W-method ($s = 4$, $p = 4, \theta = 1/2$) based on Kutta's 3/8-rule [6, p. 154]. We consider homogeneous Dirichlet boundary conditions ($\kappa = 0$, table to the left) and time-dependent boundary conditions ($\kappa = 1$, table to the right). The global error is shown in the column "PDE-GE2" and the numerically estimated order of convergence in "PDE-ORD2". This experiment suggests that order 3.25 can be attained for time-independent boundary conditions, and that the order drops to 2 for time-dependent boundary conditions.

In Table 2 we consider time-dependent inhomogeneous Dirichlet boundary conditions ($\kappa = 1$) and apply the 3-stage AMF-W method ($s = 3$, $p = 3$) based on the `W3a` method in [7, p. 573] with $\theta = 1/2$ and $\theta = 1/3$. Note that the method satisfies (10) for $\theta = 1/2$, but it satisfies only (11) for $\theta = 1/3$. The results for $\theta = 1/2$ (left table) show that order 2.25 is in general not attained by 3rd order methods for time-dependent boundary

$N+1$	PDE-GE2	PDE-ORD2	$N+1$	PDE-GE2	PDE-ORD2
4	0.1908E-01	—	4	0.1815E-01	—
8	0.4071E-02	2.229	8	0.7884E-02	1.203
16	0.1894E-02	1.104	16	0.2650E-02	1.573
32	0.5703E-03	1.732	32	0.9645E-03	1.458
64	0.1618E-03	1.817	64	0.3480E-03	1.471
128	0.4242E-04	1.932	128	0.1235E-03	1.494
256	0.1036E-04	2.034	256	0.4365E-04	1.501

TABLE 2. Statistics for $m = 3$ and $\kappa = 1$ with the AMF-W method ($s = 3$, $p = 3$) based on the `W3a` method in [7, p. 573] with $\theta = 0.5$ (left table) and with $\theta = 1/3$ (right table).

$N+1$	PDE-GE2	PDE-ORD2	$N+1$	PDE-GE2	PDE-ORD2
4	0.4806E-01	—	4	0.4802E-01	—
8	0.1223E-01	1.974	8	0.1253E-01	1.939
16	0.3451E-02	1.825	16	0.3613E-02	1.793
32	0.9982E-03	1.790	32	0.1088E-02	1.732
64	0.2858E-03	1.804	64	0.3389E-03	1.682
128	0.7964E-04	1.843	128	0.1101E-03	1.623
256	0.2121E-04	1.909	256	0.3702E-04	1.572

TABLE 3. Statistics for $m = 3$ and $\kappa = 1$ with the AMF-W-method based on the 2-stage W-methods ($p = 2$) in [10, p. 400–405] with $\theta = 0.26$ (left table) and $\theta = 1/4$ (right table).

conditions, even for $\eta = 1/2$. This experiment demonstrates that Theorem 5 of [5] does not extend to an order higher than 2. The results for $\theta = 1/3$ (right table) show that only order 1.5 is attained for time-dependent boundary conditions if (10) is not fulfilled.

In Table 3 we consider again time-dependent Dirichlet boundary conditions ($\kappa = 1$), but we apply the AMF-W-method based on the 2-stage W-methods ($p = 2$) in [10, p. 400–405] with $\theta = 0.26$ (left table) and $\theta = 1/4$ (right table). For the parameter $\theta = 1/4$ only the condition (11) (and not (10)) is fullfilled. As expected (see the comment after Theorem 3.3), we only have convergence of order 1.5. A small perturbation of this parameter to $\theta = 0.26$ reestablishes condition (10) and, as expected, PDE-convergence of order 2 is approached.

6. Local error and convergence theorems

The global error $E_n = U_n - U(t_n)$ of a AMF-W method (5), applied to the linear differential equation (3), satisfies the recursion

$$E_{n+1} = R(\tau D_1, \dots, \tau D_m)\, E_n - \nu_n, \tag{21}$$

where $R(z_1, \dots, z_m)$ is the stability function of the method, and ν_n is the local error at time t_n. To analyse the error in the Euclidean norm we use a Fourier-type analysis and diagonalise the matrices $D_{x_i x_i} = tridiag(1, -2, 1)/\Delta x_i^2$ that are present in D_i (see the paragraph after formula (3)). We let $\{\phi_j^{(x_i)}\}$ for $j = 1, \dots, N_i$ ($i = 1, \dots, m$) be an orthogonal basis such that $D_{x_i x_i}\phi_j^{(x_i)} = \lambda_j^{(x_i)}\phi_j^{(x_i)}$ with eigenvalues $\lambda_j^{(x_i)} = -4\Delta x_i^{-2}\sin^2(j\Delta x_i\pi/2)$. The set $\{\phi_{j_m}^{(x_m)} \otimes \cdots \otimes \phi_{j_1}^{(x_1)}\}$ then forms an orthonormal basis with respect to the inner product (13), and we have $D_i(\phi_{j_m}^{(x_m)} \otimes \cdots \otimes \phi_{j_1}^{(x_1)}) = d_i \lambda_{j_i}^{(x_i)}(\phi_{j_m}^{(x_m)} \otimes \cdots \otimes \phi_{j_1}^{(x_1)})$. When expanding a vector

$V \in \mathbf{R}^{n_x}$ in this basis

$$V = \sum_{j_1=1}^{n_{x_1}} \cdots \sum_{j_m=1}^{n_{x_m}} \widehat{V}_{j_1,\ldots,j_m}\, \phi_{j_m}^{(x_m)} \otimes \cdots \otimes \phi_{j_1}^{(x_1)}, \tag{22}$$

we denote the Fourier coefficients by $\widehat{V}_{j_1,\ldots,j_m}$. Recall that by Parseval's identity we have $\|V\| = \|\widehat{V}\|_2$, where $\widehat{V} = (\widehat{V}_{j_1,\ldots,j_m})_{j_1,\ldots,j_m}$. In Fourier coefficients the recursion for the global error reads

$$\widehat{E}_{n+1,j_1,\ldots,j_m} = R(z_1,\ldots,z_m)\, \widehat{E}_{n,j_1,\ldots,j_m} - \widehat{\nu}_{n,j_1,\ldots,j_m}, \tag{23}$$

where $z_i = \tau d_i \lambda_{j_i}^{(x_i)}$, $1 \le i \le m$.

6.1. **Local error**

The local error of (5) is obtained by considering U_n on the exact solution $U(t_n)$, so that $E_n = 0$ in (21). We omit the details of the straight-forward computation, because they are essentially the same as for the case $m = 2$ in [5]. In the following we use the notation

$$\begin{aligned} z^{[k]} &= z_1 + \ldots + z_{k-1}, & z &= z^{[m+1]} = z_1 + \ldots + z_m, \\ \pi^{[k]} &= (1-\theta z_1)\cdots(1-\theta z_{k-1}), & \pi &= \pi^{[m+1]} = (1-\theta z_1)\cdots(1-\theta z_m) \end{aligned} \tag{24}$$

and we consider the triangular matrix, see (8),

$$S(z_1,\ldots,z_m) := P(z_1,\ldots,z_m)^{-1} = \big(\pi I - zA - L\big)^{-1}. \tag{25}$$

The Fourier coefficients of the local error ν_n for general Dirichlet boundary conditions are then

$$\begin{aligned} \widehat{\nu}_{n,j_1,\ldots,j_m} = &\sum_{\ell\ge 1} \tau^\ell \Big(\frac{1}{\ell!} - b^\top S(z_1,\ldots,z_m)\Big(\alpha^{(\ell)} + \beta^{(\ell)} z + \gamma^{(\ell)}(\pi + \theta z - 1)\Big)\Big) \widehat{U}^{(\ell)}_{j_1,\ldots,j_m}(t_n) \\ &+ \sum_{\ell\ge 1} \tau^{\ell+1} b^\top S(z_1,\ldots,z_m)\gamma^{(\ell)}\theta \sum_{k=2}^{m} (1-\pi^{[k]})\widehat{\varphi}^{(\ell)}_{k,j_1,\ldots,j_m}(t_n). \end{aligned} \tag{26}$$

Here, $\widehat{U}^{(\ell)}_{j_1,\ldots,j_m}(t_n)$ are the Fourier coefficients of the ℓth derivative of the exact solution, and $\widehat{\varphi}^{(\ell)}_{k,j_1,\ldots,j_m}(t_n)$ are those of the ℓth derivative of

$$\varphi_k(t) = D_k U(t) + g_k(t). \tag{27}$$

Note that, in contrast to $g_k(t)$, the functions $\varphi_k(t)$ are smooth in the sense that their derivatives have bounds in the Euclidean norm that are independent of the space discretisation. The coefficients $\alpha^{(\ell)} = (\alpha_i^{(\ell)})_{i=1}^s$, $\beta^{(\ell)} = (\beta_i^{(\ell)})_{i=1}^s$, $\gamma^{(\ell)} = (\gamma_i^{(\ell)})_{i=1}^s$ are given by $\alpha_i^{(1)} = 1$, and

$$\begin{aligned} \alpha_i^{(\ell)} &= \tfrac{1}{(\ell-1)!} c_i^{\ell-1} + \theta\rho_i \tfrac{1}{(\ell-2)!}\eta^{\ell-2}, \quad \ell \ge 2, \\ \beta_i^{(\ell)} &= -\tfrac{1}{\ell!} c_i^{\ell} - \theta\rho_i \tfrac{1}{(\ell-1)!}\eta^{\ell-1}, \quad \ell \ge 1, \\ \gamma_i^{(\ell)} &= \rho_i \tfrac{1}{(\ell-1)!}\eta^{\ell-1}, \quad \ell \ge 1. \end{aligned} \tag{28}$$

In the situation of time-independent Dirichlet boundary conditions the functions $g_k(t)$, $2 \le k \le m$, are constant, so that $\varphi_k^{(\ell)}(t) = D_k U^{(\ell)}(t)$ for $\ell \ge 1$ and $2 \le k \le m$. Using the algebraic identity

$$\theta \sum_{k=2}^{m} (1-\pi^{[k]}) z_k = \pi + \theta z - 1$$

this implies, as it has been seen for $m = 2$ in [5], that the terms with factor $\gamma^{(\ell)}$ cancel. The Fourier coefficients of the local error, for time-independent boundary conditions, are thus given by

$$\widehat{\nu}_{n,j_1,\ldots,j_m} = \sum_{\ell\geq 1}\tau^\ell\Big(\frac{1}{\ell!} - b^\top S(z_1,\ldots,z_m)\big(\alpha^{(\ell)} + \beta^{(\ell)} z\big)\Big)\widehat{U}^{(\ell)}_{j_1,\ldots,j_m}(t_n). \tag{29}$$

6.2. Techniques for PDE-convergence

Since for the considered integrators the global error depends linearly on the local error, it is possible to treat summands in the local error separately. In the following ν_n and its Fourier coefficients are typically only parts of the complete local error. Lemma 6.1 and Lemma 6.2 (see [10, Section II.2.3]) are standard techniques for proving convergence. Lemma 6.3 is new and extends [5, Theorem 9] to arbitrary space dimension.

Lemma 6.1 (Lady Windermere's fan). *Assume that the local error satisfies*

$$|\widehat{\nu}_{n,j_1,\ldots,j_m}| \leq C\tau^{r+1}|\widehat{\chi}_{j_1,\ldots,j_m}(t_n)|,$$

where $\widehat{\chi}_{j_1,\ldots,j_m}(t)$ are the Fourier coefficients of some derivative of some smooth function $\chi(t)$. Under the stability assumption (9) *the global error satisfies*

$$\|E_n\| \leq C_1\tau^r \qquad \text{for} \quad n\tau \leq T.$$

The proof is based on $E_n = R^n E_0 - \sum_{j=0}^{n-1} R^{n-j-1}\nu_j$.

Lemma 6.2 (use of partial summation). *Assume that the local error satisfies*

$$|\widehat{\nu}_{n,j_1,\ldots,j_m}| \leq C\tau^r\big|1 - R(z_1,\ldots,z_m)\big| \cdot |\widehat{\chi}_{j_1,\ldots,j_m}(t_n)|, \tag{30}$$

where $z_i = \tau d_i \lambda^{(x_i)}_{j_i}$ and $\widehat{\chi}_{j_1,\ldots,j_m}(t_n)$ as in Lemma 6.1. Furthermore, assume that $\widehat{\nu}_{n+1,j_1,\ldots,j_m} - \widehat{\nu}_{n,j_1,\ldots,j_m}$ satisfies the same estimate with one additional factor τ. Under the stability assumption (9) *the global error then satisfies*

$$\|E_n\| \leq C_1\tau^r \qquad \text{for} \quad n\tau \leq T.$$

The proof is based on $E_n = R^n E_0 - \sum_{j=-1}^{n-2}(I - R^{n-1-j})(I-R)^{-1}(\nu_{j+1} - \nu_j)$ (with $\nu_{-1} = 0$) which follows from the above error relation by partial summation. Both Lemmata only require power-boundedness of $R = R(\tau D_1,\ldots,\tau D_m)$ and are not restricted to the Euclidean norm. For the verification of (30) the stability assumptions (10) and (11) are helpful.

A further technique (developed in [4] and applied for 2D problems in [5]) is required to get sharp estimates for the global error. The weak step size restriction $\tau \geq c_0 \max(\Delta x_1^2,\ldots,\Delta x_m^2)$ will be assumed to hold.

Lemma 6.3 (use of Fourier analysis). *Assume that the local error satisfies*

$$|\widehat{\nu}_{n,j_1,\ldots,j_m}| \leq C\,\tau^r \frac{|z_1|\cdots|z_k|}{\pi(z_1,\ldots,z_m)^{\gamma_1}}|\widehat{\chi}_{j_1,\ldots,j_m}(t_n)|, \tag{31}$$

where z_i $(1 \leq i \leq m)$ and $\widehat{\chi}_{j_1,\ldots,j_m}(t_n)$ are as in Lemma 6.2, and γ_1 is a non-negative constant. Moreover, assume that the difference $\widehat{\nu}_{n+1,j_1,\ldots,j_m} - \widehat{\nu}_{n,j_1,\ldots,j_m}$ satisfies the same estimate with one additional factor τ.

- *Under the Assumption (10), we have, for $n\tau \leq T$,*

$$\|E_n\| \leq C_1\tau^{p_k^*} \qquad \text{with} \qquad p_k = r + \frac{(k-2)^2}{4(k-1)}, \qquad \text{if} \quad \gamma_1 = 1 \ \text{ and } \ 2 \leq k \leq m. \tag{32}$$

If $k = 2$ and $\tau \leq c_1 \min(\Delta x_1,\ldots,\Delta x_m)$, with some $c_1 > 0$, then $\|E_n\| \leq C_1\tau^{p_2}$.

- *Under the weaker Assumption (11) we still have, for $n\tau \le T$,*

$$\|E_n\| \le C_1 \tau^{p_k^*} \qquad \text{with} \qquad p_k = r + \begin{cases} \frac{(k-2)^2-2}{2(2k-1)} & \text{if } \gamma_1 = 1 \text{ and } 1 \le k \le m, \\ \frac{(k-1)^2+1}{2(2k-1)} & \text{if } \gamma_1 = 2 \text{ and } 2 \le k \le m. \end{cases} \tag{33}$$

If $\gamma_1 = 1$ and $k = 1$, and $\tau \le c_1 \min(\Delta x_1, \dots, \Delta x_m)$, with some $c_1 > 0$, then $\|E_n\| \le C_1 \tau^{p_1}$.

The proof is technical and postponed to an appendix (Section 9). It uses the special form of $z_i = \tau d_i \lambda_{j_i}^{(x_i)}$, and the decay of Fourier coefficients.

7. Proof of Theorems 3.1 and 3.2 (time-independent boundary conditions)

For the case of time-independent boundary conditions the local error (29) remains essentially the same as for two spatial dimensions. The only difference is that $z = z_1 + \ldots + z_m$ contains more than two summands, and the matrix $S(z_1, \dots, z_m)$ depends on all m variables. It is not surprising that the statements of [5] carry over straight-forwardly to arbitrary space dimensions.

Recall that in the Euclidean norm (13) we have, for $\ell \ge 1$ and $1 \le i_1 < \ldots < i_k \le m$,

$$D_{i_1} \cdots D_{i_k} U^{(\ell)}(t) = \big(d_{i_1} \partial^2_{x_{i_1} x_{i_1}} + \mathcal{O}(\Delta x_{i_1}^2)\big) \cdots \big(d_{i_k} \partial^2_{x_{i_k} x_{i_k}} + \mathcal{O}(\Delta x_{i_k}^2)\big) \frac{\partial^\ell u(t, \vec{x})}{\partial t^\ell} = \mathcal{O}(1),$$

so that in Fourier coefficients

$$z_{i_1} \cdots z_{i_k} \, \widehat{U}^{(\ell)}_{j_1,\dots,j_m}(t) = \tau^k \, \widehat{\chi}_{j_1,\dots,j_m}(t), \tag{34}$$

where $\chi(t) = D_{i_1} \cdots D_{i_k} U^{(\ell)}(t)$. In the following we only mention which of the three convergence lemmata have to be applied, and leave the details to the reader.

7.1. **Proof of Theorem 3.1.**

For proving order $p = 1$ and order $p = 2$ only Lemma 6.1 is needed. For the proof of order $p = 3$ all three lemmata have to be used. Note that in the formulas of [5, Section 7] the product $\theta^2 z_1 z_2$ has to be replaced by $\big(\pi(z_1, \dots, z_m) + \theta(z_1 + \ldots + z_m) - 1\big)$ if $m \ge 3$. This expression is a linear combination of products $z_{i_1} \cdots z_{i_k}$, where $1 \le i_1 < \ldots < i_k \le m$ and $k \ge 2$. Because of (34) the critical case is for $k = 2$. The local error term δ_1^b in [5, Section 7.3] can be treated, similar to the two-dimensional case, by applying Lemma 6.3 with $k = 2$ and $\gamma_1 = 2$, giving an order $3 + 1/3 - \varepsilon$ for every $\varepsilon > 0$.

7.2. **Proof of Theorem 3.2.**

The proof of fractional order $p = 3 + \alpha$ (for every $\alpha < 1/4$) uses the estimates

$$\Big(|z_1|^\alpha + \ldots + |z_m|^\alpha\Big)\Big(|z_1| + \ldots + |z_m|\Big)\Big|\widehat{U}^{(\ell)}_{j_1,\dots,j_m}(t)\Big| \le \tau^{1+\alpha} \, v_{j_1,\dots,j_m} \tag{35}$$

$$\Big(|z_1|^\alpha + \ldots + |z_m|^\alpha\Big)\Big|\pi(z_1, \dots, z_m) + \theta\big(z_1 + \ldots + z_m\big) - 1\Big|\Big|\widehat{U}^{(\ell)}_{j_1,\dots,j_m}(t)\Big| \le \tau^{2+\alpha} \, w_{j_1,\dots,j_m} \tag{36}$$

with bounded $\sum_{j_1,\dots,j_m} |v_{j_1,\dots,j_m}|^2$ and $\sum_{j_1,\dots,j_m} |w_{j_1,\dots,j_m}|^2$, where we use the notation $z_i = \tau d_i \lambda_{j_i}^{(x_i)}$ as before. These estimates are a consequence of the relation (34) and of [10, Lemma III.6.5]. The proof of Theorem 3.2 extends straight-forwardly from space dimension $m = 2$ (see [5]) to arbitrary m. Apart from the term that has been shown above to give an order $3 + 1/3 - \varepsilon$ contribution (which is better than order 3.25), only the convergence estimates of Lemmata 6.1 and 6.2 have to be applied.

8. Proof of Theorem 3.3 (time-dependent boundary conditions)

For the case of time-dependent boundary conditions the extension of the convergence results to more than two spatial dimensions is not as straight-forward. We shall see that not all convergence results, valid in dimension two, carry over to higher dimension.

In the following we use the notation (24) and we write S for the matrix $S(z_1,\dots,z_m)$ of (25). We then express the local error (26) as $\widehat{\nu}_{n,j_1,\dots,j_m}=\sum_{\ell\ge1}(\delta_\ell+\omega_\ell^a+\omega_\ell^b)$, where

$$\begin{aligned}\delta_\ell &= \tau^\ell\Big(\frac{1}{\ell!}-b^\top S\Big(\alpha^{(\ell)}+\beta^{(\ell)}z+\gamma^{(\ell)}\big(\pi+\theta z-1\big)\Big)\Big)\,\widehat{U}^{(\ell)}_{j_1,\dots,j_m}(t_n),\\ \omega_\ell^a &= \tau^{\ell+1}\,b^\top S\gamma^{(\ell)}\theta^2\sum_{k=2}^m z^{[k]}\,\widehat{\varphi}^{(\ell)}_{k,j_1,\dots,j_m}(t_n),\\ \omega_\ell^b &= -\tau^{\ell+1}\,b^\top S\gamma^{(\ell)}\theta\sum_{k=2}^m(\pi^{[k]}+\theta z^{[k]}-1)\,\widehat{\varphi}^{(\ell)}_{k,j_1,\dots,j_m}(t_n),\end{aligned}\tag{37}$$

and we observe that $\omega_\ell^b=0$ for spatial dimension $m=2$, but not for $m\ge3$.

8.1. Proof of Theorem 3.3

The proof for order $p=1$ is a straight-forward extension of the case $m=2$. It requires only the application of Lemma 6.1. Similarly, also the proofs of order $p^*=2$ of the global errors corresponding to δ_ℓ and ω_ℓ^a are extended straight-forwardly. Care has to be taken for the term ω_1^b, which is not present for $m=2$. It is seen to be a linear combination of terms that can be bounded by (31) with $r=2$, $2\le k\le m$, and $\gamma_1=1$, so that the statement (32) of Lemma 6.3 applies. This completes the proof of Theorem 3.3.

8.2. Reason why the fractional order $p^*=2.25$ cannot be achieved for $m\ge3$

It is not a surprise that the impossibility for obtaining order of convergence $p^*=2.25$ is due to the term ω_1^b in the local error. For time-dependent boundary conditions we do not have an estimate (35) with $U(t)$ replaced by $\varphi_k(t)$, we only have

$$\Big(|z_1|^\alpha+\ldots+|z_m|^\alpha\Big)\Big|\widehat{\varphi}^{(\ell)}_{k,j_1,\dots,j_m}(t)\Big|\le\tau^\alpha\,v_{j_1,\dots,j_m}\tag{38}$$

with bounded $\sum_{j_1,\dots,j_m}|v_{j_1,\dots,j_m}|^2$. Similar to the case $m=2$ we can split $|z_1|=|z_1|^{1-\alpha}|z_1|^\alpha$ in ω_1^b and apply (38). This gives the desired factor $\tau^{2+\alpha}$, but the remaining expression in ω_1^b cannot be uniformly bounded by $|z|/|\pi|$. Therefore, Lemma 6.2 cannot be applied. This does not change, if we modify ω_1^b with the use of $b^\top S(0,\dots,0)\gamma^{(1)}=0$ as in the proof for the case $m=2$. In any case, the numerical experiment of Table 2 demonstrates the impossibility of obtaining $p^*=2.25$.

9. Appendix: proof of Lemma 6.3

The proof follows the reasoning of [4, Section 4.2]. Using the identity

$$E_n=-(I-R^n)(I-R)^{-1}\nu_0-\sum_{j=0}^{n-2}(I-R^{n-1-j})(I-R)^{-1}(\nu_{j+1}-\nu_j)$$

we have

$$\|E_n\|\le C\,\tau^r a(n)+C\,\tau^{r+1}\sum_{j=0}^{n-2}a(n-1-j),\tag{39}$$

where

$$a(n)=\tau^k\Big\{\sum_{j_1=1}^{n_{x_1}}\cdots\sum_{j_m=1}^{n_{x_m}}\Big(\frac{1-r_{j_1,\dots,j_m}^n}{1-r_{j_1,\dots,j_m}}\Big)^2\frac{|\lambda_{j_1}^{(x_1)}|^2\cdots|\lambda_{j_k}^{(x_k)}|^2|\widehat{\chi}_{j_1,\dots,j_m}(t_n)|^2}{(1+\theta\tau d_1|\lambda_{j_1}^{(x_1)}|)^{2\gamma_1}\cdots(1+\theta\tau d_m|\lambda_{j_m}^{(x_m)}|)^{2\gamma_1}}\Big\}^{1/2}$$

with $r_{j_1,\dots,j_m}=R\big(\tau d_1\lambda_{j_1}^{(x_1)},\dots,\tau d_m\lambda_{j_m}^{(x_m)}\big)$. It follows from [4, Lemma A.6 and Lemma A.2] that for an arbitrarily chosen $\gamma\in[0,2]$ we have

$$\Big(\frac{1-r_{j_1,\dots,j_m}^n}{1-r_{j_1,\dots,j_m}}\Big)^2\le\tau^{-\gamma}2^{2-\gamma}T^\gamma\frac{1}{(1-r_{j_1,\dots,j_m})^{2-\gamma}}\qquad\text{and}\qquad\sqrt{|\lambda_{j_1}^{(x_1)}|\cdots|\lambda_{j_m}^{(x_m)}|}\,|\widehat{\chi}_{j_1,\dots,j_m}(t_j)|\le C_2,$$

where $n\tau\le T$. This implies

$$a(n)\le C_3\tau^k\Big\{\sum_{j_1=1}^{n_{x_1}}\cdots\sum_{j_m=1}^{n_{x_m}}\frac{\tau^{-\gamma}}{(1-r_{j_1,\dots,j_m})^{2-\gamma}}\frac{\big(\prod_{l=1}^k|\lambda_{j_l}^{(x_l)}|\big)\big(\prod_{l=k+1}^m|\lambda_{j_l}^{(x_l)}|^{-1}\big)}{\prod_{l=1}^m(1+\theta\tau d_l|\lambda_{j_l}^{(x_l)}|)^{2\gamma_1}}\Big\}^{1/2}.\tag{40}$$

9.1. **Use of Assumption** (10)

Under the Assumption (10), and with $\gamma_1=1$, we get from (40) the estimate

$$a(n)\le C_4\tau^k\Big\{\tau^{-2}\sum_{j_1=1}^{n_{x_1}}\cdots\sum_{j_m=1}^{n_{x_m}}\frac{\big(\prod_{l=1}^k|\lambda_{j_l}^{(x_l)}|\big)\big(\prod_{l=k+1}^m|\lambda_{j_l}^{(x_l)}|^{-1}\big)}{\big(\sum_{l=1}^m|\lambda_{j_l}^{(x_l)}|\big)^{2-\gamma}\prod_{l=1}^m(1+\theta\tau d_l|\lambda_{j_l}^{(x_l)}|)^{\gamma}}\Big\}^{1/2}.\tag{41}$$

With the help of the arithmetic-geometric mean inequality

$$\sum_{l=1}^m|\lambda_{j_l}^{(x_l)}|\ge\sum_{l=1}^k|\lambda_{j_l}^{(x_l)}|\ge k\cdot\sqrt[k]{|\lambda_{j_1}^{(x_1)}|\cdots|\lambda_{j_k}^{(x_k)}|}$$

the latter sum of products can be bounded by a product of sums, and yields

$$a(n)\le C_5\tau^{k-1}\Big\{\prod_{l=1}^k\Big(\sum_{j_l=1}^{n_{x_l}}\frac{|\lambda_{j_l}^{(x_l)}|^{1-(2-\gamma)/k}}{\big(1+\theta\tau d_l|\lambda_{j_l}^{(x_l)}|\big)^\gamma}\Big)\prod_{l=k+1}^m\Big(\sum_{j_l=1}^{n_{x_l}}\frac{1}{|\lambda_{j_l}^{(x_l)}|}\Big)\Big\}^{1/2}.\tag{42}$$

As a consequence of [10, Lemma 6.2, p. 298], the second product in (42) is bounded, i.e., $\sum_{j_l=1}^{n_{x_l}}|\lambda_{j_l}^{(x_l)}|^{-1}=\mathcal{O}(1)$. Regarding the first product, we are in the position to apply Lemma A.5 from [4], which states that for all $\widetilde{\alpha}\ge0$ and $\widetilde{\beta}\ge0$ there exists a constant $C>0$, such that

$$\sum_{j_l=1}^{n_{x_l}}\frac{|\lambda_{j_l}|^{\widetilde{\alpha}/2}}{(1+\theta\tau d_l|\lambda_{j_l}|)^{\widetilde{\beta}}}\le\begin{cases}C\tau^{-(\widetilde{\alpha}+1)/2} & \text{if}\quad\widetilde{\alpha}+1-2\widetilde{\beta}<0,\\ C\tau^{-\widetilde{\beta}}\big(\Delta x_l\big)^{2\widetilde{\beta}-\widetilde{\alpha}-1} & \text{if}\quad\widetilde{\alpha}+1-2\widetilde{\beta}>0.\end{cases}\tag{43}$$

The first product in (42) is a product of k identical single sums of the form (43) with $\widetilde{\alpha}=2\big(1-\frac{2-\gamma}{k}\big)$ and $\widetilde{\beta}=\gamma$. With the choice $\gamma=\frac{3k-4}{2(k-1)}+\epsilon$ (with $\epsilon>0$ arbitrarily small such that $\gamma\in[0,2]$, for $k\ge2$), we have $\widetilde{\alpha}+1-2\widetilde{\beta}<0$, and from (43) and (42) we get that that $(\widetilde{\alpha}+1)/2=\frac{3k-4}{2(k-1)}+\frac{1}{k}\epsilon$ and $a(n)\le C_6\tau^{\frac{(k-2)^2}{4(k-1)}-\frac{1}{2}\epsilon}$. This proves the statement (32) under the Assumption (10).

Furthermore, when $k=2$ and $\tau \le c_1 \min(\Delta x_1, \dots, \Delta x_m)$, taking $\gamma = 1-\epsilon$ (with $\epsilon > 0$ arbitrarily small such that $\gamma \in [0,2]$), we have $\widetilde{\alpha} + 1 - 2\widetilde{\beta} > 0$. From (43) we get, for every $\gamma < 1$, that

$$\sum_{j_l=1}^{n_{x_l}} \frac{|\lambda_{j_l}|^{\gamma/2}}{(1+\theta\tau d_l|\lambda_{j_l}|)^{\gamma}} \le C\tau^{-\gamma}\big(\Delta x_l\big)^{\gamma-1} = C\tau^{-1}(\tau/\Delta x_l)^{1-\gamma} \le C\tau^{-1}.$$

This estimate inserted in (42) with $k=2$ provides $a(n) = \mathcal{O}(1)$ and $\|E_n\| \le C_1\tau^{p_2}$, with $p_2 = r$.

9.2. **Use of Assumption** (11)

Under the weaker Assumption (11), starting with (40) and using (11) the same computation as above leads to

$$a(n) \le C_4\tau^k \Bigg\{\tau^{-2}\sum_{j_1=1}^{n_{x_1}}\cdots\sum_{j_m=1}^{n_{x_m}} \frac{\big(\prod_{l=1}^k |\lambda_{j_l}^{(x_l)}|\big)\big(\prod_{l=k+1}^m |\lambda_{j_l}^{(x_l)}|^{-1}\big)}{\big(\sum_{l=1}^m |\lambda_{j_l}^{(x_l)}|\big)^{2-\gamma}\prod_{l=1}^m (1+\theta\tau d_l|\lambda_{j_l}^{(x_l)}|)^{2\gamma+2\gamma_1-4}}\Bigg\}^{1/2}, \tag{44}$$

which, analogously as in (42), can be bounded by

$$a(n) \le C_5\tau^{k-1}\Bigg\{\prod_{l=1}^k\Bigg(\sum_{j_l=1}^{n_{x_l}} \frac{|\lambda_{j_l}^{(x_l)}|^{1-(2-\gamma)/k}}{\big(1+\theta\tau d_l|\lambda_{j_l}^{(x_l)}|\big)^{2\gamma+2\gamma_1-4}}\Bigg)\Bigg\}^{1/2}. \tag{45}$$

This is again a product of factors of the form (43) with $\widetilde{\alpha} = 2\big(1-\frac{2-\gamma}{k}\big)$ and $\widetilde{\beta} = 2\gamma + 2\gamma_1 - 4$.

When $\gamma_1 = 2$, and for $2 \le k \le m$, taking $\gamma = \frac{3k-4}{2(2k-1)} + \epsilon$ ($\epsilon > 0$ arbitrary) gives $\widetilde{\alpha} + 1 - 2\widetilde{\beta} < 0$, and from (45) and the corresponding estimate in (43) we get that $(\widetilde{\alpha}+1)/2 = \frac{3k-4}{2k-1} + \frac{1}{k}\epsilon$ and $a(n) \le C_6\tau^{\frac{(k-1)^2+1}{2(2k-1)} - \frac{1}{2}\epsilon}$. This proves the second estimate in (33).

When $\gamma_1 = 1$, and for $1 \le k \le m$, taking $\gamma = \frac{7k-4}{2(2k-1)} + \epsilon$ ($\epsilon > 0$ arbitrary) gives $\widetilde{\alpha} + 1 - 2\widetilde{\beta} < 0$, and from (45) and the corresponding estimate in (43) we get that that $(\widetilde{\alpha}+1)/2 = \frac{3k-2}{2k-1} + \frac{1}{k}\epsilon$ and $a(n) \le C_6\tau^{\frac{(k-2)^2-2}{2(2k-1)} - \frac{1}{2}\epsilon}$. This proves the first estimate in (33).

Finally, when $\gamma_1 = 1$, $k=1$ and $\tau \le c_1\min(\Delta x_1, \dots, \Delta x_m)$, taking $\gamma = \frac{3}{2} - \epsilon$ (with arbitrary $\epsilon > 0$ such that $\gamma \in [0,2]$), we have $\widetilde{\alpha} + 1 - 2\widetilde{\beta} > 0$. From (43) we get, for every $\gamma < \frac{3}{2}$, that

$$\sum_{j_l=1}^{n_{x_l}} \frac{|\lambda_{j_l}|^{\gamma-1}}{(1+\theta\tau d_l|\lambda_{j_l}|)^{2\gamma-2}} \le C\tau^{-(2\gamma-2)}\big(\Delta x_l\big)^{2\gamma-3} = C\tau^{-1}(\tau/\Delta x_l)^{3-2\gamma} \le C\tau^{-1}.$$

This estimate inserted in (45) with $k=1$ provides $a(n) = \mathcal{O}(\tau^{-1/2})$ and $\|E_n\| \le C_1\tau^{p_1}$, with $p_1 = r - 1/2$. This proves the statement of Lemma 6.3. □

10. Conclusion

For the case of multidimensional linear parabolic differential equations, we have discussed PDE-convergence of AMF-W methods.

For time-independent boundary conditions, the classical conditions for order p and a stability condition that is slightly stronger than A_0-stability imply PDE-convergence (in any space dimension) of order p, if $p \le 3$. Such a statement does not hold for $p = 4$, but order $p = 3.25 - \epsilon$ (for all $\epsilon > 0$) can be achieved if a subset of order conditions for order four hold.

For time-dependent boundary conditions, the classical conditions for order p, a stability condition that it slightly stronger than A_0-stability, and a step size restriction that is common for the study of PDE-convergence imply PDE-convergence of order p, if $p \le 2$. Imposing the order conditions for order three, order $p = 2.25 - \epsilon$ (for all $\epsilon > 0$) can be achieved in space dimension $m = 2$, but not in space dimension $m \ge 3$.

References

[1] A. Arrarás, K. J. in 't Hout, W. Hundsdorfer, and L. Portero. Modified Douglas splitting methods for reaction-diffusion equations. *BIT*, 57(2):261–285, 2017.

[2] J. Douglas, Jr. On the numerical integration of $\partial^2 u/\partial x^2 + \partial^2 u/\partial y^2 = \partial u/\partial t$ by implicit methods. *J. Soc. Indust. Appl. Math.*, 3:42–65, 1955.

[3] J. Douglas, Jr. and H. H. Rachford, Jr. On the numerical solution of heat conduction problems in two and three space variables. *Trans. Amer. Math. Soc.*, 82:421–439, 1956.

[4] S. González-Pinto, E. Hairer, and D. Hernández-Abreu. Convergence in ℓ_2 and in ℓ_∞ norm of one-stage AMF-W-methods for parabolic problems. *SIAM J. Numer. Anal.*, 58(2):1117–1137, 2020.

[5] S. González-Pinto, E. Hairer, and D. Hernández-Abreu. High order PDE-convergence of AMF-W methods for 2D-linear parabolic problems. *J. Comput. Appl. Math.*, 417:Paper No. 114642, 18, 2023.

[6] S. González-Pinto, D. Hernández-Abreu, and S. Pérez-Rodríguez. W-methods to stabilize standard explicit Runge-Kutta methods in the time integration of advection-diffusion-reaction PDEs. *J. Comput. Appl. Math.*, 316:143–160, 2017.

[7] S. González-Pinto, D. Hernández-Abreu, S. Pérez-Rodríguez, and R. Weiner. A family of three-stage third order AMF-W-methods for the time integration of advection diffusion reaction PDEs. *Applied Mathematics and Computation*, 274:565–584, 2016.

[8] E. Hairer and G. Wanner. *Solving Ordinary Differential Equations II. Stiff and Differential-Algebraic Problems*. Springer Series in Computational Mathematics 14. Springer-Verlag, Berlin, 2nd edition, 1996.

[9] W. Hundsdorfer and J.G. Verwer. Stability and convergence of the Peaceman-Rachford ADI method for initial-boundary value problems. *Math. Comp.*, 53(187):81–101, 1989.

[10] W. Hundsdorfer and J.G. Verwer. *Numerical solution of time-dependent advection-diffusion-reaction equations*, volume 33 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 2003.

[11] K. J. in 't Hout and M. Wyns. Convergence of the modified Craig-Sneyd scheme for two-dimensional convection-diffusion equations with mixed derivative term. *J. Comput. Appl. Math.*, 296:170–180, 2016.

[12] J. Rang and L. Angermann. New Rosenbrock W-methods of order 3 for partial differential algebraic equations of index 1. *BIT*, 45(4):761–787, 2005.

[13] P. J. van der Houwen and B. P. Sommeijer. Approximate factorization for time-dependent partial differential equations. *J. Comput. Appl. Math.*, 128(1-2):447–466, 2001. Numerical analysis 2000, Vol. VII, Partial differential equations.