Final Version

Runge–Kutta time discretization of nonlinear parabolic equations studied via discrete maximal parabolic regularity

Peer C. Kunstmann · Buyang Li · Christian Lubich

Communicated by Hans Munthe-Kaas

Abstract For a large class of fully nonlinear parabolic equations, which include gradient flows for energy functionals that depend on the solution gradient, the semidiscretization in time by implicit Runge–Kutta methods such as the Radau IIA methods of arbitrary order is studied. Error bounds are obtained in the $W^{1,\infty}$ norm uniformly on bounded time intervals and, with an improved approximation order, in the parabolic energy norm. The proofs rely on discrete maximal parabolic regularity. This is used to obtain $W^{1,\infty}$ estimates, which are the key to the numerical analysis of these problems.

Keywords Runge–Kutta method · maximal parabolic regularity · nonlinear parabolic equation · gradient flow · stability · error bounds

Mathematics Subject Classification (2000) 65M12 · 65M15 · 65L06

1 Introduction

This paper is concerned with the stability and error analysis of implicit Runge–Kutta time discretizations of nonlinear parabolic initial-boundary value prob-

Peer C. Kunstmann
Institut für Analysis, KIT, 76131 Karlsruhe, Germany.
E-mail: peer.kunstmann@kit.edu

Buyang Li
Department of Applied Mathematics, The Hong Kong Polytechnic University,
Hung Hom, Kowloon, Hong Kong.
E-mail: buyang.li@polyu.edu.hk

Christian Lubich
Mathematisches Institut, University of Tübingen,
Auf der Morgenstelle 10, 72076 Tübingen, Germany.
E-mail: lubich@na.uni-tuebingen.de

lems for $u = u(x,t)$,

$$\frac{\partial u}{\partial t} = \nabla \cdot f(\nabla u, u), \qquad x \in \Omega,\ 0 < t \le T, \tag{1.1}$$

on a given bounded smooth domain $\Omega \subset \mathbb{R}^d$ of arbitrary dimension $d \ge 1$ and for a given final time $T > 0$, taken with homogeneous Dirichlet boundary conditions $u = 0$ on $\partial\Omega \times [0,T]$ and with given initial data $u(\cdot,0) = u_0$ on Ω.

The flux function $f : \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^d$ is assumed to be a smooth function satisfying a local ellipticity condition: for every $(p,u) \in \mathbb{R}^d \times \mathbb{R}$, the matrix

$$\partial_p f(p,u) \in \mathbb{R}^{d\times d} \text{ has a positive definite symmetric part.} \tag{1.2}$$

We do not require uniform ellipticity with respect to all $(p,u) \in \mathbb{R}^d \times \mathbb{R}$: some eigenvalues of the symmetric part $\frac{1}{2}(\partial_p f(p,u) + \partial_p f(p,u)^T)$ may tend to 0 or $+\infty$ as $|(p,u)| \to \infty$.

We will, however, assume that the initial-boundary value problem admits a sufficiently regular solution, and we ask for stability and rates of convergence of numerical discretizations in this case.

For a solution to the initial-boundary value problem that is bounded in $W^{1,\infty}(\Omega)$ (that is, both the solution and its gradient are bounded with respect to the maximum norm), we have uniform ellipticity along the exact solution by compactness in the finite-dimensional space $\mathbb{R}^d \times \mathbb{R}$, but it is by no means obvious that also the numerical approximation stays bounded in $W^{1,\infty}(\Omega)$ uniformly in the discretization parameters. Establishing such $W^{1,\infty}$ bounds for the numerical discretization is a main difficulty for this problem.

The problem (1.1) occurs in many applications, such as the following where actually $f(p,u) = f(p)$ does not depend on u:

- minimal surface flow [28,34] and the regularized models of total variation flow [10,11,25], where
$$f(p) = \frac{p}{\sqrt{\lambda^2 + |p|^2}}.$$
- More generally, with $f(p) = \nabla_p F(p)$ for a smooth convex function $F : \mathbb{R}^d \to \mathbb{R}$, (1.1) appears as the $L^2(\Omega)$ gradient flow,
$$(\partial_t u, v)_{L^2(\Omega)} = -E'(u)v \qquad \text{for all } v \text{ in a dense subspace of } H_0^1(\Omega),$$
for the energy functional $E(u) = \int_\Omega F(\nabla u)\mathrm{d}x$; see, e.g., [9, Section 9.6.3].

The problem (1.1) also includes quasilinear equations, where $f(p,u) = A(u)p$ with a positive definite matrix $A(u)$, which may degenerate as $|u| \to \infty$.

Due to the strong nonlinearity of the equation, existing works on error estimates of the time discretization of (1.1) are very limited. Feng and Prohl [11] have proved optimal-order convergence rate of the finite element solution of the regularized total variation flow with an implicit Euler scheme, under the time step size restriction $\tau = o(h^2)$, which was used to control the numerical solution in the $W^{1,\infty}$ norm via the inverse inequality. Convergence of the numerical solution was proved in [10] without time step size restriction, without explicit convergence rate. By using the methodology of [24], Li and Sun presented optimal-order L^2-norm error estimates for the finite element solution of the minimal surface flow with a linearized semi-implicit Euler scheme, without restriction on the time step size [25]. Since their proof is based on the L^2-norm error estimate, they have assumed that the order of finite elements are greater than one in order to control the $W^{1,\infty}$ norm of the numerical solution via the inverse inequality.

Existing works on the problem are all restricted to implicit Euler time discretization, with the only exception of the paper by Ostermann and Thalhammer [32], where convergence properties of implicit Runge–Kutta semidiscretizations for a class of fully nonlinear parabolic equations are analyzed in weighted Hölder spaces. While this is a very remarkable work, it is not obvious that it applies to the class of problems considered here (especially in higher dimensions), and even less so to full discretizations with finite element approximations in space.

In this paper we study semidiscretization of (1.1) in time by implicit Runge–Kutta methods such as the collocation methods based on the Radau nodes, which have excellent stability properties, allow for arbitrarily high order and can be implemented efficiently [15, Chapter IV]. To emphasize the basic techniques and to keep the paper at a reasonable length, we do not include the effect of space discretization by finite elements in our stability and error analysis. We note, however, that in considering only time discretization we cannot use inverse estimates, which are often convenient, but are restricted to quasi-uniform meshes and moreover lead to restrictions as indicated in the previous paragraph. It is thus of interest to develop techniques that do not rely on inverse estimates. Our results are new even for the case of the backward Euler time discretization. This paper may provide a foundation for further analysis of fully discrete approximations of the problem.

In Section 2 we describe the temporal semidiscretization by implicit Runge–Kutta methods and present our main results, which are error bounds in the $W^{1,\infty}$ norm and, with a higher approximation order, in the energy norm. The proof of these results forms the remainder of the paper.

Section 3 presents a sequence of auxiliary results related to maximal L^p regularity, which is the basic technique for obtaining our stability and error bounds. Discrete maximal L^p regularity was shown for the backward Euler method by Ashyralyev, Piskarev & Weis [4], for higher-order A-stable (and $A(\alpha)$-stable) multistep and Runge–Kutta time discretizations by Kovács, Li & Lubich [18], and for the θ-scheme by Kemmochi [16] and Kemmochi & Saito [17]. Discrete maximal L^p regularity up to a factor logarithmic in the step size was given by Leykekhman & Vexler [21] for discontinuous Galerkin time discretizations. The above-mentioned results relate to linear problems. Discrete maximal L^p regularity was applied to the error analysis of time discretizations of reaction-diffusion equations in [18], of Ginzburg-Landau equations in [22], and of quasilinear parabolic equations in [2]. Discrete maximal L^p regularity of semidiscrete finite element solutions of parabolic equations was investigated in

[12, 23, 27], with applications to semilinear and quasilinear parabolic equations in [13, 26].

The proof of the error bound in the $W^{1,\infty}$ norm is given in Section 4, that of the improved error bound in the energy norm in Section 5.

2 Runge–Kutta time discretization and statement of the main results

We consider the time discretization of (1.1) with constant step size $\tau > 0$ (this could be relaxed to a fixed number of changes of the step size) by an implicit Runge-Kutta method with properties that are, in particular, satisfied by the s-stage Radau IIA method [15, Section IV.5], which is the collocation method at the Radau nodes (with right-most node $c_s = 1$) and can also be viewed as a fully discretized discontinuous Galerkin method in time [3]. We require the following properties (cf. [15, Section IV.3] for these notions):

$$\begin{aligned}&\text{The Runge–Kutta method is A-stable,}\\ &\text{it has an invertible coefficient matrix } (a_{ij})_{i,j=1}^{s}\\ &\text{and its weights satisfy } b_j = a_{sj} \ \ (j = 1, \dots, s).\end{aligned} \tag{2.1}$$

We let $t_n = n\tau$ for $n \geq 0$ (as long as t_n does not exceed the final time T) and set $t_{n,i} = t_n + c_i\tau$, where $c_i = \sum_{j=1}^{s} a_{ij}$ are the nodes of the Runge–Kutta method, with $c_s = 1$ so that $t_{n+1} = t_{n,s}$.

We denote by $u_{n,i}$ $(i = 1, \dots, s)$ the internal stages and by u_n the solution approximation at the grid point t_n. The last condition in (2.1) ensures that

$$u_{n+1} = u_{n,s}. \tag{2.2}$$

The time discretization of (1.1) is then determined by the equations

$$u_{n,i} = u_n + \tau \sum_{j=1}^{s} a_{ij} \nabla \cdot f(\nabla u_{n,j}, u_{n,j}) \qquad (i = 1, \dots, s) \tag{2.3}$$

together with the Dirichlet boundary conditions $u_{n,i} = 0$ on $\partial\Omega$. These equations are to be solved subsequently for $n = 0, 1, 2, \dots$

Remark 2.1 Further finite element discretization of (2.3) can be done in the following way: find $u_{n,i}^h$ in the finite element space S_h such that

$$(u_{n,i}^h, v_h) = (u_n^h, v_h) - \tau \sum_{j=1}^{s} a_{ij} \Big(f(\nabla u_{n,j}^h, u_{n,j}^h), \nabla v_h \Big) \qquad \forall\, v_h \in S_h,$$

and $u_{n+1}^h = u_{n,s}^h$. For the efficient implementation of the fully discrete Runge–Kutta equations, using systems of linear equations of just the dimension of S_h, we refer to [15, Section IV.8]. In this paper, we focus on the time discretization (2.3).

We recall the notion of *stage order*, cf. [15, p. 226]: The Runge–Kutta method has stage order k if for each $i = 1, \dots, s$,

$$\sum_{j=1}^{s} a_{ij} c_j^{l-1} = \frac{c_i^l}{l}, \qquad l = 1, \dots, k. \tag{2.4}$$

In particular, the stage order of the s-stage Radau IIA method (as of any collocation method with polynomials of degree s) is $k = s$.

The stage order determines to what order the internal stages $u_{n,i}$ approximate the exact solution values $u(t_{n,i})$, and to what order the derivative approximations

$$\dot{u}_{n,j} := \nabla \cdot f(\nabla u_{n,j}, u_{n,j}), \qquad j = 1, \dots, s, \tag{2.5}$$

approximate the exact solution derivatives $\partial_t u(t_{n,j})$, provided the solution is sufficiently regular in time.

To simplify the notation, we define the following vectors:

$$\mathbf{u}_n := (u_{n,i})_{i=1}^s, \qquad \dot{\mathbf{u}}_n := (\dot{u}_{n,i})_{i=1}^s, \tag{2.6}$$

$$u(\mathbf{t}_n) := (u(t_{n,i}))_{i=1}^s, \qquad \mathbf{t}_n := (t_{n,i})_{i=1}^s. \tag{2.7}$$

We can now state our first main result, which in particular controls the $W^{1,\infty}(\Omega)$ norm of the internal stages uniformly over the bounded time interval.

Theorem 2.1 *Consider a Runge–Kutta method of stage order k that satisfies* (2.1)*, such as the Radau IIA method with $s = k$ stages. Assuming that the solution u of* (1.1) *is sufficiently regular, i.e.,*

$$u \in C^{k+1}\big([0,T]; L^q(\Omega)\big) \cap C\big([0,T]; W^{2,q}(\Omega)\big), \quad \textit{for some } q > d, \tag{2.8}$$

there exists a positive constant $\bar{\tau}$ (depending on f, T, $\|u\|_{C^{k+1}([0,T];L^q(\Omega))}$ and $\|u\|_{C([0,T];W^{2,q}(\Omega))}$) such that for $0 < \tau < \bar{\tau}$ the discrete problem (2.3) *admits a unique solution that satisfies*

$$\max_{0 \le n \le N} \big(\|\mathbf{u}_n - u(\mathbf{t}_n)\|_{L^\infty(\Omega)^s} + \|\nabla \mathbf{u}_n - \nabla u(\mathbf{t}_n)\|_{L^\infty(\Omega)^{ds}} \big) \le C\tau^k, \tag{2.9a}$$

$$\Big(\sum_{n=0}^{N} \tau \|\dot{\mathbf{u}}_n - \partial_t u(\mathbf{t}_n)\|_{L^q(\Omega)^s}^p + \sum_{n=0}^{N} \tau \|\mathbf{u}_n - u(\mathbf{t}_n)\|_{W^{2,q}(\Omega)^s}^p \Big)^{\frac{1}{p}} \le C_{p,q}\tau^k, \tag{2.9b}$$

$$\textit{for all} \;\; 1 < p < \infty.$$

The constants C and $C_{p,q}$ are independent of τ and N with $N\tau \le T$.

The proof of Theorem 2.1 is based on discrete maximal parabolic regularity and will be presented in Section 4. For simplicity, we carry out the proof for the special case $f(\nabla u, u) = f(\nabla u)$. The proof for the general case is similar but contains additional lower order terms, which do not pose substantial difficulties in the analysis but clutter the formulas.

Using Theorem 2.1 together with energy estimates, the order of approximation can be improved to $k+1$ in the energy norm provided that the Runge-Kutta method satisfies the following two extra conditions:

- The method is *algebraically stable*, that is,

$$\begin{aligned}&\text{the weights } b_i \text{ are all positive and}\\ &\text{the } s\times s \text{ matrix with entries } b_i a_{ij} + b_j a_{ji} - b_i b_j \text{ is positive semidefinite.}\end{aligned} \tag{2.10}$$

- The quadrature formula with weights b_i and nodes c_i has at least order $k+1$:

$$\sum_{i=1}^{s} b_i c_i^{l-1} = \frac{1}{l}, \qquad l = 1, \dots, k+1. \tag{2.11}$$

This is satisfied for the Radau IIA methods with $s \geq 2$ stages since they satisfy equations (2.11) for $l \leq 2s-1$ and are algebraically stable; see [15, Section IV.12]. We will prove the following result in Section 5.

Theorem 2.2 *Consider a Runge–Kutta method of stage order k that satisfies* (2.1)*,* (2.10) *and* (2.11)*, such as the Radau IIA method with $s = k \geq 2$ stages. Assume that the solution u of* (1.1) *is sufficiently regular, i.e., satisfies* (2.8) *and*

$$u \in H^{k+1}\big(0,T; H_0^1(\Omega)\big) \cap H^{k+2}\big(0,T; H^{-1}(\Omega)\big). \tag{2.12}$$

Then, for $0 < \tau < \bar{\tau}$ (with $\bar{\tau}$ from Theorem 2.1) the solution of the discrete problem (2.3) *satisfies*

$$\max_{1\leq n\leq N} \|u_n - u(t_n)\|_{L^2(\Omega)} + \left(\sum_{n=0}^{N} \tau \|\nabla \mathbf{u}_n - \nabla u(\mathbf{t}_n)\|^2_{L^2(\Omega)^{ds}}\right)^{\frac{1}{2}} \leq C_2 \tau^{k+1}. \tag{2.13}$$

The constant C_2 is independent of τ and N with $N\tau \leq T$.

3 Auxiliary results related to maximal L^p regularity

The key to the error bounds of Theorems 2.1 and 2.2 is to control the $W^{1,\infty}(\Omega)$ norm of the numerical solution. In this paper, this is done using the space-time Sobolev inequality, for $2/p + d/q < 1$ and $v \in W^{1,p}(0,T; L^q(\Omega)) \cap L^p(0,T; W^{2,q}(\Omega))$ such that $v(0) = 0$,

$$\|v\|_{L^\infty(0,T;W^{1,\infty}(\Omega))} \leq c_{p,q}\big(\|\partial_t v\|_{L^p(0,T;L^q(\Omega))} + \|v\|_{L^p(0,T;W^{2,q}(\Omega))}\big) \tag{3.1}$$

with a constant $c_{p,q}$ that is independent of T, together with the observation that the norm on the right-hand side is what is controlled by *maximal L^p regularity* for the solution of a linear parabolic problem with a second-order elliptic differential operator. Maximal L^p regularity is characterized by Weis [35] in terms of the *R-boundedness* of the resolvent on a sector, a property that also yields *discrete maximal ℓ^p regularity* for the Runge–Kutta time discretization uniformly in the step size [18]. In this section we present some results from this range of ideas and techniques. These results follow by suitably combining various results scattered in the literature. They will be important in the proof of Theorem 2.1 and are also of independent interest.

3.1 A Sobolev embedding

Lemma 3.1 *If $2/p + d/q < 1$, then the following embeddings hold:*

$$W^{1,p}(\mathbb{R}_+; L^q(\Omega)) \cap L^p(\mathbb{R}_+; W^{2,q}(\Omega)) \hookrightarrow L^\infty(\mathbb{R}_+; W^{1,\infty}(\Omega)), \tag{3.2}$$

$$W^{1,p}(0,T; L^q(\Omega)) \cap L^p(0,T; W^{2,q}(\Omega)) \hookrightarrow L^\infty(0,T; W^{1,\infty}(\Omega)), \tag{3.3}$$

where the second embedding is compact.

Remark 3.1 The first embedding in Lemma 3.1 implies the bound (3.1) with $c_{p,q}$ independent of T, for all $v \in W^{1,p}(0,T; L^q(\Omega)) \cap L^p(0,T; W^{2,q}(\Omega))$ such that $v(0) = 0$ (one can extend v to $W^{1,p}(\mathbb{R}_+; L^q(\Omega)) \cap L^p(\mathbb{R}_+; W^{2,q}(\Omega))$ by reflection on $[T, 2T]$ and zero extension on $[2T, \infty)$).

Proof Via Sobolev embedding, we have

$$\begin{aligned}
&W^{1,p}(\mathbb{R}_+; L^q(\Omega)) \cap L^p(\mathbb{R}_+; W^{2,q}(\Omega)) \\
&\hookrightarrow L^\infty(\mathbb{R}_+; (L^q(\Omega), W^{2,q}(\Omega))_{1-1/p,p}) \qquad \text{see [31, Proposition 1.2.10]} \\
&= L^\infty(\mathbb{R}_+; B^{2-2/p;q,p}(\Omega)) \qquad \text{by the definition of Besov spaces [1, §7.32].}
\end{aligned} \tag{3.4}$$

Hence, $W^{1,p}(\mathbb{R}_+; L^q(\Omega)) \cap L^p(\mathbb{R}_+; W^{2,q}(\Omega))$ is continuously embedded into the following space:

$$X := \{u \in L^\infty(\mathbb{R}_+; B^{2-2/p;q,p}(\Omega)) : \partial_t u \in L^p(\mathbb{R}_+; L^q(\Omega))\}. \tag{3.5}$$

If $2/p+d/q < 1$, then there exists a small $\epsilon > 0$ such that $2/p+\epsilon+d/q < 1$, and so [1, Theorem 7.34] implies that $B^{2-2/p-\epsilon;q,p}(\Omega)$ is continuously embedded into $W^{1,\infty}(\Omega)$. This proves the first embedding in Lemma 3.1.

For any fixed T, the functions in $W^{1,p}(0,T; L^q(\Omega)) \cap L^p(0,T; W^{2,q}(\Omega))$ can be boundedly extended to $W^{1,p}(\mathbb{R}_+; L^q(\Omega)) \cap L^p(\mathbb{R}_+; W^{2,q}(\Omega))$ (via reflection on $[T, 2T]$, multiplied by a smooth cut-off function $\chi(t)$ such that $\chi(t) = 1$ for $t \in [0, T]$ and $\chi = 0$ for $t \geq 2T$). Consequently, the second embedding in Lemma 3.1 is a consequence of the first embedding. The compactness of the second embedding can be seen as follows.

Since $W^{2,q}(\Omega)$ is compactly embedded into $W^{1,\infty}(\Omega)$ (cf. [1, Theorem 6.3]) and

$$B^{2-2/p;q,p}(\Omega) = (B^{2-2/p-\epsilon;q,p}(\Omega), W^{2;q}(\Omega))_{\theta,p}, \quad \text{with } \theta = \frac{\epsilon}{2/p+\epsilon},$$

the Lions–Peetre theorem ([29, Chapter V, Theorem 2.2], see also [7]) implies that $B^{2-2/p;q,p}(\Omega)$ is also compactly embedded into $W^{1,\infty}(\Omega)$. Since $B^{2-2/p;q,p}(\Omega)$ is compactly embedded into $W^{1,\infty}(\Omega)$ and $W^{1,\infty}(\Omega)$ is continuously embedded into $L^q(\Omega)$, the Aubin–Lions–Simon lemma [5, Theorem II.5.16] implies that X is compactly embedded into $L^\infty(0,T;W^{1,\infty}(\Omega))$. □

3.2 An R-boundedness result

We begin by recalling the notion of R-boundedness on L^q-spaces; see [20, formula (2.7)]. A collection $\mathcal{T}$ of operators on $L^q(\Omega)$ is *R-bounded* if and only if there is a constant C_R, called an R-bound of $\mathcal{T}$, such that any finite subcollection of operators $T_1,\dots,T_l \in \mathcal{T}$ satisfies

$$\Big\| \Big(\sum_{j=1}^{l} |T_j v_j|^2 \Big)^{\frac{1}{2}} \Big\|_{L^q(\Omega)} \le C_R \Big\| \Big(\sum_{j=1}^{l} |v_j|^2 \Big)^{\frac{1}{2}} \Big\|_{L^q(\Omega)}, \quad \forall\, v_1, v_2, ..., v_l \in L^q(\Omega).$$

We will need the following result.

Lemma 3.2 *Let the elliptic operator $A : W^{2,q}(\Omega) \cap W_0^{1,q}(\Omega) \to L^q(\Omega)$ with $1 < q < \infty$ be defined by*

$$A\varphi = \sum_{i,j=1}^{d} \alpha_{ij} \partial_i \partial_j \varphi, \tag{3.6}$$

where the coefficient functions $\alpha_{ij} : \Omega \to \mathbb{R}$ $(i,j = 1,\dots,d)$ (which can be assumed symmetric: $\alpha_{ij} = \alpha_{ji}$) satisfy the following assumptions for some positive constants K and κ:

(A1) The coefficients are bounded in a Hölder norm (with exponent $\mu > 0$):

$$\|\alpha_{ij}\|_{C^\mu(\overline{\Omega})} \le K\,;$$

(A2) The symmetric coefficient matrix (α_{ij}) satisfies the uniform ellipticity condition

$$\sum_{i,j=1}^{d} \alpha_{ij}(x)\xi_i\xi_j \ge \kappa \sum_{j=1}^{d} \xi_j^2 \qquad \forall\, x \in \Omega,\ \forall\, \xi = (\xi_j) \in \mathbb{C}^d\,. \tag{3.7}$$

Then, the collection of operators $\{z(z-A)^{-1} : |\arg z| < \theta\}$ is R-bounded on $L^q(\Omega)$ for some $\theta \in (\pi/2, \pi)$. Both the R-bound and the angle θ depend only on μ, K, κ, Ω and q.

Proof We argue by compactnesss. Fix K, μ, κ, $q \in (d, \infty)$, and an angle $\theta \in (\pi/2, \pi)$. We denote by M the set of all symmetric coefficient matrices (α_{ij}) on Ω satisfying conditions (A1) and (A2). Clearly, M is convex and closed in $\| \cdot \|_{C^\mu(\overline{\Omega})}$ but also in the sup norm on $\overline{\Omega}$. By the Arzela-Ascoli theorem, M is compact in sup norm.

For any coefficient matrix $(\alpha_{ij}) \in M$, the corresponding operator A generates an analytic semigroup by [31, Subsection 3.1.1]. This semigroup is positive, so $\max \operatorname{Re} \sigma(A)$ is an eigenvalue. By [14, Theorem 9.15] the half line $[0, \infty)$ belongs to the resolvent set of A. Thus A is invertible and generates a bounded analytic semigroup. Moreover, for some $\theta_A \in (\pi/2, \pi)$, the set $\{z(z-A)^{-1} : |\arg z| < \theta_A\}$ is R-bounded with R-bound $R(A)$ (see [33], [19, Theorem 1.1] or [20, 7.18]).

If $(\widetilde{\alpha}_{ij}) \in M$ is another coefficient matrix with corresponding operator $\widetilde{A}$, then

$$
\begin{aligned}
\|(\widetilde{A} - A)u\|_{L^q(\Omega)} &\leq \max_{ij} \|\widetilde{\alpha}_{ij} - \alpha_{ij}\|_{L^\infty(\Omega)} \|u\|_{W^{2,q}(\Omega)} \\
&\leq C_A \max_{ij} \|\widetilde{\alpha}_{ij} - \alpha_{ij}\|_{L^\infty(\Omega)} \|Au\|_{L^q(\Omega)}
\end{aligned}
$$

since A is invertible and $D(A) = D(\widetilde{A}) = W^{2,q}(\Omega) \cap W_0^{1,q}(\Omega)$, note that C_A depends on A. By the perturbation theorem for R-sectorial operators ([20, Theorem 6.5]) we find $\eta_A > 0$ such that $(\widetilde{\alpha}_{ij}) \in M$, $\|\widetilde{\alpha}_{ij} - \alpha_{ij}\|_\infty < \eta_A$ implies that for the operator $\widetilde{A}$ corresponding to $(\widetilde{\alpha}_{ij})$ the set $\{z(z-\widetilde{A})^{-1} : |\arg z| < \theta_A\}$ is R-bounded with R-bound $\leq 2R(A)$. By compactness of M we thus find finitely many matrices (α_{ij}^l) with corresponding operators A_l, $l \in F$, such that for each coefficient matrix $(\alpha_{ij}) \in M$ with corresponding operator A there is $l \in F$ with $\|\alpha_{ij} - \alpha_{ij}^l\|_\infty < \eta_{A_l}$. We conclude that, for $\theta := \min_{l \in F} \theta_{A_l}$, the set $\{z(z-A)^{-1} : |\arg| < \theta\}$ is R-bounded with R-bound $\leq 2 \max_{l \in F} R(A_l)$. $\square$

3.3 Maximal L^p regularity

Lemma 3.3 *Under the conditions of Lemma 3.2, the operator A has maximal L^p regularity for $1 < p < \infty$: for every $f \in L^p(0,T; L^q(\Omega))$ (with arbitrary $T > 0$), the solution u of the linear parabolic problem*

$$
\begin{cases}
\dfrac{\partial u}{\partial t} - Au = f & \text{in } \Omega, \\
u = 0 & \text{on } \partial\Omega,
\end{cases}
\tag{3.8}
$$

with zero initial values is bounded by

$$
\|\partial_t u\|_{L^p(0,T;L^q(\Omega))} + \|u\|_{L^p(0,T;W^{2,q}(\Omega))} \leq C_{p,q} \|f\|_{L^p(0,T;L^q(\Omega))}, \tag{3.9}
$$

where the constant $C_{p,q}$ depends only on μ, K, κ, Ω and p and q.

Proof By Lemma 3.2, the operator-valued Mikhlin multiplier theorem used in Weis' characterization of maximal L^p regularity [35, Theorem 4.2] yields the maximal L^p regularity

$$\|\partial_t u\|_{L^p(0,T;L^q(\Omega))} + \|Au\|_{L^p(0,T;L^q(\Omega))} \le C_{p,q}\|f\|_{L^p(0,T;L^q(\Omega))},$$

where $C_{p,q}$ depends only on p, q and the R-bound of Lemma 3.2.

Since $\alpha_{ij} \in C^\mu(\overline{\Omega})$, [6, Theorem 6.1 of Chapter 3] implies that the elliptic operator $A : W^{2,q}(\Omega) \cap W_0^{1,q}(\Omega) \to L^q(\Omega)$ is invertible and

$$\|u\|_{W^{2,q}(\Omega)} \le C_q\|Au\|_{L^q(\Omega)}, \tag{3.10}$$

where C_q depends only on K, κ, Ω and q. This yields the result. □

3.4 Discrete maximal ℓ^p regularity for Runge–Kutta methods

As is shown in [18, Theorem 5.1], A-stable Runge–Kutta methods with an invertible coefficient matrix preserve maximal L^p regularity, uniformly in the step size. Before we formulate the Runge–Kutta analog of Lemma 3.3, we need to introduce further notation.

For any Banach space X and any sequence $(v_n)_{n=1}^N$ with entries in X we denote, for a given step size $\tau > 0$,

$$\big\|(v_n)_{n=1}^N\big\|_{L^p(X)} := \Big(\sum_{n=1}^N \tau\|v_n\|_X^p\Big)^{1/p},$$

which is the $L^p(0, N\tau; X)$ norm of the piecewise constant function that equals v_n on the time interval $(t_{n-1}, t_n]$. We use the same notation also for sequences $(v_n)_{n=0}^N$, replacing $n = 1$ by $n = 0$ in the sum.

Considering the piecewise linear interpolant of a sequence $(v_n)_{n=1}^N$ in $W^{2,q}(\Omega)$ and the starting value $v_0 = 0$, the inequality (3.1) gives, for $2/p + d/q < 1$,

$$\begin{aligned} &\|(v_n)_{n=1}^N\|_{L^\infty(W^{1,\infty}(\Omega))} \\ &\le c_{p,q}\left(\left\|\left(\frac{v_n - v_{n-1}}{\tau}\right)_{n=1}^N\right\|_{L^p(L^q(\Omega))} + \|(v_n)_{n=1}^N\|_{L^p(W^{2,q}(\Omega))}\right). \end{aligned} \tag{3.11}$$

We now consider the Runge–Kutte time discretization of the linear parabolic problem (3.8) with step size τ,

$$u_{n,i} = u_n + \tau\sum_{j=1}^s a_{ij}\big(Au_{n,j} + f_{n,j}\big) \qquad (i = 1, \dots, s), \tag{3.12}$$

and $u_{n+1} = u_{n,s}$ for a Runge–Kutta method with (2.1). We use again the vector notation of (2.6), $\mathbf{u}_n = (u_{n,i})_{i=1}^s$ and $\mathbf{f}_n = (f_{n,i})_{i=1}^s$. We then have the following time-discrete analog of Lemma 3.3.

Lemma 3.4 *Consider a Runge–Kutta method that satisfies* (2.1)*, such as the s-stage Radau IIA method. Under the conditions of Lemma 3.2, there is discrete maximal L^p regularity for $1 < p < \infty$ uniformly in the step size $\tau > 0$: for every sequence $(\mathbf{f}_n)_{n=0}^N$ with entries in $L^q(\Omega)^s$ (with arbitrary $N \geq 1$), the numerical solution defined by* (3.12) *with zero initial value $u_0 = 0$ satisfies the bound, with $\mathbf{u}_{-1} = 0$,*

$$\left\| \left(\frac{\mathbf{u}_n - \mathbf{u}_{n-1}}{\tau} \right)_{n=0}^N \right\|_{L^p(L^q(\Omega)^s)} + \left\| (\mathbf{u}_n)_{n=0}^N \right\|_{L^p(W^{2,q}(\Omega)^s)} \leq C_{p,q} \left\| (\mathbf{f}_n)_{n=0}^N \right\|_{L^p(L^q(\Omega)^s)}, \tag{3.13}$$

where the constant $C_{p,q}$ depends only on K, κ, Ω and p and q. In particular, $C_{p,q}$ is independent of N and τ.

Proof In view of Lemma 3.2, [18, Theorem 5.1] gives the bound, with $\dot{\mathbf{u}}_n = (\dot{u}_{n,j})_{j=1}^s$ for $\dot{u}_{n,j} = Au_{n,j} + f_{n,j}$,

$$\|(\dot{\mathbf{u}}_n)_{n=0}^N\|_{L^p(L^q(\Omega)^s)} + \|(A\mathbf{u}_n)_{n=0}^N\|_{L^p(L^q(\Omega)^s)} \leq \widetilde{C}_{p,q} \left\| (\mathbf{f}_n)_{n=0}^N \right\|_{L^p(L^q(\Omega)^s)},$$

where $\widetilde{C}_{p,q}$ depends only on p, q and the R-bound of Lemma 3.2.

For the second term on the left-hand side we recall (3.10). For the first term we note that (3.12) yields

$$\left\| \left(\frac{u_{n,i} - u_n}{\tau} \right)_{i=1}^s \right\|_{L^q(\Omega)^s} \leq \gamma \left\| \left(\dot{u}_{n,j} \right)_{j=1}^s \right\|_{L^q(\Omega)^s},$$

where γ is the norm of the Runge–Kutta coefficient matrix (a_{ij}). Writing

$$u_{n,i} - u_{n-1,i} = (u_{n,i} - u_n) + (u_n - u_{n-1}) - (u_{n-1,i} - u_{n-1})$$

and noting that $u_n - u_{n-1} = u_{n-1,s} - u_{n-1}$, we find that the above inequality (for n and $n-1$) yields

$$\left\| \frac{\mathbf{u}_n - \mathbf{u}_{n-1}}{\tau} \right\|_{L^q(\Omega)^s} \leq \gamma \left\| \left(\dot{u}_{n,j} \right)_{j=1}^s \right\|_{L^q(\Omega)^s} + 2\gamma \left\| \left(\dot{u}_{n-1,j} \right)_{j=1}^s \right\|_{L^q(\Omega)^s},$$

which completes the proof of the result. □

Combining Lemma 3.4 and (3.11), we thus obtain the bound

$$\left\| (\mathbf{u}_n)_{n=0}^N \right\|_{L^\infty(W^{1,\infty}(\Omega)^s)} \leq \widehat{C}_{p,q} \left\| (\mathbf{f}_n)_{n=0}^N \right\|_{L^p(L^q(\Omega)^s)}, \tag{3.14}$$

with $\widehat{C}_{p,q} = c_{p,q} C_{p,q}$. This $W^{1,\infty}$ bound of the numerical solution is the key to proving Theorem 2.1.

3.5 Nonautonomous linear parabolic problems

Let the time-dependent elliptic operators $A(t) : W^{2,q}(\Omega) \cap W_0^{1,q}(\Omega) \to L^q(\Omega)$ for $0 \le t \le T$ be defined by

$$A(t)\varphi = \sum_{i,j=1}^{d} \alpha_{ij}(\cdot,t)\partial_i\partial_j\varphi, \tag{3.15}$$

where the coefficient functions $\alpha_{ij}(\cdot,t) : \Omega \to \mathbb{R}$ $(i,j = 1,\dots,d)$ satisfy conditions (A1) and (A2) of Lemma 3.2 uniformly for $0 \le t \le T$ and additionally the Lipschitz condition

$$\|\alpha_{ij}(\cdot,t) - \alpha_{ij}(\cdot,s)\|_{L^\infty(\Omega)} \le L\,|t-s|, \qquad 0 \le s,t \le T. \tag{3.16}$$

Lemma 3.5 *In the above situation of time-dependent elliptic operators $A(t)$, the solution of the nonautonomous linear problem* (3.8) *is bounded by* (3.9), *where the constant $C_{p,q}$ depends additionally on L and T.*

Proof For $0 \le t \le \bar t \le T$, we rewrite the differential equation as

$$\partial_t u(t) = A(\bar t)u(t) - \big(A(\bar t) - A(t)\big)u(t) + f(t)$$

and apply Lemma 3.3 for the operator $A(\bar t)$ to bound

$$\begin{aligned}\|\partial_t u\|_{L^p(0,\bar t;L^q(\Omega))} + \|u\|_{L^p(0,\bar t;W^{2,q}(\Omega))} &\le C_{p,q}\|\big(A(\bar t) - A(\cdot)\big)u\|_{L^p(0,\bar t;L^q(\Omega))} \\ &\quad + C_{p,q}\|f\|_{L^p(0,\bar t;L^q(\Omega))}.\end{aligned} \tag{3.17}$$

We denote

$$\eta(\bar t) = \|u\|^p_{L^p(0,\bar t;W^{2,q}(\Omega))}.$$

By the Lipschitz condition (3.16) and by integration by parts we obtain

$$\begin{aligned}\int_0^{\bar t} \|\big(A(\bar t) - A(t)\big)u(t)\|^p_{L^q(\Omega)}\,\mathrm{d}t &\le L^p \int_0^{\bar t} (\bar t - t)^p \|u(t)\|^p_{W^{2,q}(\Omega)}\,\mathrm{d}t \\ &= L^p p \int_0^{\bar t} (\bar t - t)^{p-1}\eta(t)\,\mathrm{d}t.\end{aligned}$$

Hence we have from (3.17)

$$\eta(\bar t) \le C\int_0^{\bar t} (\bar t - t)^{p-1}\eta(t)\,\mathrm{d}t + C\|f\|^p_{L^p(0,\bar t;L^q(\Omega))}, \qquad 0 \le \bar t \le T,$$

and a Gronwall inequality yields

$$\eta(T) \le C'\|f\|^p_{L^p(0,T;L^q(\Omega))},$$

which combined with (3.17) yields the result. $\square$

3.6 Runge–Kutta discretization of nonautonomous linear problems

With Lemma 3.4, the previous result for the nonautonomous linear problem extends to its Runge–Kutta time discretization

$$u_{n,i} = u_n + \tau \sum_{j=1}^{s} a_{ij} \big(A(t_{n,j}) u_{n,j} + f_{n,j} \big) \qquad (i = 1, \dots, s), \tag{3.18}$$

and $u_{n+1} = u_{n,s}$ for a Runge–Kutta method with (2.1).

Lemma 3.6 *Consider a Runge–Kutta method that satisfies* (2.1)*, such as the s-stage Radau IIA method. Under the conditions of Lemma 3.5, there is discrete maximal L^p regularity for $1 < p < \infty$ uniformly in the step size $\tau > 0$: for every sequence $(\mathbf{f}_n)_{n=0}^N$ with entries in $L^q(\Omega)^s$ (with arbitrary $N \geq 1$), the numerical solution defined by* (3.18) *with zero initial value $u_0 = 0$ satisfies the bound* (3.13)*, where $C_{p,q}$ is independent of N and τ with $N\tau \leq T$, but depends on T.*

Proof The result follows from Lemma 3.4 in the same way as Lemma 3.5 follows from Lemma 3.3, using a summation by parts in place of the integration by parts. □

4 Proof of Theorem 2.1

4.1 Defects and error equation

The exact solution values satisfy the Runge–Kutta relations up to a defect:

$$u(t_n + c_i\tau) = u(t_n) + \tau \sum_{j=1}^{s} a_{ij}\, \partial_t u(t_n + c_j\tau) + d_{n,i},$$

where we note that $d_{n,i}$ is the quadrature error over the interval $[t_n, t_n + c_i\tau]$ of the quadrature formula with weights a_{ij} and nodes c_j. Using Taylor expansion at t_n and the definition of the stage order (2.4) and the regularity condition (2.8), we can bound $\mathbf{d}_n = (d_{n,i})_{i=1}^s$ by

$$\|(\mathbf{d}_n)_{n=0}^N\|_{L^p(L^q(\Omega)^s)} \leq C\tau^{k+1}.$$

We rewrite the above equation as

$$u(t_n + c_i\tau) = u(t_n) + \tau \sum_{j=1}^{s} a_{ij} \left(\partial_t u(t_n + c_j\tau) - r_{n,j} \right), \tag{4.1}$$

where $\mathbf{r}_n = (r_{n,j})_{j=1}^s$ is the solution of the linear system with the invertible Runge–Kutta matrix (a_{ij}),

$$\tau \sum_{j=1}^{s} a_{ij} r_{n,j} = -d_{n,i}, \qquad \text{so that} \quad \rho := \|(\mathbf{r}_n)_{n=0}^N\|_{L^p(L^q(\Omega)^s)} \leq C\tau^k. \tag{4.2}$$

We rewrite the partial differential equation as

$$\partial_t u = \nabla \cdot f(\nabla u) = \sum_{k,l=1}^{d} f_{k,l}(\nabla u)\partial_k\partial_l u, \quad \text{with } f_{k,l} = \partial f_k/\partial p_l. \tag{4.3}$$

Comparing (2.3) and (2.5) with (4.1) and (4.3), we see that the errors

$$e_{n,i} := u_{n,i} - u(t_{n,i}) \quad \text{and} \quad \dot{e}_{n,j} := \dot{u}_{n,j} - \partial_t u(t_{n,j}) + r_{n,j} \tag{4.4}$$

satisfy the error equations (for $i, j = 1, \dots, s$)

$$e_{n,i} = e_n + \tau \sum_{j=1}^{s} a_{ij}\dot{e}_{n,j}, \qquad e_{n+1} = e_{n,s} \tag{4.5a}$$

$$\begin{aligned}\dot{e}_{n,j} &= \sum_{k,l=1}^{d} f_{k,l}(\nabla u(t_{n,j}))\partial_k\partial_l e_{n,j} \\ &+ \sum_{k,l=1}^{d} \Big(f_{k,l}(\nabla u(t_{n,j}) + \nabla e_{n,j}) - f_{k,l}(\nabla u(t_{n,j}))\Big)\partial_k\partial_l\big(u(t_{n,j}) + e_{n,j}\big) + r_{n,j}.\end{aligned} \tag{4.5b}$$

Clearly, $\mathbf{e}_n = (e_{n,i})$ is a solution of the error equations (4.5) if and only if $(u_{n,i}) = (u(t_{n,i}) + e_{n,i})$ is a solution of the Runge–Kutta equations (2.2)-(2.3).

4.2 Error bound

We first show the error bound of Theorem 2.1 under the additional assumption that the errors remain bounded by a small constant in the $W^{1,\infty}$ norm. This condition will be verified in the next subsection.

Lemma 4.1 *In the situation of Theorem 2.1, suppose that the error equations have a solution $(e_{n,i})$ for $0 \le n \le N$ and $i = 1, \dots, s$ such that*

$$\max_{0\le n\le N} \max_{1\le i\le s} \|e_{n,i}\|_{W^{1,\infty}(\Omega)} \le \mu \tag{4.6}$$

with a sufficiently small constant μ (independent of τ and N with $N\tau \le T$). Then the $O(\tau^k)$ error bounds (2.9) are satisfied.

Proof If we consider $\mathbf{g}_n = (g_{n,j})$ with

$$g_{n,j} = \sum_{k,l=1}^{d} \Big(f_{k,l}(\nabla u(t_{n,j}) + \nabla e_{n,j}) - f_{k,l}(\nabla u(t_{n,j}))\Big)\partial_k\partial_l\big(u(t_{n,j}) + e_{n,j}\big)$$

as an inhomogeneity in (4.5b), then Lemma 3.6 shows that

$$\begin{aligned}\Big\|\Big(\frac{\mathbf{e}_n - \mathbf{e}_{n-1}}{\tau}\Big)_{n=0}^{N}\Big\|_{L^p(L^q(\Omega)^s)} &+ \big\|(\mathbf{e}_n)_{n=0}^N\big\|_{L^p(W^{2,q}(\Omega)^s)} \\ &\le C\Big(\big\|(\mathbf{g}_n)_{n=0}^N\big\|_{L^p(L^q(\Omega)^s)} + \big\|(\mathbf{r}_n)_{n=0}^N\big\|_{L^p(L^q(\Omega)^s)}\Big).\end{aligned} \tag{4.7}$$

We bound, with a local Lipschitz constant L of $f_{k,l}$,

$$\|g_{n,j}\|_{L^q(\Omega)} \le L\|\nabla e_{n,j}\|_{L^\infty(\Omega)}\, \|u(t_{n,j})\|_{W^{2,q}(\Omega)} + L\|\nabla e_{n,j}\|_{L^\infty(\Omega)}\, \|e_{n,j}\|_{W^{2,q}(\Omega)} \tag{4.8}$$

so that

$$\big\|(\mathbf{g}_n)_{n=0}^N\big\|_{L^p(L^q(\Omega)^s)} \le C_1\big\|(\mathbf{e}_n)_{n=0}^N\big\|_{L^p(W^{1,\infty}(\Omega)^s)} + C_2\mu\big\|(\mathbf{e}_n)_{n=0}^N\big\|_{L^p(W^{2,q}(\Omega)^s)}.$$

Using the bound

$$\|\mathbf{e}_n\|_{W^{1,\infty}(\Omega)^s} \le \mu\|\mathbf{e}_n\|_{W^{2,q}(\Omega)^s} + C_\mu\|\mathbf{e}_n\|_{L^q(\Omega)^s},$$

we obtain

$$\big\|(\mathbf{g}_n)_{n=0}^N\big\|_{L^p(L^q(\Omega)^s)} \le C\mu\big\|(\mathbf{e}_n)_{n=0}^N\big\|_{L^p(W^{2,q}(\Omega)^s)} + C_\mu\big\|(\mathbf{e}_n)_{n=0}^N\big\|_{L^p(L^q(\Omega)^s)}.$$

If μ is sufficiently small, then the first term on the right-hand side can be absorbed in the left-hand side of (4.7), and we are left with

$$\begin{aligned}\bigg\|\Big(\frac{\mathbf{e}_n-\mathbf{e}_{n-1}}{\tau}\Big)_{n=0}^N\bigg\|_{L^p(L^q(\Omega)^s)} &+ \big\|(\mathbf{e}_n)_{n=0}^N\big\|_{L^p(W^{2,q}(\Omega)^s)} \\ &\le C\Big(\big\|(\mathbf{e}_n)_{n=0}^N\big\|_{L^p(L^q(\Omega)^s)} + \big\|(\mathbf{r}_n)_{n=0}^N\big\|_{L^p(L^q(\Omega)^s)}\Big).\end{aligned}$$

Such a bound holds not only for the final N, but for each $\bar n \le N$. We write $\mathbf{e}_n = \tau\sum_{m=0}^n(\mathbf{e}_m-\mathbf{e}_{m-1})/\tau$ and use, for $\alpha_j = \frac{1}{\tau}\|\mathbf{e}_j-\mathbf{e}_{j-1}\|_{L^q(\Omega)^s}$, the inequality

$$\bigg\|\Big(\sum_{j=0}^m \alpha_j\Big)_{m=0}^{\bar n}\bigg\|_p \le \sum_{m=0}^{\bar n}\Big\|(\alpha_j)_{j=0}^m\Big\|_p\,, \tag{4.9}$$

which is just the triangle inequality for the sum of vectors in $\mathbb{R}^{\bar n+1}$

$$\begin{pmatrix}0\\ \vdots\\ 0\\ \alpha_0\end{pmatrix} + \begin{pmatrix}0\\ \vdots\\ \alpha_0\\ \alpha_1\end{pmatrix} + \ldots + \begin{pmatrix}0\\ \alpha_0\\ \vdots\\ \alpha_{\bar n-1}\end{pmatrix} + \begin{pmatrix}\alpha_0\\ \alpha_1\\ \vdots\\ \alpha_{\bar n}\end{pmatrix}.$$

We thus obtain, for $0\le \bar n\le N$,

$$\begin{aligned}\bigg\|\Big(\frac{\mathbf{e}_n-\mathbf{e}_{n-1}}{\tau}\Big)_{n=0}^{\bar n}\bigg\|_{L^p(L^q(\Omega)^s)} &+ \big\|(\mathbf{e}_n)_{n=0}^{\bar n}\big\|_{L^p(W^{2,q}(\Omega)^s)} \\ &\le C\Bigg(\tau\sum_{m=0}^{\bar n}\bigg\|\Big(\frac{\mathbf{e}_n-\mathbf{e}_{n-1}}{\tau}\Big)_{n=0}^{m}\bigg\|_{L^p(L^q(\Omega)^s)} + \big\|(\mathbf{r}_n)_{n=0}^{\bar n}\big\|_{L^p(L^q(\Omega)^s)}\Bigg).\end{aligned}$$

Applying a discrete Gronwall inequality then yields

$$\bigg\|\Big(\frac{\mathbf{e}_n-\mathbf{e}_{n-1}}{\tau}\Big)_{n=0}^N\bigg\|_{L^p(L^q(\Omega)^s)} + \big\|(\mathbf{e}_n)_{n=0}^N\big\|_{L^p(W^{2,q}(\Omega)^s)} \le \widetilde{C}\big\|(\mathbf{r}_n)_{n=0}^N\big\|_{L^p(L^q(\Omega)^s)}, \tag{4.10}$$

and the result follows with the bound (4.2). □

4.3 Existence of the numerical solution

In this subsection, we prove the existence of a solution $\mathbf{e}_n$ for (4.5) satisfying the error bound (4.10) by using Schaefer's fixed point theorem via the arguments of the proof of Lemma 4.1, which rely on the maximal regularity properties of Section 3.

Lemma 4.2 (Schaefer's fixed point theorem [9, Chapter 9.2, Theorem 4]) *Let X be a Banach space and let $\mathcal{M} : X \to X$ be a continuous and compact map. If the set*

$$\big\{\phi \in X : \ \phi = \theta\mathcal{M}(\phi) \ \ \textit{for some} \ \ \theta \in [0,1]\big\} \tag{4.11}$$

is bounded in X, then the map $\mathcal{M}$ has a fixed point.

We define a map $\mathcal{M} : C([0,T], W^{1,\infty}(\Omega)^s) \to C([0,T], W^{1,\infty}(\Omega)^s)$ in the following way: for any given $\boldsymbol{\varphi} = (\varphi_j)_{j=1}^s \in C([0,T], W^{1,\infty}(\Omega)^s)$, we define $\mathbf{e} := \mathcal{M}\boldsymbol{\varphi}$ as the piecewise linear interpolation in time of the vectors $\mathbf{e}_n = (e_{n,i})_{i=1}^s$ for $n = 0, \dots, N$ (that is, interpolating linearly between $e_{n,i}$ and $e_{n-1,i}$ for each i), where $\mathbf{e}_n = (e_{n,i})_{i=1}^s$ are the solution of the *linear* problem

$$e_{n,i} = e_n + \tau \sum_{j=1}^s a_{ij}\dot{e}_{n,j}, \qquad e_{n+1} = e_{n,s} \tag{4.12a}$$

$$\begin{aligned}\dot{e}_{n,j} &= \sum_{k,l=1}^d f_{k,l}\big(\nabla u(t_{n,j})\big)\partial_k\partial_l e_{n,j} \\ &+ \sum_{k,l=1}^d \Big(f_{k,l}\big(\nabla u(t_{n,j}) + \beta(\varphi_j(t_{n,j}))\nabla\varphi_j(t_{n,j})\big) - f_{k,l}\big(\nabla u(t_{n,j})\big)\Big)\times \\ &\qquad\qquad \partial_k\partial_l\big(u(t_{n,j}) + e_{n,j}\big) \\ &+ r_{n,j},\end{aligned} \tag{4.12b}$$

where

$$\beta(\varphi) = \min\Big(\frac{\sqrt{\rho}}{\|\varphi\|_{W^{1,\infty}(\Omega)}}, 1\Big),$$

which has the following properties:

$$\|\beta(\varphi)\varphi\|_{W^{1,\infty}(\Omega)} \le \sqrt{\rho}, \tag{4.13a}$$

$$\beta(\varphi) = 1 \quad \text{if} \quad \|\varphi\|_{W^{1,\infty}(\Omega)} \le \sqrt{\rho}. \tag{4.13b}$$

Lemma 4.3 *The map $\mathcal{M} : C([0,T], W^{1,\infty}(\Omega)^s) \to C([0,T], W^{1,\infty}(\Omega)^s)$ is well defined, continuous and compact.*

Proof Following the lines of the proof of Lemma 4.1, with the only difference that $\|\nabla e_{n,j}\|_{L^\infty(\Omega)}$ is replaced with $\|\beta(\varphi_j(t_{n,j}))\nabla\varphi_j(t_{n,j})\|_{L^\infty(\Omega)} \le \sqrt{\rho}$ in (4.8), it is seen that $\mathcal{M}$ maps boundedly into the space $W^{1,p}(0,T; L^q(\Omega)^s) \cap L^p(0,T; W^{2,q}(\Omega)^s)$, which is compactly embedded in $C([0,T], W^{1,\infty}(\Omega)^s)$ by Lemma 3.1. The continuity of $\mathcal{M}$ is also obtained by the arguments used in the proof of Lemma 4.1. □

To apply Schaefer's fixed point theorem (Lemma 4.2), we assume that

$$\boldsymbol{\varphi} = \theta \mathcal{M} \boldsymbol{\varphi} \quad \text{for some } \theta \in [0,1].$$

Then $\mathbf{e} := \mathcal{M}\boldsymbol{\varphi}$ is the piecewise linear interpolation of the solution of the equations (4.12) with $\varphi_j = \theta e_j$. Using the same proof as that of Lemma 4.1, it is now seen that $\mathbf{e}$ satisfies $O(\rho) = O(\tau^k)$ error bounds (2.9). This implies that $\|\boldsymbol{\varphi}\|_{W^{1,\infty}(\Omega)^s} \le C\rho$ for sufficiently small τ (note that then $\beta(\varphi_j) = 1$), and hence Schaefer's fixed point theorem yields the existence of a solution to the error equations (4.5) satisfying (2.9). For sufficiently small τ, (2.9) implies (4.6) (since μ is a constant independent of the step size τ).

4.4 Uniqueness of the numerical solution

In the last subsection, we have proved for sufficiently small τ, say $0 < \tau < \bar{\tau}$, the existence of a numerical solution in a $W^{1,\infty}(\Omega)$ neighbourhood of the PDE solution with width μ, satisfying (2.9). The stability result of Lemma 4.1, that is, the bound (4.10) used with $\mathbf{r}_n = 0$, implies the local uniqueness of the Runge-Kutta solution in the $W^{1,\infty}(\Omega)$ neighbourhood of width μ.

5 Proof of Theorem 2.2

The proof is similar to previous proofs of error bounds for Runge–Kutta time discretizations of parabolic problems using energy estimates [30,8]. In particular, the same use is made of the algebraic stability condition (2.10). However, the proof differs in that here we need to invoke the $W^{1,\infty}(\Omega)$ error bounds provided by Theorem 2.1.

5.1 Defects

We denote the exact solution values $u^*_{n,i} = u(t_n + c_i\tau)$, $\dot{u}^*_{n,i} = \partial_t u(t_n + c_i\tau)$, and $u^*_n = u(t_n)$. Note that $u^*_{n+1} = u^*_{n,s}$ by our condition $c_s = 1$. We denote by $d_{n,i}$ and d_{n+1} the defects obtained on inserting the exact solution into the Runge–Kutta equations,

$$u^*_{n,i} = u^*_n + \tau \sum_{j=1}^{s} a_{ij} \dot{u}^*_{n,j} + d_{n,i}, \qquad u^*_{n+1} = u^*_n + \tau \sum_{j=1}^{s} b_j \dot{u}^*_{n,j} + d_{n+1}.$$

The defects are thus quadrature errors. By Taylor expansion at t_n and the definition of the stage order (2.4) and by condition (2.11), the defects are of the form

$$
\begin{aligned}
d_{n,i} &= \tau^k \int_{t_n}^{t_{n+1}} K_i\Big(\frac{t-t_n}{\tau}\Big)\, u^{(k+1)}(t)\,\mathrm{d}t \\
d_{n+1} &= \tau^{k+1} \int_{t_n}^{t_{n+1}} K\Big(\frac{t-t_n}{\tau}\Big)\, u^{(k+2)}(t)\,\mathrm{d}t \\
&= -\tau^k \int_{t_n}^{t_{n+1}} K'\Big(\frac{t-t_n}{\tau}\Big)\, u^{(k+1)}(t)\,\mathrm{d}t
\end{aligned}
$$

with bounded Peano kernels K_i and K. Here we assume for simplicity that all $c_i \in [0,1]$, as is the case for all methods of interest. In the following we denote by $\langle\cdot,\cdot\rangle$ the duality pairing between $H_0^1(\Omega)$ and $H^{-1}(\Omega)$, which restricted to $H_0^1(\Omega)\times L^2(\Omega)$ coincides with the $L^2(\Omega)$ inner product. We further denote

$$
|\cdot| = \|\cdot\|_{L^2(\Omega)}, \quad \|\cdot\| = \|\cdot\|_{H_0^1(\Omega)}, \quad \|\cdot\|_\star = \|\cdot\|_{H^{-1}(\Omega)}.
$$

We define $\delta \ge 0$ by setting

$$
\delta^2 = \tau \sum_{n=0}^{N}\sum_{i=1}^{s} \|d_{n,i}\|^2 + \tau\sum_{n=0}^{N}\big(\|d_{n+1}\|^2 + \|d_{n+1}/\tau\|_\star^2\big) \tag{5.1}
$$

and note that by our regularity assumption and the above defect estimates we have

$$
\delta \le C\tau^{k+1}.
$$

5.2 Error equations

The errors $e_{n,i} = u_{n,i} - u_{n,i}^*$, $\dot e_{n,i} = \dot u_{n,i} - \dot u_{n,i}^*$, and $e_n = u_n - u_n^*$ satisfy the error equations (written in the divergence form):

$$
\dot e_{n,i} = \sum_{k=1}^{d} \partial_k\big(f_k(\nabla u_{n,i}) - f_k(\nabla u_{n,i}^*)\big) \tag{5.2a}
$$

$$
e_{n,i} = e_n + \tau\sum_{j=1}^{s} a_{ij}\dot e_{n,j} - d_{n,i} \tag{5.2b}
$$

$$
e_{n+1} = e_n + \tau\sum_{i=1}^{s} b_i \dot e_{n,i} - d_{n+1}. \tag{5.2c}
$$

5.3 Energy estimate using algebraic stability

Taking the square of the $L^2(\Omega)$ norm in (5.2c) yields

$$
|e_{n+1}|^2 = \Big|e_n + \tau\sum_{i=1}^{s} b_i\dot e_{n,i}\Big|^2 - 2\Big\langle d_{n+1}, e_n + \tau\sum_{i=1}^{s} b_i\dot e_{n,i}\Big\rangle + |d_{n+1}|^2. \tag{5.3}
$$

The three terms on the right-hand side will now be estimated separately. We express e_n by (5.2b) to obtain

$$\begin{aligned}\Big|e_n + \tau\sum_{i=1}^s b_i \dot e_{n,i}\Big|^2 &= |e_n|^2 + 2\tau\sum_{i=1}^s b_i\langle \dot e_{n,i}, e_{n,i} + d_{n,i}\rangle \\ &\quad + \tau^2\sum_{i=1}^s\sum_{j=1}^s (b_ib_j - b_ia_{ij} - b_ja_{ji})\,\langle \dot e_{n,i}, \dot e_{n,j}\rangle.\end{aligned}$$

Here the last term is nonpositive by the algebraic stability condition (2.10). We next estimate the second term on the right-hand side. We have by (5.2a)

$$\begin{aligned}\langle \dot e_{n,i}, e_{n,i} + d_{n,i}\rangle &= -\sum_{k=1}^d \Big\langle f_k(\nabla u_{n,i}) - f_k(\nabla u^*_{n,i}), \partial_k e_{n,i} + \partial_k d_{n,i}\Big\rangle \qquad (5.4)\\ &= -\sum_{k,l=1}^d \Big\langle\Big(\int_0^1 f_{k,l}((1-\theta)\nabla u^*_{n,i} + \theta\nabla u_{n,i})\mathrm{d}\theta\Big)\partial_l e_{n,i}, \partial_k e_{n,i} + \partial_k d_{n,i}\Big\rangle,\end{aligned}$$

where $f_{k,l} := \partial f_k/\partial p_l$. Under the regularity condition (2.8) about the exact solution we have the bounds $\|u^*\|_{W^{1,\infty}(\Omega)} \le R$ and $\|u_{n,i}\|_{W^{1,\infty}(\Omega)} \le R$ as a consequence of (2.9a), which is already proved in Theorem 2.1. Hence there exists $\kappa_R > 0$ such that we have for all $x \in \Omega$

$$\sum_{k,l=1}^d \Big(\int_0^1 f_{k,l}((1-\theta)\nabla u^*_{n,i}(x) + \theta\nabla u_{n,i}(x))\mathrm{d}\theta\Big)\xi_k\xi_l \ge \kappa_R\sum_{l=1}^d \xi_l^2, \qquad \forall \xi = (\xi_l) \in \mathbb{R}^d,$$

and there is a positive constant K_R such that for all $x \in \Omega$

$$\Big|\int_0^1 f_{k,l}((1-\theta)\nabla u^*_{n,i}(x) + \theta\nabla u_{n,i}(x))\mathrm{d}\theta\Big| \le K_R.$$

Hence (5.4) reduces to

$$\langle \dot e_{n,i}, e_{n,i} + d_{n,i}\rangle \le -\kappa_R\|e_{n,i}\|^2 + K_R\|e_{n,i}\|\|d_{n,i}\| \le -\frac{\kappa_R}{2}\|e_{n,i}\|^2 + C\|d_{n,i}\|^2 \qquad (5.5)$$

with $C = K_R^2/(2\kappa_R)$.

With the same arguments, again invoking Theorem 2.1, we also obtain from (5.2a) (with a different constant C)

$$\|\dot e_{n,i}\|_* \le C\,\|e_{n,i}\|.$$

The second and third terms in (5.3) are estimated as (note that all $b_i > 0$)

$$\Big\langle d_{n+1}, e_n + \tau\sum_{i=1}^s b_i\dot e_{n,i}\Big\rangle \le \sqrt{\tau}\|d_{n+1}/\tau\|_*\,\sqrt{\tau}\|e_n\| + \sqrt{\tau}\|d_{n+1}\|\,\sqrt{\tau}\sum_{i=1}^s b_i\|\dot e_{n,i}\|_*,$$

and

$$|d_{n+1}|^2 \le \sqrt{\tau}\|d_{n+1}/\tau\|_* \sqrt{\tau}\|d_{n+1}\| \le \tfrac{1}{2}\tau \,\|d_{n+1}\|^2 + \tfrac{1}{2}\tau \,\|d_{n+1}/\tau\|_*^2 \,,$$

respectively.

Combining the above estimates (and noting that $e_n = e_{n-1,s}$) we obtain

$$\begin{aligned}|e_{n+1}|^2 - |e_n|^2 + \tfrac{1}{8}\kappa_R\tau \sum_{i=1}^{s} b_i \|e_{n,i}\|^2 \\ \le C\tau \sum_{i=1}^{s} \|d_{n,i}\|^2 + C\tau(\|d_{n+1}\|^2 + \|d_{n+1}/\tau\|_*^2).\end{aligned}$$

Summing up these inequalities and recalling (5.1) yields

$$|e_{n+1}|^2 + \tfrac{1}{8}\kappa_R\tau \sum_{m=0}^{n}\sum_{i=1}^{s} b_i \|e_{m,i}\|^2 \le C\delta,$$

which completes the proof. □

Acknowledgements The work of Buyang Li was partially supported by a grant from the Germany/Hong Kong Joint Research Scheme sponsored by the Research Grants Council of Hong Kong and the German Academic Exchange Service of Germany (Ref. No. G-PolyU502/16). The research stay of Buyang Li at the University of Tübingen during the preparation of this paper was funded by the Alexander von Humboldt Foundation. Peer Kunstmann and Christian Lubich were supported by Deutsche Forschungsgemeinschaft, SFB 1173. We thank Christian Power Guerra for the interpretation of (4.9) as a triangle inequality.

References

1. R. A. Adams and J. J. F. Fournier, *Sobolev Spaces*, 2nd ed., Academic Press, Amsterdam, 2003.
2. G. Akrivis, B. Li, and C. Lubich, Combining maximal regularity and energy estimates for time discretizations of quasilinear parabolic equations. *Math. Comp.*, 86 (2017), pp. 1527–1552.
3. G. Akrivis, C. Makridakis, and R. H. Nochetto, Galerkin and Runge–Kutta methods: unified formulation, a posteriori error estimates and nodal superconvergence. *Numer. Math.*, 118 (2011), pp. 429–456.
4. A. Ashyralyev, S. Piskarev, and L. Weis, On well-posedness of difference schemes for abstract parabolic equations in $L_p([0,T];E)$ spaces, *Numer. Funct. Anal. Optim.*, 23 (2002), pp. 669–693.
5. F. Boyer and P. Fabrie, *Mathematical Tools for the Study of the Incompressible Navier-Stokes Equations and Related Models*, Applied Mathematical Sciences 183, Springer, New York, 2013.
6. Y. Z. Chen and L. C. Wu, *Second Order Elliptic Equations and Elliptic Systems*, Translations of Mathematical Monographs, Volume 174, AMS, 1998.
7. F. Cobos, Real interpolation and compactness, *Revista Matematica de la Universidad Complutense de Madrid*, 2 (1989), número suplementario.
8. G. Dziuk, C. Lubich, and D. Mansour, Runge-Kutta time discretization of parabolic differential equations on evolving surfaces. *IMA J. Numer. Anal.*, 32 (2012), pp. 394–416.

9. L. C. Evans, *Partial Differential Equations*, second edition, Graduate Studies in Mathematics 19, American Mathematical Society, Providence, RI, 2010.
10. X. Feng, M. von Oehsen, and A. Prohl, Rate of convergence of regularization procedures and finite element approximations for the total variation flow. *Numer. Math.*, 100 (2005), pp. 441–456.
11. X. Feng and A. Prohl, Analysis of total variation flow and its finite element approximations. *ESAIM Math. Model. Numer. Anal.*, 37 (2003), pp. 533–556.
12. M. Geissert, *Discrete maximal L^p regularity for finite element operators*, SIAM J. Numer. Anal., 44 (2006), pp. 677–698.
13. M. Geissert, *Applications of discrete maximal L^p regularity for finite element operators*, Numer. Math., 108 (2007), pp. 121–149.
14. D. Gilbarg and N. S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, reprint of the third edition, Springer–Verlag, Berlin Heidelberg New York, 2001.
15. E. Hairer and G. Wanner, *Solving Ordinary Differential Equations II: Stiff and Differential–Algebraic Problems*, 2^{nd} revised ed., Springer–Verlag, Berlin Heidelberg, Springer Series in Computational Mathematics v. 14, 2002.
16. T. Kemmochi, Discrete maximal regularity for abstract Cauchy problems. *Studia Math.,* 234 (2016), pp. 241–263.
17. T. Kemmochi and N. Saito, *Discrete maximal regularity and the finite element method for parabolic equations.* Preprint. arXiv:1602.06864
18. B. Kovács, B. Li, and C. Lubich, A-stable time discretizations preserve maximal parabolic regularity, *SIAM J. Numer. Anal.,* 54 (2016), pp. 3600–3624.
19. P. C. Kunstmann, Maximal L_p-regularity for second order elliptic operators with uniformly continuous coefficients on domains, *Progress in Nonlinear Differential Equations and Applications*, Vol. 55, 293-305, Birkhäuser 2003.
20. P. C. Kunstmann and L. Weis, *Maximal L^p regularity for Parabolic Equations, Fourier Multiplier Theorems and H^1-Functional Calculus. Functional Analytic Methods for Evolution Equations*, Lecture Notes in Mathematics, vol. 1855 (2004), pp. 65–311.
21. D. Leykekhman and B. Vexler, Discrete maximal parabolic regularity for Galerkin finite element methods. *Numer. Math.,* 135 (2017), pp 923–952.
22. B. Li, Convergence of a decoupled mixed FEM for the dynamic Ginzburg–Landau equations in nonsmooth domains with incompatible initial data. arXiv:1605.01208
23. B. Li, Maximum-norm stability and maximal L^p regularity of FEMs for parabolic equations with Lipschitz continuous coefficients, *Numer. Math.,* 131 (2015), pp. 489-516.
24. B. Li and W. Sun, Error analysis of linearized semi-implicit Galerkin finite element methods for nonlinear parabolic equations. *Int. J. Numer. Anal. & Modeling*, 10 (2013), pp. 622–633.
25. B. Li and W. Sun, Linearized FE approximations to a nonlinear gradient flow. *SIAM J. Numer. Anal.*, 52 (2014), pp. 2623–2646.
26. B. Li and W. Sun, Regularity of the diffusion-dispersion tensor and error analysis of FEMs for a porous media flow, *SIAM J. Numer. Anal.,* 53 (2015), pp. 1418–1437.
27. B. Li and W. Sun, Maximal L^p analysis of finite element solutions for parabolic equations with nonsmooth coefficients in convex polyhedra, *Math. Comp.*, 86 (2017), pp. 1071–1102.
28. A. Lichnewsky and R. Temam, Pseudosolutions of the time-dependent minimal surface problem. *J. Differential Equations*, 30 (1978), pp. 340–364.
29. J.-L. Lions and J. Peetre, Sur une classe d'espaces d'interpolation. *Publications mathématiques de l'IHES*, 19 (1964), pp. 5–68.
30. C. Lubich and A. Ostermann, Runge–Kutta approximations of quasi-linear parabolic equations, *Math. Comp.*, 64 (1995), pp. 601–627.
31. A. Lunardi, *Analytic Semigroups and Optimal Regularity in Parabolic Problems.* Birkhäuser Verlag, Basel (1995).
32. A. Ostermann and M. Thalhammer, Convergence of Runge–Kutta methods for nonlinear parabolic equations, *Appl. Numer. Math.*, 42 (2002), pp. 367–380.
33. J. Prüss and H. Sohr, Imaginary powers of elliptic second order differential operators in L^p-spaces, *Hiroshima Math. J.*, 23 (1993), pp. 161-192.
34. E. Spadaro and U. Stefanelli, A variational view at the time-dependent minimal surface equation. *J. Evolution Equations*, 11 (2011), pp. 793–809.
35. L. Weis, Operator-valued Fourier multiplier theorems and maximal L_p-regularity, *Math. Ann.*, 319 (2001), pp. 735–758.