DFG-Schwerpunktprogramm 1324

„Extraktion quantifizierbarer Information aus komplexen Systemen"

On the convergence analysis of Rothe's method

P.A. Cioica, S. Dahlke, N. Döhring, U. Friedrich, S. Kinzel, F. Lindner, T. Raasch, K. Ritter, R.L. Schilling

Preprint 124

Edited by

AG Numerik/Optimierung
Fachbereich 12 - Mathematik und Informatik
Philipps-Universität Marburg
Hans-Meerwein-Str.
35032 Marburg

DFG-Schwerpunktprogramm 1324

„Extraktion quantifizierbarer Information aus komplexen Systemen"

On the convergence analysis of Rothe's method

P.A. Cioica, S. Dahlke, N. Döhring, U. Friedrich, S. Kinzel, F. Lindner, T. Raasch, K. Ritter, R.L. Schilling

Preprint 124

On the convergence analysis of Rothe's method*

P.A. Cioica, S. Dahlke, N. Döhring, U. Friedrich, S. Kinzel,
F. Lindner, T. Raasch, K. Ritter, R.L. Schilling

Abstract This paper is about the convergence analysis of the horizontal method of lines for deterministic and stochastic parabolic evolution equations. We use uniform discretizations in time and nonuniform (adaptive) discretizations in space. The space discretization methods are assumed to converge up to a given tolerance ε when applied to the resulting elliptic subproblems. Typical examples are adaptive finite element or wavelet methods. We investigate how the tolerances ε in each time step have to be tuned so that the overall scheme converges with the same order as in the case of exact evaluations of the elliptic subproblems. We show that the analysis can be applied to rather general classes of deterministic parabolic problems and arbitrary S-stage discretization schemes. Moreover, also stochastic evolution equations can be treated, at least, if the linearly-implicit Euler scheme is the method of choice. We present a detailed analysis of the case of adaptive wavelet discretizations in space. Using concepts from regularity theory for partial differential equations and from nonlinear approximation theory, we determine an upper bound for the degrees of freedom for the overall scheme that are needed to adaptively approximate the solution up to a prescribed tolerance.

MSC 2010: Primary: 35K90, 60H15, 65J08, 65M20, 65M22, 65T60; secondary: 41A65, 46E35.
Key Words: Deterministic and stochastic parabolic evolution equations, horizontal method of lines, S-stage methods, adaptive wavelet methods, Besov spaces, nonlinear approximation, Brownian motion, stochastic partial differential equations.

1 Introduction

This paper is concerned with the numerical treatment of deterministic and stochastic evolution equations of parabolic type. Such equations describe

*July 11, 2012. This work has been supported by the Deutsche Forschungsgemeinschaft (DFG, grants DA 360/12-2, DA 360/13-2, RI 599/4-2, SCHI 419/5-2), a doctoral scholarship of the Philipps-Universität Marburg, and the LOEWE Center for Synthetic Microbiology (Synmikro), Marburg.

diffusion processes and they are very often used for the mathematical modelling of biological, chemical and physical processes. There are two principally different approaches: the vertical method of lines and the horizontal method of lines. The former starts with an approximation first in space and then in time. We refer to [34, 38, 46] for detailed information. The latter starts with a discretization first in time and then in space; it is also known as Rothe's method. It has been studied in [8, 32, 40, 42]. These references are indicative and by no means complete.

In this paper, we concentrate on Rothe's method for the following reasons. In the horizontal method of lines, the parabolic equation can be interpreted as an abstract Cauchy problem, i.e., as an ordinary deterministic or stochastic differential equation in some suitable function spaces. This immediately provides a way to employ adaptive strategies. Indeed, in time direction we might use one of the well known ODE-solvers or an SDE-solver with step size control. This solver must be based on an implicit discretization scheme since the equation under consideration is usually stiff. Then, in each time step, a system of elliptic equations has to be solved. To this end, as a second level of adaptivity, well-established adaptive numerical schemes based, e.g., on wavelets or finite elements, can be used. We refer to [11,12,20] for the wavelet case, and [2–4, 7, 28–31, 35, 47, 48] for the finite element case. As before, these lists are not complete.

Although the combination of Rothe's method with adaptive strategies is natural, a rigorous convergence analysis seems to be still in its infancy. For deterministic parabolic equations and finite element discretization in space, the most far reaching results have been obtained by [40]. In the stochastic setting Rothe's method with exact evaluation of the elliptic subproblems, has been considered in [8, 32], and explicit convergence rates have been established in [13,33,45]. Results for full discretizations of stochastic equations have been given in, e.g., [26].

Not very much seems to be known for fully adaptive schemes. This paper can be seen as a first step in this direction. We still use uniform discretizations in time, but for the space discretization we use an arbitrary (non-uniform and adaptive) discretization scheme that allows to compute an approximation to the exact solution up to a prescribed accuracy. To treat the convergence problem, we start with the observation that at an abstract level Rothe's method can be re-written as a consecutive application of two operators, the inverse of a (linear) elliptic differential operator and a (nonlinear) evaluation operator. Adaptivity enters via distributing degrees of freedom in each time step for the inexact evaluation of both operators up to a given tolerance. Obviously, we need to know whether the whole scheme still converges with all these perturbations and how the tolerances in each time step have to be tuned to obtain convergence and corresponding convergence orders. These aspects are studied in Section 2.

In the subsequent sections we apply this analysis to the following concrete situations: in Section 3, we study deterministic equations of the form

$$u'(t) = Au(t) + f(t, u(t)), \quad u(0) = u_0$$

and, in Section 4, we consider stochastic evolution equations of the form

$$\mathrm{d}u(t) = \Big(Au(t) + f(u(t))\Big)\,\mathrm{d}t + B(u(t))\,\mathrm{d}W(t), \quad u(0) = u_0.$$

In practical applications, usually A is a differential operator, f a linear or nonlinear forcing term and, in the stochastic case, $B(u(t))\,\mathrm{d}W(t)$ describes additive or multiplicative noise. Formally, both equations are special cases of the following general problem

$$\mathrm{d}u(t) = F(t, u(t))\,\mathrm{d}t + B(u(t))\,\mathrm{d}W(t), \quad u(0) = u_0. \tag{1}$$

(Clearly, the deterministic setting can be obtained by setting $B \equiv 0$.)

This paper is organized as follows. In Section 2 we state the abstract setting of Rothe's method and derive sufficient conditions for convergence in the case of inexact operator evaluations. In particular, we obtain conditions on the tolerances of the elliptic subproblems which guarantee the overall convergence; moreover, we derive abstract complexity estimates. In Section 3 we apply the whole machinery to deterministic parabolic equations. Then in Section 4 we study the stochastic counterpart. Finally, in Section 5 we apply and substantiate the abstract analysis by considering the case that implementable adaptive wavelet methods are used for the numerical treatment of the elliptic stage equations.

2 Abstract description of Rothe's method

We begin with an example that motivates our perspective on the analysis of Rothe's method. The setting and notation will be given in Subsection 2.2, and in Subsection 2.3 we state and prove one of our main results, that is an abstract convergence proof.

2.1 Motivation

To introduce our abstract setting of Rothe's method, let us consider the heat equation

$$\left.\begin{aligned} u'(t) &= \Delta u(t) + f(t, u(t)) && \text{on } \mathcal{O},\ t \in (0, T], \\ u(0) &= u_0 && \text{on } \mathcal{O}, \\ u &= 0 && \text{on } \partial\mathcal{O},\ t \in (0, T], \end{aligned}\right\} \tag{2}$$

where $\mathcal{O} \subset \mathbb{R}^d$, $d \geq 1$, denotes a bounded Lipschitz domain. We discretize this equation by means of a linearly-implicit Euler scheme with uniform time steps. Let $K \in \mathbb{N}$ be the number of subdivisions of the time interval $[0, T]$, where the step size will be denoted by $\tau := T/K$. The k-th point in time is denoted by $t_k := \tau k$, $k \in \{0, \dots, K\}$. The linearly-implicit Euler scheme, starting at u_0, is given by

$$\frac{u_{k+1} - u_k}{\tau} = \Delta u_{k+1} + f(t_k, u_k),$$

i.e.,

$$(I - \tau\Delta)u_{k+1} = u_k + \tau f(t_k, u_k), \tag{3}$$

for $k = 0, \dots, K-1$. If we assume that the elliptic problem

$$L_\tau v := (I - \tau\Delta)v = w \quad \text{on } \mathcal{O}, \quad v|_{\partial\mathcal{O}} = 0,$$

can be solved exactly, then we see that one step of the linearly-implicit Euler scheme (3) can be written as

$$u_{k+1} = L_\tau^{-1} R_{\tau,k}(u_k), \tag{4}$$

where

$$R_{\tau,k}(v) := v + \tau f(t_k, v)$$

and L_τ is a boundedly invertible operator between suitable Hilbert spaces. That is, we can look at this equation in a Gelfand triple setting ($H_0^1(\mathcal{O})$, $L_2(\mathcal{O})$, $H^{-1}(\mathcal{O})$) with L_τ as an operator from $H_0^1(\mathcal{O})$ to $H^{-1}(\mathcal{O})$. We may also consider (4) in $L_2(\mathcal{O})$, since $H_0^1(\mathcal{O})$ is embedded in $L_2(\mathcal{O})$ and $L_2(\mathcal{O})$ is embedded in $H^{-1}(\mathcal{O})$, provided that $R_{\tau,k} : L_2(\mathcal{O}) \to L_2(\mathcal{O})$ is well defined.

Having the above simple example in mind, we observe that the fundamental form of (4) essentially remains the same even if we introduce

- more sophisticated discretizations in time, e.g., as outlined below and in Section 3;
- more complicated source terms, including, e.g., stochastic components, see Section 4.

We are going to exploit this observation in the following subsections.

2.2 Setting and assumptions

In many applications not only one-stage approximation methods, such as the linearly-implicit Euler scheme, are used, but also more sophisticated S-stage schemes. Compared to one-stage schemes, S-stage schemes can lead to higher convergence orders of the approximation in time direction, see also Section 3 for further details. Therefore, in this subsection we state a variant of (4) that provides an abstract interpretation of S-stage schemes, where $S \in \mathbb{N}$.

As above, we begin with a uniform discretization of the time interval $[0,T]$ with $K \in \mathbb{N}$ subdivisions, step size $\tau := T/K$, and $t_k := k\tau$ for $k \in \{0,\dots,K\}$. Taking an abstract point of view, we introduce separable real Hilbert spaces $\mathcal{H}$, $\mathcal{G}$, and consider a function $u : [0,T] \to \mathcal{H}$. Furthermore, let $L_{\tau,i}$ be a family of, possibly unbounded, linear operators which have bounded inverses $L_{\tau,i}^{-1} : \mathcal{G} \to \mathcal{H}$, and let

$$R_{\tau,k,i} : \underbrace{\mathcal{H} \times \ldots \times \mathcal{H}}_{i} \to \mathcal{G} \tag{5}$$

be a family of (nonlinear) evaluation operators for $k \in \{0,\dots,K-1\}$ and $i = 1,\dots,S$. As the norm on the Cartesian product in (5) we set

$$\|(v_1,\dots,v_i)\|_{\mathcal{H}\times\ldots\times\mathcal{H}} := \sum_{l=1}^{i} \|v_l\|_{\mathcal{H}}.$$

Remark 2.1. **(i)** The function $u : [0,T] \to \mathcal{H}$ is understood to be a solution of a (deterministic or stochastic) parabolic partial differential equation of the form (1).

(ii) In most cases $L_{\tau,i}^{-1}$ is not given explicitly and, for this reason, we need an efficient numerical scheme for its evaluation. The situation is completely different with $R_{\tau,k,i}$, which is usually given explicitly and does not require the solution of an operator equation for its evaluation. Concrete examples of these operators will be presented and studied in Sections 3 and 4.

(iii) In a Gelfand triple setting (V,U,V^*) typical choices for the spaces $\mathcal{H}$ and $\mathcal{G}$ are $\mathcal{H} = V$, $\mathcal{G} = V^*$ or $\mathcal{H} = \mathcal{G} = U$. However, also a more general setting such as

$$V \subseteq \mathcal{H} \subseteq U \subseteq V^* \subseteq \mathcal{G}.$$

is possible. Observe that our motivating example from Subsection 2.1 fits in this setting with $H_0^1(\mathcal{O}) = \mathcal{H} \subseteq L_2(\mathcal{O})$ and $\mathcal{G} = H^{-1}(\mathcal{O})$. In Section 4 an even more general setting will be studied.

Starting from the given value $u_0 := u(0) \in \mathcal{H}$, we define the *abstract exact S-stage scheme* iteratively by

$$\left.\begin{aligned} u_{k+1} &:= \sum_{i=1}^{S} w_{k,i}, \\ w_{k,i} &:= L_{\tau,i}^{-1} R_{\tau,k,i}(u_k, w_{k,1},\dots,w_{k,i-1}), \quad i = 1,\dots,S, \end{aligned}\right\} \tag{6}$$

for $k = 0,\dots,K-1$. One step of this iteration can be described as an application of the operator

$$\left.\begin{aligned} E_{\tau,k,k+1} &: \mathcal{H} \to \mathcal{H}, \\ & v \mapsto \sum_{i=1}^{S} w_{k,i}(v), \\ w_{k,i}(v) &:= L_{\tau,i}^{-1} R_{\tau,k,i}(v, w_{k,1}(v),\dots,w_{k,i-1}(v)), \quad i = 1,\dots,S. \end{aligned}\right\} \tag{7}$$

If we define the family of operators

$$E_{\tau,j,k} := \begin{cases} E_{\tau,k-1,k} \circ \ldots \circ E_{\tau,j,j+1}, & j < k \\ I, & j = k, \end{cases} \tag{8}$$

then the output of the exact S-stage scheme (6) is given by the sequence

$$u_k = E_{\tau,0,k}(u_0), \quad k = 0, \ldots, K. \tag{9}$$

The convergence analysis which we present relies on a crucial technical assumption on the operators defined in (8).

Assumption 2.2. For all $0 \leq j, k \leq K$ the operators

$$E_{\tau,j,k} : \mathcal{H} \to \mathcal{H} \text{ are globally Lipschitz continuous}$$

with Lipschitz constants $C^{\text{Lip}}_{\tau,j,k}$.

Remark 2.3. Assumption 2.2 is relatively mild, as it is usually fulfilled in the applications we have in mind. Concrete examples will be given at the end of this section, as well as in Sections 3 and 4.

We call the sequence (9) the *output of the exact S-stage scheme*, since the operators involved in the definition of $E_{\tau,0,k}$ are evaluated exactly. In practical applications this is very often not possible; the operators $L^{-1}_{\tau,i}$ and $R_{\tau,k,i}$ can only be evaluated inexactly and, at best, this is possible up to given tolerances. Therefore, as a start, we make the following

Assumption 2.4. For all $\tau > 0$, $k \in \{0, \ldots, K-1\}$, and for any prescribed tolerance ε_k and arbitrary $v \in \mathcal{H}$, we have an approximation $\tilde{E}_{\tau,k,k+1}(v)$ of $E_{\tau,k,k+1}(v)$ at hand, such that

$$\|E_{\tau,k,k+1}(v) - \tilde{E}_{\tau,k,k+1}(v)\|_{\mathcal{H}} \leq \varepsilon_k$$

with a known upper bound $M_{\tau,k}(\varepsilon_k, v) < \infty$ for the degrees of freedom needed to achieve the prescribed tolerance ε_k.

For simplicity, we make the following

Assumption 2.5. The initial value is given exactly, i.e.,

$$\tilde{u}_0 := u(0).$$

Remark 2.6. The case where Assumption 2.5 does not hold, i.e., $\tilde{u}_0 \neq u(0)$, can be handled in a similar way. However, this only increases *notational* complexity.

Given an approximation scheme satisfying Assumption 2.4 and using Assumption 2.5, the *abstract inexact* variant of (6) is defined by

$$\left.\begin{aligned} \tilde{u}_0 &:= u(0), \\ \tilde{u}_{k+1} &:= \tilde{E}_{\tau,k,k+1}(\tilde{u}_k) \quad \text{for } k = 0, \ldots, K-1. \end{aligned}\right\} \tag{10}$$

We will show in Theorem 2.17 how to tune the tolerances $(\varepsilon_k)_{k=0,\dots,K-1}$ in such a way that the scheme (10) has the same qualitative properties as the exact scheme (6). As in (8), we define

$$\tilde{E}_{\tau,j,k} := \begin{cases} \tilde{E}_{\tau,k-1,k} \circ \dots \circ \tilde{E}_{\tau,j,j+1}, & j < k \\ I, & j = k. \end{cases}$$

Consequently, the output of the inexact S-stage scheme (10) is given by

$$\tilde{u}_k = \tilde{E}_{\tau,0,k}(u(0)), \quad k = 0, \dots, K.$$

Now, we are faced with the following problems. In practice, the Lipschitz constants $C^{\text{Lip}}_{\tau,j,k}$ of $E_{\tau,j,k}$, given by Assumption 2.2, might be hard to estimate directly. As we shall see in Section 5, the individual operators $L^{-1}_{\tau,i}R_{\tau,k,i}$, $i = 1, \dots, S$, are much easier to handle. Moreover, a direct approximation scheme for $E_{\tau,j,k}$, as required by Assumption 2.4, might be hard to get. Nevertheless, very often, one has convergent numerical schemes for the individual operators $L^{-1}_{\tau,i}R_{\tau,k,i}$. Therefore, with these observations in mind, we are now going to state the corresponding assumptions for these individual operators.

Assumption 2.7. For $k = 0, \dots, K-1$ and $i = 1, \dots, S$ the operators

$$L^{-1}_{\tau,i}R_{\tau,k,i} : \underbrace{\mathcal{H} \times \dots \times \mathcal{H}}_{i} \to \mathcal{H} \text{ are globally Lipschitz continuous}$$

with Lipschitz constants $C^{\text{Lip}}_{\tau,k,(i)}$.

Remark 2.8. Note that, on the one hand, Assumption 2.2 is slightly more general than Assumption 2.7, since it is easy to see that a composition of non-Lipschitz continuous operators can be Lipschitz continuous. On the other hand, Assumption 2.7 implies Assumption 2.2. This is a consequence of the fact, that, if we introduce the constants

$$C'_{\tau,k,(i)} := \prod_{l=i+1}^{S} (1 + C^{\text{Lip}}_{\tau,k,(l)}) \tag{11}$$

for $k = 0, \dots, K-1$ and $i = 0, \dots, S$, we can estimate the Lipschitz constants $C^{\text{Lip}}_{\tau,j,k}$ of $E_{\tau,j,k}$ as follows:

$$C^{\text{Lip}}_{\tau,j,k} \leq \prod_{r=j}^{k-1} \left(C'_{\tau,r,(0)} - 1 \right), \quad 0 \leq j \leq k \leq K. \tag{12}$$

This will be worked out in detail in the proof of Theorem 2.20.

The analogue to Assumption 2.4 is

Assumption 2.9. For all $\tau > 0$, $k \in \{0, \dots, K-1\}$, $i \in \{1, \dots, S\}$, there exists a numerical scheme that, for any prescribed tolerance ε and arbitrary $v_0, \dots, v_{i-1} \in \mathcal{H}$, yields an approximation $[v]_\varepsilon$ of

$$v := L_{\tau,i}^{-1} R_{\tau,k,i}(v_0, \dots, v_{i-1}),$$

such that

$$\|v - [v]_\varepsilon\|_\mathcal{H} \leq \varepsilon$$

with a known upper bound $M_{\tau,k,i}(\varepsilon, v) < \infty$ for the degrees of freedom needed to achieve the prescribed tolerance ε.

For any numerical scheme satisfying Assumption 2.9, and given tolerances $\varepsilon_{k,i}$, $k = 0, \dots, K-1$, $i = 1, \dots, S$, the corresponding *inexact* variant of (6) is defined by

$$\left.\begin{aligned} \tilde{u}_0 &:= u(0), \\ \tilde{u}_{k+1} &:= \sum_{i=1}^{S} \tilde{w}_{k,i}, \\ \tilde{w}_{k,i} &:= [L_{\tau,i}^{-1} R_{\tau,k,i}(\tilde{u}_k, \tilde{w}_{k,1}, \dots, \tilde{w}_{k,i-1})]_{\varepsilon_{k,i}}, \quad i = 1, \dots, S, \end{aligned}\right\} \tag{13}$$

for $k = 0, \dots, K-1$. Note that (13) is consistent with (10), since it corresponds to the specific choice

$$\left.\begin{aligned} \tilde{E}_{\tau,k,k+1} &: \mathcal{H} \to \mathcal{H}, \\ & v \mapsto \sum_{i=1}^{S} \tilde{w}_{k,i}(v), \\ \tilde{w}_{k,i}(v) &:= [L_{\tau,i}^{-1} R_{\tau,k,i}(v, \tilde{w}_{k,1}(v), \dots, \tilde{w}_{k,i-1}(v))]_{\varepsilon_{k,i}}, \quad i = 1, \dots, S. \end{aligned}\right\} \tag{14}$$

In Theorem 2.22 we will show how to tune the tolerances in the scheme (13) in such a way that the approximation of u in $\mathcal{H}$ has the same qualitative properties as the exact scheme (6).

Remark 2.10. **(i)** For $\tilde{E}_{\tau,k,k+1}$ as in (14) and arbitrary $v \in \mathcal{H}$, the estimate

$$\|E_{\tau,k,k+1}(v) - \tilde{E}_{\tau,k,k+1}(v)\|_\mathcal{H} \leq \sum_{i=1}^{S} C'_{\tau,k,(i)} \varepsilon_{k,i} \tag{15}$$

holds with $C'_{\tau,k,(i)}$ given by (11). Thus, for any prescribed tolerance ε_k, if Assumptions 2.9 and 2.7 are fulfilled, we can choose $\varepsilon_{k,i}$, $i = 1, \dots, S$, in such a way that the error we make by applying $\tilde{E}_{\tau,k,k+1}$ from (14) instead of $E_{\tau,k,k+1}$ is bounded by ε_k, uniformly in $\mathcal{H}$. In this sense Assumption 2.9 implies Assumption 2.4. Detailed arguments for the validity of estimate (15), will be given in the proof of Theorem 2.20.

(ii) We do not specify the numerical scheme $[\,\cdot\,]_\varepsilon$ at this point. It might be based on, e.g., a spectral method, an (adaptive) finite element scheme, or an adaptive wavelet solver. The latter case will be discussed in detail in Section 5. There, $M_{\tau,k,i}(\varepsilon, v)$ will be an upper bound for the number of elements in the wavelet basis that is needed to achieve the desired tolerance.

(iii) Later on, in Section 5, we will assume that $R_{\tau,k,i}$ can be evaluated exactly. This, of course, may not always be possible. Especially not in the stochastic setting of Section 4, see, e.g., Remark 4.11. We postpone the analysis of these additional difficulties to a forthcoming paper.

2.3 Controlling the error of the inexact schemes

We want to use the schemes described in Subsection 2.2 to compute approximations to a solution $u : [0,T] \to \mathcal{H}$ of a (deterministic or stochastic) parabolic partial differential equation.

Assumption 2.11. There exists a unique solution $u : [0,T] \to \mathcal{H}$ to the problem under consideration, i.e., to (1).

Remark 2.12. Of course, the type of such solutions depends on the form of the specific parabolic partial differential equation. We avoid, on purpose, a detailed discussion of these issues in this subsection. Further information can be found in Section 3, Remark 3.9 and in Section 4, Proposition 4.7.

The analysis presented in this section is based on the following central

Assumption 2.13. The *exact* scheme (6) converges to $u(T)$ with order $\delta > 0$, i.e., for some constant $C_{\text{exact}} > 0$,

$$\|u(T) - E_{\tau,0,K}(u(0))\|_{\mathcal{H}} \le C_{\text{exact}}\, \tau^{\delta}.$$

Remark 2.14. Error estimates as in Assumption 2.13 are quite natural and hold very often, see Section 3 and the references therein, in particular, [42, Theorem 6.2]. We also refer to Section 4, Proposition 4.15 and Remark 4.16, as well as to [45, Theorem 3.2].

The main goal of this subsection is to state conditions how to tune the tolerances in the inexact schemes (10) and (13) so that they still converge to u and inherit the approximation order of Assumption 2.13.

At first, we give an estimate for the error propagation of the scheme (10) measured in the norm of $\mathcal{H}$.

Theorem 2.15. *Suppose that Assumptions 2.2, 2.4, 2.5, and 2.11 hold. Let $(u_k)_{k=0}^K$, $K \in \mathbb{N}$, be the output of the exact scheme* (6)*, and let $(\tilde{u}_k)_{k=0}^K$ be the output of the scheme* (10) *with given tolerances ε_k, $k = 0, \ldots, K-1$. Then, for all $0 \le k \le K$,*

$$\|u(t_k) - \tilde{u}_k\|_{\mathcal{H}} \le \|u(t_k) - u_k\|_{\mathcal{H}} + \sum_{j=0}^{k-1} C^{\text{Lip}}_{\tau,j+1,k}\, \varepsilon_j.$$

Proof. The triangle inequality yields

$$\|u(t_k) - \tilde{u}_k\|_{\mathcal{H}} \leq \|u(t_k) - u_k\|_{\mathcal{H}} + \|u_k - \tilde{u}_k\|_{\mathcal{H}},$$

so it remains to estimate the second term. For simplicity we write $E_{j,k} := E_{\tau,j,k}$. Using $u_0 = \tilde{u}_0$ and writing $u_k - \tilde{u}_k$ as a telescopic sum, we get

$$\begin{aligned} u_k - \tilde{u}_k &= (E_{0,k}(\tilde{u}_0) - E_{1,k}\tilde{E}_{0,1}(\tilde{u}_0)) \\ &\qquad + (E_{1,k}\tilde{E}_{0,1}(\tilde{u}_0) - E_{2,k}\tilde{E}_{0,2}(\tilde{u}_0)) \\ &\qquad\quad \dots \\ &\qquad + (E_{k-1,k}\tilde{E}_{0,k-1}(\tilde{u}_0) - \tilde{E}_{0,k}(\tilde{u}_0)) \\ &= \sum_{j=0}^{k-1} \big(E_{j,k}\tilde{E}_{0,j}(u_0) - E_{j+1,k}\tilde{E}_{0,j+1}(u_0)\big). \end{aligned}$$

Another application of the triangle inequality yields

$$\|u_k - \tilde{u}_k\|_{\mathcal{H}} \leq \sum_{j=0}^{k-1} \|E_{j,k}\tilde{E}_{0,j}(u_0) - E_{j+1,k}\tilde{E}_{0,j+1}(u_0)\|_{\mathcal{H}}.$$

Due to the Lipschitz continuity of $E_{\tau,j,k}$, cf. Assumption 2.2, each term in the sum can be estimated from above by

$$\begin{aligned} &\|E_{j,k}\tilde{E}_{0,j}(u_0) - E_{j+1,k}\tilde{E}_{0,j+1}(u_0)\|_{\mathcal{H}} \\ &\quad = \|E_{j+1,k}E_{j,j+1}\tilde{E}_{0,j}(u_0) - E_{j+1,k}\tilde{E}_{0,j+1}(u_0)\|_{\mathcal{H}} \\ &\quad \leq C^{\mathrm{Lip}}_{\tau,j+1,k}\|E_{j,j+1}\tilde{E}_{0,j}(u_0) - \tilde{E}_{0,j+1}(u_0)\|_{\mathcal{H}}. \end{aligned} \tag{16}$$

With $\tilde{E}_{0,j}(u_0) = \tilde{u}_j$ and using Assumption 2.4, we observe

$$\|E_{j,j+1}\tilde{E}_{0,j}(u_0) - \tilde{E}_{0,j+1}(u_0)\|_{\mathcal{H}} = \|E_{j,j+1}(\tilde{u}_j) - \tilde{E}_{j,j+1}(\tilde{u}_j)\|_{\mathcal{H}} \leq \varepsilon_j. \quad \square$$

Remark 2.16. In the description of our abstract setting we have chosen the spaces $\mathcal{H}$ and $\mathcal{G}$ to be the same in all time steps. In Section 4 we will encounter the situation where a proper definition of the operators $L_{\tau,1}^{-1}$ and $R_{\tau,k,1}$ may require different Hilbert spaces $\mathcal{H}_k$ and $\mathcal{G}_k$ for each $0 \leq k \leq K$, $K \in \mathbb{N}$. If we guarantee the Lipschitz continuity of the mappings $E_{\tau,j,k} : \mathcal{H}_j \to \mathcal{H}_k$ with corresponding Lipschitz constants $C^{\mathrm{Lip}}_{\tau,j,k}$, Theorem 2.15 stays true with $\mathcal{H}$ replaced by $\mathcal{H}_k$, cf. Section 4, Observation 4.12.

Based on Theorem 2.15 we are now able to state the conditions on the tolerances $(\varepsilon_k)_{k=0,\dots,K-1}$ such that for the scheme (10) our main goal is achieved.

Theorem 2.17. *Suppose that Assumptions 2.2, 2.4, 2.5, and 2.11 hold. Let Assumption 2.13 hold for some $\delta > 0$. If we consider the case of inexact operator evaluations as described in* (10) *and choose*

$$\varepsilon_k \leq (C^{\mathrm{Lip}}_{\tau,k+1,K})^{-1}\, \tau^{1+\delta},$$

$k = 0, \dots, K-1$, *then we get*

$$\|u(T) - \tilde{E}_{\tau,0,K}(u(0))\|_{\mathcal{H}} \leq (C_{\text{exact}} + T)\,\tau^{\delta}.$$

Proof. Applying Theorem 2.15, Assumption 2.13 and $K = T/\tau$, we obtain

$$\begin{aligned}
\|u(t_K) - \tilde{u}_K\|_{\mathcal{H}} &\leq \|u(t_K) - u_K\|_{\mathcal{H}} + \sum_{k=0}^{K-1} C^{\text{Lip}}_{\tau,k+1,K}\,\varepsilon_k \\
&\leq C_{\text{exact}}\,\tau^{\delta} + \sum_{k=0}^{K-1} C^{\text{Lip}}_{\tau,k+1,K}\,(C^{\text{Lip}}_{\tau,k+1,K})^{-1}\,\tau^{1+\delta} \\
&= C_{\text{exact}}\,\tau^{\delta} + K\tau^{1+\delta} = (C_{\text{exact}} + T)\,\tau^{\delta}. \qquad \square
\end{aligned}$$

One of the final goals of our analysis is to provide upper estimates for the overall complexity of the resulting scheme. As a first step, in this section, we provide a quite abstract version, which is a direct consequence of Theorem 2.17.

Corollary 2.18. *Suppose that the assumptions of Theorem 2.17 are satisfied. Choose*

$$\varepsilon_k := (C^{\text{Lip}}_{\tau,k+1,K})^{-1}\,\tau^{1+\delta},$$

for $k = 0, \dots, K-1$, *then* $\tilde{E}_{\tau,0,K}(u_0)$ *requires at most*

$$M_{\tau,T}(\delta, (\varepsilon_k)) := \sum_{k=0}^{K-1} M_{\tau,k}(\varepsilon_k, E_{\tau,k,k+1}(\tilde{u}_k))$$

degrees of freedom.

Remark 2.19. At this point, without specifying an approximation scheme and therefore without a concrete knowledge of $M_{\tau,k}(\varepsilon, \cdot)$, Corollary 2.18 might not look very deep. Nevertheless, it will be filled with content in Section 5. There, we will discuss the specific case of adaptive wavelet solvers for which concrete estimates for $M_{\tau,k}(\varepsilon, \cdot)$ are available.

The next step is to play the same game for the inexact scheme (13). We start again by controlling the error propagation.

Theorem 2.20. *Suppose that Assumptions 2.5, 2.7, 2.9, and 2.11 hold. Let* $(u_k)_{k=0}^{K}$, $K \in \mathbb{N}$, *be the output of the exact scheme* (6), *and let* $(\tilde{u}_k)_{k=0}^{K}$ *be the output of the inexact scheme* (13) *with prescribed tolerances* $\varepsilon_{k,i}$, $k = 0, \dots, K-1$, $i = 1, \dots, S$. *Then, for all* $0 \leq k \leq K$,

$$\|u(t_k) - \tilde{u}_k\|_{\mathcal{H}} \leq \|u(t_k) - u_k\|_{\mathcal{H}} + \sum_{j=0}^{k-1} \Big(\prod_{l=j+1}^{k-1} (C'_{\tau,l,(0)} - 1) \Big) \sum_{i=1}^{S} C'_{\tau,j,(i)} \varepsilon_{j,i}.$$

Proof. We just have to repeat the proof of Theorem 2.15 with the special choice (14) for $\tilde{E}_{\tau,k,k+1}$ and change two things. First, instead of the exact Lipschitz constants $C_{\tau,j+1,k}$ in (16), we can use their estimates (12) presented in Remark 2.8(i). Second, in the last step of the proof of Theorem 2.15, we may estimate the error we make when using $\tilde{E}_{\tau,j,j+1}$ instead of $E_{\tau,j,j+1}$ as in Remark 2.10(i). Thus, to finish the proof we have to show that the estimates (12) and (15) hold.

We start with (12). Note that, it is enough to show that

$$C^{\mathrm{Lip}}_{\tau,k,k+1} \leq C'_{\tau,k,(0)} - 1, \quad 0 \leq k \leq K-1, \tag{17}$$

since, obviously,

$$C^{\mathrm{Lip}}_{\tau,j,k} \leq \prod_{r=j}^{k-1} C^{\mathrm{Lip}}_{\tau,r,r+1}, \quad 0 \leq j \leq k \leq K.$$

Thus, let us prove that (17) is true, if Assumption 2.7 holds. To this end, we fix $k \in \{0, \ldots, K-1\}$ as well as arbitrary $u, v \in \mathcal{H}$. Using (7) and the triangle inequality, we obtain

$$\|E_{\tau,k,k+1}(u) - E_{\tau,k,k+1}(v)\|_{\mathcal{H}} \leq \sum_{i=1}^{S} \|w_{k,i}(u) - w_{k,i}(v)\|_{\mathcal{H}}. \tag{18}$$

Applying Assumption 2.7, we get for each $i \in \{1, \ldots, S\}$:

$$\|w_{k,i}(u) - w_{k,i}(v)\|_{\mathcal{H}} \leq C^{\mathrm{Lip}}_{\tau,k,(i)} \Big(\|u-v\|_{\mathcal{H}} + \sum_{l=1}^{i-1} \|w_{k,l}(u) - w_{k,l}(v)\|_{\mathcal{H}} \Big).$$

Hence, for $r = 0, \ldots, S-1$, we have

$$\begin{aligned} \sum_{i=1}^{r+1} \|w_{k,i}(u) - w_{k,i}(v)\|_{\mathcal{H}} \leq \big(1 + C^{\mathrm{Lip}}_{\tau,k,(r+1)}\big) \sum_{i=1}^{r} \|w_{k,i}(u) - w_{k,i}(v)\|_{\mathcal{H}} \\ + C^{\mathrm{Lip}}_{\tau,k,(r+1)} \|u-v\|_{\mathcal{H}}. \end{aligned} \tag{19}$$

By induction, it is easy to show, that inequalities of the form $e_{r+1} \leq a_r e_r + b_r$, with $e_0 = 0$, imply

$$e_r \leq \sum_{j=1}^{r} b_{j-1} \prod_{l=j+1}^{r} a_{l-1}. \tag{20}$$

In our situation, this fact leads to the estimate

$$\sum_{i=1}^{S} \|w_{k,i}(u) - w_{k,i}(v)\|_{\mathcal{H}} \leq \sum_{i=1}^{S} C^{\mathrm{Lip}}_{\tau,k,(i)} \|u-v\|_{\mathcal{H}} \prod_{l=i+1}^{S} \big(1 + C^{\mathrm{Lip}}_{\tau,k,(l)}\big),$$

since (19) holds for $r = 0, \dots, S-1$. Furthermore, we can use the equality

$$\sum_{i=1}^{S} C^{\mathrm{Lip}}_{\tau,k,(i)} \prod_{l=i+1}^{S} (1 + C^{\mathrm{Lip}}_{\tau,k,(l)}) = \prod_{i=1}^{S} (1 + C^{\mathrm{Lip}}_{\tau,k,(i)}) - 1 = C'_{\tau,k,(0)} - 1$$

to obtain

$$\sum_{i=1}^{S} \|w_{k,i}(u) - w_{k,i}(v)\|_{\mathcal{H}} \leq \big(C'_{\tau,k,(0)} - 1\big) \|u - v\|_{\mathcal{H}}.$$

Together with (18), this proves (17).

Finally, let us move to the estimate (15). Fix $k \in \{0, \dots, K-1\}$ and let $\tilde{E}_{\tau,k,k+1}$ be given by (14) with the prescribed tolerances $\varepsilon_{k,i}$, $i = 1, \dots, S$, from our assertion. Then, for arbitrary $v \in \mathcal{H}$, we have

$$\|E_{\tau,k,k+1}(v) - \tilde{E}_{\tau,k,k+1}(v)\|_{\mathcal{H}} \leq \sum_{i=1}^{S} \|w_{k,i}(v) - \tilde{w}_{k,i}(v)\|_{\mathcal{H}}. \tag{21}$$

Using the triangle inequality, as well as Assumption 2.7, we obtain for every $i = 1, \dots, S$,

$$\begin{aligned}
&\|w_{k,i}(v) - \tilde{w}_{k,i}(v)\|_{\mathcal{H}} \\
&= \Big\| L_{\tau,i}^{-1} R_{\tau,k,i}(v, w_{k,1}(v), \dots, w_{k,i-1}(v)) \\
&\qquad\qquad - \big[L_{\tau,i}^{-1} R_{\tau,k,i}(v, \tilde{w}_{k,1}(v), \dots, \tilde{w}_{k,i-1}(v))\big]_{\varepsilon_{k,i}} \Big\|_{\mathcal{H}} \\
&\leq \Big\| L_{\tau,i}^{-1} R_{\tau,k,i}(v, w_{k,1}(v), \dots, w_{k,i-1}(v)) - L_{\tau,i}^{-1} R_{\tau,k,i}(v, \tilde{w}_{k,1}(v), \dots, \tilde{w}_{k,i-1}(v)) \Big\|_{\mathcal{H}} \\
&\quad + \Big\| L_{\tau,i}^{-1} R_{\tau,k,i}(v, \tilde{w}_{k,1}(v), \dots, \tilde{w}_{k,i-1}(v)) \\
&\qquad\qquad - \big[L_{\tau,i}^{-1} R_{\tau,k,i}(v, \tilde{w}_{k,1}(v), \dots, \tilde{w}_{k,i-1}(v))\big]_{\varepsilon_{k,i}} \Big\|_{\mathcal{H}} \\
&\leq C^{\mathrm{Lip}}_{\tau,k,(i)} \sum_{l=1}^{i-1} \|w_{k,l}(v) - \tilde{w}_{k,l}(v)\|_{\mathcal{H}} + \varepsilon_{k,i}.
\end{aligned}$$

Thus, for $r = 0, \dots, S-1$,

$$\sum_{i=1}^{r+1} \|w_{k,i}(v) - \tilde{w}_{k,i}(v)\|_{\mathcal{H}} \leq \big(1 + C^{\mathrm{Lip}}_{\tau,k,(r+1)}\big) \sum_{i=1}^{r} \|w_{k,i}(v) - \tilde{w}_{k,i}(v)\|_{\mathcal{H}} + \varepsilon_{k,i}.$$

Arguing as above, cf. (20), we get

$$\sum_{i=1}^{S} \|w_{k,i}(v) - \tilde{w}_{k,i}(v)\|_{\mathcal{H}} \leq \sum_{i=1}^{S} \varepsilon_{k,i} \prod_{l=i+1}^{S} (1 + C^{\mathrm{Lip}}_{\tau,k,(l)}) = \sum_{i=1}^{S} C'_{\tau,k,(i)} \varepsilon_{k,i}.$$

Together with (21), this proves (15). □

Remark 2.21. On the one hand, by construction, Theorem 2.20 is slightly weaker than Theorem 2.15. On the other hand, from the practical point of view, Theorem 2.20 is more realistic for the following reasons. As already outlined above, in many cases, estimates for the Lipschitz constants according to Assumption 2.7 are available, while *direct* estimates for the whole composition as in Assumption 2.2 might be hard to get and are only available in very specific situations. Moreover, let us once again recall the fact that approximation schemes for $E_{\tau,j,k}$ are hard to get directly, whereas this is much easier for the individual operator equations.

Based on Theorem 2.20, we are able to state the conditions on the tolerances $\varepsilon_{k,i}$, $k = 0, \dots, K-1$, $i = 1, \dots, S$, such that the scheme (13) converges with the desired order. We put

$$C''_{\tau,k} := \prod_{l=k+1}^{K-1} \Big(C'_{\tau,l,(0)} - 1 \Big) \tag{22}$$

for $k = 0, \dots, K-1$.

Theorem 2.22. *Suppose that Assumptions 2.5, 2.7, 2.9, and 2.11 hold. Let Assumption 2.13 hold for some $\delta > 0$. If we consider the case of inexact operator evaluations as described in* (13) *and choose*

$$\varepsilon_{k,i} \leq \frac{1}{S} \Big(C''_{\tau,k} C'_{\tau,k,(i)} \Big)^{-1} \tau^{1+\delta}, \tag{23}$$

then we get

$$\|u(T) - \tilde{u}_K\|_{\mathcal{H}} \leq (C_{\text{exact}} + T)\, \tau^{\delta}. \tag{24}$$

Proof. Applying Theorem 2.20, Assumption 2.13, and choosing $\varepsilon_{k,i}$ as in (23), we obtain

$$\begin{aligned} \|u(t_K) - \tilde{u}_K\|_{\mathcal{H}} &\leq \|u(t_K) - u_K\|_{\mathcal{H}} + \sum_{k=0}^{K-1} \sum_{i=1}^{S} C''_{\tau,k} C'_{\tau,k,(i)}\, \varepsilon_{k,i} \\ &= (C_{\text{exact}} + T)\, \tau^{\delta}. \end{aligned}$$

□

Remark 2.23. **(i)** Let us take a closer look at condition (23). The number of factors in $C''_{\tau,k}$ is proportional to $K - k$, so that the tolerances are allowed to grow with k (if all factors in $C''_{\tau,k}$ are greater than or equal to 1, which is usually the case). This means that the stage equations at earlier time steps have to be solved with higher accuracy compared to those towards the end of the iteration. This corresponds to the well known smoothing character of parabolic equations. Furthermore, the number of factors in $C'_{\tau,k,(i)}$ is proportional to $S - i$, but independent of k. Consequently, also the early stages have to be solved with higher accuracy compared to the later ones.

(ii) In Theorem 2.22, i.e., (23) a specific choice for the tolerances $\varepsilon_{k,i}$, $k = 0, \ldots, K-1$, $i = 1, \ldots, S$, has been used. Essentially, it is an equilibrium strategy. However, also alternative choices are possible. Indeed, an inspection of the proof of Theorem 2.22 shows that any choice of $\varepsilon_{k,i}$ satisfying

$$\sum_{i=1}^{S} C'_{\tau,k,(i)} \varepsilon_{k,i} \leq (C''_{\tau,k})^{-1}\, \tau^{1+\delta}$$

would also be sufficient.

(iii) In practical applications, it would be natural to use the additional flexibility for the choice of $\varepsilon_{k,i}$ as outlined in (ii) to minimize the overall computational cost of the method, given by

$$M_{\tau,T}(\delta) := M_{\tau,T}(\delta, (\varepsilon_{k,i})_{k,i}) := \sum_{k=0}^{K-1} \sum_{i=1}^{S} M_{\tau,k,i}(\varepsilon_{k,i}, \hat{w}_{k,i}), \tag{25}$$

where for $k = 0, \ldots, K-1$, $i = 1, \ldots, S$,

$$\hat{w}_{k,i} := L_{\tau,i}^{-1} R_{\tau,k,i}(\tilde{u}_k, \tilde{w}_{k,1}, \ldots, \tilde{w}_{k,i-1}), \tag{26}$$

and $M_{\tau,k,i}(\varepsilon_{k,i}, \hat{w}_{k,i})$ as in Assumption 2.9. We will omit the dependency on $(\varepsilon_{k,i})_{k,i}$, when the tolerances are clear from the context. This leads to the abstract minimization problem

$$\min_{(\varepsilon_{k,i})_{k,i}} \sum_{k=0}^{K-1} \sum_{i=1}^{S} M_{\tau,k,i}(\varepsilon_{k,i}, \hat{w}_{k,i}) \quad \text{subject to} \quad \sum_{k=0}^{K-1} \sum_{i=1}^{S} C''_{\tau,k} C'_{\tau,k,(i)}\, \varepsilon_{k,i} \leq T\tau^{\delta}.$$

We conclude this section with first applications of Theorem 2.17.

Example 2.24. Let us continue the example from the very beginning of this section and consider Eq. (2) in the Gelfand triple $(H_0^1(\mathcal{O}), L_2(\mathcal{O}), H^{-1}(\mathcal{O}))$. We want to interpret the linearly-implicit Euler scheme as an abstract one-stage method with $\mathcal{H} = \mathcal{G} = L_2(\mathcal{O})$. To this end, let

$$\Delta_{\mathcal{O}}^{D} : D(\Delta_{\mathcal{O}}^{D}) \subseteq L_2(\mathcal{O}) \to L_2(\mathcal{O}),$$

denote the *Dirichlet-Laplacian* with domain

$$D(\Delta_{\mathcal{O}}^{D}) := \Big\{ u \in H_0^1(\mathcal{O}) : \Delta u := \sum_{i=1}^{d} \frac{\partial^2}{\partial x_i^2} u \in L_2(\mathcal{O}) \Big\},$$

which is defined as in Appendix A.1, starting with the elliptic, symmetric and bounded bilinear form

$$\begin{aligned} a : H_0^1(\mathcal{O}) \times H_0^1(\mathcal{O}) &\to \mathbb{R} \\ (u, v) &\mapsto a(u,v) := \int_{\mathcal{O}} \langle \nabla u, \nabla v \rangle \, \mathrm{d}x. \end{aligned} \tag{27}$$

Moreover, we pick a smooth initial value $u_0 \in D(\Delta_{\mathcal{O}}^D)$, and consider a continuously differentiable function

$$f : [0,T] \times L_2(\mathcal{O}) \to L_2(\mathcal{O}),$$

which we assume to be Lipschitz continuous in the second variable, uniformly in $t \in [0,T]$. We denote the Lipschitz constant by $C^{\mathrm{Lip},f}$. Since $\Delta_{\mathcal{O}}^D$ generates a strongly continuous contraction semigroup on $L_2(\mathcal{O})$ (cf. Appendix A.1), Eq. (2) has a unique classical solution, see, e.g. [43, Theorems 6.1.5 and 6.1.7]. Thus, there exists a unique continuous function $u : [0,T] \to L_2(\mathcal{O})$, continuously differentiable in $(0,T]$, taking values in $D(\Delta_{\mathcal{O}}^D)$, and fulfilling

$$u(0) = u_0, \quad \text{as well as} \quad u'(t) = \Delta_{\mathcal{O}}^D u(t) + f(t,u(t)), \text{ for } t \in (0,T).$$

In this setting, we can state the exact linearly-implicit Euler scheme (3) in the form of an abstract one-stage scheme as follows: We set $\mathcal{H} := \mathcal{G} := L_2(\mathcal{O})$, and for $\tau = T/K > 0$ we define the operators

$$\begin{aligned} L_{\tau,1}^{-1} : L_2(\mathcal{O}) &\to L_2(\mathcal{O}) \\ v &\mapsto L_{\tau,1}^{-1} v := (I - \tau\Delta_{\mathcal{O}}^D)^{-1} v, \end{aligned}$$

as well as

$$\begin{aligned} R_{\tau,k,1} : L_2(\mathcal{O}) &\to L_2(\mathcal{O}) \\ v &\mapsto R_{\tau,k,1}(v) := v + \tau f(t_k, v), \end{aligned}$$

for $k = 0, \ldots, K-1$. Then the exact linearly-implicit Euler scheme perfectly fits into (6) with $S = 1$.

Under our assumptions on the initial value u_0 and the forcing term f, this scheme converges to the exact solution of Eq. (2) with order $\delta = 1$, i.e., there exists a constant $C_{\mathrm{exact}} > 0$, such that

$$\|u(T) - u_K\|_{L_2(\mathcal{O})} \le C_{\mathrm{exact}} \tau^1,$$

see for instance [14]. Therefore, Assumption 2.13 is satisfied.

Assumption 2.2 can be verified by the following argument: It is well known that for any $\tau > 0$, the operator $L_{\tau,1}^{-1}$ defined above is bounded with norm less than or equal to one (cf. Appendix A.1). Because of the Lipschitz continuity of f, for each $k \in \{0, \ldots, K-1\}$, the composition

$$E_{\tau,k,k+1} := L_{\tau,1}^{-1} R_{\tau,k,1} : L_2(\mathcal{O}) \to L_2(\mathcal{O})$$

is Lipschitz continuous with Lipschitz constant

$$C_{\tau,k,k+1}^{\mathrm{Lip}} \le 1 + \tau C^{\mathrm{Lip},f}.$$

Thus, if we define $E_{\tau,j,k} : L_2(\mathcal{O}) \to L_2(\mathcal{O})$ for $0 \le j \le k \le K$ as in (8), these operators are Lipschitz continuous with Lipschitz constants

$$C^{\mathrm{Lip}}_{\tau,j,k} \le (1 + \tau C^{\mathrm{Lip},f})^{k-j},$$

i.e., Assumption 2.2 is fulfilled. Furthermore, these constants can be estimated uniformly for all j, k and τ, since

$$1 \le C^{\mathrm{Lip}}_{\tau,j,k} \le (1 + \tau C^{\mathrm{Lip},f})^{K} \le \exp(TC^{\mathrm{Lip},f}).$$

Now, let us assume that we have an approximation $\tilde{E}_{\tau,k,k+1}(v)$, $v \in L_2(\mathcal{O})$, such that Assumption 2.4 is fulfilled. We want to use the abstract results from above and present a concrete way to choose the tolerances $(\varepsilon_k)_{k=0}^{K-1}$, so that the output $(\tilde{u}_k)_{k=0}^{K}$ of the inexact linearly-implicit Euler scheme (10) converges to the exact solution with the same order $\delta = 1$.

Therefore, if we choose

$$\varepsilon_k \le \frac{\tau^2}{\exp(TC^{\mathrm{Lip},f})} \qquad \text{for } k = 0, \ldots, K-1,$$

we can conclude from Theorem 2.17 that the inexact linearly-implicit Euler-scheme (10) converges to the exact solution of Eq. (2) with order $\delta = 1$, i.e., it holds the following estimate

$$\|u(T) - \tilde{u}_K\|_{L_2(\mathcal{O})} \le (C_{\mathrm{exact}} + T)\,\tau^1,$$

for all $K \in \mathbb{N}$.

Example 2.25. In the situation from Example 2.24, let us consider a specific form of $f : (0,T] \times L_2(\mathcal{O}) \to L_2(\mathcal{O})$, namely

$$(t, v) \mapsto f(t, v) := \bar{f}(v),$$

where $\bar{f} : \mathbb{R} \to \mathbb{R}$ is continuously differentiable with bounded, strictly negative derivative, i.e., there exists a constant $\bar{B} > 0$, so that

$$-\bar{B} < \frac{\mathrm{d}}{\mathrm{d}x}\bar{f}(x) < 0 \text{ for all } x \in \mathbb{R}.$$

Then, for arbitrary $v_1, v_2 \in L_2(\mathcal{O})$ we get for any $k = 0, \ldots, K-1$,

$$\begin{aligned}
\|L^{-1}_{\tau,1} R_{\tau,k,1}(v_1) - L^{-1}_{\tau,1} R_{\tau,k,1}(v_2)\|_{L_2(\mathcal{O})} \qquad&\\
\le \|R_{\tau,k,1}(v_1) - R_{\tau,k,1}(v_2)\|_{L_2(\mathcal{O})}&\\
= \|v_1 + \tau\bar{f}(v_1) - (v_2 + \tau\bar{f}(v_2))\|_{L_2(\mathcal{O})}&\\
\le \sup_{x\in\mathbb{R}} \left| 1 + \tau\frac{\mathrm{d}}{\mathrm{d}x}\bar{f}(x) \right| \|v_1 - v_2\|_{L_2(\mathcal{O})}.&
\end{aligned}$$

Thus, if $\tau < 2/\bar{B}$, we have a contraction. For $K \in \mathbb{N}$ big enough, and $\varepsilon_k \leq \tau^2$, $k = 0, \ldots, K-1$, we can argue as in Example 2.24 to show that

$$\|u(T) - \tilde{u}_K\|_{L_2(\mathcal{O})} \leq (C_{\text{exact}} + T)\,\tau^1,$$

i.e., the inexact linearly-implicit Euler scheme (10) again converges to the exact solution of Eq. (2) with order $\delta = 1$ (asymptotically), but for much larger values of ε_k, thus, with much less degrees of freedom.

3 Application to deterministic evolution equations

In this section we substantiate our abstract convergence analysis to the case when Rothe's method is applied to deterministic parabolic evolution equations. We focus on S-stage methods for the discretization in time.

We want to compute solutions $u : (0,T] \to V$ to initial value problems of the form

$$u'(t) = F\big(t, u(t)\big), \quad t \in (0,T], \quad u(0) = u_0, \tag{28}$$

where $F : [0,T] \times V \to V^*$ is a nonlinear right-hand side and $u_0 \in V$ is some initial value. Consequently, we consider the Gelfand triple setting (V, U, V^*).

Essentially this section consists of two parts. First of all, we show that a general S-stage scheme for deterministic equations of the form (28) fits nicely into the abstract setting as outlined in Section 2 with $\mathcal{H} = V$ and $\mathcal{G} = V^*$. However, the error estimates for the discretization in time are often formulated in the norm of U, since then a higher order of convergence might be achieved, see, e.g., Theorem 3.8 below. Therefore, in the second part, we analyse the case $\mathcal{H} = \mathcal{G} = U$.

In their most general form, S-stage methods are given by

$$u_{k+1} = u_k + \tau \sum_{i=1}^{S} m_i y_{k,i}, \quad k = 0, 1, \ldots, K-1, \tag{29}$$

with S linear *stage equations*

$$(I - \tau\gamma_{i,i} J) y_{k,i} = F\Big(t_k + a_i\tau, u_k + \tau \sum_{j=1}^{i-1} a_{i,j} y_{k,j}\Big) + \sum_{j=1}^{i-1} c_{i,j} y_{k,j} + \tau\gamma_i g, \tag{30}$$

and

$$a_i := \sum_{j=1}^{i-1} a_{i,j} \sum_{l=1}^{j} \frac{\gamma_{j,l}}{\gamma_{j,j}}, \quad \gamma_i := \sum_{l=1}^{i} \gamma_{i,l}, \tag{31}$$

for $1 \leq i \leq S$. By J and g we denote (approximations of) the partial derivatives $F_u(t_k, u_k)$ and $F_t(t_k, u_k)$, respectively. This particular choice for a_i ensures that J does not enter the right-hand side of (30). The parameters $a_{i,j}$, $c_{i,j}$, $\gamma_{i,j}$ and $m_i \neq 0$ have to be suitably chosen according to the desired properties of the scheme.

Remark 3.1. If $J = F_u(t_k, u_k)$ and $g = F_t(t_k, u_k)$ are the exact derivatives, the corresponding scheme is also known as a method of *Rosenbrock* type. However, this specific choice of J and g is not needed to derive a convergent time discretization scheme. In the larger class of *W–methods*, J and g are allowed to be approximations to the exact Jacobians. Often one chooses $g = 0$. This is done at the price of a significantly lower order of convergence and a substantially more complicated stability analysis.

First of all, we consider the case $\mathcal{H} = V$, $\mathcal{G} = V^*$. The scheme (29) immediately fits into the abstract setting of Section 2 if we interpret the term u_k as the solution to an additional 0th stage equation given by the identity operator I on V. If we define

$$\begin{aligned} L_{\tau,i} : V &\to V^*, \\ v &\mapsto (I - \tau\gamma_{i,i}J)v \end{aligned} \tag{32}$$

and use the right-hand side of the stage equations (30) to define the operators

$$\begin{aligned} &R_{\tau,k,i} : V \times \ldots \times V \to V^*, \\ &(v_0, \ldots, v_{i-1}) \mapsto \tau m_i \Big(F\big(t_k + a_i\tau, v_0 + \sum_{j=1}^{i-1} \frac{a_{i,j}}{m_j} v_j\big) + \sum_{j=1}^{i-1} \frac{c_{i,j}}{\tau m_j} v_j + \tau\gamma_i g \Big), \end{aligned} \tag{33}$$

for $k = 0, \ldots, K-1$ and $i = 1, \ldots, S$, the scheme (29),(30),(31) is related to the abstract Rothe method (6) as follows.

Observation 3.2. *For $k = 0, \ldots, K-1$ and $i = 1, \ldots, S$ let $L_{\tau,i}$ and $R_{\tau,k,i}$ be defined by* (32) *and* (33)*, respectively, and set $L_{\tau,0}^{-1} R_{\tau,k,0} := I_{V\to V}$. Then the linearly-implicit S-stage scheme* (29),(30),(31) *is an abstract $(S+1)$-stage scheme in the sense of* (6) *with $\mathcal{H} = V$, $\mathcal{G} = V^*$. We have*

$$\begin{aligned} u_{k+1} &:= \sum_{i=0}^{S} w_{k,i}, \\ w_{k,i} &:= L_{\tau,i}^{-1} R_{\tau,k,i}(u_k, w_{k,1}, \ldots, w_{k,i-1}), \quad i = 0, \ldots, S, \end{aligned}$$

for $k = 0, \ldots, K-1$.

Remark 3.3. Of course, since the operators $R_{\tau,k,i}$ are derived from the right-hand side F, it might happen that they contain, e.g., nontrivial partial differential operators. Nevertheless, even in this case these differential operators are only *applied* to the current iteration and do not require the numerical solution of an operator equation. Therefore, the operators $R_{\tau,k,i}$ can still be interpreted as evaluation operators.

Before we move to a detailed discussion of the case $\mathcal{H} = \mathcal{G} = U$, let us look at an example, where a simple one-stage scheme of the form (29),(30),(31) with $\mathcal{H} = V$ and $\mathcal{G} = V^*$ is translated into a scheme with $\mathcal{H} = \mathcal{G} = U$.

Example 3.4. Let $\mathcal{O} \subseteq \mathbb{R}^d$ be a bounded Lipschitz domain. Consider the heat equation with zero Dirichlet boundary conditions in the Gelfand triple $(H_0^1(\mathcal{O}), L_2(\mathcal{O}), H^{-1}(\mathcal{O}))$, i.e., Eq. (28) with $F : [0,T] \times H_0^1(\mathcal{O}) \to H^{-1}(\mathcal{O})$, $F(t,u) = \Delta_{\mathcal{O}}^D u + f(t,u)$, and assume that f fulfils the conditions from Example 2.24. The scheme (29),(30),(31) with $S = 1$, $\gamma_{1,1} = m_1 = 1$, $J = \Delta_{\mathcal{O}}^D : H_0^1(\mathcal{O}) \to H^{-1}(\mathcal{O})$, and $g = 0$ leads to

$$u_{k+1} = u_k + \tau\big(I - \tau\Delta_{\mathcal{O}}^D\big)^{-1}\big(\Delta_{\mathcal{O}}^D u_k + f(t_k,u_k)\big), \quad k = 0,\dots,K-1.$$

This scheme perfectly fits into the setting of Section 2 and can be rewritten as a 2-stage scheme of the form (6) with $\mathcal{H} = V$ and $\mathcal{G} = V^*$, cf. Observation 3.2 above. However, since the Dirichlet-Laplacian is not bounded on $L_2(\mathcal{O})$, it can not be understood immediately as an S-stage scheme of the form (6) with $\mathcal{H} = \mathcal{G} = L_2(\mathcal{O})$. But, a short computation shows that it can be rewritten as

$$u_{k+1} = \big(I - \tau\Delta_{\mathcal{O}}^D\big)^{-1}\big(u_k + \tau f(t_k,u_k)\big), \quad k = 0,\dots,K-1.$$

Thus, if we start with $u_0 \in D(\Delta_{\mathcal{O}}^D)$, and consider the Dirichlet Laplacian as an unbounded operator on $L_2(\mathcal{O})$, the last scheme can be interpreted as an abstract one-stage scheme of the form (6) with $\mathcal{H} = \mathcal{G} = U$. It is just the linearly-implicit Euler scheme for the heat equation we have already discussed in Example 2.24. As we have seen there, it converges with rate $\delta = 1$. It is worth noting that this result stays true for a wider class of operators A instead of $\Delta_{\mathcal{O}}^D$, see [14] for details.

The next step is to discuss the case $\mathcal{H} = \mathcal{G} = U$ in detail. In order to avoid technical difficulties, we restrict the discussion to the case of semi-linear problems with a right-hand side of the form

$$F : [0,T] \times V \to V^*, \quad F(t,u) := A(t)u + f(t,u), \tag{34}$$

where $A(t)$ is given for all $t \in (0,T)$ in the sense of Appendix A.1. Furthermore, we will focus on W-methods with the specific choice

$$J(t_k) := A(t_k), \quad g := 0, \tag{35}$$

in (30). We restrict our analysis to these methods for the following reasons. First, the linearly-implicit Euler scheme, which is the most important example and the method of choice for stochastic evolution equations, is a W-method and not a Rosenbrock method. Second, the choice of J in (35) avoids the evaluation of the Jacobian in each time step, which might be numerically costly.

In our setting, the overall convergence rate that can be expected is limited by the convergence rate of the exact scheme, cf. Theorem 2.22 and Assumption 2.13. Therefore, to obtain a reasonable result, it is clearly necessary to discuss the approximation properties of the exact S-stage scheme.

In general the rates achievable by W-methods with $S = 1$ are limited. To the best of our knowledge, the most far reaching result concerning W-methods with $S > 1$ have been derived by Lubich and Ostermann [42]. For the readers convenience, we discuss their results as far as it is needed for our purposes. To do so, we need to introduce the following definitions and assumptions.

Definition 3.5. The method (29),(30),(31) is called *$A(\theta)$-stable* if the related stability function

$$R(z) := 1 + z\,\mathbf{m}^\top\Big(\mathbf{I} - (c_{i,j})_{i,j=1}^S - z\big(\operatorname{diag}(\gamma_{i,i}) + (a_{i,j})_{i,j=1}^S\big)\Big)^{-1}\mathbf{1},$$

where $\mathbf{1}^\top := (1,\ldots,1)^\top$ and $\mathbf{m}^\top := (m_1,\ldots,m_S)^\top$, fulfills

$$|R(z)| \leq 1 \quad \text{for all } z \in \mathbb{C} \text{ with } |\arg(z)| \geq \pi - \theta.$$

If, additionally, the limit $|R(\infty)| := \lim_{|z|\to\infty}|R(z)| < 1$, the method is called *strongly $A(\theta)$-stable.*

Definition 3.6. We say that the scheme (29),(30),(31) is of *order* $p \in \mathbb{N}$, if the error of the method, when applied to ordinary differential equations defined on open subsets of $\mathbb{R}^d$ with sufficiently smooth right-hand sides, satisfies $\|u(t_k) - u_k\|_{\mathbb{R}^d} \leq C_{\mathrm{ord}}\tau^p$, uniformly on bounded time intervals.

Moreover, the results in [42] rely on the following assumption.

Assumption 3.7. Let $C_{\mathrm{offset}} \geq 0$ and denote $\hat{J}(t) := A(t) + C_{\mathrm{offset}}I$.

(i) For both instances $G(t) := F_u(t, u(t))$ and $G(t) := \hat{J}(t)$ it holds that $G(t) : V \to V^*$, $t \in [0, T]$, is a uniformly bounded family of linear operators in $\mathcal{L}(V, V^*)$. Each $G(t)$ is boundedly invertible and the family $G(t)^{-1}$, $t \in [0, T]$, is uniformly bounded in $\mathcal{L}(V^*, V)$.

(ii) There exist constants $\phi < \pi/2$, C_i^{sect}, $i = 1, 2$ such that for all $t \in [0, T]$ and $z \in \mathbb{C}$ with $|\arg(z)| \leq \pi - \phi$ the operators $zI - F_u(t, u(t))$ and $zI - \hat{J}(t)$ are invertible, and their resolvents are bounded on V, i.e.,

$$\|\big(zI - F_u(t, u(t))\big)^{-1}\|_{\mathcal{L}(V,V)} \leq \frac{C_1^{\mathrm{sect}}}{|z|}, \quad \|\big(zI - \hat{J}(t)\big)^{-1}\|_{\mathcal{L}(V,V)} \leq \frac{C_2^{\mathrm{sect}}}{|z|}.$$

(iii) The mapping $t \mapsto F_u(t, u(t)) \in \mathcal{L}(V, V^*)$ is sufficiently often differentiable on $[0, T]$ and fulfills the Lipschitz condition

$$\|F_u(t, u(t)) - F_u(t', u(t'))\|_{\mathcal{L}(V,V^*)} \leq C_u^F|t - t'| \quad \text{for } 0 \leq t \leq t' \leq T.$$

(iv) The following bounds hold uniformly for v varying in bounded subsets of V and $0 \leq t \leq T$:

$$\big\|F_{tu}(t, v)w\big\|_{V^*} \leq C_{tu}^F\|w\|_V, \quad \big\|F_{uu}(t, v)[w_1, w_2]\big\|_{V^*} \leq C_{uu}^F\|w_1\|_V\|w_2\|_V.$$

(v) There exists a splitting

$$f_u(t, u(t)) =: S_k^{(l)} + S_k^{(r)} \tag{36}$$

and constants $\mu < 1$, $\beta \geq \mu$ (positive), $C_k^{(l)}$ (sufficiently small) as well as $C_{k,\mu}^{(r)}$, $C_{k,\beta}$, $C_k^{(l)}$, and $C_{k,\beta}^{(r)}$, such that

$$\begin{aligned}
\|S_k^{(l)}\|_{\mathcal{L}(V,V^*)} &\leq C_k^{(l)}, \\
\|S_k^{(r)}\hat{J}^{-\mu}(t_k)\|_{\mathcal{L}(V^*,V^*)} &\leq C_{k,\mu}^{(r)}, \\
\|\hat{J}^{\beta}(t_k)\big(F_u(t_k,u(t_k))\big)^{-\beta}\|_{\mathcal{L}(V,V)} &\leq C_{k,\beta}, \\
\|\hat{J}^{\beta}(t_k)S_k^{(l)}\hat{J}^{-\beta}(t_k)\|_{\mathcal{L}(V,V^*)} &\leq C_k^{(l)}, \\
\|S_k^{(r)}\hat{J}^{-\beta}(t_k)\|_{\mathcal{L}(V^*,V^*)} &\leq C_{k,\beta}^{(r)}.
\end{aligned}$$

Now, with these formulations at hand, the main result in [42], Theorem 6.2, reads as follows.

Theorem 3.8. *Suppose that the solution u of Eq.* (28) *is unique and that its temporal derivatives are sufficiently regular. Let Assumption 3.7 hold. Suppose that the scheme* (29),(30),(31) *is a W-method of order $p \geq 2$ that is strongly $A(\theta)$-stable with $\theta > \phi$ with $\phi < \pi/2$ as in Assumption 3.7(ii). Let $\beta \in [0,1]$ be as in Assumption 3.7(v) such that $D(A(t)^{\beta})$ is independent of t (with uniformly equivalent norms), $A^{\beta}u' \in L_2(0,T;V)$. Then the error provided by the numerical solution u_k, $k = 0,\dots,K$ is bounded in $\tau \leq \tau_0$ by*

$$\begin{aligned}
&\Big(\tau\sum_{k=0}^{K}\|u_k - u(t_k)\|_V^2\Big)^{1/2} + \max_{0\leq k\leq K}\|u_k - u(t_k)\|_U \\
&\quad\leq C_1^{\mathrm{conv}}\tau^{1+\beta}\Big(C_2^{\mathrm{conv}} + C_1^{\mathrm{conv}}C_k^{(l)}\Big)C_k^{(l)}\Big(\int_0^T\|A^{\beta}u'(t)\|_V^2\,\mathrm{d}t\Big)^{1/2} \\
&\quad+ C_1^{\mathrm{conv}}\tau^2\Big(\int_0^T\|A^{\beta}u'(t)\|_V^2\,\mathrm{d}t + \int_0^T\|u''(t)\|_V^2\,\mathrm{d}t + \int_0^T\|u'''(t)\|_{V^*}^2\,\mathrm{d}t\Big)^{1/2}.
\end{aligned} \tag{37}$$

The constants C_1^{conv}, C_2^{conv}, and τ_0 depend on the concrete choice of the W-method, the constants in the Assumptions, and on T. The maximal time step size τ_0 depends in addition on the size of the integral terms in (37).

Remark 3.9. As in Theorem 3.8, throughout this section, we assume that a unique solution to Eq. (28) exists, i.e., Assumption 2.11 holds. This is the starting point for our convergence analysis of inexact S-stage schemes. Thus, we will not discuss the solvability and uniqueness theory for PDEs in detail. However, since in the forthcoming Examples, we will use the results from [42], let us briefly recall which solution concept is used there in the following standard situation: Consider a linear operator $A : V \to V^*$ fulfilling the conditions from Appendix A.1, and assume that F in (28) has the form $F(t,u) := Au + f(t)$. Then, a weak formulation of Eq. (28) is: find

$$u \in C([0,T];U) \cap L_2(0,T;V),$$

such that

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle u(t), v\rangle_U = \langle Au(t), v\rangle_{V^*\times V} + \langle f(t), v\rangle_U \qquad \text{for all } v\in V, t\in(0,T].$$

Before we continue our analysis, let us present a well-known W-method which fulfils the assumptions of Theorem 3.8.

Example 3.10. As a W-method with $S = 2$ we present the following scheme taken from [49]. There it has been shown that this method is of order $p = 2$ and strongly $A(\theta)$ stable with $\theta = \pi/2$. It is sometimes called *ROS2* in the literature and is given by

$$u_{k+1} = u_k + \frac{3}{2}\tau y_{k,1} + \frac{1}{2}\tau y_{k,2},$$

where

$$\begin{aligned}
y_{k,1} &= \big(I - \tau\frac{1}{2+\sqrt{2}}A(t_k)\big)^{-1}\big(A(t_k)u_k + f(t_k,u_k)\big),\\
y_{k,2} &= \big(I - \tau\frac{1}{2+\sqrt{2}}A(t_k)\big)^{-1}\Big(A(t_k+\tau)(u_k+\tau y_{k,1})\\
&\qquad\qquad\qquad\qquad + f(t_k+\tau, u_k+\tau y_{k,1}) - 2y_{k,1}\Big).
\end{aligned}$$

It fits into the setting of (29),(30),(31) with $m_1 = 3/2$, $m_2 = 1/2$, $\gamma_{1,1} = \gamma_{2,2} = (2+\sqrt{2})^{-1}$, $a_{2,1} = 1$ and $c_{2,1} = -2$.

To avoid technical difficulties we will restrict the setting of (34) for the remainder of this section to the special case

$$F : [0,T]\times V \to V^*, \quad F(t,u) := Au + f(t), \tag{38}$$

where $A : V \to V^*$ is given in the sense of Appendix A.1, and $f : [0,T] \to U$ is a continuously differentiable function. In this case, as already mentioned in Example 2.24, Eq. (28) has a unique classical solution, provided $u_0 \in D(A;U)$, see e.g. [43, Corollary 2.5]. It is worth noting that this unique solution is a also a weak solution in the sense of [42], see also Remark 3.9.

Using the abstract results from Section 2, we will analyse the inexact S-stage method corresponding to the W-method with

$$J := A \quad \text{and} \quad g := 0. \tag{39}$$

In the sequel, we will restrict the discussion to the case $S = 2$. This is not a major restriction for the following reason: In Theorem 3.8, the maximal convergence order of W-methods is bounded by $\delta = 1+\beta$, where $\beta\in[0,1]$. In Example 3.13 we will show that an F of the form (38) fulfils Assumption 3.7 with $\beta = 1$. If we additionally impose the asserted regularity assumptions with $\beta = 1$, see (45) below, then we can apply Theorem 3.8 with $\beta = 1$ to the $ROS2$-method from Example 3.10 above (which is a 2-stage method), and get the optimal order in this context.

The structure (38) of the right hand side F in (28), allows the following reformulation of the W-method with (J, g) as in (39) (a proof can be found in Appendix A.2).

Lemma 3.11. *Consider the S-stage W-method given by* (29),(30),(31) *with $S = 2$ and F and (J, g) as in* (38) *and* (39)*, respectively. Then, if $\gamma_{i,i} \neq 0$, for $i = 1, 2$, we have*

$$u_{k+1} = \Big(1 - \frac{m_1}{\gamma_{1,1}} - \frac{m_2}{\gamma_{2,2}}(1 - \frac{a_{2,1}}{\gamma_{1,1}})\Big)u_k + \Big(\tau m_1 - \tau m_2 \frac{a_{2,1}}{\gamma_{2,2}}\Big)v_{k,1} + \tau m_2 v_{k,2},$$

where

$$v_{k,1} = L_{\tau,1}^{-1}\Big(\frac{1}{\tau\gamma_{1,1}}u_k + f(t_k)\Big),$$
$$v_{k,2} = L_{\tau,2}^{-1}\Big(\big(\frac{1}{\tau\gamma_{2,2}}(1 - \frac{a_{2,1}}{\gamma_{1,1}}) - \frac{c_{2,1}}{\tau\gamma_{1,1}}\big)u_k + \big(\frac{a_{2,1}}{\gamma_{2,2}} + c_{2,1}\big)v_{k,1} + f(t_k + a_2\tau)\Big).$$

Observation 3.12. *Note that, if $\gamma_{i,i} \neq 0$, for $i = 1, 2$, and $m_1\gamma_{2,2} \neq m_2 a_{2,1}$, the scheme under consideration perfectly fits into the setting of Section 2 with $\mathcal{H} = \mathcal{G} = U$. It can be written in the form of the abstract Rothe method* (6)*. More precisely,*

$$\left.\begin{aligned} u_{k+1} &= \sum_{i=0}^{2} w_{k,i}, \\ w_{k,i} &:= L_{\tau,i}^{-1} R_{\tau,k,i}(u_k, w_{k,1}, \ldots, w_{k,i-1}), \quad i = 0, 1, 2, \end{aligned}\right\} \tag{40}$$

with

$$\begin{aligned} L_{\tau,i}^{-1} &: U \to U, \\ &v \mapsto \big(I - \tau\gamma_{i,i}A\big)^{-1}v, \quad \text{for } i = 1, 2, \end{aligned} \tag{41}$$

the evaluation operators

$$\begin{aligned} R_{\tau,k,1} &: U \to U, \\ &v \mapsto \Big(\frac{m_1}{\gamma_{1,1}} - \frac{m_2 a_{2,1}}{\gamma_{2,2}\gamma_{1,1}}\Big)v + \tau\Big(m_1 - m_2\frac{a_{2,1}}{\gamma_{2,2}}\Big)f(t_k), \end{aligned} \tag{42}$$

and

$$\begin{aligned} R_{\tau,k,2} &: U \times U \to U, \\ (v_0, v_1) &\mapsto \Big(\frac{m_2}{\gamma_{2,2}}(1 - \frac{a_{2,1}}{\gamma_{1,1}}) - \frac{c_{2,1}m_2}{\gamma_{1,1}}\Big)v_0 \\ &\quad + \frac{m_2 a_{2,1} + m_2\gamma_{2,2}c_{2,1}}{m_1\gamma_{2,2} - m_2 a_{2,1}}v_1 + \tau m_2 f(t_k + a_2\tau), \end{aligned} \tag{43}$$

as well as a 0th step given by

$$L_{\tau,0}^{-1}R_{\tau,k,0} : U \to U, \qquad v \mapsto \Big(1 - \frac{m_1}{\gamma_{1,1}} - \frac{m_2}{\gamma_{2,2}}\big(1 - \frac{a_{2,1}}{\gamma_{1,1}}\big)\Big)v. \tag{44}$$

This is an immediate consequence of Lemma 3.11.

An easy computation, together with the fact that $L_{\tau,1}^{-1}$ and $L_{\tau,2}^{-1}$ are contractions on U (cf. Appendix A.1), yield the Lipschitz constant

$$C_{\tau,k,(1)}^{\mathrm{Lip}} = \left|\frac{m_1}{\gamma_{1,1}} - \frac{m_2 a_{2,1}}{\gamma_{2,2}\gamma_{1,1}}\right|,$$

of $L_{\tau,1}^{-1}R_{\tau,k,1}$. Simultaneously, the Lipschitz constant of $L_{\tau,2}^{-1}R_{\tau,k,2}$ can be estimated as follows:

$$C_{\tau,k,(2)}^{\mathrm{Lip}} \leq \max\left\{\left|\frac{m_2}{\gamma_{2,2}}\big(1 - \frac{a_{2,1}}{\gamma_{1,1}}\big) - \frac{m_2 c_{2,1}}{\gamma_{1,1}}\right|, \left|\frac{m_2 a_{2,1} + m_2\gamma_{2,2}c_{2,1}}{m_1\gamma_{2,2} - m_2 a_{2,1}}\right|\right\}.$$

Note that both constants are independent of k and τ.

Example 3.13. As a first step towards the case of inexact operator evaluations we need to check the applicability of Theorem 3.8 in the current setting (38), (39). Therefore, we now check Assumption 3.7. We begin by choosing $C_{\text{offset}} = 0$. As a consequence it holds that $\hat{J} = F_u(t, u(t)) = A$, independently of t. Assumption 3.7(i) holds by the assumptions on A, i.e., Appendix A.1. This, together with the ellipticity assumption (97) already implies Assumption 3.7(ii), see [39]. Further, $A = F_u(t, v)$ is independent of (t, v), and as a consequence Assumption 3.7(iii) and (iv) hold with $C_u^F = C_{tu}^F = C_{uu}^F = 0$. Finally, since J is the exact Jacobian, it is possible to choose $S_k^{(l)} = S_k^{(r)} = 0$ in (36), such that Assumption 3.7(v) holds with $C_k^{(l)} = C_{k,\mu}^{(r)} = C_{k,\beta}^{(r)} = 0$, $C_{k,\beta} = 1$ and $\beta = 1$. Concerning the W-method (29),(30),(31) we assume it to be of order $p \geq 2$ and strongly $A(\theta)$-stable with $\theta > \phi$, where ϕ is as in Assumption 3.7(ii). E.g., the scheme from Example 3.10 could be employed. If for the solution of Eq. (28) with F as in (38) the regularity assumptions

$$Au', u'' \in L_2(0,T;V), \; u''' \in L_2(0,T;V^*) \tag{45}$$

hold, then we can apply Theorem 3.8. Using $C_k^{(l)} = 0$ and $\beta = 1$, the convergence result (37) reads as

$$\Big(\tau\sum_{k=0}^{K}\|u_k - u(t_k)\|_V^2\Big)^{1/2} + \max_{0\leq k\leq K}\|u_k - u(t_k)\|_U$$
$$\leq C_1^{\mathrm{conv}}\tau^2\Big(\int_0^T \|Au'(t)\|_V^2\,\mathrm{d}t + \int_0^T \|u''(t)\|_V^2\,\mathrm{d}t + \int_0^T \|u'''(t)\|_{V^*}^2\,\mathrm{d}t\Big)^{1/2}.$$

That means, the error measured in the norm $\|\cdot\|_U$ is of order $\delta = 2$.

Example 3.14. We want to apply our general convergence results for the case of inexact solution of the stage equations, i.e., Theorem 2.22. Here we consider the special case when the method $ROS2$ from Example 3.10 is employed. We begin by presenting the method in its reformulation on $\mathcal{H} = \mathcal{G} = U$, as given in Observation 3.12. Inserting the coefficients

$$m_1 = \frac{3}{2},\ m_2 = \frac{1}{2},\ \gamma_{1,1} = \gamma_{2,2} = (2+\sqrt{2})^{-1},\ a_{2,1} = 1,\ \text{and}\ c_{2,1} = -2$$

into (40),(41), (42), (43), and (44) yields

$$\begin{aligned} u_{k+1} &= \sum_{i=0}^{2} w_{k,i}, \\ w_{k,i} &:= L_{\tau,i}^{-1} R_{\tau,k,i}(u_k, w_{k,1}, \ldots, w_{k,i-1}), \quad i = 0,1,2, \end{aligned}$$

where the 0th stage vanishes, i.e., $L_{\tau,0}^{-1} R_{\tau,k,0} \equiv 0$,

$$\begin{aligned} L_{\tau,1}^{-1} = L_{\tau,2}^{-1} &: U \to U \\ & v \mapsto \Big(I - \tau \frac{1}{2+\sqrt{2}} A\Big)^{-1} v, \end{aligned}$$

and the evaluation operators are given by

$$\begin{aligned} R_{\tau,k,1} &: U \to U, \\ & v \mapsto -\frac{\sqrt{2}}{2} v + \tau \frac{1-\sqrt{2}}{2} f(t_k), \end{aligned}$$

and

$$\begin{aligned} R_{\tau,k,2} &: U \times U \to U, \\ & (v_0, v_1) \mapsto -\frac{\sqrt{2}}{2} v_0 + \frac{\sqrt{2}}{1-\sqrt{2}} v_1 + \tau \frac{1}{2} f(t_k + \tau). \end{aligned}$$

This scheme perfectly fits into the abstract Rothe method (6) with $S = 2$. As in Observation 3.12, by a simple calculation we get the following estimates for the Lipschitz constants of $L_{\tau,i}^{-1} R_{\tau,k,i}$, $i = 1,2$:

$$C_{\tau,k,(1)}^{\mathrm{Lip}} = \frac{\sqrt{2}}{2}, \quad \text{and} \quad C_{\tau,k,(2)}^{\mathrm{Lip}} \leq \max\Big\{\frac{\sqrt{2}}{2}, \frac{-\sqrt{2}}{1-\sqrt{2}}\Big\} \leq \frac{\sqrt{2}}{2}.$$

As in Example 3.13, we assume that the exact solution u satisfies (45). Furthermore, we assume we have a method at hand, such that Assumption 2.9 is satisfied. Then, by Theorem 3.8 and Theorem 2.22, if we choose the tolerances $\varepsilon_{k,i}$, for $k = 0, \ldots, K-1$ and $i = 1,2$, so that they satisfy

$$\varepsilon_{k,i} \leq \frac{1}{2} \tau^3 \Big(\frac{1}{2} + \sqrt{2}\Big)^{K-k-1} \prod_{l=i+1}^{2} \Big(1 + \frac{\sqrt{2}}{2}\Big),$$

the corresponding inexact 2-stage scheme (13) converges with order $\delta = 2$. The computational cost can be estimated by

$$\sum_{k=0}^{K-1} M_{\tau,k,1}(\varepsilon_{k,1}, \hat{w}_{k,1}) + M_{\tau,k,2}(\varepsilon_{k,2}, \hat{w}_{k,2})$$

with $M_{\tau,k,i}(\cdot,\cdot)$ as in Assumption 2.9 and $\hat{w}_{k,i}$ as in Remark 2.23(iii).

Remark 3.15. For methods of Rosenbrock type, i.e., under the assumption that we use exact Jacobians J and g, a result similar to Theorem 3.8 holds. In [42, Theorem 5.2] it is shown that for methods of order $p \geq 3$ and under certain additional regularity assumptions on the exact solution u of Eq. (28) the error can be bounded similar to (37) with rate $\tau^{2+\beta}$, $\beta \in [0, 1]$.

4 Application to stochastic evolution equations

In this section we apply Rothe's method to a class of semi-linear parabolic SPDEs. We use the stochastic analogue of the linearly-implicit Euler scheme (3) in time, and the spatial discretization in every time step is a (possibly nonlinear) blackbox-solver $[\,\cdot\,]_\varepsilon$, e.g., an adaptive wavelet solver as described in Subsection 5.3. As before, we interpret parabolic SPDEs as ordinary SDEs in a suitable function space U. We consider a separable real Hilbert space U and the U-valued SDE

$$\mathrm{d}u(t) = \big(Au(t) + f(u(t))\big)\,\mathrm{d}t + B(u(t))\,\mathrm{d}W(t), \quad u(0) = u_0, \tag{46}$$

on the time interval $[0, T]$. Here $u = (u(t))_{t\in[0,T]}$ is a U-valued stochastic process, $A : D(A) \subset U \to U$ is a densely defined, strictly negative definite, self-adjoint, linear operator such that zero belongs to the resolvent set and the inverse A^{-1} is compact on U. The forcing term $f : D((-A)^\varrho) \to D((-A)^{\varrho-\sigma})$ and $B : D((-A)^\varrho) \to \mathcal{L}(\ell_2, D((-A)^{\varrho-\beta}))$ are Lipschitz continuous maps for suitable constants ϱ, σ and β; and finally, $W = (W(t))_{t\in[0,T]}$ is a cylindrical Wiener process on the sequence space $\ell_2 = \ell_2(\mathbb{N})$. Details are given in Subsection 4.1. Our setting is based on the one considered in Printems [45] where the convergence of semi-discretizations in time is investigated. This is why, in contrast to the previous sections, the forcing term f does not depend on the time variable t. Compared with [45] we allow the spatial regularity of the whole setting to be 'shifted' in terms of the additional parameter ϱ. In concrete applications to parabolic SPDEs, this will lead to estimates of the discretization error in terms of the numerically important energy norm, cf. Example 4.10, provided that the initial condition u_0 and the forcing terms f and B are sufficiently regular.

4.1 Setting and assumptions

Let us describe the setting systematically and in detail.

Assumption 4.1. The operator $A : D(A) \subset U \to U$ is linear, densely defined, strictly negative definite and self-adjoint. Zero belongs to the resolvent set of A and the inverse $A^{-1} : U \to U$ is compact. There exists an $\alpha > 0$ such that $(-A)^{-\alpha}$ is a trace class operator on U.

To simplify notation, the separable real Hilbert space U is always assumed to be infinite-dimensional. It follows that A enjoys a spectral decomposition of the form

$$Av = \sum_{j\in\mathbb{N}} \lambda_j \langle v, e_j\rangle_U e_j, \qquad v \in D(A),$$

where $(e_j)_{j\in\mathbb{N}}$ is an orthonormal basis of U consisting of eigenvectors of A with strictly negative eigenvalues $(\lambda_j)_{j\in\mathbb{N}}$ such that

$$0 > \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_j \to -\infty, \qquad j \to \infty.$$

For $s \geq 0$ we set

$$D((-A)^s) := \Big\{ v \in U \,:\, \sum_{j=1}^{\infty} \big|(-\lambda_j)^s \langle v, e_j\rangle_U\big|^2 < \infty \Big\}, \tag{47}$$

$$(-A)^s v := \sum_{j\in\mathbb{N}} (-\lambda_j)^s \langle v, e_j\rangle_U e_j, \qquad v \in D((-A)^s), \tag{48}$$

so that $D((-A)^s)$, endowed with the norm $\|\cdot\|_{D((-A)^s)} = \|(-A)^s \cdot\|_U$, is a Hilbert space; by construction this norm is equivalent to the graph norm of $(-A)^s$. For $s < 0$ we define $D((-A)^s)$ as the completion of U with respect to the norm $\|\cdot\|_{D((-A)^s)}$, defined on U by $\|v\|^2_{D((-A)^s)} = \sum_{j\in\mathbb{N}} \big|(-\lambda_j)^s \langle v, e_j\rangle_U\big|^2$. Thus, $D((-A)^s)$ can be considered as a space of formal sums

$$v = \sum_{j\in\mathbb{N}} v^{(j)} e_j \quad \text{such that} \quad \sum_{j\in\mathbb{N}} \big|(-\lambda_j)^s v^{(j)}\big|^2 < \infty$$

with coefficients $v^{(j)} \in \mathbb{R}$. Generalizing (48) in the obvious way, we obtain operators $(-A)^s$, $s \in \mathbb{R}$, which map $D((-A)^r)$ isometrically onto $D((-A)^{r-s})$ for all $r \in \mathbb{R}$.

The trace class condition in Assumption 4.1 can now be reformulated as the requirement that there exists an $\alpha > 0$ such that

$$\operatorname{Tr}(-A)^{-\alpha} = \sum_{j\in\mathbb{N}} (-\lambda_j)^{-\alpha} < \infty. \tag{49}$$

Before we state our conditions on the forcing terms f and B, let us consider an example for an operator fulfilling Assumption 4.1.

Example 4.2. Let $\mathcal{O}$ be a bounded open subset of $\mathbb{R}^d$, set $U := L_2(\mathcal{O})$ and let $A := \Delta_{\mathcal{O}}^D$ be the Dirichlet-Laplacian on $\mathcal{O}$ from Example 2.24, i.e.,

$$\Delta_{\mathcal{O}}^D : D(\Delta_{\mathcal{O}}^D) \subseteq L_2(\mathcal{O}) \to L_2(\mathcal{O})$$

with domain

$$D(\Delta_{\mathcal{O}}^D) = \Big\{ u \in H_0^1(\mathcal{O}) : \Delta u := \sum_{i=1}^d \frac{\partial^2}{\partial x_i^2} u \in L_2(\mathcal{O}) \Big\}.$$

Note that this definition of the domain of the Dirichlet-Laplacian is consistent with the definition of $D((-\Delta_{\mathcal{O}}^D)^s)$ for $s = 1$ in (47), see e.g. [41, Remark 1.13] for details. This linear operator fulfils Assumption 4.1 for all $\alpha > d/2$: It is densely defined, self-adjoint, and strictly negative definite, since it has been introduced in complete analogy to the variational operator $\tilde{A}$ from Appendix A.1, starting with the symmetric, elliptic and bounded bilinear form (27). Furthermore, due to the Rellich-Kondrachev theorem (see, e.g. [1, Chapter VI]), it possesses a compact inverse $(\Delta_{\mathcal{O}}^D)^{-1} : L_2(\mathcal{O}) \to L_2(\mathcal{O})$. Moreover, Weyl's law states that

$$-\lambda_j \asymp j^{2/d}, \qquad j \in \mathbb{N},$$

see [6], implying that (49) holds for all $\alpha > d/2$.

Concerning the forcing terms f and B we assume the following.

Assumption 4.3. For certain smoothness parameters

$$\varrho \geq 0, \qquad \sigma < 1 \quad \text{and} \quad \beta < \frac{1-\alpha}{2} \tag{50}$$

(α as in Assumption 4.1) f and B map $D((-A)^\varrho)$ to $D((-A)^{\varrho-\sigma})$ and $D((-A)^\varrho)$ to $\mathcal{L}(\ell_2; D((-A)^{\varrho-\beta}))$, respectively. Furthermore, they are globally Lipschitz continuous, that is, there exist positive constants C_f^{Lip} and C_B^{Lip} such that for all $v, w \in D((-A)^\varrho)$,

$$\|f(v) - f(w)\|_{D((-A)^{\varrho-\sigma})} \leq C_f^{\mathrm{Lip}} \|v - w\|_{D((-A)^\varrho)},$$

and

$$\|B(v) - B(w)\|_{\mathcal{L}(\ell_2; D((-A)^{\varrho-\beta}))} \leq C_B^{\mathrm{Lip}} \|v - w\|_{D((-A)^\varrho)}.$$

Remark 4.4. (i) The parameters σ and β in Assumption 4.3 are allowed to be negative.

(ii) Assumption 4.3 follows the lines of [45] ('shifted' by $\varrho \geq 0$). The linear growth conditions (3.5) and (3.7) in [45] follow from the (global) Lipschitz continuity of the mappings f and B.

Finally, we describe the noise and the initial condition in (46). For the notion of a normal filtration we refer to [44].

Assumption 4.5. The noise $W = (W(t))_{t\in[0,T]}$ is a cylindrical Wiener process on ℓ_2 with respect to a normal filtration $(\mathscr{F}_t)_{t\in[0,T]}$. The underlying probability space $(\Omega, \mathscr{F}, \mathbb{P})$ is complete. For ϱ as in Assumption 4.3, the initial condition u_0 in Eq. (46) belongs to the space $L_2(\Omega, \mathscr{F}_0, \mathbb{P}; D((-A)^\varrho)))$.

Let $(e^{tA})_{t\geq 0}$ be the strongly continuous semigroup of contractions on U which is generated by A. We call a *mild solution* to Eq. (46) a predictable process $u : \Omega \times [0, T] \to D((-A)^\varrho)$ with

$$\sup_{t\in[0,T]} \mathbb{E}\|u(t)\|^2_{D((-A)^\varrho)} < \infty \tag{51}$$

such that for every $t \in [0, T]$ the equality

$$u(t) = e^{tA}u_0 + \int_0^t e^{(t-s)A} f(u(s))\,\mathrm{d}s + \int_0^t e^{(t-s)A} B(u(s))\,\mathrm{d}W(s) \tag{52}$$

holds $\mathbb{P}$-almost surely in $D((-A)^\varrho)$. The first integral in (52) is meant to be a $D((-A)^\varrho)$-valued Bochner integral for $\mathbb{P}$-almost every $\omega \in \Omega$; the second integral is a $D((-A)^\varrho)$-valued stochastic integral as defined, e.g., in [15, 44].

Remark 4.6. (i) Both integrals in (52) exist due to (51) and Assumptions 4.1, 4.3. For example, considering the stochastic integral in (52), we know that it exists as an element of $L_2(\Omega, \mathscr{F}_t, \mathbb{P}; D((-A)^\varrho))$ if the integral

$$\int_0^t \mathbb{E}\big\|e^{(t-s)A} B(u(s))\big\|^2_{\mathcal{L}_{\mathrm{HS}}(\ell_2;D((-A)^\varrho))}\,\mathrm{d}s$$

is finite, where $\mathcal{L}_{\mathrm{HS}}(\ell_2; D((-A)^\varrho))$ denotes the space of Hilbert-Schmidt operators from ℓ_2 to $D((-A)^\varrho)$. The integrand of the last integral can be estimated from above by

$$\mathrm{Tr}(-A)^{-\alpha}\big\|(-A)^{\beta+\alpha/2}e^{(t-s)A}\big\|^2_{\mathcal{L}(D((-A)^\varrho))}\mathbb{E}\big\|(-A)^{-\beta}B(u(s))\big\|^2_{\mathcal{L}(\ell_2;D((-A)^\varrho))},$$

and we have

$$\big\|(-A)^{\beta+\alpha/2}e^{(t-s)A}\big\|^2_{\mathcal{L}(D((-A)^\varrho))} \leq C(t-s)^{-(2\beta+\alpha)}$$

with $2\beta + \alpha < 1$. Moreover,

$$\mathbb{E}\big\|(-A)^{-\beta}B(u(s))\big\|^2_{\mathcal{L}(\ell_2;D((-A)^\varrho))} \leq C\Big(1 + \sup_{r\in[0,T]} \mathbb{E}\|u(r)\|^2_{D((-A)^\varrho)}\Big).$$

The last estimate follows from the global Lipschitz property of the mapping $B : D((-A)^\varrho) \to \mathcal{L}(\ell_2; D((-A)^{\varrho-\beta}))$.

(ii) For the case $\varrho = 0$ existence and uniqueness of a mild solution to Eq. (46) has been stated in [45, Proposition 3.1]. The proof consists of a modification of the proof of Theorem 7.4 in [15] — a contraction argument in $L_\infty([0, T]; L_2(\Omega; U))$. For the general case $\varrho \geq 0$ existence and uniqueness can be proved analogously, see [36, Theorem 5.1]. Alternatively, the case $\varrho > 0$ can be traced back to the case $\varrho = 0$ as described in the proof of Proposition 4.7 below.

Proposition 4.7. *Let Assumptions 4.1, 4.3 and 4.5 be fulfilled. Then, Eq.* (46) *has a unique mild solution, i.e., there exists a unique (up to modifications) predictable stochastic process* $u : \Omega \times [0,T] \to D((-A)^{\varrho})$ *with* $\sup_{t\in[0,T]} \mathbb{E}\|u(t)\|^2_{D((-A)^{\varrho})} < \infty$ *such that, for every* $t \in [0,T]$*, Eq.* (52) *holds* $\mathbb{P}$*-almost surely.*

Proof. If Assumptions 4.1, 4.3 and 4.5 are fulfilled for $\varrho = 0$, Eq. (46) fits into the setting of [45] (the Hilbert space U is denoted by H there). By Proposition 3.1 in [45] there exists a unique mild solution u to Eq. (46). Now suppose that Assumptions 4.1, 4.3 and 4.5 hold for some $\varrho > 0$. Set

$$\hat{U} := D((-A)^{\varrho}), \quad D(\hat{A}) := D((-A)^{\varrho+1})$$

and consider the unbounded operator $\hat{A}$ on $\hat{U}$ given by

$$\hat{A} : D(\hat{A}) \subset \hat{U} \to \hat{U},\ v \mapsto \hat{A}v := Av.$$

Note that $\hat{A}$ fulfills Assumption 4.1 with A, $D(A)$ and U replaced by $\hat{A}$, $D(\hat{A})$ and $\hat{U}$, respectively. Defining the spaces $D((-\hat{A})^s)$ analogously to the spaces $D((-A)^s)$, we have $D((-A)^{\varrho+s}) = D((-\hat{A})^s)$, $s \in \mathbb{R}$, so that Assumptions 4.3 and 4.5 can be reformulated with ϱ, $D((-A)^{\varrho})$, $D((-A)^{\varrho-\sigma})$ and $D((-A)^{\varrho-\beta})$ replaced by $\hat{\varrho} := 0$, $D((-\hat{A})^{\hat{\varrho}})$, $D((-\hat{A})^{\hat{\varrho}-\sigma})$ and $D((-\hat{A})^{\hat{\varrho}-\beta})$, respectively. Thus, the equation

$$\mathrm{d}u(t) = \big(\hat{A}u(t) + f(u(t))\big)\,\mathrm{d}t + B(u(t))\,\mathrm{d}W(t), \quad u(0) = u_0, \tag{53}$$

fits into the setting of [45] (now $\hat{U}$ corresponds to the space H there), so that, by Proposition 3.1 in [45], there exists a unique mild solution u to Eq. (53). Since the operators $e^{tA} \in \mathcal{L}(U)$ and $e^{t\hat{A}} \in \mathcal{L}(\hat{U})$ coincide on $\hat{U} \subset U$, it is clear that any mild solution to Eq. (53) is a mild solution to Eq. (46) and vice versa. □

Remark 4.8. If the initial condition u_0 belongs to $L_p(\Omega, \mathscr{F}_0, \mathbb{P}; D((-A)^{\varrho}))$ $\subset L_2(\Omega, \mathscr{F}_0, \mathbb{P}; D((-A)^{\varrho}))$ for some $p > 2$, the solution u even satisfies $\sup_{t\in[0,T]} \mathbb{E}\|u(t)\|^p_{D((-A)^{\varrho})} < \infty$. This is a consequence of the Burkholder-Davis-Gundy inequalities, cf. [15, Theorem 7.4] or [45, Proposition 3.1]. Analogous improvements are valid for the estimates in Propositions 4.15 and 4.19 below.

We finish this subsection with concrete examples for stochastic PDEs that fit into our setting.

Example 4.9. Let $\mathcal{O}$ be an open and bounded subset of $\mathbb{R}^d$, $U := L_2(\mathcal{O})$, and let $A = \Delta^D_{\mathcal{O}}$ be the Dirichlet-Laplacian on $\mathcal{O}$ as described in Example 4.2. We consider examples for stochastic PDEs in dimension $d = 1$ and $d \geq 2$.

First, let $\mathcal{O} \subset \mathbb{R}^1$ be one-dimensional and consider the problem

$$\left.\begin{aligned} \mathrm{d}u(t,x) &= \Delta_x u(t,x)\,\mathrm{d}t + g(u(t,x))\,\mathrm{d}t + h(u(t,x))\;\mathrm{d}W_1(t,x), \\ &\qquad\qquad (t,x) \in [0,T] \times \mathcal{O}, \\ u(t,x) &= 0, \qquad\quad (t,x) \in [0,T] \times \partial\mathcal{O}, \\ u(0,x) &= u_0(x), \quad\; x \in \mathcal{O}. \end{aligned}\right\} \tag{54}$$

where $u_0 \in L_2(\mathcal{O})$, $g : \mathbb{R} \to \mathbb{R}$ is globally Lipschitz continuous, $h : \mathbb{R} \to \mathbb{R}$ is bounded and globally Lipschitz continuous, and $W_1 = (W_1(t))_{t\in[0,T]}$ is a Wiener process (with respect to a normal filtration on a complete probability space) whose Cameron–Martin space is some space of functions on $\mathcal{O}$ that is continuously embedded in $L_2(\mathcal{O})$, e.g., W_1 is a cylindrical Wiener process on $L_2(\mathcal{O})$. Let $(\psi_k)_{k\in\mathbb{N}}$ be an arbitrary orthonormal basis of the Cameron–Martin space of W_1 and set

$$\begin{aligned} f(v)(x) &:= g(v(x)), && v \in L_2(\mathcal{O}),\ x \in \mathcal{O}, \\ \big(B(v)\mathbf{a}\big)(x) &:= h(v(x)) \sum_{k\in\mathbb{N}} a_k \psi_k(x), && v \in L_2(\mathcal{O}),\ \mathbf{a} = (a_k)_{k\in\mathbb{N}} \in \ell_2,\ x \in \mathcal{O}. \end{aligned} \tag{55}$$

Then, Eq. (46) is an abstract version of problem (54), and the mappings f and B are globally Lipschitz continuous (and thus linearly growing) from $D((-A)^0) = L_2(\mathcal{O})$ to $L_2(\mathcal{O})$ and from $D((-A)^0)$ to $\mathcal{L}(\ell_2; L_2(\mathcal{O}))$, respectively. It follows that Assumptions 4.1, 4.3 and 4.5 are fulfilled for $1/2 < \alpha < 1$ (compare Example 4.2) and $\varrho = \sigma = \beta = 0$.

Now let $\mathcal{O} \subset \mathbb{R}^d$ be d-dimensional, $d \geq 2$, and consider the problem (54) where $u_0 \in L_2(\mathcal{O})$, $g : \mathbb{R} \to \mathbb{R}$ is globally Lipschitz continuous, $h : \mathbb{R} \to \mathbb{R}$ is constant (additive noise), and $W_1 = (W_1(t))_{t\in[0,T]}$ is a Wiener process whose Cameron–Martin space is some space of functions on $\mathcal{O}$ that is continuously embedded in $D((-A)^{-\beta})$ for some $\beta < 1/2 - d/4$. One easily sees that the mappings f and B, defined as in (55), are globally Lipschitz continuous (and thus linearly growing) from $D((-A)^0) = L_2(\mathcal{O})$ to $L_2(\mathcal{O})$ and from $D((-A)^0)$ to $\mathcal{L}(\ell_2; D((-A)^{-\beta}))$, respectively. It follows that Assumptions 4.1, 4.3 and 4.5 are fulfilled for $\beta < 1/2 - d/4$, $d/2 < \alpha < 1 - 2\beta$, and $\varrho = \sigma = 0$. Alternatively, we could assume h to be sufficiently smooth and replace $h(u(t,x))$ in problem (54) by, e.g., $h\big(\int_{\mathcal{O}} k(x,y)u(y)\,\mathrm{d}y\big)$ with a sufficiently smooth kernel $k : \mathcal{O} \times \mathcal{O} \to \mathbb{R}$.

Example 4.10. As in Examples 4.2 and 4.9, let $A = \Delta_{\mathcal{O}}^D$ be the Dirichlet-Laplacian on an open and bounded domain $\mathcal{O} \subset \mathbb{R}^d$. From the numerical point of view, we are especially interested in stochastic PDEs of type (46) with $\varrho = 1/2$. In this case the solution process takes values in the space $D((-A)^{1/2}) = H_0^1(\mathcal{O})$, and, as we will see later in Proposition 4.15 and Theorem 4.17, we obtain estimates for the approximation error in terms of the energy norm

$$\|v\|_{D((-\Delta_{\mathcal{O}}^D)^{1/2})} = \langle \nabla v, \nabla v \rangle_{L_2(\mathcal{O})}^{1/2}, \qquad v \in H_0^1(\mathcal{O}).$$

The energy norm is crucial because error estimates for numerical solvers of elliptic problems (which we want to apply in each time step) are usually expressed in terms of this norm, compare Section 5 where we consider adaptive wavelet solvers with optimal convergence rates.

First, let $\mathcal{O} \subset \mathbb{R}^1$ be one-dimensional, and consider the problem (54) where $u_0 \in H_0^1(\mathcal{O})$, $g : \mathbb{R} \to \mathbb{R}$ is globally Lipschitz continuous, $h : \mathbb{R} \to \mathbb{R}$ is linear or constant, and $W_1 = (W_1(t))_{t\in[0,T]}$ is a Wiener process whose Cameron–Martin space is some space of functions on $\mathcal{O}$ that is continuously embedded in $D((-A)^{1/2-\beta})$ for some nonnegative $\beta < 1/4$, so that W_1 takes values in a bigger Hilbert space, say, in $D((-A)^{-1/4})$. (The embedding $D((-A)^{1/2-\beta}) \hookrightarrow D((-A)^{-1/4})$ is Hilbert–Schmidt since (49) is fulfilled for $\alpha > 1/2$, compare Example 4.2.) Take an arbitrary orthonormal basis $(\psi_k)_{k\in\mathbb{N}}$ of the Cameron–Martin space of W_1, and define $f(v)$ and $B(v)$ for $v \in H_0^1(\mathcal{O})$ analogously to (55). Then, Eq. (46) is an abstract version of problem (54), and the mappings f and B are globally Lipschitz continuous (and thus linearly growing) from $D((-A)^{1/2}) = H_0^1(\mathcal{O})$ to $D((-A)^0) = L_2(\mathcal{O})$ and from $D((-A)^{1/2})$ to $\mathcal{L}(\ell_2; D((-A)^{1/2-\beta}))$, respectively. The mapping properties of B follow from the inequalities $\|vw\|_{L_2(\mathcal{O})} \leq \|v\|_{H_0^1(\mathcal{O})}\|w\|_{L_2(\mathcal{O})}$ and $\|vw\|_{H_0^1(\mathcal{O})} \leq C\|v\|_{H_0^1(\mathcal{O})}\|w\|_{H_0^1(\mathcal{O})}$ (a consequence of the Sobolev embedding $H^1(\mathcal{O}) \hookrightarrow L_\infty(\mathcal{O})$ in dimension 1) and interpolation since $D((-A)^{1/2-\beta}) = [L_2(\mathcal{O}), D((-A)^{1/2})]_{1-2\beta}$. Thus, Assumptions 4.1, 4.3 and 4.5 are fulfilled for $\varrho = \sigma = 1/2$, $0 \leq \beta < 1/4$ and $1/2 < \alpha < 1 - 2\beta$.

Now let $\mathcal{O} \subset \mathbb{R}^d$ be d-dimensional and consider problem (54) where $u_0 \in H_0^1(\mathcal{O})$, $g : \mathbb{R} \to \mathbb{R}$ is globally Lipschitz continuous, $h : \mathbb{R} \to \mathbb{R}$ is constant, and $W_1 = (W_1(t))_{t\in[0,T]}$ is a Wiener process whose Cameron–Martin space is continuously embedded in $D((-A)^{1/2-\beta})$ for some $\beta < 1/2 - d/4$. Then, the mappings f and B, defined analogously to the one dimensional case, are globally Lipschitz continuous (and thus linearly growing) from $D((-A)^{1/2}) = H_0^1(\mathcal{O})$ to $D((-A)^0) = L_2(\mathcal{O})$ and from $D((-A)^{1/2})$ to $\mathcal{L}(\ell_2; D((-A)^{1/2-\beta}))$ respectively. It follows that Assumptions 4.1, 4.3 and 4.5 are fulfilled for $\varrho = \sigma = 1/2$, $\beta < 1/2 - d/4$ and $1 < \alpha < 1 - 2\beta$. As in Example 4.9 we could alternatively assume $h : \mathbb{R} \to \mathbb{R}$ to be sufficiently smooth and replace $h(u(t,x))$ in problem (54) by, e.g., $h\big(\int_{\mathcal{O}} k(x,y)u(y)\,\mathrm{d}y\big)$ with a sufficiently smooth kernel $k : \mathcal{O} \times \mathcal{O} \to \mathbb{R}$. We do not go into details here.

4.2 Semi-discretization in time

From now on, let Assumptions 4.1, 4.3 and 4.5 be fulfilled.

For the time discretization of the (mild) solution $u = (u(t))_{t\in[0,T]}$ to Eq. (46) we use the stochastic analogue of the linearly-implicit Euler scheme (3), i.e., for $K \in \mathbb{N}$ and $\tau = T/K$ we consider discretizations $(u_k)_{k=0}^K$ given by the initial condition u_0 in (46) and

$$\left.\begin{aligned} u_{k+1} = (I - \tau A)^{-1}\big(u_k + \tau f(u_k) + \sqrt{\tau} B(u_k)\chi_k\big), \\ k = 0, \ldots, K-1. \end{aligned}\right\} \tag{56}$$

We use the abbreviation

$$\chi_k := \chi_k^K := \frac{1}{\sqrt{\tau}} \left(W\big(t_{k+1}^K\big) - W\big(t_k^K\big)\right).$$

Note that each χ_k, $k = 0, \ldots, K-1$, is an $\mathscr{F}_{t_{k+1}}$-measurable Gaussian white noise on ℓ_2, i.e., a linear isometry from ℓ_2 to $L_2(\Omega, \mathscr{F}_{t_{k+1}}, \mathbb{P})$ such that for each $\mathbf{a} \in \ell_2$ the real valued random variable $\chi_k(\mathbf{a})$ is centred Gaussian with variance $\|\mathbf{a}\|_{\ell_2}^2$. We write $\chi_k(\mathbf{a}) \sim \mathcal{N}(0, \|\mathbf{a}\|_{\ell_2})$ for short. Moreover, for each $k = 0, 1, \ldots, K-1$, the sub-σ-field of $\mathscr{F}$ generated by $\{\chi_k(\mathbf{a}) : \mathbf{a} \in \ell_2\}$ is independent of $\mathscr{F}_{t_k}$.

We explain in which way the scheme (56) has to be understood. Let G be a separable real Hilbert space such that $D((-A)^{\varrho-\beta})$ is embedded into G via a Hilbert–Schmidt embedding. Then, for all $k = 0, \ldots, K-1$ and for all $\mathscr{F}_{t_k}$-measurable, $D((-A)^{\varrho})$-valued, square integrable random variables $v \in L_2(\Omega, \mathscr{F}_{t_k}, \mathbb{P}; D((-A)^{\varrho}))$, the term $B(v)\chi_k$ can be interpreted as an $\mathscr{F}_{t_{k+1}}$-measurable, square integrable, G-valued random variable in the sense

$$B(v)\chi_k := L_2(\Omega, \mathscr{F}_{t_{k+1}}, \mathbb{P}; G)\text{-}\lim_{J\to\infty} \sum_{j=1}^{J} \chi_k(\mathbf{b}_j) B(v)\mathbf{b}_j \tag{57}$$

where $\{\mathbf{b}_j\}_{j\in\mathbb{N}}$ is an orthonormal basis of ℓ_2. This definition is independent of the specific choice of the orthonormal basis $\{\mathbf{b}_j\}_{j\in\mathbb{N}}$. Note that the stochastic independence of $\{\chi_k(\mathbf{a}) : \mathbf{a} \in \ell_2\}$ and $\mathscr{F}_{t_k}$ is important at this point. We have

$$\mathbb{E}\|B(v)\chi_k\|_G^2 = \mathbb{E}\|B(v)\|_{\mathcal{L}_{\mathrm{HS}}(\ell_2;G)}^2, \tag{58}$$

the last term being finite due to the Lipschitz continuity of B by Assumption 4.3 (see also Remark 4.4) and the fact that the embedding $D((-A)^{\varrho-\beta}) \hookrightarrow G$ is Hilbert–Schmidt. Let us explicitly set

$$G := D((-A)^{\varrho - \max(\sigma, \beta+\alpha/2)}).$$

The condition $\mathrm{Tr}(-A)^{-\alpha} < \infty$ in Assumption 4.1 yields that the embedding $D((-A)^{\varrho-\beta}) \hookrightarrow D((-A)^{\varrho-\beta-\alpha/2})$ is Hilbert–Schmidt, and the embedding $D((-A)^{\varrho-\beta-\alpha/2}) \hookrightarrow D((-A)^{\varrho-\max(\sigma,\beta+\alpha/2)})$ is clearly continuous. Thus, we have indeed a Hilbert–Schmidt embedding $D((-A)^{\varrho-\beta}) \hookrightarrow G$. For all $k = 0, \ldots, K-1$ and $v \in L_2(\Omega, \mathscr{F}_{t_k}, \mathbb{P}; D((-A)^{\varrho}))$ we consider the term $B(v)\chi_k$ as an element in the space

$$L_2(\Omega, \mathscr{F}_{t_{k+1}}, \mathbb{P}; G) = L_2\big(\Omega, \mathscr{F}_{t_{k+1}}, \mathbb{P}; D((-A)^{\varrho-\max(\sigma,\beta+\alpha/2)})\big).$$

Next, due to the Lipschitz continuity of f by Assumption 4.3 (see also Remark 4.4), we also know that for all $v \in L_2(\Omega, \mathscr{F}_{t_k}, \mathbb{P}; D((-A)^\varrho))$ the term $f(v)$ is an element in $L_2(\Omega, \mathscr{F}_{t_k}, \mathbb{P}; G)$. Finally, as a consequence of Lemma 4.13 below and the fact that $\max(\sigma, \beta+\alpha/2) < \max(\sigma, 1/2) < 1$ due to (50), the operator $(I - \tau A)^{-1}$ is continuous from G to $D((-A)^\varrho)$. It follows that the discretizations $(u_k)_{k=0}^K$ are uniquely determined by (56) and that every u_k belongs to the space $L_2(\Omega, \mathscr{F}_{t_k}, \mathbb{P}; D((-A)^\varrho))$.

Remark 4.11. In practice, one has to truncate the noise expansion (57) and one will approximate $B(v)\chi_k \in L_2(\Omega, \mathscr{F}_{t_{k+1}}, \mathbb{P}; G)$ by a finite sum $\sum_{j=1}^J \chi_k(\mathbf{b}_j)B(v)\mathbf{b}_j \in L_2(\Omega, \mathscr{F}_{t_{k+1}}, \mathbb{P}; D((-A)^{\varrho-\beta}))$, $J \in \mathbb{N}$. However, in this paper we assume the right hand sides of the elliptic equations in each time step to be given exactly, cf. Section 5, Assumption 5.9. We postpone the analysis of these additional truncation errors for stochastic right hand sides to a forthcoming paper.

Now we can embed the scheme (56) into the abstract setting of Section 2. For measurability reasons the spaces $\mathcal{H}$ and $\mathcal{G}$ have to depend on the time step k, i.e., we consider spaces $\mathcal{H} = \mathcal{H}_k$ and $\mathcal{G} = \mathcal{G}_k$.

Observation 4.12. *Let*

$$\left.\begin{aligned}
\mathcal{H}_k &:= L_2(\Omega, \mathscr{F}_{t_k}, \mathbb{P}; D((-A)^\varrho)), && k = 0, \ldots, K,\\
\mathcal{G}_k &:= L_2(\Omega, \mathscr{F}_{t_k}, \mathbb{P}; G), && k = 1, \ldots, K,\\
R_{\tau,k} &: \mathcal{H}_k \to \mathcal{G}_{k+1} && \\
&\quad v \mapsto R_{\tau,k}(v) := v + \tau f(v) + \sqrt{\tau} B(v)\chi_k, && k = 0, \ldots, K-1,\\
L_\tau^{-1} &: \mathcal{G}_k \to \mathcal{H}_k && \\
&\quad v \mapsto L_\tau^{-1} v := (I - \tau A)^{-1} v, && k = 1, \ldots, K.
\end{aligned}\right\} \tag{59}$$

With these definitions at hand, the linearly-implicit Euler scheme (56) *can be written in the form of the abstract S-stage scheme* (6) *with $S = 1$, $L_{\tau,1}^{-1} := L_\tau^{-1}$ and $R_{\tau,k,1} := R_{\tau,k}$, for $k = 0, \ldots, K-1$. We have*

$$u_{k+1} = L_\tau^{-1} R_{\tau,k}(u_k), \quad k = 0, \ldots, K-1. \tag{60}$$

The fact that the spaces $\mathcal{H} = \mathcal{H}_k$ and $\mathcal{G} = \mathcal{G}_k$ depend on the time step k does not cause any problems when using results from Section 2, as far as the corresponding assumptions are fulfilled, see also Remark 2.16.

The following Lemma is helpful not only for the preceding argument but also for estimates further down.

Lemma 4.13. *Let $\tau > 0$ and $r \in \mathbb{R}$. The operator $I - \tau A$ is a homeomorphism from $D((-A)^r)$ to $D((-A)^{r-1})$. For $n \in \mathbb{N}$ we have the following operator norm estimates for $(I - \tau A)^{-n}$, considered as an operator from $D((-A)^{r-s})$ to $D((-A)^r)$, $s \le 1$:*

$$\|(I - \tau A)^{-n}\|_{\mathcal{L}(D((-A)^{r-s}), D((-A)^r))} \le \begin{cases} s^s \left(1 - \frac{s}{n}\right)^{(n-s)} (n\tau)^{-s} & , 0 < s \le 1 \\ (-\lambda_1)^s (1 - \tau\lambda_1)^{-n} & , s \le 0. \end{cases}$$

Proof. The bijectivity of $I - \tau A : D((-A)^r) \to D((-A)^{r-1})$ is almost obvious. Its proof is left to the reader. The bicontinuity follows from the continuity of the inverse as shown below (case $s = 1$) and the bounded inverse theorem. Concerning the operator norm estimates, we use Parseval's identity and the spectral properties of A to obtain

$$\begin{aligned}
&\sup_{\|v\|_{D((-A)^{r-s})}=1} \|(I-\tau A)^{-n} v\|^2_{D((-A)^r)} \\
&\qquad = \sup_{\|w\|_U=1} \|(I-\tau A)^{-n}(-A)^{s-r} w\|^2_{D((-A)^r)} \\
&\qquad = \sup_{\|w\|_U=1} \|(-A)^r (I-\tau A)^{-n}(-A)^{s-r} w\|^2_U \\
&\qquad = \sup_{\|w\|_U=1} \sum_{k\in\mathbb{N}} \left|(-\lambda_k)^s (1-\tau\lambda_k)^{-n} \langle w, e_k \rangle_U\right|^2 .
\end{aligned}$$

If $s \leq 0$, the last expression is equal to $(-\lambda_1)^{2s}(1-\tau\lambda_1)^{-2n}$. If $0 < s \leq 1$, an upper bound is given by the square of

$$\sup_{x>0} \frac{x^s}{(1+\tau x)^n} = s^s \Big(1 - \frac{s}{n}\Big)^{(n-s)} (n\tau)^{-s}. \qquad \square$$

Remark 4.14. Without additional assumptions on B or a truncation of the noise expansion (57), the operator $R_{\tau,k}$ cannot easily be traced back to a family of operators

$$R_{\tau,k,\omega} : D((-A)^\varrho) \to G,\ \omega \in \Omega,$$

in the sense that for $v \in \mathcal{H}_k = L_2(\Omega, \mathscr{F}_{t_k}, \mathbb{P}; D((-A)^\varrho))$ the image $R_{\tau,k}(v)$ is determined by

$$(R_{\tau,k}(v))(\omega) = R_{\tau,k,\omega}(v(\omega)) \quad \text{for } \mathbb{P}\text{-almost all } \omega \in \Omega. \tag{61}$$

However, this is possible, for instance, if for all $v \in D((-A)^\varrho)$ the operator $B(v) : \ell_2 \to D((-A)^{\varrho-\beta})$ has a continuous extension $B(v) : U_0 \to D((-A)^{\varrho-\beta})$ to a bigger Hilbert space U_0 such that ℓ_2 is embedded into U_0 via a Hilbert–Schmidt embedding. Another instance where a representation of the form (61) is possible is the case where the mapping $B : D((-A)^\varrho) \to \mathcal{L}(\ell_2; D((-A)^{\varrho-\beta}))$ is constant, i.e., the case of additive noise. We take a closer look at the latter case, writing $B \in \mathcal{L}(\ell_2; D((-A)^{\varrho-\beta}))$ for short. We fix a version of each of the $\mathbb{P}$-almost surely determined, G-valued random variables $B\chi_k = B\chi_k^K$, $k = 0, 1, \ldots, K-1$, $K \in \mathbb{N}$, and set

$$R_{\tau,k,\omega}(v) := v + f(v) + (B\chi_k)(\omega), \qquad \omega \in \Omega,\ v \in D((-A)^\varrho).$$

It is clear that (61) holds for all $v \in L_2(\Omega, \mathscr{F}_{t_k}, \mathbb{P}; D((-A)^\varrho))$, and we have the following alternative interpretation of the scheme (56) in the case of additive noise within the abstract setting of Section 2.

$$
\left.
\begin{aligned}
\mathcal{H} &:= D((-A)^{\varrho})),\\
\mathcal{G} &:= G = D((-A)^{\varrho-\max(\sigma,\beta+\alpha/2)}),\\
R_{\tau,k,\omega} &: \mathcal{H} \to \mathcal{G}\\
&\qquad v \mapsto R_{\tau,k,\omega}(v) := v + \tau f(v) + \sqrt{\tau}(B\chi_k)(\omega),\\
L_\tau^{-1} &: \mathcal{G} \to \mathcal{H}\\
&\qquad v \mapsto L_\tau^{-1} v := (I - \tau A)^{-1} v,
\end{aligned}
\right\}
\qquad (59_\omega)
$$

$k = 0, \ldots, K-1$. With these definitions, the abstract scheme (6) in Section 2 with $S = 1$ describes the stochastic scheme (56) in an ω-wise sense.

We close this subsection with an extension of the error estimate for the linearly-implicit Euler scheme in [45]. In this way, we verify Assumption 2.13 for the scheme (56) in its abstract form (60), see Remark 4.16(i) below.

Proposition 4.15. *Let Assumptions 4.1, 4.3, and 4.5 be fulfilled. Let* $(u_k)_{k=0}^K$ *be the time discretization of the mild solution* $(u(t))_{t\in[0,T]}$ *to Eq.* (46)*, given by the scheme* (56)*. Then, for every*

$$\delta < \min(1-\sigma, (1-\alpha)/2 - \beta),$$

we have for all $1 \le k \le K$

$$\Big(\mathbb{E}\|u(t_k) - u_k\|^2_{D((-A)^{\varrho})}\Big)^{1/2} \le C\left(\tau^\delta + \frac{1}{k}\Big(\mathbb{E}\|u_0\|^2_{D((-A)^{\varrho})}\Big)^{1/2}\right),$$

where the constant $C > 0$ *depends only on* δ*,* A*,* B*,* f*,* α*,* β*,* σ *and* T*.*

Proof. We argue as in the proof of Proposition 4.7 and consider the equation

$$\mathrm{d}u(t) = \big(\hat{A}u(t) + f(u(t))\big)\,\mathrm{d}t + B(u(t))\,\mathrm{d}W(t), \quad u(0) = u_0, \qquad (62)$$

where the operator $\hat{A}: \ D(\hat{A}) \subset \hat{U} \to \hat{U}$ is defined by $\hat{U} := D((-A)^{\varrho})$, $D(\hat{A}) := D((-A)^{\varrho+1})$, and $\hat{A}v := Av$, $v \in D(\hat{A})$. Eq. (62) fits into the setting of [45], and its mild solution $u = (u(t))_{t\in[0,T]}$ coincides with the mild solution to Eq. (46), compare the proof of Proposition 4.7. For $K \in \mathbb{N}$ let $(\hat{u}_k)_{k=0}^K$ be given by the linearly-implicit Euler scheme

$$
\begin{aligned}
\hat{u}_0 &= u_0,\\
\hat{u}_{k+1} &= (I - \tau\hat{A})^{-1}\big(\hat{u}_k + \tau f(\hat{u}_k) + \sqrt{\tau}B(\hat{u}_k)\chi_k\big), \quad k = 0, \ldots, K-1.
\end{aligned}
$$

By Theorem 3.2 in [45] we have, for all $\delta < \min(1-\sigma, (1-\alpha)/2 - \beta)$,

$$\Big(\mathbb{E}\|u(t_k) - \hat{u}_k\|^2_{\hat{U}}\Big)^{1/2} \le C\left(\tau^\delta + \frac{1}{k}\Big(\mathbb{E}\|u_0\|^2_{\hat{U}}\Big)^{1/2}\right),$$

$1 \leq k \leq K$. The proof in [45] reveals that the constant $C > 0$ depends only on δ, $\hat{A}$, B, f, α, β, σ and T. The assertion of Proposition 4.15 follows from the fact that the natural extensions and restrictions of the operators $(I - \tau\hat{A})^{-1}$ and $(I - \tau A)^{-1}$ to the spaces $D((-\hat{A})^s) = D((-A)^{s+\varrho})$, $s \in \mathbb{R}$, coincide, so that $\hat{u}_k = u_k$ for all $0 \leq k \leq K$, $K \in \mathbb{N}$. □

Remark 4.16. **(i)** If $k \geq K^\delta$ $(\delta > 0)$, then $1/k \leq T^{-\delta}\tau^\delta$, and we obtain

$$\Big(\mathbb{E}\|u(t_k) - u_k\|^p_{D((-A)^\varrho)}\Big)^{1/p} \leq C_{\text{exact}}\,\tau^\delta \tag{63}$$

with a constant $C_{\text{exact}} > 0$ that depends only on δ, u_0, A, B, f, α, β, σ and T. Since δ is always smaller than 1, it follows in particular that (63) holds for $k = K$. Using the definitions in (59) and the notation introduced in Section 2 this means that Assumption 2.13 is fulfilled, i.e., we have

$$\|u(T) - E_{\tau,0,K}(u_0)\|_{\mathcal{H}_K} \leq C_{\text{exact}}\,\tau^\delta,$$

where the Euler scheme operator $E_{\tau,0,K} : \mathcal{H}_0 \to \mathcal{H}_K$ is given by the composition $(L_\tau^{-1}R_{\tau,K-1}) \circ (L_\tau^{-1}R_{\tau,K-2}) \circ \cdots \circ (L_\tau^{-1}R_{\tau,0})$.

(ii) The proof of Proposition 4.15 is based on an application of Theorem 3.2 in [45] to Eq. (53). The careful reader might have observed that the parameter σ, which corresponds to the parameter s in [45], is assumed to be positive in [45]. However, a closer look at the estimates in the proof of Theorem 3.2 in [45] reveals that the result can be extended to negative values of σ and s, respectively. Alternatively, one can argue that if $\sigma \leq 0$ then $D((-\hat{A})^{-\sigma})$ is continuously embedded into, say, $D((-\hat{A})^{-1/2})$, so that Eq. (53) fits into the setting of [45] if f is considered as a mapping from $\hat{U} = D((-\hat{A})^0)$ to $D((-\hat{A})^{-1/2})$. We refer to [13] where the Euler scheme for stochastic evolution equations is considered in a more general setting than in [45].

4.3 Discretization in time and space

So far we have verified the existence and uniqueness of a mild solution to Eq. (46) as well as the convergence of the exactly evaluated Euler scheme (56) with rate $\delta < \min(1 - \sigma, (1 - \alpha)/2 - \beta)$. We now turn to the corresponding inexact scheme where the iterated stationary problems in (56), respectively (60), are solved only approximately. As in Sections 2 and 3 we are interested in how the tolerances for the spatial approximation errors in each time step have to be chosen to achieve the same order of convergence as for the exact scheme. Throughout this subsection we use the definitions given in Observation 4.12 to interprete the setting described in the previous Subsections 4.1 and 4.2 in terms of the abstract framework of Section 2.

As in Section 2, Assumption 2.9, we assume that we have a numerical scheme that, for all $w \in \mathcal{H}_k$, $k = 0, \ldots, K - 1$, and for every prescribed tolerance $\varepsilon > 0$, provides us with an approximation $[v]_\varepsilon$ of

$$v = L_\tau^{-1} R_{\tau,k}(w)$$

such that

$$\|v - [v]_\varepsilon\|_{\mathcal{H}_{k+1}} = \Big(\mathbb{E}\|v - [v]_\varepsilon\|^2_{D((-A)^\varrho)}\Big)^{1/2} \leq \varepsilon.$$

We think of $[v]_\varepsilon$ as the result of an ω-wise application of some deterministic solver with error $\leq \varepsilon$ in $D((-A)^\varrho)$, e.g., an adaptive wavelet solver as described in Subsection 5.3. In contrast to Sections 2 and 3 we pass on any assumptions and discussions concerning the degrees of freedom involved, since in the stochastic case this is a fairly different matter that deserves an independent treatment in a future paper. However, we note that all results from Section 2 that do not involve assertions concerning the degrees of freedom remain valid in the present setting with obvious modifications.

Given prescribed tolerances ε_k, $k = 0, \ldots, K-1$, for the spatial approximation errors in each time step, we consider the inexact version of the linearly-implicit Euler scheme (60)

$$\left.\begin{aligned} \tilde{u}_0 &= u_0, \\ \tilde{u}_{k+1} &= [L_\tau^{-1} R_{\tau,k}(\tilde{u}_k)]_{\varepsilon_k}, \quad k = 0, \ldots, K-1, \end{aligned}\right\} \tag{64}$$

which is the analogue to scheme (13) from Section 2 with $S = 1$ and $\varepsilon_{k,1} := \varepsilon_k$. We already know from the considerations in Section 2 that sufficient conditions how to tune the tolerances ε_k in the inexact scheme (64) to obtain the same order of convergence as for the exact scheme (60) can be described in terms of the Lipschitz constants $C^{\mathrm{Lip}}_{\tau,j,k}$ of the operators

$$E_{\tau,j,k} = (L_\tau^{-1} R_{\tau,k-1}) \circ (L_\tau^{-1} R_{\tau,k-2}) \circ \cdots \circ (L_\tau^{-1} R_{\tau,j}) : \mathcal{H}_j \to \mathcal{H}_k,$$

$1 \leq j \leq k \leq K$, $K \in \mathbb{N}$. In the present setting we are able to show that the constants $C^{\mathrm{Lip}}_{\tau,j,k}$ are bounded uniformly in j, k and τ, see Proposition 4.18. Together with the arguments in Section 2 and Proposition 4.15 this will lead to the following main result of this Section.

Theorem 4.17. *Let Assumptions 4.1, 4.3 and 4.5 be fulfilled. Let* $(u(t))_{t\in[0,T]}$ *be the unique mild solution to Eq.* (46) *and let*

$$\delta < \min(1-\sigma, (1-\alpha)/2 - \beta).$$

If one chooses

$$\varepsilon_k \leq \tau^{1+\delta}$$

for all $k = 0, \ldots, K-1$, $K \in \mathbb{N}$, *then the output* $\tilde{u}_K$ *of the inexact linearly-implicit Euler scheme* (64) *converges to* $u(T)$ *with rate* δ, *i.e., we have*

$$\Big(\mathbb{E}\|u(T) - \tilde{u}_K\|^2_{D((-A)^\varrho)}\Big)^{1/2} \leq C\tau^\delta$$

with a constant C *depending only on* u_0, δ, A, B, f, α, β, σ *and* T.

The verification of Theorem 4.17 is based on the estimate of the Lipschitz constants $C^{\mathrm{Lip}}_{\tau,j,k}$ in Proposition 4.18 below. Since its proof is rather lengthy we postpone it to the end of this subsection.

Proposition 4.18. *Let Assumptions 4.1, 4.3 and 4.5 be fulfilled. There exists a finite constant $C > 0$, depending only on A, B, f, α, β, σ and T, such that*

$$C^{\mathrm{Lip}}_{\tau,j,k} \leq C \quad \text{for all } \ 1 \leq j \leq k \leq K, \ K \in \mathbb{N}.$$

With this result at hand we obtain, as a next step towards the verification of Theorem 4.17, the following error estimate for the inexact scheme.

Proposition 4.19. *Let Assumptions 4.1, 4.3 and 4.5 be fulfilled. Let $(u(t))_{t\in[0,T]}$ be the unique mild solution to Eq. (46). Let $(\tilde{u}_k)_{k=0}^K$ be the discretization of $(u(t))_{t\in[0,T]}$ in time and space given by the inexact linearly-implicit Euler scheme (64), where ε_k, $k = 0, \ldots, K-1$, are prescribed tolerances for the spatial approximation errors in each time step. Then, for every $1 \leq k \leq K$, $K \in \mathbb{N}$, and for every*

$$\delta < \min(1-\sigma, (1-\alpha)/2 - \beta)$$

we have

$$\Big(\mathbb{E}\|u(t_k) - \tilde{u}_k\|^2_{D((-A)^\varrho)}\Big)^{1/2} \leq C \left(\tau^\delta + \frac{1}{k}\Big(\mathbb{E}\|u_0\|^2_{D((-A)^\varrho)}\Big)^{1/2} + \sum_{j=0}^{k-1} \varepsilon_j \right)$$

with a constant C that depends only on δ, A, B, f, α, β, σ and T.

Proof. Arguing as in the proof of Theorem 2.15, compare Remark 2.16, we obtain the general error estimate

$$\begin{aligned}\Big(\mathbb{E}\|u(t_k) - \tilde{u}_k\|^2_{D((-A)^\varrho)}\Big)^{1/2} &= \|u(t_k) - \tilde{u}_k\|_{\mathcal{H}_k} \\ &\leq \|u(t_k) - u_k\|_{\mathcal{H}_k} + \sum_{j=0}^{k-1} C^{\mathrm{Lip}}_{\tau,j+1,k} \varepsilon_j,\end{aligned} \tag{65}$$

where $(u_k)_{k=0}^K$ is the discretization of $(u(t))_{t\in[0,T]}$ in time given by the exact linearly-implicit Euler scheme (56), respectively (60). Note that the analogues to Assumptions 2.4 and 2.5 in Theorem 2.15 are fulfilled due to the construction of the inexact scheme (64); the analogues to Assumptions 2.2 and 2.11 follow from Assumptions 4.1, 4.3 and 4.5 via Propositions 4.7 and 4.18. Alternatively, we could have used a modified version of Theorem 2.20 with $S = 1$ and $\varepsilon_{j,1} := \varepsilon_j$. The assertion of the proposition follows directly from (65), Proposition 4.15 and Proposition 4.18. □

The proof of Theorem 4.17 is now straightforward, similar to the argumentation in the proofs of Theorems 2.17 and 2.22.

Proof of Theorem 4.17. The assertion follows from Proposition 4.19 and the elementary estimates

$$\frac{1}{K} \leq \frac{1}{K^\delta} = T^{-\delta}\tau^\delta \quad \text{and} \quad \sum_{j=0}^{K-1} \varepsilon_k \leq T\tau^\delta. \qquad \square$$

It remains to verify the estimate of the Lipschitz constants $C^{\mathrm{Lip}}_{\tau,j,k}$ of the operators $E_{\tau,j,k}$ stated in Proposition 4.18. The proof is based on a Gronwall argument.

Proof of Proposition 4.18. Fix $1 \leq j \leq k \leq K$ and observe that, by induction over k,

$$\begin{aligned} E_{\tau,j,k}(v) &= L_\tau^{-(k-j)} v \\ &\quad + \sum_{i=0}^{k-j-1} L_\tau^{-(k-j)+i} \Big(\tau f\big(E_{\tau,j,j+i}(v)\big) + \sqrt{\tau} B\big(E_{\tau,j,j+i}(v)\big)\chi_{j+i} \Big) \end{aligned}$$

for all $v \in \mathcal{H}_j$. Therefore, for all $v, w \in \mathcal{H}_j$, we have

$$\begin{aligned} &\|E_{\tau,j,k}(v) - E_{\tau,j,k}(w)\|_{\mathcal{H}_k} \\ &\quad \leq \|L_\tau^{-(k-j)} v - L_\tau^{-(k-j)} w\|_{\mathcal{H}_k} \\ &\qquad + \sum_{i=0}^{k-j-1} \tau \Big\| L_\tau^{-(k-j)+i} \Big(f\big(E_{\tau,j,j+i}(v)\big) - f\big(E_{\tau,j,j+i}(w)\big) \Big) \Big\|_{\mathcal{H}_k} \\ &\qquad + \Big\| \sum_{i=0}^{k-j-1} \sqrt{\tau} L_\tau^{-(k-j)+i} \Big(B\big(E_{\tau,j,j+i}(v)\big) - B\big(E_{\tau,j,j+i}(w)\big) \Big)\chi_{j+i} \Big\|_{\mathcal{H}_k} \\ &\quad =: (I) + (II) + (III). \end{aligned} \tag{66}$$

We estimate each of the terms (I), (II) and (III) separately.

By Lemma 4.13 and the trivial fact that $\|v - w\|_{\mathcal{H}_k} = \|v - w\|_{\mathcal{H}_j}$ for all $v, w \in \mathcal{H}_j$, we have

$$\begin{aligned} (I) &\leq \big\|L_\tau^{-1}\big\|^{k-j}_{\mathcal{L}(D((-A)^\varrho))} \|v - w\|_{\mathcal{H}_k} \\ &\leq (1 - \tau\lambda_1)^{-(k-j)} \|v - w\|_{\mathcal{H}_k} \\ &\leq \|v - w\|_{\mathcal{H}_j}. \end{aligned} \tag{67}$$

Concerning the term (II) in (66), let us first concentrate on the case $\sigma \in (0,1)$. We use the Lipschitz condition on f in Assumption 4.3 and

Lemma 4.13 to obtain

$$\begin{aligned}(II) &\le \sum_{i=0}^{k-j-1} \tau \big\| L_\tau^{-(k-j)+i}(-A)^\sigma \big\|_{\mathcal{L}(D((-A)^\varrho))} \\ &\qquad\qquad \times C_f^{\mathrm{Lip}} \big\| E_{\tau,j,j+i}(v) - E_{\tau,j,j+i}(w) \big\|_{\mathcal{H}_{j+i}} \\ &\le \sum_{i=0}^{k-j-1} \tau \frac{\sigma^\sigma}{(\tau(k-j-i))^\sigma} C_f^{\mathrm{Lip}} C_{\tau,j,j+i}^{\mathrm{Lip}} \|v-w\|_{\mathcal{H}_j} \\ &\le C_f^{\mathrm{Lip}} \sum_{i=0}^{k-j-1} \frac{\tau}{(\tau(k-j-i))^\sigma} C_{\tau,j,j+i}^{\mathrm{Lip}} \|v-w\|_{\mathcal{H}_j} .\end{aligned} \tag{68}$$

For the case that $\sigma \le 0$ we get with similar arguments

$$\begin{aligned}(II) &\le C_f^{\mathrm{Lip}} \sum_{i=0}^{k-j-1} \frac{\tau(-\lambda_1)^\sigma}{(1-\tau\lambda_1)^n} C_{\tau,j,j+i}^{\mathrm{Lip}} \|v-w\|_{\mathcal{H}_j} \\ &\le C_f^{\mathrm{Lip}} (-\lambda_1)^\sigma \sum_{i=0}^{k-j-1} \tau C_{\tau,j,j+i}^{\mathrm{Lip}} \|v-w\|_{\mathcal{H}_j} .\end{aligned} \tag{69}$$

Let us now look at the term (III) in (66). Using the independence of the stochastic increments χ_{j+i} and Eq. (58), we get

$$\begin{aligned}(III)^2 &= \sum_{i=0}^{k-j-1} \tau \mathbb{E} \Big\| L_\tau^{-(k-j)+i} \Big(B\big(E_{\tau,j,j+i}(v)\big) - B\big(E_{\tau,j,j+i}(w)\big)\Big)\chi_{j+i} \Big\|_{D((-A)^\varrho)}^2 \\ &\le \sum_{i=0}^{k-j-1} \tau \big\| L_\tau^{-(k-j)+i} \big\|_{\mathcal{L}(D((-A)^{\varrho-\beta-\alpha/2}),D((-A)^\varrho))}^2 \\ &\qquad \times \mathbb{E} \Big\| \Big(B\big(E_{\tau,j,j+i}(v)\big) - B\big(E_{\tau,j,j+i}(w)\big)\Big)\chi_{j+i} \Big\|_{D((-A)^{\varrho-\beta-\alpha/2})}^2 \\ &\le \sum_{i=0}^{k-j-1} \tau \big\| L_\tau^{-(k-j)+i} \big\|_{\mathcal{L}(D((-A)^{\varrho-\beta-\alpha/2}),D((-A)^\varrho))}^2 \\ &\qquad \times \mathbb{E} \big\| B\big(E_{\tau,j,j+i}(v)\big) - B\big(E_{\tau,j,j+i}(w)\big) \big\|_{\mathcal{L}_{\mathrm{HS}}(\ell_2; D((-A)^{\varrho-\beta-\alpha/2}))}^2 .\end{aligned}$$

Concentrating first on the case $\beta + \alpha/2 > 0$, we continue by using the Lipschitz condition on B in Assumption 4.3 and Lemma 4.13 to obtain

$$\begin{aligned}(III)^2 &\le \sum_{i=0}^{k-j-1} \tau \frac{(\beta+\alpha/2)^{2\beta+\alpha}}{(\tau(k-j-i))^{2\beta+\alpha}} \mathrm{Tr}(-A)^{-\alpha} \\ &\qquad \times (C_B^{\mathrm{Lip}})^2 \, \mathbb{E} \, \|E_{\tau,j,j+i}(v) - E_{\tau,j,j+i}(w)\|_{D((-A)^\varrho)}^2 \\ &\le (C_B^{\mathrm{Lip}})^2 \, \mathrm{Tr}(-A)^{-\alpha} \\ &\qquad \times \sum_{i=0}^{k-j-1} \frac{\tau}{(\tau(k-j-i))^{2\beta+\alpha}} (C_{\tau,j,j+i}^{\mathrm{Lip}})^2 \|v-w\|_{\mathcal{H}_j}^2 .\end{aligned} \tag{70}$$

In the case $\beta + \alpha/2 \leq 0$ the same arguments lead to

$$
\begin{aligned}
(\mathit{III})^2 \leq \sum_{i=0}^{k-j-1} \tau \frac{(-\lambda_1)^{2\beta+\alpha}}{(1-\tau\lambda_1)^{2n}} \operatorname{Tr}(-A)^{-\alpha} \\
\times (C_B^{\mathrm{Lip}})^2 \mathbb{E} \left\| E_{\tau,j,j+i}(v) - E_{\tau,j,j+i}(w) \right\|^2_{D((-A)^{\varrho})} \qquad (71)\\
\leq (C_B^{\mathrm{Lip}})^2 \operatorname{Tr}(-A)^{-\alpha} (-\lambda_1)^{2\beta+\alpha} \sum_{i=0}^{k-j-1} \tau (C^{\mathrm{Lip}}_{\tau,j,j+i})^2 \|v-w\|^2_{\mathcal{H}_j}.
\end{aligned}
$$

Now we have to consider four different cases.

Case 1. $\sigma \in (0,1)$ and $\beta + \alpha/2 \in (0, 1/2)$. The combination of (66), (67), (68) and (70) yields

$$
\begin{aligned}
C^{\mathrm{Lip}}_{\tau,j,k} \leq 1 + C_f^{\mathrm{Lip}} \sum_{i=0}^{k-j-1} \frac{\tau}{(\tau(k-j-i))^{\sigma}} C^{\mathrm{Lip}}_{\tau,j,j+i} \\
+ C_B^{\mathrm{Lip}} (\operatorname{Tr}(-A)^{-\alpha})^{1/2} \left(\sum_{i=0}^{k-j-1} \frac{\tau}{(\tau(k-j-i))^{2\beta+\alpha}} (C^{\mathrm{Lip}}_{\tau,j,j+i})^2 \right)^{1/2}.
\end{aligned}
\qquad (72)
$$

Note that inductively we obtain in particular the finiteness of all $C^{\mathrm{Lip}}_{\tau,j,j+i}$. Next, we estimate the two sums over i on the right hand side of (72) via Hölder's inequality. Set

$$
q := \frac{1}{\min(1-\sigma, (1-\alpha)/2 - \beta)} + 2 > 2.
$$

Hölder's inequality with exponents $q/(q-1)$ and q yields

$$
\begin{aligned}
&\sum_{i=0}^{k-j-1} \frac{\tau}{(\tau(k-j-i))^{\sigma}} C^{\mathrm{Lip}}_{\tau,j,j+i} \\
&\leq \left(\sum_{i=0}^{k-j-1} \frac{\tau}{(\tau(k-j-i))^{\frac{\sigma q}{q-1}}} \right)^{\frac{q-1}{q}} \left(\sum_{i=0}^{k-j-1} \tau (C^{\mathrm{Lip}}_{\tau,j,j+i})^q \right)^{\frac{1}{q}} \\
&\leq \left(\sum_{i=1}^{K} \frac{\tau}{(\tau i)^{\frac{\sigma q}{q-1}}} \right)^{\frac{q-1}{q}} \left(\sum_{i=0}^{k-j-1} \tau (C^{\mathrm{Lip}}_{\tau,j,j+i})^q \right)^{\frac{1}{q}} \qquad (73)\\
&\leq \left(\int_0^T t^{-\frac{\sigma q}{q-1}} \, \mathrm{d}t \right)^{\frac{q-1}{q}} \left(\sum_{i=0}^{k-j-1} \tau (C^{\mathrm{Lip}}_{\tau,j,j+i})^q \right)^{\frac{1}{q}},
\end{aligned}
$$

where the integral in the last line is finite since $\frac{\sigma q}{q-1} = \frac{\sigma}{1-1/q} < \frac{\sigma}{1-(1-\sigma)} = 1$.

Similarly, applying Hölder's inequality with exponents $q/(q-2)$ and $q/2$,

$$\sum_{i=0}^{k-j-1} \frac{\tau}{(\tau(k-j-i))^{2\beta+\alpha}} (C^{\mathrm{Lip}}_{\tau,j,j+i})^2$$

$$\leq \left(\sum_{i=0}^{k-j-1} \frac{\tau}{(\tau(k-j-i))^{\frac{(2\beta+\alpha)q}{q-2}}} \right)^{\frac{q-2}{q}} \left(\sum_{i=0}^{k-j-1} \tau (C^{\mathrm{Lip}}_{\tau,j,j+i})^q \right)^{\frac{2}{q}} \tag{74}$$

$$\leq \left(\int_0^T t^{-\frac{(2\beta+\alpha)q}{q-2}} \,\mathrm{d}t \right)^{\frac{q-2}{q}} \left(\sum_{i=0}^{k-j-1} \tau (C^{\mathrm{Lip}}_{\tau,j,j+i})^q \right)^{\frac{2}{q}}.$$

The integral in the last line is finite since $\frac{(2\beta+\alpha)q}{q-2} = \frac{(2\beta+\alpha)}{1-2/q} < \frac{(2\beta+\alpha)}{1-(1-\alpha-2\beta)} = 1$.

Combining (72), (73), (74) and using the equivalence of norms in $\mathbb{R}^3$, we obtain

$$(C^{\mathrm{Lip}}_{\tau,j,k})^q \leq C_0 \left(1 + \sum_{i=0}^{k-j-1} \tau (C^{\mathrm{Lip}}_{\tau,j,j+i})^q \right), \tag{75}$$

with a constant C_0 that depends only on A, f, B, α, β, σ and T. Since (75) holds for arbitrary $K \in \mathbb{N}$ and $1 \leq j \leq k \leq K$, we can apply a discrete version of Gronwall's lemma and obtain

$$(C^{\mathrm{Lip}}_{\tau,j,k})^q \leq e^{(k-j)\tau C_0} C_0 \leq e^{T C_0} C_0.$$

for all $1 \leq j \leq k \leq K$, $K \in \mathbb{N}$ and $\tau = T/K$. It follows that the assertion of the proposition holds in this first case with

$$C := (e^{TC_0} C_0)^{1/q}.$$

Case 2. $\sigma \leq 0$ and $\beta + \alpha/2 \leq 0$. A combination of (66) with (67), (69), and (71) leads to

$$C^{\mathrm{Lip}}_{\tau,j,k} \leq 1 + C^{\mathrm{Lip}}_f (-\lambda_1)^\sigma \sum_{i=0}^{k-j-1} \tau C^{\mathrm{Lip}}_{\tau,j,j+i}$$
$$+ C^{\mathrm{Lip}}_B (\mathrm{Tr}(-A)^{-\alpha})^{1/2} (-\lambda_1)^{\beta+\alpha/2} \Big(\sum_{i=0}^{k-j-1} \tau (C^{\mathrm{Lip}}_{\tau,j,j+i})^2 \Big)^{1/2}.$$

Applying Hölder's inequality with exponent $q_2 := 2$ to estimate the first sum over i on the right hand side, we get

$$C^{\mathrm{Lip}}_{\tau,j,k} \leq 1 + C^{\mathrm{Lip}}_f (-\lambda_1)^\sigma T^{1/2} \Big(\sum_{i=0}^{k-j-1} \tau (C^{\mathrm{Lip}}_{\tau,j,j+i})^2 \Big)^{1/2}$$
$$+ C^{\mathrm{Lip}}_B (\mathrm{Tr}(-A)^{-\alpha})^{1/2} (-\lambda_1)^{\beta+\alpha/2} \Big(\sum_{i=0}^{k-j-1} \tau (C^{\mathrm{Lip}}_{\tau,j,j+i})^2 \Big)^{1/2},$$

which leads to

$$(C^{\mathrm{Lip}}_{\tau,j,k})^2 \leq C\Big(1 + \sum_{i=0}^{k-j-1} \tau (C^{\mathrm{Lip}}_{\tau,j,j+i})^2\Big),$$

where the constant $C \in (0,\infty)$ depends only on A, f, B, α, β, σ and T. As in Case 1, an application of Gronwall's lemma proves the assertion in this second case.

Case 3. $\sigma \in (0,1)$ and $\beta + \alpha/2 \leq 0$. In this situation, we combine (66) with (67), (68) and (71) to get

$$\begin{aligned} C^{\mathrm{Lip}}_{\tau,j,k} &\leq 1 + C^{\mathrm{Lip}}_f \sum_{i=0}^{k-j-1} \frac{\tau}{(\tau(k-j-i))^\sigma} C^{\mathrm{Lip}}_{\tau,j,j+i} \\ &\quad + C^{\mathrm{Lip}}_B (\mathrm{Tr}(-A)^{-\alpha})^{1/2} (-\lambda_1)^{\beta+\alpha/2} \Big(\sum_{i=0}^{k-j-1} \tau (C^{\mathrm{Lip}}_{\tau,j,j+i})^2 \Big)^{1/2}. \end{aligned}$$

Setting

$$q_3 := \frac{1}{1-\sigma} + 2$$

and following the line of argumentation from the first case with q_3 instead of q we reach our goal also in this situation.

Case 4. $\sigma \leq 0$ and $\beta + \alpha/2 \in (0, 1/2)$. Combine (66), (67), (69) and (70) to get

$$\begin{aligned} C^{\mathrm{Lip}}_{\tau,j,k} &\leq 1 + C^{\mathrm{Lip}}_f (-\lambda_1)^\sigma \sum_{i=0}^{k-j-1} \tau C^{\mathrm{Lip}}_{\tau,j,j+i} \\ &\quad + C^{\mathrm{Lip}}_B (\mathrm{Tr}(-A)^{-\alpha})^{1/2} \left(\sum_{i=0}^{k-j-1} \frac{\tau}{(\tau(k-j-i))^{2\beta+\alpha}} (C^{\mathrm{Lip}}_{\tau,j,j+i})^2 \right)^{1/2}. \end{aligned}$$

Arguing as in the third case with

$$q_4 := \frac{1}{1/2 - (\beta + \alpha/2)} + 2$$

instead of q_3, we get the estimate we need to finish the proof. □

5 Spatial approximation by wavelet methods

For the inexact Rothe method in Section 2 we assumed, cf. Assumption 2.9, that we have a solver which enables us to compute the solution of the subproblem arising at the k-th time step and i-th stage up to a given tolerance $\varepsilon_{k,i}$. In the wavelet setting, this goal can be achieved by employing *adaptive* strategies. Indeed, as will be explained in Subsection 5.3, there exist adaptive strategies based on wavelets that are guaranteed to converge for a large range of problems. Moreover, they are *asymptotically optimal* in the sense that they asymptotically realize the same convergence order as best m-term wavelet approximation and the computational effort is proportional to the degrees m of freedom.

We start in Subsection 5.1 by introducing the wavelet setting as far as it is needed for our purposes. In Subsection 5.2, we combine estimates for optimal wavelet solvers with the complexity results for the inexact Rothe method as stated in Section 2. For deterministic equations of the form (77), we derive estimates on the degrees of freedom that are needed to guarantee that the inexact scheme converges with the same order as the exact scheme. In Subsection 5.3 we give an outline on how an optimal adaptive wavelet solver can be constructed in practice.

Throughout this section, $\mathcal{O} \subseteq \mathbb{R}^d$, $d \geq 1$, will denote a bounded Lipschitz domain.

5.1 Wavelet setting

Let us briefly recall the wavelet setting. In general, a *wavelet* basis $\Psi = \{\psi_\mu : \mu \in \mathcal{J}\}$ is a Riesz basis for $L_2(\mathcal{O})$. The indices $\mu \in \mathcal{J}$ typically encode several types of information, namely the *scale* (often denoted by $|\mu|$), the *spatial location* as well as the *type* of the wavelet. For instance, on the real line, μ can be identified with a pair of integers (j, k), where $j = |\mu|$ denotes the dyadic refinement level and $2^{-j}k$ the location of the wavelet.

We will ignore any explicit dependence on the type of the wavelet from now on, since this only produces additional constants. Hence, we frequently use $\mu = (j, k)$ and

$$\mathcal{J} = \big\{(j, k) \,:\, j \geq j_0, k \in \mathcal{J}_j\big\},$$

where $\mathcal{J}_j$ is some countable index set and $|(j, k)| = j$. Moreover,

$$\tilde{\Psi} = \{\tilde{\psi} \,:\, \mu \in \mathcal{J}\}$$

denotes the *dual wavelet basis*, which is biorthogonal to Ψ, i.e.,

$$\langle \psi_\mu, \tilde{\psi}_\nu \rangle = \delta_{\mu,\nu}, \quad \mu, \nu \in \mathcal{J}.$$

We will not discuss any technical description of the basis Ψ. Instead, we assume that the domain $\mathcal{O}$ enables us to construct a wavelet basis Ψ with the following properties:

(W1) The wavelets are *local* in the sense that there exist two constants $c_{\mathrm{loc}}, C_{\mathrm{loc}} > 0$, independent of $\mu \in \mathcal{J}$, such that

$$c_{\mathrm{loc}} 2^{-|\mu|} \leq \mathrm{diam}(\mathrm{supp}\,\psi_\mu) \leq C_{\mathrm{loc}} 2^{-|\mu|}, \quad \mu \in \mathcal{J}.$$

(W2) The wavelets satisfy the *cancellation property*, i.e., for some parameter $\tilde{m} \in \mathbb{N}$,

$$|\langle v, \psi_\mu \rangle_{L_2(\mathcal{O})}| \leq C_{\mathrm{can}} 2^{-|\mu|(\frac{d}{2}+\tilde{m})} |v|_{W^{\tilde{m}}(L_\infty(\operatorname{supp}\psi_\mu))}$$

for $|\mu| > j_0$, with a constant $C_{\mathrm{can}} > 0$, which does not depend on v and μ.

(W3) The wavelet basis induces characterizations of Besov spaces $B^s_q(L_p(\mathcal{O}))$, i.e., there exist constants $c_{\mathrm{norm}}, C_{\mathrm{norm}} > 0$, independent of v, such that

$$\begin{aligned} c_{\mathrm{norm}} \|v\|^q_{B^s_q(L_p(\mathcal{O}))} &\leq \sum_{j=j_0}^{\infty} 2^{j(s+d(\frac{1}{2}-\frac{1}{p}))q} \Big(\sum_{\mu \in \mathcal{J}_j} |\langle v, \tilde{\psi}_\mu \rangle_{L_2(\mathcal{O})}|^p \Big)^{q/p} \\ &\leq C_{\mathrm{norm}} \|v\|^q_{B^s_q(L_p(\mathcal{O}))} \end{aligned} \tag{76}$$

holds for $0 < p, q < \infty$ and all s with $s_1 > s > d(1/p-1)_+$ for some parameter s_1.

In (W3) the upper bound s_1 depends, in particular, on the smoothness and the approximation properties of the wavelet basis.

From now on we will make the following

Assumption 5.1. There exists a biorthogonal wavelet basis for $L_2(\mathcal{O})$ that satisfies the properties (W1), (W2), (W3) for a sufficiently large range of parameters s_1, s, p, q and $\tilde{m}$.

Remark 5.2. (i) The norm equivalence (76) and the fact that $B^s_2(L_2(\mathcal{O})) = H^s(\mathcal{O})$ imply that a simple rescaling immediately yields a Riesz basis for $H^s(\mathcal{O})$ with $0 < s < s_1$. We will also assume that Dirichlet boundary conditions can be included, so that a characterization of the type (76) also holds for $H^s_0(\mathcal{O})$.

(ii) Suitable constructions of wavelets on domains can be found, e.g., in [9, 23–25]. For a detailed discussion we refer to [10].

5.2 Complexity estimates for a wavelet Rothe method

In this subsection we study Rothe schemes based on wavelets. To keep the technical difficulties at a reasonable level, we restrict ourselves to deterministic parabolic evolution equations of the form

$$u'(t) = A(t)u(t) + f(t, u(t)), \quad t \in (0,T], \quad u(0) = u_0. \tag{77}$$

Throughout, $A : (0,T] \times V \to V^*$ and $f : (0,T] \times U \to U$, where (V, U, V^*) is a Gelfand triple, with $V = H^\nu_0(\mathcal{O})$, $U = L_2(\mathcal{O})$, and $V^* = H^{-\nu}(\mathcal{O})$, $\nu > 0$. In Subsection 5.2.1 we combine the abstract analysis presented in Section 2 with complexity estimates for adaptive wavelet solvers. To this end, among other things, regularity estimates for the exact solution in specific Besov spaces are essential. Then, in Subsection 5.2.2, the analysis is further substantiated for the concrete case of the heat equation. In particular, we show that in the wavelet setting a concrete estimate of the overall complexity can be derived.

5.2.1 Complexity estimates using adaptive wavelet solvers

In this subsection, we derive estimates for the degrees of freedom that are needed to guarantee that the inexact wavelet scheme converges with the same order as the exact scheme. We split our analysis into two parts. In the first part, we concentrate on the (rather theoretical) case, where the solutions of the stage equations are approximated by using best m-term wavelet approximation. As we will see, this task is closely related to smoothness estimates of the exact solutions in specific scales of Besov spaces. Unfortunately, best m-term approximation is not implementable in our case, since the solutions to the elliptic sub-problems are not known explicitly and therefore the m largest wavelet coefficients cannot be extracted directly. Therefore, in the second part, we turn our attention to the case, where the stage equations are solved numerically by using an implementable wavelet solver which is asymptotically optimal, i.e., it realizes the same convergence order as best m-term wavelet approximation. In Theorem 5.13 we show that the complexity estimates derived in the first part immediately extend to this case.

We are in the setting of Section 2 with $\mathcal{H} = H^t(\mathcal{O})$ for some smoothness parameter $t \geq 0$. In particular, we assume that an exact scheme (6) is given which satisfies Assumption 2.7 and 2.13.

Now, as the first step, we apply best m-term wavelet approximation in order to derive an approximation scheme that fulfills Assumption 2.9. We refer to the survey article [27] for a detailed discussion of best m-term wavelet approximation and other nonlinear approximation schemes. The error of best m-term wavelet approximation in $H^t(\mathcal{O})$ is defined as

$$\sigma_{m,t}(v) := \inf \Big\{ \big\| v - \sum_{\mu \in \Lambda} c_\mu \psi_\mu \big\|_{H^t(\mathcal{O})} \; : \; c_\mu \in \mathbb{R}, \Lambda \subset \mathcal{J}, \#\Lambda = m \Big\}.$$

The following theorem can be derived from [27, Section 7.7], see also [17, Chapter 7].

Theorem 5.3. *Let $t \geq 0$ and $v \in B^s_q(L_q(\mathcal{O}))$, where*

$$\frac{1}{q} = \frac{s-t}{d} + \frac{1}{2}, \quad s > t. \tag{78}$$

Furthermore, let Assumption 5.1 hold with $s_1 > s$. Then the error of best m-term wavelet approximation in $H^t(\mathcal{O})$ can be estimated as follows:

$$\sigma_{m,t}(v) \leq C_{\mathrm{nlin}} \|v\|_{B^s_q(L_q(\mathcal{O}))} m^{-\frac{s-t}{d}}, \tag{79}$$

with a constant $C_{\mathrm{nlin}} > 0$, which does not depend on v or m.

Remark 5.4. **(i)** In the case that Ψ is an orthnormal wavelet basis a best m-term approximation to a function v can be derived by selecting the m biggest wavelet coefficients in the wavelet expansion of v. In the biorthogonal case, choosing the m biggest coefficients yields a best m-term approximation up to a constant. In this sense best m-term approximation is an approximation scheme that fulfills Assumption 2.9.

(ii) The scale of Besov spaces $B^s_q(L_q(\mathcal{O}))$, where the parameters (s, q) are linked by (78) for a $t \geq 0$ is sometimes called the *nonlinear approximation line for approximation in* $H^t(\mathcal{O})$. In Fig. 1 it is displayed in a DeVore-Triebel diagram for the case $t = 0$, i.e., $H^t(\mathcal{O}) = L_2(\mathcal{O})$ and for the case $t = 1$.

Now, consider the inexact scheme (13) based on best m-term approximation in each stage. We can apply Theorem 2.22 and derive an estimate for the degrees of freedom needed to compute a solution up to a tolerance $(C_{\text{exact}} + T)\, \tau^\delta$.

Lemma 5.5. *Suppose that Assumptions 2.5, 2.7, and 2.11 hold. Let Assumption 2.13 hold for some $\delta > 0$. Further, suppose that Assumptions 5.1 holds, and let the inexact scheme* (13) *be based on best m-term approximation with the tolerances given by*

$$\varepsilon_{k,i} := \Big(S\, C''_{\tau,k} C'_{\tau,k,(i)} \Big)^{-1} \tau^{1+\delta}. \tag{80}$$

Let the exact solutions $\hat{w}_{k,i}$ of the stage equations in (13)*, be given by* (26)*, and assume that all $\hat{w}_{k,i}$ are contained in the same Besov space $B^s_q(L_q(\mathcal{O}))$ with $1/q = (s-t)/d + 1/2$. Then we have* (24)*, i.e.,*

$$\|u(T) - \tilde{u}_K\|_{\mathcal{H}} \leq (C_{\text{exact}} + T)\, \tau^\delta,$$

and the number of the degrees of freedom $M_{\tau,T}(\delta)$, given by (25)*, that are needed for the computation of $\{\tilde{u}_k\}_{k=0}^K$ is bounded from above by*

$$M_{\tau,T}(\delta) \leq \sum_{k=0}^{K-1} \sum_{i=1}^{S} \left\lceil C_{\text{nlin}}^{\frac{d}{s-t}} \, \|\hat{w}_{k,i}\|^{\frac{d}{s-t}}_{B^s_q(L_q(\mathcal{O}))} \Big(\Big(S\, C''_{\tau,k} C'_{\tau,k,(i)} \Big)^{-1} \tau^{1+\delta} \Big)^{-\frac{d}{s-t}} \right\rceil,$$

with C_{nlin} as in (79)*, and where $\lceil \cdot \rceil$ denotes the upper Gauss-bracket.*

Proof. We are in the setting of Theorem 2.22. By Theorem 5.3 we may, for each state equation, choose $m \in \mathbb{N}_0$ as the smallest possible integer, such that

$$\sigma_{m,t}(\hat{w}_{k,i}) \leq C_{\text{nlin}} \|\hat{w}_{k,i}\|_{B^s_q(L_q(\mathcal{O}))} m^{-\frac{s-t}{d}} \leq \varepsilon_{k,i},$$

holds, that is

$$m = \left\lceil \big(C_{\text{nlin}} \|\hat{w}_{k,i}\|_{B^s_q(L_q(\mathcal{O}))} \big)^{\frac{d}{s-t}} \varepsilon_{k,i}^{-\frac{d}{s-t}} \right\rceil.$$

Using (80) and summing over k and i completes the proof. □

Lemma 5.5 shows that we need estimates for the Besov norms of the exact solutions of the stage equations. We can provide an estimate in the following setting.

Lemma 5.6. *Let Assumption 2.7 hold with $\mathcal{H} = H^t(\mathcal{O})$ and define $C'_{\tau,j,(i)}$ as in* (11)*. Let $L_{\tau,i}^{-1} \in \mathcal{L}(L_2(\mathcal{O}), B_q^s(L_q(\mathcal{O})))$ with $1/q = (s-t)/d + 1/2$, $i = 1, \dots, S$, and assume that the operators $R_{\tau,k,i} : L_2(\mathcal{O}) \times \ldots \times L_2(\mathcal{O}) \to L_2(\mathcal{O})$ are Lipschitz continuous with Lipschitz constants $C^{\mathrm{Lip},R}_{\tau,k,(i)}$ for all $k = 0, \dots, K-1$, $i = 1, \dots S$. With $C'_{\tau,j,(i)}$ as in* (11) *we define*

$$
\begin{aligned}
C^{\mathrm{Bes}}_{k,i} := \Big(& \prod_{l=1}^{i-1} \big(1+\max\{C^{\mathrm{Lip}}_{\tau,k,(l)}, \|L_{\tau,l}^{-1} R_{\tau,k,l}(0,\dots,0)\|_{L_2(\mathcal{O})}\}\big)\big(1+\|u_k\|_{L_2(\mathcal{O})}\big) \\
& + \prod_{l=1}^{i-1} \big(1 + C^{\mathrm{Lip}}_{\tau,k,(l)}\big) \sum_{j=0}^{k-1} \Big(\prod_{n=j+1}^{k-1} \big(C'_{\tau,n,(0)} - 1\big)\Big) \sum_{r=1}^{S} C'_{\tau,j,(r)} \varepsilon_{j,r} \\
& + \sum_{j=1}^{i-1} \varepsilon_{k,j} \prod_{l=j+1}^{i-1} \big(1 + C^{\mathrm{Lip}}_{\tau,k,(l)}\big)\Big).
\end{aligned}
\tag{81}
$$

Then all $\hat{w}_{k,i}$ as defined in (26) *are contained in $B_q^s(L_q(\mathcal{O}))$, $1/q = (s-t)/d + 1/2$, and their norms can be estimated by*

$$
\begin{aligned}
\|\hat{w}_{k,i}\|_{B_q^s(L_q(\mathcal{O}))} \le \|L_{\tau,i}^{-1}\|_{\mathcal{L}(L_2(\mathcal{O}), B_q^s(L_q(\mathcal{O}))} & \\
\times \max\{C^{\mathrm{Lip},R}_{\tau,k,(i)}, \|R_{\tau,k,i}(0,\dots,0)\|_{L_2(\mathcal{O})}\}\, C^{\mathrm{Bes}}_{k,i}&.
\end{aligned}
$$

Proof. The proof is similar to the proof of Theorem 2.20. It can be found in Appendix A.2. □

Remark 5.7. The assumption $L_{\tau,i}^{-1} \in \mathcal{L}(L_2(\mathcal{O}), B_q^s(L_q(\mathcal{O})))$ with $1/q = (s-t)/d+1/2$, and the Lipschitz continuity of $R_{\tau,k,i} : L_2(\mathcal{O}) \times \ldots \times L_2(\mathcal{O}) \to L_2(\mathcal{O})$ imply Assumption 2.7 with $\mathcal{H} = H^t(\mathcal{O})$. However, this Lipschitz constant may not be optimal.

Now, the combination of Lemma 5.5 and 5.6 yields the first main result, i.e., the complexity estimate for the (theoretical) case that best m-term approximations are used for the solution of the S-stage equations.

Theorem 5.8. *Let the assumptions of the Lemmas 5.5 and 5.6 be satisfied and $C''_{\tau,k}$ be as in* (22)*. Then*

$$
\begin{aligned}
& M_{\tau,T}(\delta) \\
& \le \sum_{k=0}^{K-1} \sum_{i=1}^{S} \Big\lceil C_{\mathrm{nlin}}^{\frac{d}{s-t}} \Big(\max\{C^{\mathrm{Lip},R}_{\tau,k,(i)}, \|R_{\tau,k,i}(0,\dots,0)\|_{L_2(\mathcal{O})}\}\, C^{\mathrm{Bes}}_{k,i}\Big)^{\frac{d}{s-t}} \\
& \qquad \times \big(\|L_{\tau,i}^{-1}\|_{\mathcal{L}(L_2(\mathcal{O}), B_q^s(L_q(\mathcal{O})))}\big)^{\frac{d}{s-t}} \Big(\Big(S\, C''_{\tau,k} C'_{\tau,k,(i)}\Big)^{-1} \tau^{1+\delta}\Big)^{-\frac{d}{s-t}} \Big\rceil.
\end{aligned}
\tag{82}
$$

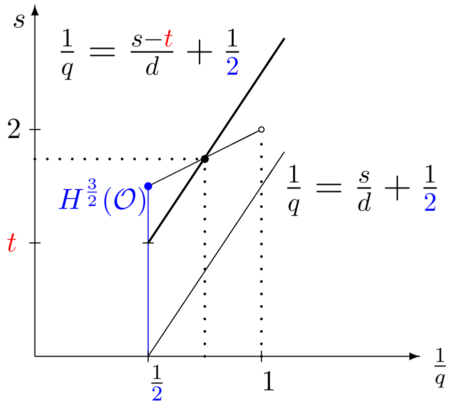

Figure 1: DeVore-Triebel diagram, $d = 3$

As outlined above, the next step is to discuss the complexity of Rothe's method in the case that *implementable* numerical wavelet schemes instead of the (non implementable) best m-term approximation are employed for the stage equations. Therefore, we make the following assumptions.

Assumption 5.9. The operators $R_{\tau,k,i}$ can be evaluated exactly.

Remark 5.10. A detailed study of the case where Assumption 5.9 is not satisfied, will be presented in a forthcoming paper, see also Remark 2.10(iii).

Assumption 5.11. There exists an implementable asymptotically optimal numerical wavelet scheme for the elliptic stage equations arising in (13). That is, if the best m-term approximation in $H^t(\mathcal{O})$ converges with the rate $m^{-(s-t)/d}$, for some $s > t > 0$, then the scheme computes finite index sets $\Lambda_l \subset \mathcal{J}$ and coefficients $(c_\mu)_{\mu\in\Lambda_l}$ with

$$\big\|L_{\tau,i}^{-1}v - \sum_{\mu\in\Lambda_l} c_\mu\psi_\mu\big\|_{H^t(\mathcal{O})} \leq C^{\mathrm{asym}}_{\tau,i,s,t}(L_{\tau,i}^{-1}v)\,(\#\Lambda_l)^{-\frac{s-t}{d}} \tag{83}$$

for some constant $C^{\mathrm{asym}}_{\tau,i,s,t}(L_{\tau,i}^{-1}v)$. Further, for all $\varepsilon > 0$ there exists an $l(\varepsilon)$ such that

$$\big\|L_{\tau,i}^{-1}v - \sum_{\mu\in\Lambda_l} c_\mu\psi_\mu\big\|_{H^t(\mathcal{O})} \leq \varepsilon, \quad l \geq l(\varepsilon),$$

and such that

$$\#\Lambda_{l(\varepsilon)} \leq C^{\mathrm{asym}}_{\tau,i,s,t}(L_{\tau,i}^{-1}v)\,\varepsilon^{-\frac{d}{s-t}}.$$

As a prototype of an adaptive wavelet method we will discuss in Subsection 5.3 a scheme that has been derived in [11]. It satisfies an optimality estimate of the form (83) for the energy norm (94). However, since the energy norm is equivalent to some Sobolev norm $\|\cdot\|_{H^t}$, cf. (95), the estimate (83) also holds for this case. Moreover, it has been shown in [11] that the constant is of a specific form, which is similar to (79). Therefore, we specify Assumption 5.11 in the following way.

Assumption 5.12. The constant $C^{\mathrm{asym}}_{\tau,i,s,t}(L^{-1}_{\tau,i}v)$ in (83) is of the form

$$C^{\mathrm{asym}}_{\tau,i,s,t}(L^{-1}_{\tau,i}v) = \hat{C}^{\mathrm{asym}}_{\tau,i} \|L^{-1}_{\tau,i}v\|_{B^s_q(L_q(\mathcal{O}))}, \quad \frac{1}{q} = \frac{s-t}{d} + \frac{1}{2},$$

with a constant $\hat{C}^{\mathrm{asym}}_{\tau,i}$ independent of $L^{-1}_{\tau,i}v$.

In this setting we are immediately able to state our main result.

Theorem 5.13. *Let the assumptions of the Lemmas 5.5 and 5.6 be satisfied. If an optimal numerical wavelet scheme, that satisfies Assumption 5.12, is used for the numerical solution of the stage equations, then the necessary degrees of freedom can be estimated as in Theorem 5.8 with $\hat{C}^{\mathrm{asym}}_{\tau,i}$ instead of C_{nlin}, i.e.,*

$$\begin{aligned} &M_{\tau,T}(\delta) \\ &\leq \sum_{k=0}^{K-1} \sum_{i=1}^{S} \Big\lceil \big(\hat{C}^{\mathrm{asym}}_{\tau,i}\big)^{\frac{d}{s-t}} \Big(\max\{C^{\mathrm{Lip},R}_{\tau,k,(i)}, \|R_{\tau,k,i}(0,\ldots,0)\|_{L_2(\mathcal{O})}\}\, C^{\mathrm{Bes}}_{k,i} \Big)^{\frac{d}{s-t}} \\ &\times \big(\|L^{-1}_{\tau,i}\|_{\mathcal{L}(L_2(\mathcal{O}),B^s_q(L_q(\mathcal{O})))}\big)^{\frac{d}{s-t}} \Big(\Big(S\, C''_{\tau,k} C'_{\tau,k,(i)} \Big)^{-1} \tau^{1+\delta} \Big)^{-\frac{d}{s-t}} \Big\rceil. \end{aligned} \tag{84}$$

Remark 5.14. The constant $\hat{C}^{\mathrm{asym}}_{\tau,i}$ depends on the concrete design of the adaptive method at hand. As an example this constant may depend on the design of the routines **APPLY**, **RHS** and **COARSE**, as presented in Subsection 5.3. Moreover the value of $\hat{C}^{\mathrm{asym}}_{\tau,i}$ depends on the equivalence constants of the energy norm and the Sobolev norm in (95). Therefore this constant may grow as τ gets small. However, the reader should observe that this is an intrinsic problem and not caused by our approach.

Now the question arises if and how the Besov norms of the exact solutions of the stage equations $\hat{w}_{k,i}$, cf. (26), and moreover all the constants involved in (82) and (84) can be estimated. In the next subsection we will present a detailed study for the most important model problem, that is the linearly-implicit Euler scheme applied to the heat equation.

5.2.2 Complexity estimates for the heat equation

In this subsection, we conclude the discussion of Example 2.24. We derive a concrete estimate of the overall complexity for the heat equation in the wavelet setting. To this end, we apply the linearly-implicit Euler scheme to the heat equation

$$\begin{aligned} u'(t) &= \Delta^D_{\mathcal{O}} u(t) + f(t,u(t)) && \text{on } \mathcal{O},\ t \in (0,T], \\ u(0) &= u_0 && \text{on } \mathcal{O}, \\ u &= 0 && \text{on } \partial\mathcal{O},\ t \in (0,T], \end{aligned}$$

on a bounded Lipschitz domain $\mathcal{O} \subset \mathbb{R}^d$, and consider the case $\mathcal{H} = \mathcal{G} = U = L_2(\mathcal{O})$. The operators $L_{\tau,1}$, $R_{\tau,k,1}$ are given by

$$L_{\tau,1} = (I - \tau \Delta_{\mathcal{O}}^D), \quad R_{\tau,k,1} = (I + \tau f(t_k, \cdot)). \tag{85}$$

The first step is to estimate the Besov regularity of the solutions to the stage equations. To this end, the mapping properties of $L_{\tau,i}^{-1}$ with respect to the adaptivity scale of Besov spaces have to be evaluated, see Lemma 5.16. Recall that for special cases bounds for the Lipschitz constant of $R_{\tau,k,1} : L_2(\mathcal{O}) \to L_2(\mathcal{O})$ have already been proven in the in the Examples 2.24, 2.25. In Example 2.24 we assumed $f : (0, T] \times L_2(\mathcal{O}) \to L_2(\mathcal{O})$ to be defined by a continuously differentiable function $\bar{f} : (0, T] \times \mathbb{R} \to \mathbb{R}$ that is bounded and Lipschitz continuous in the second argument, uniformly in t. This implies that f is globally Lipschitz continuous in u with a Lipschitz constant $C^{\,\mathrm{Lip},f}$ that is independent of t. Furthermore, it has been shown that in this setting $R_{\tau,k,i}$ is also Lipschitz continuous with constant

$$C_{\tau,k,(i)}^{\mathrm{Lip},R} \le 1 + \tau C^{\,\mathrm{Lip},f}.$$

In Example 2.25 we considered the special case $f(t, v) = \bar{f}(v)$ with a continuously differentiable function $\bar{f} : \mathbb{R} \to \mathbb{R}$. There it has been shown that if the derivative of $\bar{f}$ is bounded and strictly negative, i.e., there exists a constant $\bar{B} > 0$, so that

$$-\bar{B} < \frac{\mathrm{d}}{\mathrm{d}x}\bar{f}(x) < 0 \text{ for all } x \in \mathbb{R}$$

it is possible to estimate

$$C_{\tau,k,(i)}^{\mathrm{Lip},R} \le \sup_{x \in \mathbb{R}} \left| 1 + \tau \frac{\mathrm{d}}{\mathrm{d}x}\bar{f}(x) \right| < 1,$$

if $\tau < 2/\bar{B}$, i.e., $R_{\tau,k,1}$ is a contraction if τ is small enough.

The following result is the basis for our regularity estimates for $L_{\tau,i}^{-1}$. We put

$$C_{\mathrm{Bes},\varepsilon}^{\mathrm{Lap}} := \|(\Delta_{\mathcal{O}}^D)^{-1}\|_{\mathcal{L}(L_2(\mathcal{O}), B_1^{2-\varepsilon}(L_1(\mathcal{O}))}$$

and

$$C_{\mathrm{Sob}}^{\mathrm{Lap}} := \|(\Delta_{\mathcal{O}}^D)^{-1}\|_{\mathcal{L}(L_2(\mathcal{O}), H^{3/2}(\mathcal{O}))}.$$

Lemma 5.15. *Let $\varepsilon > 0$. Then the operator $(I - \tau\Delta_{\mathcal{O}}^D)^{-1}$ is contained in $\mathcal{L}(L_2(\mathcal{O}), B_1^{2-\varepsilon}(L_1(\mathcal{O})))$ and in $\mathcal{L}(L_2(\mathcal{O}), H^{3/2}(\mathcal{O}))$. The respective operator norms can be estimated by*

$$\|(I - \tau\Delta_{\mathcal{O}}^D)^{-1}\|_{\mathcal{L}(L_2(\mathcal{O}), B_1^{2-\varepsilon}(L_1(\mathcal{O})))} \le \frac{1}{\tau} C_{\mathrm{Bes},\varepsilon}^{\mathrm{Lap}} \tag{86}$$

and

$$\|(I - \tau\Delta_{\mathcal{O}}^D)^{-1}\|_{\mathcal{L}(L_2(\mathcal{O}), H^{3/2}(\mathcal{O}))} \le \frac{1}{\tau} C_{\mathrm{Sob}}^{\mathrm{Lap}}, \tag{87}$$

respectively.

Proof. We start by proving (86). The observation

$$(I-\tau\Delta_{\mathcal{O}}^D)^{-1} = (-\tau\Delta_{\mathcal{O}}^D)^{-1}(I-(I-\tau\Delta_{\mathcal{O}}^D)^{-1})$$

leads to

$$\begin{aligned}&\|(I-\tau\Delta_{\mathcal{O}}^D)^{-1}\|_{\mathcal{L}(L_2(\mathcal{O}),B_1^{2-\varepsilon}(L_1(\mathcal{O})))}\\ &\quad\leq \tau^{-1}\|(\Delta_{\mathcal{O}}^D)^{-1}\|_{\mathcal{L}(L_2(\mathcal{O}),B_1^{2-\varepsilon}(L_1(\mathcal{O}))}\|(I-(I-\tau\Delta_{\mathcal{O}}^D)^{-1})\|_{\mathcal{L}(L_2(\mathcal{O}))}.\end{aligned}$$

It has been shown in [19] that $(\Delta_{\mathcal{O}}^D)^{-1}\in\mathcal{L}(L_2(\mathcal{O}),B_1^{2-\varepsilon}(L_1(\mathcal{O}))$, see also [22, Corollary 1] for details. In order to estimate the second term, we proceed as in Lemma 4.13, i.e., we have

$$\begin{aligned}\|(I-(I-\tau\Delta_{\mathcal{O}}^D)^{-1})\|^2_{\mathcal{L}(L_2(\mathcal{O}))} &= \sup_{\|v\|_{L_2(\mathcal{O})}=1}\sum_{k\in\mathbb{N}}|(1-(1-\tau\lambda_k)^{-1})\langle v,e_k\rangle_{L_2(\mathcal{O})}|^2\\ &= \sup_{\|v\|_{L_2(\mathcal{O})}=1}\sum_{k\in\mathbb{N}}\left|\frac{-\tau\lambda_k}{1-\tau\lambda_k}\langle v,e_k\rangle_{L_2(\mathcal{O})}\right|^2\\ &\leq \sup_{\|v\|_{L_2(\mathcal{O})}=1}\sum_{k\in\mathbb{N}}\left|\langle v,e_k\rangle_{L_2(\mathcal{O})}\right|^2\\ &=1\end{aligned}$$

The estimate (87) follows in a similar fashion by using the fundamental result $(\Delta_{\mathcal{O}}^D)^{-1}\in\mathcal{L}(L_2(\mathcal{O}),H^{3/2}(\mathcal{O}))$ from [37, Theorem B]. □

We are now ready to prove the regularity estimate for $L_{\tau,i}^{-1}$ in $B_q^s(L_q(\mathcal{O}))$, where $1/q=(s-t)/d+1/2$. We put

$$C_{\text{inter}}^{\text{Lap}}(\theta) := (C_{\text{Sob}}^{\text{Lap}})^{1-\theta}(C_{\text{Bes},\varepsilon}^{\text{Lap}})^{\theta},\quad \theta\in(0,1).$$

Lemma 5.16. *Let the assumptions of Lemma 5.15 hold. Let $d\geq 2$ and $t\geq 0$ be such that*

$$\theta := \frac{3-2t}{d-1+2\varepsilon}\in(0,1). \tag{88}$$

Then

$$(I-\tau\Delta_{\mathcal{O}}^D)^{-1}\in\mathcal{L}(L_2(\mathcal{O}),B_q^s(L_q(\mathcal{O}))),\quad \textit{with } s=\frac{3d-2t+4\varepsilon t}{2d-2+4\varepsilon},$$

where $1/q=(s-t)/d+1/2$. Its norm can be bounded in the following way

$$\|(I-\tau\Delta_{\mathcal{O}}^D)^{-1}\|_{\mathcal{L}(L_2(\mathcal{O}),B_q^s(L_q(\mathcal{O})))}\leq\frac{1}{\tau}C_{\text{inter}}^{\text{Lap}}(\theta). \tag{89}$$

Proof. The proof is based on interpolation properties of Besov spaces, see [5] for details. For real interpolation it holds that

$$\big(B_{p_0}^{s_0}(L_{p_0}(\mathcal{O})),B_{p_1}^{s_1}(L_{p_1}(\mathcal{O}))\big)_{\bar\theta,p} = B_p^s(L_p(\mathcal{O}))$$

in the sense of equivalent (quasi) norms, provided that the parameters satisfy

$$0<\bar{\theta}<1, \quad s=(1-\bar{\theta})s_0+\bar{\theta}s_1, \quad \frac{1}{p}=\frac{1-\bar{\theta}}{p_0}+\frac{\bar{\theta}}{p_1} \tag{90}$$

and $s_0, s_1 \in \mathbb{R}$, $0<p_0,p_1<\infty$. Furthermore, if (90) holds, a linear operator $\mathcal{T}$ that is contained in $\mathcal{L}(L_2(\mathcal{O}), B^{s_0}_{p_0}(L_{p_0}(\mathcal{O})))$ and $\mathcal{L}(L_2(\mathcal{O}), B^{s_1}_{p_1}(L_{p_1}(\mathcal{O})))$ is also an element of $\mathcal{L}(L_2(\mathcal{O}), B^{s}_{p}(L_{p}(\mathcal{O})))$. Its norm can be estimated by

$$\|\mathcal{T}\|_{\mathcal{L}(L_2(\mathcal{O}),B^s_p(L_p(\mathcal{O})))} \leq \|\mathcal{T}\|^{1-\bar{\theta}}_{\mathcal{L}(L_2(\mathcal{O}),B^{s_0}_{p_0}(L_{p_0}(\mathcal{O})))} \|\mathcal{T}\|^{\bar{\theta}}_{\mathcal{L}(L_2(\mathcal{O}),B^{s_1}_{p_1}(L_{p_1}(\mathcal{O})))}.$$

Observe that $H^{3/2}(\mathcal{O}) = B^{3/2}_2(L_2(\mathcal{O}))$ and that we can apply Lemma 5.15. We need to determine the value for $\bar{\theta}$, such that the resulting interpolation space lies on the nonlinear approximation line $1/p=(s-t)/d-1/2$. This is the case for $\bar{\theta}=(3-2t)/(d-1+2\varepsilon)$, cf. Fig. 1. □

Remark 5.17. Our findings for the discretization of the heat equation by means of the linearly-implicit Euler scheme carry over to discretizations with $S>1$ stages. For the case $S=2$ the operators $L_{\tau,i}$, $R_{\tau,k,i}$, $i=1,2$, are provided by Lemma 3.11 and similar to (85), e.g.,

$$L_{\tau,i} = \ (I-\tau\gamma_{i,i}\Delta^D_{\mathcal{O}}), \quad i=1,2.$$

Lemma 5.16 can be reformulated with $\tau\gamma_{i,i}$ replacing τ, and the Lipschitz continuity of $R_{\tau,k,i}$ can be established directly as before.

We are now able to give specific bounds for the degrees of freedom needed to compute the solution of the heat equation by means of the linearly-implicit Euler scheme. In this setting, we can apply Theorem 2.22 to the case when best m-term approximation is used in each step of the inexact scheme (13), and the following estimate holds.

Theorem 5.18. *Let the assumptions of the Lemmas 5.5, 5.6 and 5.16 hold. Further let τ be small enough such that*

$$(1+\tau C^{\mathrm{Lip},f})^{-1}\tau\|f(0)\|_{L_2(\mathcal{O})} \leq 1,$$

and let C_{nlin}, θ, and $C^{\mathrm{Lap}}_{\mathrm{inter}}$ be given by (79), (88), *and* (89), *respectively. Set $C_{\mathrm{bound}}(u) := \sup_{t\in[0,T]}\|u(t)\|_{L_2(\mathcal{O})}$ and $C_{\mathrm{short}}(\tau) := C_{\mathrm{nlin}}C^{\mathrm{Lap}}_{\mathrm{inter}}(1+\tau C^{\mathrm{Lip},f})$. Let C_{exact} be given as in Assumption 2.13. In the setting of Example 2.24, if best m-term wavelet approximation for the spatial approximation of the stage equations is applied, then the degrees of freedom needed to compute a solution up to a tolerance $(C_{\mathrm{exact}}+T)\,\tau$ can be estimated from above by*

$$M_{\tau,T} \leq T\tau^{-1} + \frac{1}{2}\Big(2C_{\mathrm{short}}(\tau)\Big)^{\frac{2}{\theta}} \left(T^{\frac{2}{\theta}+1}\tau^{-(\frac{2}{\theta}+1)} + C_{\mathrm{lim}}(\tau)\tau^{-(\frac{6}{\theta}+1)}\right),$$

with

$$C_{\lim}(\tau) := \Big(\big(1 + C_{\text{bound}}(u) + C_{\text{exact}}\tau\big)\Big)^{\frac{2}{\theta}} \Big(\tau \frac{\big(1+\tau C^{\text{Lip},f}\big)^{\frac{2}{\theta}T\tau^{-1}} - 1}{1 - \big(1+\tau C^{\text{Lip},f}\big)^{-\frac{2}{\theta}}}\Big),$$

where it holds that

$$\lim_{\tau\to 0} C_{\lim}(\tau) = \Big(\big(1 + C_{\text{bound}}(u)\big)\Big)^{\frac{2}{\theta}} \frac{\theta}{2}(C^{\text{Lip},f})^{-1}\Big(\exp\big(C^{\text{Lip},f}\frac{2}{\theta}T\big) - 1\Big).$$

Proof. We apply Theorem 5.8 with $S = 1$ and $\delta = 1$. In the setting of Example 2.24 it holds that

$$C^{\text{Lip},R}_{\tau,k,(1)} = 1 + \tau C^{\text{Lip},f}, \quad C'_{\tau,k,(1)} = 1, \quad C'_{\tau,k,(0)} = 2 + \tau C^{\text{Lip},f},$$

independent of k. Thus (22) reads as $C''_{\tau,k} = (1 + \tau C^{\text{Lip},f})^{K-k-1}$ and (81) can be simplified to

$$C^{\text{Bes}}_{k,i} = 1 + \|u_k\|_{L_2(\mathcal{O})} + k(C^{\text{Lip},R}_{\tau,k,(1)})^{k-K}\tau^2.$$

The norm of $\|u_k\|_{L_2(\mathcal{O})}$ can be bounded as follows. By Assumption 2.13 we have $\|u(t_k) - u_k\|_{L_2(\mathcal{O})} \leq C_{\text{exact}}\tau$ and as a consequence

$$\|u_k\|_{L_2(\mathcal{O})} \leq \|u(t_k) - u_k\|_{L_2(\mathcal{O})} + \|u(t_k)\|_{L_2(\mathcal{O})} \leq C_{\text{exact}}\tau + C_{\text{bound}}(u).$$

where $C_{\text{bound}}(u)$ is finite since $[0,T]$ is compact and u is continuous. Using the bound (89) the estimate for the degrees of freedom (82) is given by

$$\begin{aligned} M_{\tau,T} &\leq \sum_{k=0}^{K-1} \bigg\lceil \Big(C_{\text{short}}\Big((C^{\text{Lip},R}_{\tau,k,(1)})^{K-k}\tau^{-3}(1 + C_{\text{bound}}(u) + C_{\text{exact}}\tau) + k\Big)\Big)^{\frac{2}{\theta}} \bigg\rceil \\ &\leq K + \sum_{k=0}^{K-1} \Big(C_{\text{short}}\Big((C^{\text{Lip},R}_{\tau,k,(1)})^{K-k}\tau^{-3}(1+C_{\text{bound}}(u)+C_{\text{exact}}\tau) + k\Big)\Big)^{\frac{2}{\theta}}. \end{aligned}$$

An application of Jensen's inequality and the geometric series formula yields

$$\begin{aligned} M_{\tau,T} &\leq K + C_{\text{short}}^{\frac{2}{\theta}} 2^{\frac{2}{\theta}-1} \\ &\quad \times \sum_{k=0}^{K-1} \bigg(\Big((C^{\text{Lip},R}_{\tau,k,(1)})^{K-k}\tau^{-3}(1 + C_{\text{bound}}(u) + C_{\text{exact}}\tau)\Big)^{\frac{2}{\theta}} + k^{\frac{2}{\theta}}\bigg) \\ &\leq K\Big(1 + \frac{1}{2}(2C_{\text{short}}K)^{\frac{2}{\theta}}\Big) \\ &\quad + \tau^{-\frac{6}{\theta}}\frac{1}{2}\Big(2C_{\text{short}}\big(1 + C_{\text{bound}}(u) + C_{\text{exact}}\tau\big)\Big)^{\frac{2}{\theta}} \frac{(1+\tau C^{\text{Lip},f})^{\frac{2}{\theta}K} - 1}{1 - (1+\tau C^{\text{Lip},f})^{-\frac{2}{\theta}}}. \end{aligned}$$

The proof is finalized by the insertion of $K = T\tau^{-1}$ and the observations

$$\lim_{\tau\to 0}(1+\tau C^{\mathrm{Lip},f})^{\frac{2}{\theta}T\tau^{-1}} - 1 = \exp\big(C^{\mathrm{Lip},f}\frac{2}{\theta}T\big) - 1,$$
$$\lim_{\tau\to 0}\frac{\tau}{1-(1+\tau C^{\mathrm{Lip},f})^{-\frac{2}{\theta}}} = \frac{1}{\frac{2}{\theta}C^{\mathrm{Lip},f}}. \qquad \square$$

Now, we want turn to the case when an optimal numerical wavelet scheme is used for the numerical solution of the stage equations in (13). As we shall see in Subsection 5.3 the implementable wavelet schemes we have in mind are optimal with respect to the energy norm (94). In our setting it is induced by L_τ and equivalent to the Sobolev norm $H^1(\mathcal{O})$. For this reason, we now state the estimate for the degrees of freedom for the case of the Sobolev norm $H^1(\mathcal{O})$, i.e., $t = 1$.

Theorem 5.19. *Let the assumptions of Theorem 5.18 hold, and let Ψ be a wavelet basis fulfilling Assumption 5.1. We employ an implementable asymptotically optimal numerical scheme, based on Ψ, such that Assumption 5.12 holds for $t = 1$. Using the abbreviation $\hat{C}_{\mathrm{short}}(\tau) := \hat{C}^{\mathrm{asym}}_{\tau,1}C^{\mathrm{Lap}}_{\mathrm{inter}}(1+\tau C^{\mathrm{Lip},f})$ the degrees of freedom needed can be estimated by*

$$M_{\tau,T} \le T\tau^{-1} + \frac{1}{2}\big(2\hat{C}_{\mathrm{short}}(\tau)\big)^{\frac{2}{\theta}}\Big(T^{\frac{2}{\theta}+1}\tau^{-(\frac{2}{\theta}+1)} + \hat{C}_{\mathrm{lim}}(\tau)\tau^{-(\frac{6}{\theta}+1)}\Big), \quad (91)$$

with

$$\hat{C}_{\mathrm{lim}}(\tau) := \Big(\big(1 + C_{\mathrm{bound}}(u) + C_{\mathrm{exact}}\tau\big)\Big)^{\frac{2}{\theta}}\Big(\tau\frac{\big(1+\tau C^{\mathrm{Lip},f}\big)^{\frac{2}{\theta}T\tau^{-1}} - 1}{1-\big(1+\tau C^{\mathrm{Lip},f}\big)^{-\frac{2}{\theta}}}\Big)$$

and

$$\theta := \frac{1}{d-1+2\varepsilon}. \quad (92)$$

It holds that

$$\lim_{\tau\to 0}\hat{C}_{\mathrm{lim}}(\tau) = \Big(\big(1 + C_{\mathrm{bound}}(u)\big)\Big)^{\frac{2}{\theta}}\frac{\theta}{2}(C^{\mathrm{Lip},f})^{-1}\Big(\exp\big(C^{\mathrm{Lip},f}\frac{2}{\theta}T\big) - 1\Big).$$

Remark 5.20. (i) The calculation above shows that, among other things, the overall complexity of the resulting scheme heavily depends on the Besov smoothness of the exact solutions to the stage equations. Due to the Lipschitz character of the domain $\mathcal{O}$, and since we are working in the L_2-setting, this Besov regularity is limited by $s = 2$. However, for more specific domains, i.e., polygonal domains in $\mathbb{R}^2$ and smoother right-hand sides, much higher Besov smoothness can be achieved, see, e.g., [16, 21] for details. Therefore, for polygonal domains and smoother source terms f we expect that also in our case higher Besov smoothness for the solutions of the stage equations arises, yielding a much lower overall complexity. The details will be discussed in a forthcoming paper.

(ii) Let us further discuss the asymptotic behavior of $M_{\tau,T}$ as τ tends to zero. For simplicity, let us consider the case $d = 2$, then we can choose θ arbitrary close to 1. Asymptotical optimal schemes are usually described in the energy norm induced by the operator $L_{\tau,1}$, whith a constant in the analog to (83) that is independent of $L_{\tau,1}$, see, e.g., [11]. With the notation as (95) the following consideration for the energy norm induced by $L_{\tau,1}$

$$\langle (I + \tau\Delta_{\mathcal{O}}^{D})u, u\rangle_{L_2(\mathcal{O})} \geq \langle u, u\rangle_{L_2(\mathcal{O})} + \tau\, c_{\mathrm{energy}}^2(\Delta_{\mathcal{O}}^{D})\, \|u\|_{H^1(\mathcal{O})}^2,$$

implies $c_{\mathrm{energy}}(I + \tau\Delta_{\mathcal{O}}^{D}) \geq \tau^{\frac{1}{2}} c_{\mathrm{energy}}(\Delta_{\mathcal{O}}^{D})$, so that we can conclude

$$\hat{C}_{\tau,1}^{\mathrm{asym}} = \hat{C}_1\, \tau^{-\frac{1}{2}}$$

with some constant $\hat{C}_1$ independent of τ. In this case (91) reads as

$$M_{\tau,T} \leq T\tau^{-1} + \frac{1}{2}\big(2\hat{C}_1 C_{\mathrm{inter}}^{\mathrm{Lap}}(1 + \tau C^{\mathrm{Lip},f})\big)^2 \left(T^3\tau^{-4} + \hat{C}_{\mathrm{lim}}(\tau)\tau^{-8+\varepsilon'}\right),$$

i.e., for small τ the last term is dominating and therefore the number of degrees of freedom behaves as $\tau^{-8+\varepsilon'}$.

5.3 Adaptive wavelet schemes for elliptic problems

We show how wavelets can be used for the adaptive numerical treatment of elliptic operator equations. To be specific, we are interested in equations of the form

$$\mathcal{A}u = f, \tag{93}$$

where we will assume $\mathcal{A}$ to be a boundedly invertible operator from some Hilbert space V into its normed dual V^*, i.e.,

$$c_{\mathrm{ell}}\|v\|_V \leq \|\mathcal{A}v\|_{V^*} \leq C_{\mathrm{ell}}\|v\|_V, \quad v \in V.$$

Consequently, we are again in a Gelfand triple setting (V, U, V^*). We will only discuss some basic ideas. For further information, the reader is referred to [11, 12, 18]. In our setting, that is the setting of the Rothe method, the operator $\mathcal{A}$ will be one of the operators $L_{\tau,i}$ that arise in the treatment of the stage equations. Therefore, in the applications we have in mind V will always be one of the Sobolev space $H^t(\mathcal{O})$ or $H_0^t(\mathcal{O})$.

We will focus on the special case where

$$a(v, w) := \langle \mathcal{A}v, w\rangle_{V^*\times V}$$

defines a continuous, symmetric and elliptic bilinear form on V in the sense of (97). Then, of course, $\mathcal{A}$ corresponds to the operator $-A$ in (98). In this setting the bilinear form induces a norm on V, the *energy norm*, by setting

$$\|\cdot\| := a(\cdot,\cdot)^{\frac{1}{2}}. \tag{94}$$

It is equivalent to the Sobolev norm, i.e.,

$$c_{\mathrm{energy}}\|\cdot\|_{H^t(\mathcal{O})} \le \|\cdot\| \le C_{\mathrm{energy}}\|\cdot\|_{H^t(\mathcal{O})}. \tag{95}$$

Usually, operator equations of the form (93) are solved by a Galerkin scheme, i.e., one defines an increasing sequence of finite dimensional approximation spaces $S_{\Lambda_l} := \mathrm{span}\{\eta_\mu : \mu \in \Lambda_l\}$, where $S_{\Lambda_l} \subset S_{\Lambda_{l+1}}$, and projects the problem onto these spaces, i.e.,

$$\langle \mathcal{A} u_{\Lambda_l}, v\rangle_{V^*\times V} = \langle f, v\rangle_{V^*\times V} \quad \text{for all} \quad v \in S_{\Lambda_l}.$$

To compute the current Galerkin approximation, one has to solve a linear system

$$\mathbf{G}_{\Lambda_l}\mathbf{c}_{\Lambda_l} = \mathbf{f}_{\Lambda_l},$$

with $\mathbf{G}_{\Lambda_l} := (\langle \mathcal{A}\eta_{\mu'}, \eta_\mu\rangle_{V^*\times V})_{\mu,\mu'\in\Lambda_l}$, $(\mathbf{f}_\Lambda)_\mu := \langle f, \eta_\mu\rangle_{V^*\times V},\ \mu \in \Lambda_l$.

It is a somewhat delicate task to choose the approximation spaces in the right way. Doing it in an arbitrary way might result in a very inefficient scheme. A natural idea is to use an *adaptive* scheme, i.e., an updating strategy which essentially consists of the following steps

solve	–	estimate	–	refine
$\mathbf{G}_{\Lambda_l}\mathbf{c}_{\Lambda_l} = \mathbf{f}_{\Lambda_l}$		$\lVert u - u_{\Lambda_l}\rVert = ?$ a posteriori error estimator		add functions if necessary.

The second step is highly nontrivial since the exact solution u is unknown, so that clever *a posteriori* error estimators are needed. An equally challenging task is to show that the refinement strategy leads to a convergent scheme and to estimate its order of convergence, if possible. In recent years, it has been shown that both tasks can be solved if wavelets are used as basis functions for the Galerkin scheme as we will now explain.

The first step is to transform (93) into a discrete problem. From the norm equivalences (76) it is easy to see that (93) is equivalent to

$$\mathbf{A}\mathbf{u} = \mathbf{f},$$

where $\mathbf{A} := \mathbf{D}^{-1}\langle \mathcal{A}\Psi, \Psi\rangle_{V^*\times V}^\top \mathbf{D}^{-1}$, $\mathbf{u} := \mathbf{D}\mathbf{c}$, $\mathbf{f} := \mathbf{D}^{-1}\langle f, \Psi\rangle_{V^*\times V}^\top$, and $\mathbf{D} := (2^{-s|\mu|}\,\delta_{\mu,\mu'})_{\mu,\mu'\in\mathcal{J}}$. Computing a Galerkin approximation amounts to solving the system

$$\mathbf{A}_\Lambda \mathbf{u}_\Lambda = \mathbf{f}_\Lambda := \mathbf{f}|_\Lambda, \qquad \mathbf{A}_\Lambda := (2^{-s(|\mu|+|\nu|)}\langle \psi_\mu, \mathcal{A}\psi_\nu\rangle_{V^*\times V})_{\mu,\nu\in\Lambda}.$$

Now, ellipticity and the norm equivalences (76) yield

$$\begin{aligned}\|\mathbf{u} - \mathbf{u}_\Lambda\|_{\ell_2(\mathcal{J})} &\le c_{\mathrm{dis}}\|\mathbf{A}(\mathbf{u} - \mathbf{u}_\Lambda)\|_{\ell_2(\mathcal{J})}\\ &\le C_{\mathrm{dis}}\|\mathbf{f} - \mathbf{A}(\mathbf{u}_\Lambda)\|_{\ell_2(\mathcal{J})}\\ &= C_{\mathrm{dis}}\|\mathbf{r}_\Lambda\|_{\ell_2(\mathcal{J})},\end{aligned}$$

so that the $\ell_2(\mathcal{J})$-norm of the *residual* $\mathbf{r}_\Lambda$ serves as an *a posteriori* error estimator. Each individual coefficient $(\mathbf{r}_\Lambda)_\mu$ can be viewed as a local error indicator. Therefore a natural adaptive strategy would consist in catching the bulk of the residual, i.e., to choose the new index set $\hat{\Lambda}$ such that

$$\|\mathbf{r}_\Lambda|_{\hat{\Lambda}}\|_{\ell_2(\mathcal{J})} \geq \zeta \|\mathbf{r}_\Lambda\|_{\ell_2(\mathcal{J})}, \quad \text{for some} \quad \zeta \in (0,1).$$

However, such a scheme cannot be implemented since the residual involves infinitely many coefficients. To transform this idea into an implementable scheme, the following three subroutines can be utilized

(S1) $\mathbf{RHS}[\varepsilon, \mathbf{g}] \to \mathbf{g}_\varepsilon$ determines for $\mathbf{g} \in \ell_2(\mathcal{J})$ a finitely supported $\mathbf{g}_\varepsilon \in \ell_2(\mathcal{J})$ such that

$$\|\mathbf{g} - \mathbf{g}_\varepsilon\|_{\ell_2(\mathcal{J})} \leq \varepsilon.$$

(S2) $\mathbf{APPLY}[\varepsilon, \mathbf{G}, \mathbf{v}] \to \mathbf{w}_\varepsilon$ determines for $\mathbf{G} \in \mathcal{L}(\ell_2(\mathcal{J}))$ and for a finitely supported $\mathbf{v} \in \ell_2(\mathcal{J})$ a finitely supported $\mathbf{w}_\varepsilon \in \ell_2(\mathcal{J})$ such that

$$\|\mathbf{G}\mathbf{v} - \mathbf{w}_\varepsilon\|_{\ell_2(\mathcal{J})} \leq \varepsilon.$$

(S3) $\mathbf{COARSE}[\varepsilon, \mathbf{v}] \to \mathbf{v}_\varepsilon$ determines for a finitely supported $\mathbf{v} \in \ell_2(\mathcal{J})$ a finitely supported $\mathbf{v}_\varepsilon \in \ell_2(\mathcal{J})$ with at most m significant coefficients, such that

$$\|\mathbf{v} - \mathbf{v}_\varepsilon\|_{\ell_2(\mathcal{J})} \leq \varepsilon. \tag{96}$$

Moreover, $m \leq C m_{\min}$ holds, $m_{\min}$ being the minimal number of entries for which (96) is valid.

Then, employing the key idea outlined above, we get the following fundamental algorithm:

Algorithm 5.21. $\mathbf{SOLVE}[\varepsilon, \mathbf{A}, \mathbf{f}] \to \mathbf{u}_\varepsilon$

$\Lambda_0 := \emptyset$; $\mathbf{r}_{\Lambda_0} := \mathbf{f}$; $\varepsilon_0 := \|\mathbf{f}\|_{\ell_2(\mathcal{J})}$; $j := 0$; $u_0 := 0$;
while $\varepsilon_j > \varepsilon$ **do**
  $\varepsilon_{j+1} := 2^{-(j+1)}\|\mathbf{f}\|_{\ell_2(\mathcal{J})}$; $\Lambda_{j,0} := \Lambda_j$; $\mathbf{u}_{j,0} := \mathbf{u}_j$;
  for $l = 1, ..., L$ **do**
    Compute Galerkin approximation $\mathbf{u}_{\Lambda_{j,l-1}}$ for $\Lambda_{j,l-1}$;
    Compute
      $\tilde{\mathbf{r}}_{\Lambda_{j,l-1}} := \mathbf{RHS}[C_1^{\text{tol}}\varepsilon_{j+1}, \mathbf{f}] - \mathbf{APPLY}[C_1^{\text{tol}}\varepsilon_{j+1}, \mathbf{A}, \mathbf{u}_{\Lambda_{j,l-1}}]$;
    Compute smallest set $\Lambda_{j,l}$,
      such that, $\|\tilde{\mathbf{r}}_{\Lambda_{j,l-1}}|_{\Lambda_{j,l}}\|_{\ell_2(\mathcal{J})} \geq \frac{1}{2}\|\tilde{\mathbf{r}}_{\Lambda_{j,l-1}}\|_{\ell_2(\mathcal{J})}$;
  end for
  $\mathbf{COARSE}[C_2^{\text{tol}}\varepsilon_{j+1}, \mathbf{u}_{\Lambda_{j,L}}] \to (\Lambda_{j+1}, \mathbf{u}_{j+1})$;
  $j := j + 1$;
end while

Remark 5.22. In [11], it has been shown that Algorithm 5.21 exactly fits into the setting of Assumptions 5.11 and 5.12. Let us denote by $\Lambda_\varepsilon \subset \mathcal{J}$ the final index set when Algorithm 5.21 terminates (the method of updating ε_j ensures termination). Then Algorithm 5.21 has the following properties.

(P1) Algorithm 5.21 is guaranteed to converge for a huge class of problems, in particular for the differential operators $L_{\tau,i}$ that we have in mind. Denoting with $H^t(\mathcal{O})$ the Sobolev space according to (95), we have

$$\|u - \sum_{\mu \in \Lambda_\varepsilon} c_\mu \psi_\mu\|_{H^t(\mathcal{O})} \leq C(u)\varepsilon.$$

(P2) Algorithm 5.21 is asymptotically optimal in the sense of Assumption 5.12, i.e., with $1/q = (s-t)/d + 1/2$, we have

$$\big\|u - \sum_{\mu \in \Lambda_\varepsilon} c_\mu \psi_\mu\big\|_{H^t(\mathcal{O})} \leq \hat{C}^{\mathrm{asym}} \|u\|_{B^s_q(L_q(\mathcal{O}))} (\#\Lambda_\varepsilon)^{-\frac{(s-t)}{d}}.$$

Remark 5.23. (i) We will not discuss the concrete numerical realization of the three fundamental subroutines in detail. The subroutine **COARSE** consists of a thresholding step, whereas **RHS** essentially requires the computation of a best m-term approximation. The most complicated building block is **APPLY**. Let us just mention that its existence can be established for elliptic operators with Schwartz kernels by using the cancellation property of wavelets.

(ii) In Algorithm 5.21, C_1^{tol} and C_2^{tol} denote some suitably chosen constants whose concrete values depend on the problem under consideration. The parameter L has to be chosen in a suitable way. We refer again to [11] for details.

(iii) It has been shown in [11] that Algorithm 5.21 has the additional property that the number of arithmetic operations stays proportional to the number of unknowns, that is, the number of floating point operations needed to compute $\mathbf{u}_\varepsilon$ is bounded by $C_{\mathrm{supp}} \# \operatorname{supp} \mathbf{u}_\varepsilon$.

A.1 Variational operators

In the preceding sections, we very often considered the same problem on different spaces, e.g., we switched from an operator equation defined on V to the same equation defined on U. In this section we want to clarify in more detail why this is justified.

Let $(V, \langle\cdot,\cdot\rangle_V)$ be a separable real Hilbert space. Furthermore, let

$$a(\cdot,\cdot) : V \times V \to \mathbb{R}$$

be a continuous, symmetric and elliptic bilinear form. This means that there exist two constants $c_{\mathrm{ell}}, C_{\mathrm{ell}} > 0$, such that for arbitrary $u, v \in V$, the bilinear form fulfills the following conditions:

$$c_{\mathrm{ell}} \|u\|_V^2 \leq a(u,u), \qquad a(u,v) = a(v,u), \qquad |a(u,v)| \leq C_{\mathrm{ell}} \|u\|_V \|v\|_V. \tag{97}$$

Then, by the Lax-Milgram theorem, the operator

$$\begin{aligned} A : V &\to V^* \\ v &\mapsto Av := -a(v,\cdot) \end{aligned} \tag{98}$$

is boundededly invertible. Let us now assume that V is densely embedded into a real Hilbert space $(U, \langle\cdot,\cdot\rangle_U)$ via a linear embedding j. We write

$$V \overset{j}{\hookrightarrow} U.$$

Furthermore, we identify the Hilbert space U with its topological dual space U^* via the Riesz isomorphism $U \ni u \mapsto \Phi u := \langle u, \cdot\rangle_U \in U^*$. The adjoint map $j^* : U^* \to V^*$ of j embeds U^* densely into the topological dual V^* of V. All in all we have a so called Gelfand triple (V, U, V^*),

$$V \overset{j}{\hookrightarrow} U \overset{\Phi}{\equiv} U^* \overset{j^*}{\hookrightarrow} V^*.$$

Using $\langle\cdot,\cdot\rangle_{V^*\times V}$ to denote the dual pairs of V and V^*, we have

$$\langle j(v_1), j(v_2)\rangle_U = \langle j^* \Phi\, j(v_1), v_2\rangle_{V^*\times V} \qquad \text{for all } v_1, v_2 \in V. \tag{99}$$

In this setting, we can consider the operator $A : V \to V^*$ as an unbounded operator on the intermediate space U. More precisely, set

$$D(A) := D(A;U) := \{u \in V \,:\, Au \in j^*\Phi(U)\},$$

and define the operator

$$\begin{aligned} \tilde{A} : D(\tilde{A}) := j(D(A;U)) \subseteq U &\to U \\ u &\mapsto \tilde{A}u := \Phi^{-1} j^{*-1} A\, j^{-1} u. \end{aligned}$$

Such an (unbounded) linear operator is sometimes called *variational*. It is densely defined, since U^* is densely embedded in V^*. Furthermore, the symmetry of the bilinear form $a(\cdot,\cdot)$ implies that $\tilde{A}$ is self-adjoint. At the same time, it is strictly negative definite, because of the ellipticity of a. Moreover, since $A : V \to V^*$ is boundedly invertible, the operator $\tilde{A}^{-1} : U \to U$, defined by $\tilde{A}^{-1} := jA^{-1}j^*\Phi$ is the bounded inverse of $\tilde{A}$. It is compact if the embedding j of V in U is compact.

Let us now fix $\tau > 0$ and consider the bilinear form

$$\begin{aligned} a_\tau : V \times V &\to \mathbb{R} \\ (u,v) &\mapsto a_\tau(u,v) := \tau\langle j(u), j(v)\rangle_U + a(u,v), \end{aligned}$$

which is also continuous, symmetric and elliptic in the sense of (99). Obviously, for $u, v \in V$, we have the identity

$$\begin{aligned} a_\tau(u,v) &= \tau\langle j^*\Phi j(u), v\rangle_{V^*\times V} - \langle Au, v\rangle_{V^*\times V} \\ &= \langle(\tau j^*\Phi j - A)u, v\rangle_{V^*\times V}, \end{aligned}$$

so that applying again the Lax-Milgram theorem, we can conclude that $(\tau j^*\Phi j - A) : V \to V^*$ is boundedely invertible. Therefore, the operator

$$\begin{aligned} (\tau I - \tilde{A}) : D(\tilde{A}) \subseteq U \to U \\ u \mapsto (\tau I - \tilde{A})u := \tau u - \tilde{A}u, \end{aligned}$$

which coincides with $\Phi^{-1}j^{*-1}(\tau j^*\Phi j - A)j^{-1}$ on $D(\tilde{A})$, possesses a bounded inverse $(\tau I - \tilde{A})^{-1} = j(\tau j^*\Phi j - A)^{-1}j^*\Phi : U \to U$. Thus, the resolvent set $\varrho(\tilde{A})$ of $\tilde{A}$ contains all $\tau \geq 0$. In particular, for any $\tau > 0$, the range of the operator $(\tau I - \tilde{A})$ is all of U. Since, furthermore, $\tilde{A}$ is dissipative, the Lumer-Phillips theorem leads to the fact, that $\tilde{A}$ generates a strongly continuous semigroup $\{e^{t\tilde{A}}\}_{t\geq 0}$ of contractions on U, see, e.g. [43, Theorem 1.4.3]. Thus, an application of the Hille-Yosida theorem (see, e.g. [43, Theorem 1.3.1]) leads to the fact that the operator $L_\tau^{-1} := (I - \tau\tilde{A})^{-1} = \tau(\tau I - \tilde{A})^{-1} : U \to U$ is a contraction for each $\tau > 0$.

By an abuse of notation, we sometimes write A instead of $\tilde{A}$.

A.2 Proofs of Lemma 3.11 and Lemma 5.6

Proof of Lemma 3.11. By (38) and (39) the stage equations (30) are given by

$$\begin{aligned} \big(I - \tau\gamma_{1,1}A\big)w_{k,1} &= Au_k + f(t_k), \\ \big(I - \tau\gamma_{2,2}A\big)w_{k,2} &= A(u_k + \tau a_{2,1}w_{k,1}) + f(t_k + a_2\tau) + c_{2,1}w_{k,1}. \end{aligned}$$

We begin with an application of the following basic observation, that

$$I = \big(I - CA\big)^{-1}\big(I - CA\big)$$

implies

$$\big(I - CA\big)^{-1}A = -\frac{1}{C}I + \frac{1}{C}\big(I - CA\big)^{-1}.$$

It follows that

$$\begin{aligned} w_{k,1} &= \Big(\big(-\frac{1}{\tau\gamma_{1,1}}I + \frac{1}{\tau\gamma_{1,1}}\big(I - \tau\gamma_{1,1}A\big)^{-1}\Big)u_k + \big(I - \tau\gamma_{1,1}A\big)^{-1}f(t_k) \\ &= -\frac{1}{\tau\gamma_{1,1}}u_k + L_{\tau,1}^{-1}\Big(\frac{1}{\tau\gamma_{1,1}}u_k + f(t_k)\Big). \end{aligned}$$

We denote

$$v_{k,1} = L_{\tau,1}^{-1}\Big(\frac{1}{\tau\gamma_{1,1}}u_k + f(t_k)\Big).$$

A similar computation for the second stage equation yields

$$\begin{aligned}
w_{k,2} &= \Big(-\frac{1}{\tau\gamma_{2,2}}I + \frac{1}{\tau\gamma_{2,2}}\big(I - \tau\gamma_{2,2}A\big)^{-1}\Big)\Big(u_k + \tau a_{2,1}w_{k,1}\Big) \\
&\quad + \big(I - \tau\gamma_{2,2}A\big)^{-1}\Big(f(t_k + a_2\tau) + c_{2,1}w_{k,1}\Big) \\
&= -\frac{1}{\tau\gamma_{2,2}}\Big((1 - \frac{a_{2,1}}{\gamma_{1,1}})u_k + \tau a_{2,1}v_{k,1}\Big) \\
&\quad + L_{\tau,2}^{-1}\Big(\frac{1}{\tau\gamma_{2,2}}((1 - \frac{a_{2,1}}{\gamma_{1,1}})u_k + \tau a_{2,1}v_{k,1}) \\
&\qquad\qquad + f(t_k + a_2\tau) + c_{2,1}\big(\frac{-1}{\tau\gamma_{1,1}}u_k + v_{k,1}\big)\Big) \\
&= \frac{-1}{\tau\gamma_{2,2}}(1 - \frac{a_{2,1}}{\gamma_{1,1}})u_k - \frac{a_{2,1}}{\gamma_{2,2}}v_{k,1} \\
&\quad + L_{\tau,2}^{-1}\Big((\frac{1}{\tau\gamma_{2,2}}(1 - \frac{a_{2,1}}{\gamma_{1,1}}) - \frac{c_{2,1}}{\tau\gamma_{1,1}})u_k \\
&\qquad\qquad + (\frac{a_{2,1}}{\gamma_{2,2}} + c_{2,1})v_{k,1} + f(t_k + a_2\tau)\Big).
\end{aligned}$$

We denote

$$v_{k,2} = L_{\tau,2}^{-1}\Big((\frac{1}{\tau\gamma_{2,2}}(1 - \frac{a_{2,1}}{\gamma_{1,1}}) - \frac{c_{2,1}}{\tau\gamma_{1,1}})u_k + \big(\frac{a_{2,1}}{\gamma_{2,2}} + c_{2,1}\big)v_{k,1} + f(t_k + a_2\tau)\Big)\Big)$$

and arrive at

$$\begin{aligned}
u_{k+1} &= u_k + \tau m_1\Big(\frac{-1}{\tau\gamma_{1,1}}u_k + v_{k,1}\Big) \\
&\qquad\qquad + \tau m_2\Big(\frac{-1}{\tau\gamma_{2,2}}(1 - \frac{a_{2,1}}{\gamma_{1,1}})u_k - \frac{a_{2,1}}{\gamma_{2,2}}v_{k,1} + v_{k,2}\Big) \\
&= \Big(1 - \frac{m_1}{\gamma_{1,1}} - \frac{m_2}{\gamma_{2,2}}\big(1 - \frac{a_{2,1}}{\gamma_{1,1}}\big)\Big)u_k + \Big(\tau m_1 - \tau m_2\frac{a_{2,1}}{\gamma_{2,2}}\Big)v_{k,1} \\
&\qquad\qquad + \tau m_2 v_{k,2}.
\end{aligned}$$

□

Proof of Lemma 5.6. We start with the estimation

$$\begin{aligned}
\|\hat{w}_{k,i}\|_{B^s_q(L_q(\mathcal{O}))} &= \|L_{\tau,i}^{-1}R_{\tau,k,i}(\tilde{u}_k, \tilde{w}_{k,1}, \ldots, \tilde{w}_{k,i-1})\|_{B^s_q(L_q(\mathcal{O}))} \\
&\le \|L_{\tau,i}^{-1}\|_{\mathcal{L}(L_2(\mathcal{O}), B^s_q(L_q(\mathcal{O})))}\|R_{\tau,k,i}(\tilde{u}_k, \tilde{w}_{k,1}, \ldots, \tilde{w}_{k,i-1})\|_{L_2(\mathcal{O})}.
\end{aligned}$$

The Lipschitz continuity of $R_{\tau,k,i}$ implies the linear growth property

$$
\begin{aligned}
&\|R_{\tau,k,i}(\tilde{u}_k, \tilde{w}_{k,1}, \ldots, \tilde{w}_{k,i-1})\|_{L_2(\mathcal{O})} \\
&\quad \le C^{\mathrm{Lip,R}}_{\tau,k,(i)} \Big(\|\tilde{u}_k\|_{L_2(\mathcal{O})} + \sum_{j=1}^{i-1} \|\tilde{w}_{k,j}\|_{L_2(\mathcal{O})} \Big) + \|R_{\tau,k,i}(0,\ldots,0)\|_{L_2(\mathcal{O})} \\
&\quad \le \max\Big\{ C^{\mathrm{Lip,R}}_{\tau,k,(i)}, \|R_{\tau,k,i}(0,\ldots,0)\|_{L_2(\mathcal{O})} \Big\} \\
&\qquad\qquad \times \Big(1 + \|\tilde{u}_k\|_{L_2(\mathcal{O})} + \sum_{j=1}^{i-1} \|\tilde{w}_{k,j}\|_{L_2(\mathcal{O})} \Big) \\
&\quad \le \max\Big\{ C^{\mathrm{Lip,R}}_{\tau,k,(i)}, \|R_{\tau,k,i}(0,\ldots,0)\|_{L_2(\mathcal{O})} \Big\} \\
&\qquad\qquad \times \Big(1 + \|u_k\|_{L_2(\mathcal{O})} + \sum_{j=1}^{i-1} \|w_{k,j}\|_{L_2(\mathcal{O})} \\
&\qquad\qquad\qquad + \|u_k - \tilde{u}_k\|_{L_2(\mathcal{O})} + \sum_{j=1}^{i-1} \|w_{k,j} - \tilde{w}_{k,j}\|_{L_2(\mathcal{O})} \Big).
\end{aligned}
$$

As before, the Lipschitz continuity of $L^{-1}_{\tau,i} R_{\tau,k,i}$ implies

$$
\begin{aligned}
\|w_{k,i}\|_{L_2(\mathcal{O})} &= \|L^{-1}_{\tau,i} R_{\tau,k,i}(u_k, w_{k,1}, \ldots, w_{k,i-1})\|_{L_2(\mathcal{O})} \\
&\le \max\Big\{ C^{\mathrm{Lip}}_{\tau,k,(i)}, \|L^{-1}_{\tau,i} R_{\tau,k,i}(0,\ldots,0)\|_{L_2(\mathcal{O})} \Big\} \\
&\qquad \times \Big(1 + \|u_k\|_{L_2(\mathcal{O})} + \sum_{j=1}^{i-1} \|w_{k,j}\|_{L_2(\mathcal{O})} \Big).
\end{aligned}
$$

By induction, we estimate

$$
\begin{aligned}
&1 + \|u_k\|_{L_2(\mathcal{O})} + \sum_{j=1}^{i-1} \|w_{k,j}\|_{L_2(\mathcal{O})} \\
&\quad \le \prod_{l=1}^{i-1} \Big(1 + \max\Big\{ C^{\mathrm{Lip}}_{\tau,k,(l)}, \|L^{-1}_{\tau,l} R_{\tau,k,l}(0,\ldots,0)\|_{L_2(\mathcal{O})} \Big\} \Big) \big(1 + \|u_k\|_{L_2(\mathcal{O})} \big).
\end{aligned}
$$

Note that

$$
\|\tilde{w}_{k,i} - \hat{w}_{k,i}\|_{L_2(\mathcal{O})} \le \|\tilde{w}_{k,i} - \hat{w}_{k,i}\|_{H^t(\mathcal{O})} \le \varepsilon_{k,i}.
$$

This enables us to follow similar lines as in the proof of Theorem 2.20. We estimate

$$
\begin{aligned}
&\|u_k - \tilde{u}_k\|_{L_2(\mathcal{O})} + \sum_{j=1}^{i-1} \|w_{k,j} - \tilde{w}_{k,j}\|_{L_2(\mathcal{O})} \\
&\le (1 + C^{\mathrm{Lip}}_{\tau,k,(i-1)})\Big(\|u_k - \tilde{u}_k\|_{L_2(\mathcal{O})} + \sum_{j=1}^{i-2} \|w_{k,j} - \tilde{w}_{k,j}\|_{L_2(\mathcal{O})}\Big) \\
&\qquad + \Big\| L^{-1}_{\tau,i-1} R_{\tau,k,i-1}(\tilde{u}_k, \tilde{w}_{k,1}, \dots, \tilde{w}_{k,i-2}) \\
&\qquad\qquad - \big[L^{-1}_{\tau,i-1} R_{\tau,k,i-1}(\tilde{u}_k, \tilde{w}_{k,1}, \dots, \tilde{w}_{k,i-2})\big]_{\varepsilon_{k,i-1}} \Big\|_{L_2(\mathcal{O})} \\
&\le (1 + C^{\mathrm{Lip}}_{\tau,k,(i-1)})\Big(\|u_k - \tilde{u}_k\|_{L_2(\mathcal{O})} + \sum_{j=1}^{i-2} \|w_{k,j} - \tilde{w}_{k,j}\|_{L_2(\mathcal{O})}\Big) + \varepsilon_{k,i-1}
\end{aligned}
$$

and conclude by induction

$$
\begin{aligned}
&\|u_k - \tilde{u}_k\|_{L_2(\mathcal{O})} + \sum_{j=1}^{i-1} \|w_{k,j} - \tilde{w}_{k,j}\|_{L_2(\mathcal{O})} \\
&\quad \le \big(\prod_{l=1}^{i-1} (1 + C^{\mathrm{Lip}}_{\tau,k,(l)})\big) \|u_k - \tilde{u}_k\|_{L_2(\mathcal{O})} + \sum_{j=1}^{i-1} \varepsilon_{k,j} \prod_{l=j+1}^{i-1} (1 + C^{\mathrm{Lip}}_{\tau,k,(l)}).
\end{aligned}
$$

The proof is finished by

$$
\|u_k - \tilde{u}_k\|_{L_2(\mathcal{O})} \le \sum_{j=0}^{k-1} \big(\prod_{l=j+1}^{k-1} (C'_{\tau,l,(0)} - 1)\big) \sum_{i=1}^{S} C'_{\tau,j,(i)} \varepsilon_{j,i},
$$

which is shown as in Theorem 2.20. □

References

[1] R.A. Adams, *Sobolev Spaces*, Pure and Applied Mathematics, 65. A Series of Monographs and Textbooks. New York-San Francisco-London: Academic Press, 1975.

[2] I. Babuška, *Advances in the p and h-p versions of the finite element method. A survey*, Numerical mathematics, Proc. Int. Conf., Singapore 1988, ISNM, Int. Ser. Numer. Math. **86** (1988), 31–46.

[3] I. Babuška and W.C. Rheinboldt, *A survey of a posteriori error estimators and adaptive approaches in the finite element method*, Finite element methods, Proc. China-France Symp., Beijing/China (1983), 1–56.

[4] R.E. Bank and A. Weiser, *Some a posteriori error estimators for elliptic partial differential equations*, Math. Comput. **44** (1985), 283–301.

[5] J. Bergh and J. Löfström, *Interpolation Spaces. An Introduction*, Springer, Berlin, 1976.

[6] M.Sh. Birman and M.Z. Solomyak, *On the asymptotic spectrum of "non-smooth" elliptic equations*, Funct. Anal. Appl. **5** (1971), 56–57.

[7] F.A. Bornemann, B. Erdmann, and R. Kornhuber, *A posteriori error estimates for elliptic problems in two and three space dimensions*, SIAM J. Numer. Anal. **33** (1996), 1188–1204.

[8] H. Breckner and W. Grecksch, *Approximation of solutions of stochastic evolution equations by Rothe's method*, Reports of the Institute of Optimization and Stochastics **13** (1997), Martin-Luther-Univ. Halle-Wittenberg, Fachbereich Mathematik und Informatik.

[9] C. Canuto, A. Tabacco, and K. Urban, *The wavelet element method. II: Realization and additional features in 2D and 3D*, Appl. Comput. Harmon. Anal. **8** (2000), 123–165.

[10] A. Cohen, *Wavelet Methods in Numerical Analysis*, North-Holland/Elsevier, Amsterdam, 2000.

[11] A. Cohen, W. Dahmen, and R.A. DeVore, *Adaptive wavelet methods for elliptic operator equations: Convergence rates*, Math. Comput. **70** (2001), 27–75.

[12] ______, *Adaptive wavelet methods. II: Beyond the elliptic case*, Found. Comput. Math. **2** (2002), 203–245.

[13] P. Cox and J.M.A.M van Neerven, *Pathwise Hölder convergence of the implicit Euler scheme for semi-linear SPDEs with multiplicative noise*, Preprint (2012), (arXiv:1201.4465v1).

[14] M. Crouzeix and V. Thomée, *On the discretization in time of semilinear parabolic equations with nonsmooth initial data*, Math. Comput. **49** (1987), 359–377.

[15] G. Da Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, Encyclopedia of Mathematics and Its Applications. 44. Cambridge: Cambridge University Press, 1992.

[16] S. Dahlke, *Besov regularity for elliptic boundary value problems in polygonal domains*, Appl. Math. Lett. **12** (1999), 31–36.

[17] S. Dahlke, W. Dahmen, and R.A. DeVore, *Nonlinear approximation and adaptive techniques for solving elliptic operator equations*, Multiscale wavelet methods for partial differential equations (W. Dahmen, A. Kurdila, and P. Oswald, eds.), Academic Press, San Diego, 1997, pp. 237–284.

[18] S. Dahlke, W. Dahmen, R. Hochmuth, and R. Schneider, *Stable multiscale bases and local error estimation for elliptic problems*, Appl. Numer. Math. **23** (1997), 21–47.

[19] S. Dahlke and R.A. DeVore, *Besov regularity for elliptic boundary value problems*, Commun. Partial Differ. Equations **22** (1997), 1–16.

[20] S. Dahlke, M. Fornasier, T. Raasch, R. Stevenson, and M. Werner, *Adaptive frame methods for elliptic operator equations: The steepest descent approach*, IMA J. Numer. Anal. **27** (2007), 717–740.

[21] S. Dahlke, E. Novak, and W. Sickel, *Optimal approximation of elliptic problems by linear and nonlinear mappings. I*, J. Complexity **22** (2006), 29–49.

[22] S. Dahlke and W. Sickel, *On Besov regularity of solutions to nonlinear elliptic partial differential equations*, Rev. Mat. Complut. (2011), 1–31, doi 10.1007/s13163-012-0093-z.

[23] W. Dahmen and R. Schneider, *Wavelets with complementary boundary conditions – functions spaces on the cube*, Result. Math. **34** (1998), 255–293.

[24] ———, *Composite wavelet bases for operator equations*, Math. Comput. **68** (1999), 1533–1567.

[25] ———, *Wavelets on manifolds. I: Construction and domain decomposition*, SIAM J. Math. Anal. **31** (1999), 184–230.

[26] A. Debussche and J. Printems, *Weak order for the discretization of the stochastic heat equation*, Math. Comput. **78** (2009), 845–863.

[27] R.A. DeVore, *Nonlinear approximation*, Acta Numerica **7** (1998), 51–150.

[28] K. Eriksson, *An adaptive finite element method with efficient maximum norm error control for elliptic problems*, Math. Models Methods Appl. Sci. **4** (1994), 313–329.

[29] K. Eriksson and C. Johnson, *Adaptive finite element methods for parabolic problems. I: A linear model problem*, SIAM J. Numer. Anal. **28** (1991), 43–77.

[30] ———, *Adaptive finite element methods for parabolic problems. II: Optimal error estimates in $L_\infty L_2$ and $L_\infty L_\infty$*, SIAM J. Numer. Anal. **32** (1995), 706–740.

[31] K. Eriksson, C. Johnson, and S. Larsson, *Adaptive finite element methods for parabolic problems. VI: Analytic semigroups*, SIAM J. Numer. Anal. **35** (1998), 1315–1325.

[32] W. Grecksch and C. Tudor, *Stochastic Evolution Equations. A Hilbert Space Approach*, Akademie Verlag, Berlin, 1995.

[33] I. Gyöngy and D. Nualart, *Implicit scheme for stochastic parabolic partial differential equations driven by space-time white noise*, Potential Anal. **7** (1997), 725–757.

[34] M. Hanke-Bourgeois, *Foundations of Numerical Mathematics and Scientific Computing*, Vieweg+Teubner, Wiesbaden, 2009.

[35] P. Hansbo and C. Johnson, *Adaptive finite element methods in computational mechanics*, Comput. Methods Appl. Mech. Eng. **101** (1992), 143–181.

[36] A. Jentzen and P.E. Kloeden, *Taylor Approximations for Stochastic Partial Differential Equations*, Regional Conference Series in Applied Mathematics 83, SIAM, Philadelphia, 2011.

[37] D. Jerison and C.E. Kenig, *The inhomogeneous Dirichlet problem in Lipschitz domains*, J. Funct. Anal. **130** (1995), 161–219.

[38] C. Johnson, *Numerical Solution of Partial Differential Equations by the Finite Element Method*, Dover Publications, Mineola, 2009.

[39] T. Kato, *Perturbation Theory for Linear Operators. 2nd corr. print. of the 2nd ed.*, Springer, Berlin, 1984.

[40] J. Lang, *Adaptive Multilevel Solution of Nonlinear Parabolic PDE Systems. Theory, Algorithm, and Applications*, Springer, Berlin, 2001.

[41] F. Lindner, *Approximation and Regularity of Stochastic PDEs*, Berichte aus der Mathematik. Aachen: Shaker Verlag; Dresden: Univ. Dresden, Fakultät Mathematik und Naturwissenschaften (Diss.), 2011.

[42] C. Lubich and A. Ostermann, *Linearly implicit time discretization of nonlinear parabolic equations*, IMA J. Numer. Anal. **15** (1995), 555–583.

[43] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Applied Mathematical Sciences, 44. New York: Springer-Verlag, 1983.

[44] C. Prévôt and M. Röckner, *A Concise Course on Stochastic Partial Differential Equations*, Springer, Berlin, 2007.

[45] J. Printems, *On the discretization in time of parabolic stochastic partial differential equations*, Math. Model. Numer. Anal. **35** (2001), 1055–1078.

[46] V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, Springer, Berlin, 2006.

[47] R. Verfürth, *A posteriori error estimation and adaptive mesh-refinement techniques*, J. Comput. Appl. Math. **50** (1994), 67–83.

[48] ______, *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refinement Techniques*, Wiley-Teubner Series Advances in Numerical Mathematics. Chichester: Wiley. Stuttgart: B. G. Teubner, 1996.

[49] J.G. Verwer, E.J. Spee, J.G. Blom, and W. Hundsdorfer, *A second-order Rosenbrock method applied to photochemical dispersion problems*, SIAM J. Sci. Comput. **20** (1999), 1456–1480.

Petru A. Cioica, Stephan Dahlke, Ulrich Friedrich, and Stefan Kinzel
Philipps-Universität Marburg
FB Mathematik und Informatik, AG Numerik/Optimierung
Hans-Meerwein-Strasse
35032 Marburg, Germany
{cioica, dahlke, friedrich, kinzel}@mathematik.uni-marburg.de

Nicolas Döhring and Klaus Ritter
TU Kaiserslautern
Department of Mathematics, Computational Stochastics Group
Erwin-Schrödinger-Strasse
67663 Kaiserslautern, Germany
{doehring, ritter}@mathematik.uni-kl.de

Felix Lindner and René L. Schilling
TU Dresden
FR Mathematik, Institut für Mathematische Stochastik
01062 Dresden, Germany
{felix.lindner, rene.schilling}@tu-dresden.de

Thorsten Raasch
Johannes Gutenberg-Universität Mainz
Institut für Mathematik, AG Numerische Mathematik
Staudingerweg 9
55099 Mainz, Germany
raasch@uni-mainz.de

Preprint Series DFG-SPP 1324

http://www.dfg-spp1324.de

Reports

[1] R. Ramlau, G. Teschke, and M. Zhariy. A Compressive Landweber Iteration for Solving Ill-Posed Inverse Problems. Preprint 1, DFG-SPP 1324, September 2008.

[2] G. Plonka. The Easy Path Wavelet Transform: A New Adaptive Wavelet Transform for Sparse Representation of Two-dimensional Data. Preprint 2, DFG-SPP 1324, September 2008.

[3] E. Novak and H. Woźniakowski. Optimal Order of Convergence and (In-) Tractability of Multivariate Approximation of Smooth Functions. Preprint 3, DFG-SPP 1324, October 2008.

[4] M. Espig, L. Grasedyck, and W. Hackbusch. Black Box Low Tensor Rank Approximation Using Fibre-Crosses. Preprint 4, DFG-SPP 1324, October 2008.

[5] T. Bonesky, S. Dahlke, P. Maass, and T. Raasch. Adaptive Wavelet Methods and Sparsity Reconstruction for Inverse Heat Conduction Problems. Preprint 5, DFG-SPP 1324, January 2009.

[6] E. Novak and H. Woźniakowski. Approximation of Infinitely Differentiable Multivariate Functions Is Intractable. Preprint 6, DFG-SPP 1324, January 2009.

[7] J. Ma and G. Plonka. A Review of Curvelets and Recent Applications. Preprint 7, DFG-SPP 1324, February 2009.

[8] L. Denis, D. A. Lorenz, and D. Trede. Greedy Solution of Ill-Posed Problems: Error Bounds and Exact Inversion. Preprint 8, DFG-SPP 1324, April 2009.

[9] U. Friedrich. A Two Parameter Generalization of Lions' Nonoverlapping Domain Decomposition Method for Linear Elliptic PDEs. Preprint 9, DFG-SPP 1324, April 2009.

[10] K. Bredies and D. A. Lorenz. Minimization of Non-smooth, Non-convex Functionals by Iterative Thresholding. Preprint 10, DFG-SPP 1324, April 2009.

[11] K. Bredies and D. A. Lorenz. Regularization with Non-convex Separable Constraints. Preprint 11, DFG-SPP 1324, April 2009.

[12] M. Döhler, S. Kunis, and D. Potts. Nonequispaced Hyperbolic Cross Fast Fourier Transform. Preprint 12, DFG-SPP 1324, April 2009.

[13] C. Bender. Dual Pricing of Multi-Exercise Options under Volume Constraints. Preprint 13, DFG-SPP 1324, April 2009.

[14] T. Müller-Gronbach and K. Ritter. Variable Subspace Sampling and Multi-level Algorithms. Preprint 14, DFG-SPP 1324, May 2009.

[15] G. Plonka, S. Tenorth, and A. Iske. Optimally Sparse Image Representation by the Easy Path Wavelet Transform. Preprint 15, DFG-SPP 1324, May 2009.

[16] S. Dahlke, E. Novak, and W. Sickel. Optimal Approximation of Elliptic Problems by Linear and Nonlinear Mappings IV: Errors in L_2 and Other Norms. Preprint 16, DFG-SPP 1324, June 2009.

[17] B. Jin, T. Khan, P. Maass, and M. Pidcock. Function Spaces and Optimal Currents in Impedance Tomography. Preprint 17, DFG-SPP 1324, June 2009.

[18] G. Plonka and J. Ma. Curvelet-Wavelet Regularized Split Bregman Iteration for Compressed Sensing. Preprint 18, DFG-SPP 1324, June 2009.

[19] G. Teschke and C. Borries. Accelerated Projected Steepest Descent Method for Nonlinear Inverse Problems with Sparsity Constraints. Preprint 19, DFG-SPP 1324, July 2009.

[20] L. Grasedyck. Hierarchical Singular Value Decomposition of Tensors. Preprint 20, DFG-SPP 1324, July 2009.

[21] D. Rudolf. Error Bounds for Computing the Expectation by Markov Chain Monte Carlo. Preprint 21, DFG-SPP 1324, July 2009.

[22] M. Hansen and W. Sickel. Best m-term Approximation and Lizorkin-Triebel Spaces. Preprint 22, DFG-SPP 1324, August 2009.

[23] F.J. Hickernell, T. Müller-Gronbach, B. Niu, and K. Ritter. Multi-level Monte Carlo Algorithms for Infinite-dimensional Integration on $\mathbb{R}^{\mathbb{N}}$. Preprint 23, DFG-SPP 1324, August 2009.

[24] S. Dereich and F. Heidenreich. A Multilevel Monte Carlo Algorithm for Lévy Driven Stochastic Differential Equations. Preprint 24, DFG-SPP 1324, August 2009.

[25] S. Dahlke, M. Fornasier, and T. Raasch. Multilevel Preconditioning for Adaptive Sparse Optimization. Preprint 25, DFG-SPP 1324, August 2009.

[26] S. Dereich. Multilevel Monte Carlo Algorithms for Lévy-driven SDEs with Gaussian Correction. Preprint 26, DFG-SPP 1324, August 2009.

[27] G. Plonka, S. Tenorth, and D. Roşca. A New Hybrid Method for Image Approximation using the Easy Path Wavelet Transform. Preprint 27, DFG-SPP 1324, October 2009.

[28] O. Koch and C. Lubich. Dynamical Low-rank Approximation of Tensors. Preprint 28, DFG-SPP 1324, November 2009.

[29] E. Faou, V. Gradinaru, and C. Lubich. Computing Semi-classical Quantum Dynamics with Hagedorn Wavepackets. Preprint 29, DFG-SPP 1324, November 2009.

[30] D. Conte and C. Lubich. An Error Analysis of the Multi-configuration Time-dependent Hartree Method of Quantum Dynamics. Preprint 30, DFG-SPP 1324, November 2009.

[31] C. E. Powell and E. Ullmann. Preconditioning Stochastic Galerkin Saddle Point Problems. Preprint 31, DFG-SPP 1324, November 2009.

[32] O. G. Ernst and E. Ullmann. Stochastic Galerkin Matrices. Preprint 32, DFG-SPP 1324, November 2009.

[33] F. Lindner and R. L. Schilling. Weak Order for the Discretization of the Stochastic Heat Equation Driven by Impulsive Noise. Preprint 33, DFG-SPP 1324, November 2009.

[34] L. Kämmerer and S. Kunis. On the Stability of the Hyperbolic Cross Discrete Fourier Transform. Preprint 34, DFG-SPP 1324, December 2009.

[35] P. Cerejeiras, M. Ferreira, U. Kähler, and G. Teschke. Inversion of the noisy Radon transform on $SO(3)$ by Gabor frames and sparse recovery principles. Preprint 35, DFG-SPP 1324, January 2010.

[36] T. Jahnke and T. Udrescu. Solving Chemical Master Equations by Adaptive Wavelet Compression. Preprint 36, DFG-SPP 1324, January 2010.

[37] P. Kittipoom, G. Kutyniok, and W.-Q Lim. Irregular Shearlet Frames: Geometry and Approximation Properties. Preprint 37, DFG-SPP 1324, February 2010.

[38] G. Kutyniok and W.-Q Lim. Compactly Supported Shearlets are Optimally Sparse. Preprint 38, DFG-SPP 1324, February 2010.

[39] M. Hansen and W. Sickel. Best m-Term Approximation and Tensor Products of Sobolev and Besov Spaces – the Case of Non-compact Embeddings. Preprint 39, DFG-SPP 1324, March 2010.

[40] B. Niu, F.J. Hickernell, T. Müller-Gronbach, and K. Ritter. Deterministic Multi-level Algorithms for Infinite-dimensional Integration on $\mathbb{R}^{\mathbb{N}}$. Preprint 40, DFG-SPP 1324, March 2010.

[41] P. Kittipoom, G. Kutyniok, and W.-Q Lim. Construction of Compactly Supported Shearlet Frames. Preprint 41, DFG-SPP 1324, March 2010.

[42] C. Bender and J. Steiner. Error Criteria for Numerical Solutions of Backward SDEs. Preprint 42, DFG-SPP 1324, April 2010.

[43] L. Grasedyck. Polynomial Approximation in Hierarchical Tucker Format by Vector-Tensorization. Preprint 43, DFG-SPP 1324, April 2010.

[44] M. Hansen und W. Sickel. Best m-Term Approximation and Sobolev-Besov Spaces of Dominating Mixed Smoothness - the Case of Compact Embeddings. Preprint 44, DFG-SPP 1324, April 2010.

[45] P. Binev, W. Dahmen, and P. Lamby. Fast High-Dimensional Approximation with Sparse Occupancy Trees. Preprint 45, DFG-SPP 1324, May 2010.

[46] J. Ballani and L. Grasedyck. A Projection Method to Solve Linear Systems in Tensor Format. Preprint 46, DFG-SPP 1324, May 2010.

[47] P. Binev, A. Cohen, W. Dahmen, R. DeVore, G. Petrova, and P. Wojtaszczyk. Convergence Rates for Greedy Algorithms in Reduced Basis Methods. Preprint 47, DFG-SPP 1324, May 2010.

[48] S. Kestler and K. Urban. Adaptive Wavelet Methods on Unbounded Domains. Preprint 48, DFG-SPP 1324, June 2010.

[49] H. Yserentant. The Mixed Regularity of Electronic Wave Functions Multiplied by Explicit Correlation Factors. Preprint 49, DFG-SPP 1324, June 2010.

[50] H. Yserentant. On the Complexity of the Electronic Schrödinger Equation. Preprint 50, DFG-SPP 1324, June 2010.

[51] M. Guillemard and A. Iske. Curvature Analysis of Frequency Modulated Manifolds in Dimensionality Reduction. Preprint 51, DFG-SPP 1324, June 2010.

[52] E. Herrholz and G. Teschke. Compressive Sensing Principles and Iterative Sparse Recovery for Inverse and Ill-Posed Problems. Preprint 52, DFG-SPP 1324, July 2010.

[53] L. Kämmerer, S. Kunis, and D. Potts. Interpolation Lattices for Hyperbolic Cross Trigonometric Polynomials. Preprint 53, DFG-SPP 1324, July 2010.

[54] G. Kutyniok and W.-Q Lim. Shearlets on Bounded Domains. Preprint 54, DFG-SPP 1324, July 2010.

[55] A. Zeiser. Wavelet Approximation in Weighted Sobolev Spaces of Mixed Order with Applications to the Electronic Schrödinger Equation. Preprint 55, DFG-SPP 1324, July 2010.

[56] G. Kutyniok, J. Lemvig, and W.-Q Lim. Compactly Supported Shearlets. Preprint 56, DFG-SPP 1324, July 2010.

[57] A. Zeiser. On the Optimality of the Inexact Inverse Iteration Coupled with Adaptive Finite Element Methods. Preprint 57, DFG-SPP 1324, July 2010.

[58] S. Jokar. Sparse Recovery and Kronecker Products. Preprint 58, DFG-SPP 1324, August 2010.

[59] T. Aboiyar, E. H. Georgoulis, and A. Iske. Adaptive ADER Methods Using Kernel-Based Polyharmonic Spline WENO Reconstruction. Preprint 59, DFG-SPP 1324, August 2010.

[60] O. G. Ernst, A. Mugler, H.-J. Starkloff, and E. Ullmann. On the Convergence of Generalized Polynomial Chaos Expansions. Preprint 60, DFG-SPP 1324, August 2010.

[61] S. Holtz, T. Rohwedder, and R. Schneider. On Manifolds of Tensors of Fixed TT-Rank. Preprint 61, DFG-SPP 1324, September 2010.

[62] J. Ballani, L. Grasedyck, and M. Kluge. Black Box Approximation of Tensors in Hierarchical Tucker Format. Preprint 62, DFG-SPP 1324, October 2010.

[63] M. Hansen. On Tensor Products of Quasi-Banach Spaces. Preprint 63, DFG-SPP 1324, October 2010.

[64] S. Dahlke, G. Steidl, and G. Teschke. Shearlet Coorbit Spaces: Compactly Supported Analyzing Shearlets, Traces and Embeddings. Preprint 64, DFG-SPP 1324, October 2010.

[65] W. Hackbusch. Tensorisation of Vectors and their Efficient Convolution. Preprint 65, DFG-SPP 1324, November 2010.

[66] P. A. Cioica, S. Dahlke, S. Kinzel, F. Lindner, T. Raasch, K. Ritter, and R. L. Schilling. Spatial Besov Regularity for Stochastic Partial Differential Equations on Lipschitz Domains. Preprint 66, DFG-SPP 1324, November 2010.

[67] E. Novak and H. Woźniakowski. On the Power of Function Values for the Approximation Problem in Various Settings. Preprint 67, DFG-SPP 1324, November 2010.

[68] A. Hinrichs, E. Novak, and H. Woźniakowski. The Curse of Dimensionality for Monotone and Convex Functions of Many Variables. Preprint 68, DFG-SPP 1324, November 2010.

[69] G. Kutyniok and W.-Q Lim. Image Separation Using Shearlets. Preprint 69, DFG-SPP 1324, November 2010.

[70] B. Jin and P. Maass. An Analysis of Electrical Impedance Tomography with Applications to Tikhonov Regularization. Preprint 70, DFG-SPP 1324, December 2010.

[71] S. Holtz, T. Rohwedder, and R. Schneider. The Alternating Linear Scheme for Tensor Optimisation in the TT Format. Preprint 71, DFG-SPP 1324, December 2010.

[72] T. Müller-Gronbach and K. Ritter. A Local Refinement Strategy for Constructive Quantization of Scalar SDEs. Preprint 72, DFG-SPP 1324, December 2010.

[73] T. Rohwedder and R. Schneider. An Analysis for the DIIS Acceleration Method used in Quantum Chemistry Calculations. Preprint 73, DFG-SPP 1324, December 2010.

[74] C. Bender and J. Steiner. Least-Squares Monte Carlo for Backward SDEs. Preprint 74, DFG-SPP 1324, December 2010.

[75] C. Bender. Primal and Dual Pricing of Multiple Exercise Options in Continuous Time. Preprint 75, DFG-SPP 1324, December 2010.

[76] H. Harbrecht, M. Peters, and R. Schneider. On the Low-rank Approximation by the Pivoted Cholesky Decomposition. Preprint 76, DFG-SPP 1324, December 2010.

[77] P. A. Cioica, S. Dahlke, N. Döhring, S. Kinzel, F. Lindner, T. Raasch, K. Ritter, and R. L. Schilling. Adaptive Wavelet Methods for Elliptic Stochastic Partial Differential Equations. Preprint 77, DFG-SPP 1324, January 2011.

[78] G. Plonka, S. Tenorth, and A. Iske. Optimal Representation of Piecewise Hölder Smooth Bivariate Functions by the Easy Path Wavelet Transform. Preprint 78, DFG-SPP 1324, January 2011.

[79] A. Mugler and H.-J. Starkloff. On Elliptic Partial Differential Equations with Random Coefficients. Preprint 79, DFG-SPP 1324, January 2011.

[80] T. Müller-Gronbach, K. Ritter, and L. Yaroslavtseva. A Derandomization of the Euler Scheme for Scalar Stochastic Differential Equations. Preprint 80, DFG-SPP 1324, January 2011.

[81] W. Dahmen, C. Huang, C. Schwab, and G. Welper. Adaptive Petrov-Galerkin methods for first order transport equations. Preprint 81, DFG-SPP 1324, January 2011.

[82] K. Grella and C. Schwab. Sparse Tensor Spherical Harmonics Approximation in Radiative Transfer. Preprint 82, DFG-SPP 1324, January 2011.

[83] D.A. Lorenz, S. Schiffler, and D. Trede. Beyond Convergence Rates: Exact Inversion With Tikhonov Regularization With Sparsity Constraints. Preprint 83, DFG-SPP 1324, January 2011.

[84] S. Dereich, M. Scheutzow, and R. Schottstedt. Constructive quantization: Approximation by empirical measures. Preprint 84, DFG-SPP 1324, January 2011.

[85] S. Dahlke and W. Sickel. On Besov Regularity of Solutions to Nonlinear Elliptic Partial Differential Equations. Preprint 85, DFG-SPP 1324, January 2011.

[86] S. Dahlke, U. Friedrich, P. Maass, T. Raasch, and R.A. Ressel. An adaptive wavelet method for parameter identification problems in parabolic partial differential equations. Preprint 86, DFG-SPP 1324, January 2011.

[87] A. Cohen, W. Dahmen, and G. Welper. Adaptivity and Variational Stabilization for Convection-Diffusion Equations. Preprint 87, DFG-SPP 1324, January 2011.

[88] T. Jahnke. On Reduced Models for the Chemical Master Equation. Preprint 88, DFG-SPP 1324, January 2011.

[89] P. Binev, W. Dahmen, R. DeVore, P. Lamby, D. Savu, and R. Sharpley. Compressed Sensing and Electron Microscopy. Preprint 89, DFG-SPP 1324, March 2011.

[90] P. Binev, F. Blanco-Silva, D. Blom, W. Dahmen, P. Lamby, R. Sharpley, and T. Vogt. High Quality Image Formation by Nonlocal Means Applied to High-Angle Annular Dark Field Scanning Transmission Electron Microscopy (HAADF-STEM). Preprint 90, DFG-SPP 1324, March 2011.

[91] R. A. Ressel. A Parameter Identification Problem for a Nonlinear Parabolic Differential Equation. Preprint 91, DFG-SPP 1324, May 2011.

[92] G. Kutyniok. Data Separation by Sparse Representations. Preprint 92, DFG-SPP 1324, May 2011.

[93] M. A. Davenport, M. F. Duarte, Y. C. Eldar, and G. Kutyniok. Introduction to Compressed Sensing. Preprint 93, DFG-SPP 1324, May 2011.

[94] H.-C. Kreusler and H. Yserentant. The Mixed Regularity of Electronic Wave Functions in Fractional Order and Weighted Sobolev Spaces. Preprint 94, DFG-SPP 1324, June 2011.

[95] E. Ullmann, H. C. Elman, and O. G. Ernst. Efficient Iterative Solvers for Stochastic Galerkin Discretizations of Log-Transformed Random Diffusion Problems. Preprint 95, DFG-SPP 1324, June 2011.

[96] S. Kunis and I. Melzer. On the Butterfly Sparse Fourier Transform. Preprint 96, DFG-SPP 1324, June 2011.

[97] T. Rohwedder. The Continuous Coupled Cluster Formulation for the Electronic Schrödinger Equation. Preprint 97, DFG-SPP 1324, June 2011.

[98] T. Rohwedder and R. Schneider. Error Estimates for the Coupled Cluster Method. Preprint 98, DFG-SPP 1324, June 2011.

[99] P. A. Cioica and S. Dahlke. Spatial Besov Regularity for Semilinear Stochastic Partial Differential Equations on Bounded Lipschitz Domains. Preprint 99, DFG-SPP 1324, July 2011.

[100] L. Grasedyck and W. Hackbusch. An Introduction to Hierarchical (H-) Rank and TT-Rank of Tensors with Examples. Preprint 100, DFG-SPP 1324, August 2011.

[101] N. Chegini, S. Dahlke, U. Friedrich, and R. Stevenson. Piecewise Tensor Product Wavelet Bases by Extensions and Approximation Rates. Preprint 101, DFG-SPP 1324, September 2011.

[102] S. Dahlke, P. Oswald, and T. Raasch. A Note on Quarkonial Systems and Multilevel Partition of Unity Methods. Preprint 102, DFG-SPP 1324, September 2011.

[103] A. Uschmajew. Local Convergence of the Alternating Least Squares Algorithm For Canonical Tensor Approximation. Preprint 103, DFG-SPP 1324, September 2011.

[104] S. Kvaal. Multiconfigurational time-dependent Hartree method for describing particle loss due to absorbing boundary conditions. Preprint 104, DFG-SPP 1324, September 2011.

[105] M. Guillemard and A. Iske. On Groupoid C*-Algebras, Persistent Homology and Time-Frequency Analysis. Preprint 105, DFG-SPP 1324, September 2011.

[106] A. Hinrichs, E. Novak, and H. Woźniakowski. Discontinuous information in the worst case and randomized settings. Preprint 106, DFG-SPP 1324, September 2011.

[107] M. Espig, W. Hackbusch, A. Litvinenko, H. Matthies, and E. Zander. Efficient Analysis of High Dimensional Data in Tensor Formats. Preprint 107, DFG-SPP 1324, September 2011.

[108] M. Espig, W. Hackbusch, S. Handschuh, and R. Schneider. Optimization Problems in Contracted Tensor Networks. Preprint 108, DFG-SPP 1324, October 2011.

[109] S. Dereich, T. Müller-Gronbach, and K. Ritter. On the Complexity of Computing Quadrature Formulas for SDEs. Preprint 109, DFG-SPP 1324, October 2011.

[110] D. Belomestny. Solving optimal stopping problems by empirical dual optimization and penalization. Preprint 110, DFG-SPP 1324, November 2011.

[111] D. Belomestny and J. Schoenmakers. Multilevel dual approach for pricing American style derivatives. Preprint 111, DFG-SPP 1324, November 2011.

[112] T. Rohwedder and A. Uschmajew. Local convergence of alternating schemes for optimization of convex problems in the TT format. Preprint 112, DFG-SPP 1324, December 2011.

[113] T. Görner, R. Hielscher, and S. Kunis. Efficient and accurate computation of spherical mean values at scattered center points. Preprint 113, DFG-SPP 1324, December 2011.

[114] Y. Dong, T. Görner, and S. Kunis. An iterative reconstruction scheme for photoacoustic imaging. Preprint 114, DFG-SPP 1324, December 2011.

[115] L. Kämmerer. Reconstructing hyperbolic cross trigonometric polynomials by sampling along generated sets. Preprint 115, DFG-SPP 1324, February 2012.

[116] H. Chen and R. Schneider. Numerical analysis of augmented plane waves methods for full-potential electronic structure calculations. Preprint 116, DFG-SPP 1324, February 2012.

[117] J. Ma, G. Plonka, and M.Y. Hussaini. Compressive Video Sampling with Approximate Message Passing Decoding. Preprint 117, DFG-SPP 1324, February 2012.

[118] D. Heinen and G. Plonka. Wavelet shrinkage on paths for scattered data denoising. Preprint 118, DFG-SPP 1324, February 2012.

[119] T. Jahnke and M. Kreim. Error bound for piecewise deterministic processes modeling stochastic reaction systems. Preprint 119, DFG-SPP 1324, March 2012.

[120] C. Bender and J. Steiner. A-posteriori estimates for backward SDEs. Preprint 120, DFG-SPP 1324, April 2012.

[121] M. Espig, W. Hackbusch, A. Litvinenkoy, H.G. Matthiesy, and P. Wähnert. Efficient low-rank approximation of the stochastic Galerkin matrix in tensor formats. Preprint 121, DFG-SPP 1324, May 2012.

[122] O. Bokanowski, J. Garcke, M. Griebel, and I. Klompmaker. An adaptive sparse grid semi-Lagrangian scheme for first order Hamilton-Jacobi Bellman equations. Preprint 122, DFG-SPP 1324, June 2012.

[123] A. Mugler and H.-J. Starkloff. On the convergence of the stochastic Galerkin method for random elliptic partial differential equations. Preprint 123, DFG-SPP 1324, June 2012.

[124] P.A. Cioica, S. Dahlke, N. Döhring, U. Friedrich, S. Kinzel, F. Lindner, T. Raasch, K. Ritter, and R.L. Schilling. On the convergence analysis of Rothe's method. Preprint 124, DFG-SPP 1324, July 2012.