Fully discrete, quasi non-conforming approximation of evolution equations

Luigi C. Berselli · Alex Kaltenbach · Michael Růžička

Abstract In this paper we consider a fully discrete approximation of an abstract evolution equation, by means of a quasi non-conforming space approximation and finite differences in time (Rothe-Galerkin method). The main result is the convergence of the discrete solutions to weak solutions of the continuous problem. Hence, the result can be interpreted either as a justification of the numerical method or as an alternative way of constructing weak solutions.

We set the problem in the very general and abstract setting of pseudo-monotone operators, which allows for a unified treatment of several evolution problems. Nevertheless, the paradigmatic example—which fits into our setting and which originated our research—is represented by the \( p \)-Navier-Stokes equations, since the quasi non-conforming approximation allows to handle problems with prescribed divergence.

Our abstract results for pseudo-monotone operators allow to show convergence just by verifying a few natural assumptions on the monotone operator (and its compact perturbation) time-by-time and on the discretization spaces. Hence, applications and extensions to several other evolution problems can be easily performed. The results of some numerical experiments are reported in the final section.

Keywords Fully discrete · Pseudo-monotone operator · Evolution equation
1 Introduction

We consider the numerical approximation of an abstract evolution equation

\[
\begin{align*}
\frac{dx}{dt}(t) + A(t)(x(t)) &= f(t) \quad \text{in } V^*, \\
x(0) &= x_0 \quad \text{in } H,
\end{align*}
\]

by means of a quasi non-conforming Rothe-Galerkin scheme. Here, \(V \to H \cong H^* \to V^*\) is a given evolution triple, \(I := (0, T)\) a finite time horizon, \(x_0 \in H\) an initial value, \(f \in L^p(I, V^*), 1 < p < \infty\), a right-hand side and \(A(t) : V \to V^*, t \in I\), a family of operators. If not specified differently, we will denote in boldface elements of Bochner spaces (as the solution \(x(t)\) and the external source \(f(t)\)), to highlight the difference with elements belonging to standard Banach spaces, denoted by usual symbols.

In order to make (1.1) accessible for non-conforming approximation methods, we will additionally require that there exists a further evolution triple \(X \to Y \cong Y^* \to X^*\), such that \(V \subseteq X\) with \(\| \cdot \|_V = \| \cdot \|_X\) in \(V\) and \(H \subseteq Y\) with \(\langle \cdot, \cdot \rangle_H = \langle \cdot, \cdot \rangle_Y\) in \(H\), and extensions \(\hat{A}(t) : X \to X^*, t \in I,\) and \(\mathbf{f} \in L^p(I, X^*)\) of \(\{A(t)\}_{t \in I}\) and \(f\), resp., i.e., \((\hat{A}(t)v, w)_X = (A(t)v, w)_Y\) and \((\mathbf{f}(t), v)_X = (f(t), v)_Y\) for all \(v, w \in V\) and almost every \(t \in I\). For sake of readability we set \(\hat{A}(t) := \hat{A}(t)\) and \(\mathbf{f}(t) := \mathbf{f}(t)\) for almost every \(t \in I\).

The main aim of this paper is to extend the abstract framework of Bochner pseudo-monotone operators to handle also problems coming from the analysis of incompressible non-Newtonian fluids. As a prototypical application of the abstract results we will consider a fully discrete Rothe-Galerkin scheme for the unsteady \(p\)-Navier-Stokes equations. This is a system describing the unsteady motion in the time interval \(I = (0, T)\). The motion is governed by the following initial boundary value problem

\[
\begin{align*}
\partial_t u - \text{div} S(\cdot, \cdot, \nabla u) + \text{div}(u \otimes u) + \nabla \pi &= \mathbf{f} \quad \text{in } Q_T, \\
\text{div} u &= 0 \quad \text{in } Q_T, \\
u &= 0 \quad \text{on } \Gamma_T, \\
u(0) &= u_0 \quad \text{in } \Omega.
\end{align*}
\]

Here, \(Q_T := I \times \Omega\) denotes a time-space cylinder, \(\Gamma_T := I \times \partial \Omega\), \(u : Q_T \to \mathbb{R}^d\) denotes the velocity, \(\mathbf{f} : Q_T \to \mathbb{R}^d\) is a given external force, \(u_0 : \Omega \to \mathbb{R}^d\) an initial condition, \(\pi : Q_T \to \mathbb{R}\) the pressure and \(\nabla u := \frac{1}{2} (\nabla u + \nabla u^\top)\) the symmetric gradient.

The mapping \(S : Q_T \times \mathbb{M}_{\text{sym}}^{d \times d} \to \mathbb{M}_{\text{sym}}^{d \times d}\) is supposed to possess a \((p, \delta)\)-structure, i.e., for some \(p \in (1, \infty)\) and \(\delta \geq 0\), the following properties are satisfied:

(S.1) \(S : Q_T \times \mathbb{M}_{\text{sym}}^{d \times d} \to \mathbb{M}_{\text{sym}}^{d \times d}\) is a Carathéodory mapping.\(^1\)

(S.2) \(|S(t, x, A)| \leq \alpha(\delta + |A|)^{p-2}|A| + \beta\) for all \(A \in \mathbb{M}_{\text{sym}}^{d \times d}\), a.e. \((t, x)^\top \in Q_T\). \((\alpha > 0, \beta \geq 0)\).

\(^1\) \(\mathbb{M}_{\text{sym}}^{d \times d}\) is the vector space of all symmetric \(d \times d\) matrices \(A = (A_{ij})_{i,j = 1, \ldots, d}\). We equip \(\mathbb{M}_{\text{sym}}^{d \times d}\) with the scalar product \(A : B := \sum_{i,j = 1}^{d} A_{ij} B_{ij}\) and the norm \(|A| := (A : A)^{\frac{1}{2}}\). By \(a \cdot b\) we denote the usual scalar product in \(\mathbb{R}^d\) and by \(|a|\) we denote the Euclidean norm.

\(^2\) \(S(\cdot, \cdot, A) : Q_T \to \mathbb{M}_{\text{sym}}^{d \times d}\) is Lebesgue measurable for every \(A \in \mathbb{M}_{\text{sym}}^{d \times d}\) and \(S(t, x, \cdot) : \mathbb{M}_{\text{sym}}^{d \times d} \to \mathbb{M}_{\text{sym}}^{d \times d}\) is continuous for almost every \((t, x)^\top \in Q_T\).
Moreover, the operator family $(S(t, x, A) : A \geq c_0(\delta + |A|)^{p-2}|A|^2 - c_1$ for all $A \in M^{d \times d}_{\mathrm{sym}},$ a.e. $(t, x) \in Q_T$ ($c_0 > 0, c_1 \geq 0$).

$(S_4)$ $(S(t, x, A) - S(t, x, B)) : (A - B) \geq 0$ for all $A, B \in M^{d \times d}_{\mathrm{sym}},$ a.e. $(t, x) \in Q_T$.

We define for $p > \frac{3d + 2}{d+2}$ the function spaces $X := W_0^{1,p}(\Omega)^d, Y := L^2(\Omega)^d, V := W_0^{1, p, \mathrm{div}}(\Omega)$ as the closure of $\mathcal{V} := \{ v \in C_0^\infty(\Omega)^d \mid \mathrm{div} v = 0 \}$ in $X,$ $H := L^2_{\mathrm{div}}(\Omega)$ as the closure of $\mathcal{V}$ in $Y,$ and the families of operators $S(t), B : X \rightarrow X^*, t \in I,$ for all $u, v \in X$ and almost every $t \in I$ via

$$
\langle S(t)u, v \rangle_X := \int_{\Omega} S(t, \cdot, Du) : Dv \, dx \quad \text{and} \quad \langle Bu, v \rangle_X := -\int_{\Omega} u \otimes u : Dv \, dx.
$$

Then, $(1.2)$ for $u_0 \in H$ and $f \in L^p(I, X^*)$ can be re-written as the abstract evolution equation

$$
\frac{du}{dt}(t) + S(t)(u(t)) + B(u(t)) = f(t) \quad \text{in } V^*,
$$

where, in the notation of $(1.1),$ we set $A(t) := S(t) + B : X \rightarrow X^*$ for almost every $t \in I.$

As the construction of finite element spaces $(V_n)_{n \in \mathbb{N}},$ which meet the divergence constraint, i.e., satisfy $V_n \subseteq V$ for all $n \in \mathbb{N},$ highly restricts the flexibility of the approximation, one usually forgoes to work with divergence-free finite element spaces and imposes a discrete divergence constraint instead, naturally suggesting the usage of non-conforming spaces, i.e., $V_n \not\subseteq V,$ but the spaces $V_n$ are immersed in a larger ambient space $X.$

1.1 The numerical scheme

A quasi non-conforming Rothe-Galerkin approximation of the initial value problem $(1.1)$ usually consists of two parts:

The first part is a spatial discretization, often called Galerkin approximation, which consists in the approximation of $V$ by a sequence of closed subspaces $(V_n)_{n \in \mathbb{N}}$ of $X.$ We emphasize that we do not require $(V_n)_{n \in \mathbb{N}}$ to be a sequence of subspaces of $V,$ which motivates the prefix non-conforming. Hence, we do not have $V_n \subseteq V$ and $V_n \uparrow V$ (approximation from below). The prefix quasi indicates that in contrast to a fully non-conforming spatial approximation, where the norm on $V_n$ depends on $n,$ here the subspaces $V_n$ are equipped with the norm of a space $X$ such that

$$
V_n \not\subseteq V, \quad V \subseteq X, \quad V_n \subseteq X.
$$

So in this case we have, under appropriate assumptions, a sort of approximation from above of the space $V.$

The second part is a temporal discretization, also called Rothe scheme, which consists in the approximation of the unsteady problem $(1.1)$ by a sequence of piece-wise constant, steady problems. This is achieved by replacing the time derivative $\frac{d}{dt}$ by so-called backwards difference quotients, which are for a given step-size $\tau := \frac{T}{K} > 0,$ where $K \in \mathbb{N},$ and a given finite sequence $(x^k)_{k=0,...,K} \subseteq X$ defined via

$$
d_\tau x^k := \frac{1}{\tau}(x^k - x^{k-1}) \quad \text{in } X \quad \text{for all } k = 1, \ldots, K.
$$

Moreover, the operator family $A(t) : X \rightarrow X^*, t \in I,$ and the right-hand side $f \in L^p(I, X^*)$ need to be discretized and this is obtained by means of the Clement 0-order quasi interpolant, i.e., for
a given step-size $\tau = \frac{T}{K} > 0$, where $K \in \mathbb{N}$, we replace them piece-wise by their local temporal means, i.e., by $[A]_{k}^{\tau}: X \to X^{*}$, $k = 1, \ldots, K$, and $([f]_{k}^{\tau})_{k=1,\ldots,K} \subseteq X^{*}$, resp., for every $k = 1, \ldots, K$ and $x \in X$ given via

$$[A]_{k}^{\tau}x := \int_{\tau(k-1)}^{\tau k} A(t)x \, dt, \quad [f]_{k}^{\tau} := \int_{\tau(k-1)}^{\tau k} f(t) \, dt \quad \text{in } X^{*}.$$ 

Altogether, using these two levels of approximation, we formulate the following fully discrete or Rothe-Galerkin scheme of the evolution problem (1.1):

\[ \text{Algorithm 1.5 (quasi non-conforming Rothe-Galerkin scheme)} \]

For given $K,n \in \mathbb{N}$ and $x_{n}^{0} \in V_{n}$ the sequence of iterates $(x_{n})_{k=0,\ldots,K} \subseteq V_{n}$ is given via the implicit scheme for $\tau = \frac{T}{K}$

\begin{equation}
(d_{\tau} x_{n}^{k},v_{n})_{Y} + ([A]_{k}^{\tau} x_{n}^{k},v_{n})_{X} = ([f]_{k}^{\tau},v_{n})_{X} \quad \text{for all } v_{n} \in V_{n}.
\end{equation}

As soon as the existence of the iterates $(x_{n}^{k})_{k=0,\ldots,K} \subseteq V_{n}$, solving the scheme (1.6), is proved for a sufficiently small step-size $\tau = \frac{T}{K} \in (0,\tau_{0})$, where $\tau_{0} > 0$, and $K,n \in \mathbb{N}$, one can check whether the resulting family of piece-wise constant interpolants $\varpi_{n}^{\tau} \in L^{\infty}(I,V_{n})$, $K,n \in \mathbb{N}$ with $\tau = \frac{T}{K} \in (0,\tau_{0})$ (cf. (5.1)), converges towards a weak solution of (1.1), at least in an appropriate weak sense.

The main abstract result is then the following one (see Section 6 for notation and proofs), showing the convergence of the fully discrete approximate solutions.

\[ \text{Theorem 1.7 Let Assumption (6.1) be satisfied and set } \tau_{0} := \frac{1}{4\lambda 1}. \text{ If } (\varpi_{n})_{n\in\mathbb{N}} := (\varpi_{n}^{\tau_{n}})_{n\in\mathbb{N}} \subseteq L^{\infty}(I,X) \text{, where } \tau_{n} = \frac{T}{K_{n}} \text{ and } K_{n},m_{n} \to \infty \text{ (} n \to \infty \text{), is an arbitrary diagonal sequence of piece-wise constant interpolants } \varpi_{n}^{\tau} \in S^{1}(I,X), \text{ } K,n \in \mathbb{N} \text{ with } \tau = \frac{T}{K} \in (0,\tau_{0}) \text{, from Proposition 6.8. Then, there exists a not relabelled subsequence and a weak limit } \varpi \in L^{p}(I,V) \cap \chi L^{\infty}(I,H) \text{ such that}
\]

\begin{equation}
\varpi_{n} \to \varpi \quad \text{in } L^{p}(I,X),
\end{equation}

\begin{equation}
\varpi_{n} \overset{*}{\to} \varpi \quad \text{in } L^{\infty}(I,Y), \quad (n \to \infty).
\end{equation}

Furthermore, it follows that $\varpi \in \mathcal{W}_{e}^{1,p,q}(I,V,H)$ is a weak solution of the initial value problem (1.1).

Traditionally, the verification of the weak convergence of a Rothe-Galerkin scheme like (1.6) to a weak solution of the evolution equation (1.1) causes a certain effort, and in the case of quasi non-conforming approximations like (1.6) to the best knowledge of the author’s there are no abstract results guaranteeing the weak convergence of such a scheme. Therefore, this article’s purpose is (i) to give general and easily verifiable assumptions on both the operator family $A(t): X \to X^{*}$, $t \in I$, and the sequence of approximative spaces $(V_{n})_{n\in\mathbb{N}}$ which provide both the existence of iterates $(x_{n}^{k})_{k=0,\ldots,K} \subseteq V_{n}$, solving (1.6), for a sufficiently small step-size $\tau = \frac{T}{K} \in (0,\tau_{0})$, where $\tau_{0} > 0$, and $K,n \in \mathbb{N}$; (ii) to prove the stability of the scheme, i.e., the boundedness of the piece-wise constant interpolants $\varpi_{n}^{\tau} \in L^{\infty}(I,V_{n})$, $K,n \in \mathbb{N}$ with $\tau = \frac{T}{K} \text{, given via (5.1), in } L^{p}(I,X) \cap \chi L^{\infty}(I,Y)$; (iii) to show the weak convergence of a diagonal subsequence $(\varpi_{n}^{\tau_{n}})_{n\in\mathbb{N}} \subseteq L^{\infty}(I,X)$, where $\tau_{n} = \frac{T}{K_{n}}$ and $K_{n},m_{n} \to \infty \text{ (} n \to \infty \text{), towards a weak solution of (1.1)}$.

A common approach is to require that $(V_{n})_{n\in\mathbb{N}}$ forms a conforming approximation of $V$, i.e., satisfies the following two conditions:

\[ \text{A common approach is to require that } (V_{n})_{n\in\mathbb{N}} \text{ forms a conforming approximation of } V, \text{ i.e., satisfies the following two conditions:}
\]
(C.1) $(V_n)_{n \in \mathbb{N}}$ is an increasing sequence of closed subspaces of $V$, i.e., $V_n \subseteq V_{n+1} \subseteq V$ for all $n \in \mathbb{N}$. (C.2) $\bigcup_{n \in \mathbb{N}} V_n$ is dense in $V$.

In particular, (C.1) and (C.2) allow us to choose $X = V$ above. Surprisingly, even in the conforming case there are only few contributions with a rigorous convergence analysis of fully discrete Rothe-Galerkin schemes towards weak solutions. Most authors consider only semi-discrete schemes, i.e., either a pure Rothe scheme (cf. [27]) or a pure Galerkin scheme (cf. [18], [36], [30]). Much more results are concerned with explicit convergence rates for more regular data and more regular solutions (cf. [3], [28], [25], [17], [15], [26], [13], [5], [16], [4], [10]). Concerning the convergence analysis of a conforming Rothe-Galerkin scheme we are only aware of the early contribution [1] which treats the porous media equations and the recent contributions [34], [31] treating the $p$-Navier-Stokes equations and [6] dealing with a similar setting as the present paper. In fact, if $A(t) : V \to V^*$, $t \in I$, satisfies appropriate assumptions, e.g., [23, condition (C.1)–(C.4)], one can easily verify the existence of the iterates $(x^n_k)_{k=0,\ldots,K} \subseteq V_n$, for sufficiently small $\tau = \frac{T}{k} \in (0,\tau_0)$, where $\tau_0 > 0$, and $K,n \in \mathbb{N}$, solving (1.6), the boundedness of the corresponding family of piece-wise constant interpolants $\bar{x}_n^k \in L^\infty(I,V)$, $K,n \in \mathbb{N}$ with $\tau = \frac{T}{k} \in (0,\tau_0)$, in $L^p(I,V) \cap L^\infty(I,H)$, and the existence of a diagonal subsequence $(\bar{x}_n^k)_{n \in \mathbb{N}} := (\bar{x}_{n,m}^k)_{n \in \mathbb{N}} \subseteq L^\infty(I,V)$, where $\tau_n = \frac{T}{K_n}$ and $K_n,m_n \to \infty$ ($n \to \infty$), and an element $\bar{x} \in L^p(I,V) \cap L^\infty(I,H)$, such that

$$\bar{x}_n^k \to \bar{x} \quad \text{in } L^p(I,V) \quad (n \to \infty),$$

$$\bar{x}_n^k \to \bar{x} \quad \text{in } L^\infty(I,H) \quad (n \to \infty),$$

$$\limsup_{n \to \infty} \langle A\bar{x}_n^k, \bar{x}_n^k - \bar{x} \rangle_{L^p(I,V)} \leq 0,$$

where $A : L^p(I,V) \cap L^\infty(I,H) \to (L^p(I,V))^*$ denotes the induced operator, which is for every $\bar{x} \in L^p(I,V) \cap L^\infty(I,H)$ and $y \in L^p(I,V)$ given via $\langle A\bar{x}, y \rangle_{L^p(I,V)} := \int_I (A(t)(\bar{x}(t)), y(t))_V dt$. Apart from that, using methods in [6], we can extract from the scheme (1.6) by means of (1.8) and (1.9) the additional convergence

$$\bar{x}_n(t) \to \bar{x}(t) \quad \text{in } H \quad (n \to \infty) \quad \text{for a.e. } t \in I.$$  

In this context, in [23] it is proved that $A : L^p(I,V) \cap L^\infty(I,H) \to (L^p(I,V))^*$ is Bochner pseudo-monotone, i.e., from (1.8)–(1.11) it follows for every $y \in L^p(I,V)$

$$\langle A\bar{x}, \bar{x} - y \rangle_{L^p(I,V)} \leq \liminf_{n \to \infty} \langle A\bar{x}_n, \bar{x}_n - y \rangle_{L^p(I,V)},$$

which in a standard manner leads to $A\bar{x}_n \to A\bar{x}$ in $(L^p(I,V))^*$ ($n \to \infty$), and therefore to the weak convergence of the scheme (1.6).

Without the conditions (C.1) and (C.2), i.e., $V \neq X$ and $V_n \nsubseteq V$ for all $n \in \mathbb{N}$, but the family $A(t) : X \to X^*$, $t \in I$, still satisfying appropriate assumptions, e.g., [23, condition (C.1)–(C.4)], the situation changes dramatically. Even though we can probably prove the existence of iterates $(x^n_k)_{k=0,\ldots,K} \subseteq V_n$, solving (1.6), for sufficiently small $\tau = \frac{T}{k} \in (0,\tau_0)$, where $\tau_0 > 0$, and $K,n \in \mathbb{N}$, the boundedness of the corresponding family of piece-wise constant interpolants $\bar{x}_n^k \in L^\infty(I,X)$, $K,n \in \mathbb{N}$ with $\tau = \frac{T}{k} \in (0,\tau_0)$, in $L^p(I,X) \cap L^\infty(I,Y)$, and the weak convergence of a diagonal subsequence $(\bar{x}_n^k)_{n \in \mathbb{N}} := (\bar{x}_{n,m}^k)_{n \in \mathbb{N}} \subseteq L^\infty(I,X)$, where $\tau_n = \frac{T}{K_n}$ and $K_n,m_n \to \infty$ ($n \to \infty$), to a weak limit $\bar{x} \in L^p(I,X) \cap L^\infty(I,Y)$, we can solely expect that

$$\bar{x}_n^k \to \bar{x} \quad \text{in } L^p(I,X),$$

$$\bar{x}_n^k \to \bar{x} \quad \text{in } L^\infty(I,Y), \quad (n \to \infty).$$  

(1.12)
Without any further assumptions on the spatial approximation \((V_n)_{n \in \mathbb{N}}\) it is not even clear, whether the weak limit lies in the right function space, i.e., whether \(\bar{\mathbf{x}} \in L^p(I, V) \cap L^\infty(I, H)\). In addition, we are neither aware of whether an inequality like
\[
\limsup_{n \to \infty} \langle A\mathbf{x}_n, \mathbf{x}_n - \bar{\mathbf{x}} \rangle_{L^p(I, X)} \leq 0
\]
is satisfied, nor whether we can extract from (1.6) by means of (1.12) an additional convergence similar to (1.11). To guarantee the latter, we make the following assumptions on \((V_n)_{n \in \mathbb{N}}\):

- \textbf{(QNC.1)} There exists a dense subset \(D \subseteq V\), such that for each \(v \in D\) there exists elements \(v_n \in V_n, n \in \mathbb{N}\), such that \(v_n \to v\) in \(X\) \((n \to \infty)\).

- \textbf{(QNC.2)} For each sequence \(\mathbf{x}_n \in L^p(I, V_{m_n})\), \(n \in \mathbb{N}\), where \((m_n)_{n \in \mathbb{N}} \subseteq \mathbb{N}\) with \(m_n \to \infty\) \((n \to \infty)\), from \(\mathbf{x}_n \to \mathbf{x}\) in \(L^p(I, X)\) \((n \to \infty)\), it follows that \(\mathbf{x} \in L^p(I, V)\).

In fact, using (QNC.1) and (QNC.2) in this simpler setting, we are able to derive from (1.12)\(_1\) that \(\bar{\mathbf{x}} \in L^p(I, V) \cap L^\infty(I, H)\), the inequality (1.13) and the additional convergence
\[
P_H(\mathbf{x}_n(t)) \to P_H(\mathbf{x}(t)) \quad \text{in} \quad H \quad (n \to \infty) \quad \text{for a.e.} \quad t \in I,
\]
where \(P_H : Y \to H\) denotes the orthogonal projection of \(Y\) into \(H\). Note that we have no information, whether \(\mathbf{x}_n(t) \to \mathbf{x}(t)\) in \(Y\) \((n \to \infty)\) for almost every \(t \in I\). In consequence, we cannot fall back on the approaches of [23], [6]. However, using anew (QNC.1) and (QNC.2), we are able to adapt and extend the methods in [23], [6], and deduce from (1.12)–(1.14) that for all \(y \in L^p(I, X)\)
\[
\langle A\mathbf{x}, \mathbf{x} - y \rangle_{L^p(I, X)} \leq \liminf_{n \to \infty} \langle A\mathbf{x}_n, \mathbf{x}_n - y \rangle_{L^p(I, X)};
\]
which leads to \(A\mathbf{x}_n \to A\mathbf{x}\) in \((L^p(I, X))^*\) \((n \to \infty)\), i.e., the convergence of the scheme (1.6).

### 1.2 The example of the \(p\)-Navier-Stokes equations

The case of the \(p\)-Navier-Stokes equations is the prototypical example that motivated the above abstract setting. The discretely divergence-free finite element approximation, introduced below, fits into the abstract setting of the previous section. Hence, the corresponding convergence Theorem 7.4 follows just by checking that the hypotheses of the abstract result are satisfied.

Let \(Z := L^p(\Omega)\) and for a given family of shape regular triangulations (cf. [11]) \((\mathcal{T}_h)_{h>0}\) of our polygonal Lipschitz domain \(\Omega\) and given \(m, \ell \in \mathbb{N}_0\), we denote by\(^3\) \(X_h \subset \mathcal{P}_m(\mathcal{T}_h)^d \cap X\) and \(Z_h \subset \mathcal{P}_\ell(\mathcal{T}_h) \cap Z\) appropriate finite element spaces. Note that we consider always continuous approximations for the velocity, while we allow for discontinuous approximations for the pressure. In addition, we define for \(h > 0\) the \textit{discretely divergence free finite element spaces}
\[
V_h := \{ \mathbf{v}_h \in X_h \mid \langle \text{div} \mathbf{v}_h, \eta_h \rangle_Z = 0 \quad \text{for all} \quad \eta_h \in Z_h \}.
\]
For a null sequence \((h_n)_{n \in \mathbb{N}} \subseteq (0, \infty)\) and \(V_n := V_{h_n}, n \in \mathbb{N}\), one usually formulates the following algorithm of a time-space discrete approximation of (1.2):

\(^3\) \(\mathcal{P}_m(\mathcal{T}_h), m \in \mathbb{N}_0\), denotes the space of possibly discontinuous scalar functions, which are polynomials of degree at most \(m\) on each simplex \(K \in \mathcal{T}_h\).
Algorithm 1.15 For given $K, n \in \mathbb{N}$ and $u^k_n \in V_n$ the sequence of iterates $(u^k_n)_{k=0,\ldots,K} \subseteq V_n$ is given via the implicit Rothe-Galerkin scheme for $\tau = \frac{T}{K}$

$$
(d_n u^k_n, v_n)_Y + (|S|^r_k u^k_n, v_n)_X + (\hat{B} u^k_n, v_n)_X = (|f|^r_k, v_n)_X \quad \text{for all } v_n \in V_n,
$$

(1.16)

where $\hat{B} : X \to X^*$ is given via $\langle \hat{B} u, w \rangle_X := \frac{1}{2} \int_{\Omega} u \otimes v : Du - \frac{1}{2} \int_{\Omega} u \otimes u : Dv$ for all $u, v \in X$.

The operator $\hat{B}$ can be viewed as an extension of $B$, as $\langle \hat{B} u, v \rangle_X = \langle B u, v \rangle_X$ for all $u, v \in V$, which in contrast to $B$ fulfills $\langle \hat{B} u, u \rangle_X = 0$ for all $u \in X$, and therefore guarantees the stability of the scheme (1.16).

The sequence $(V_n)_{n \in \mathbb{N}}$ violates the conditions (C.1) and (C.2). However, $(V_n)_{n \in \mathbb{N}}$ perfectly fits into the framework of quasi non-conforming approximations (cf. Proposition 3.2). To be more precise, we will see that the assumptions (QNC.1) and (QNC.2) on the discrete spaces $(V_n)_{n \in \mathbb{N}}$ are often fulfilled, e.g., if the following assumption on the existence of appropriate projection operators with respect to the discrete spaces $X_h$ and $Z_h$ is satisfied:

Assumption 1.17 (Projection operators) We assume that for every $h > 0$ it holds $\mathcal{P}_1(\mathcal{T}_h)^d \subset X_h$, $\mathcal{R} \subset Z_h$, and that there exist linear projection operators $\Pi^\text{div}_h : X \to X_h$ and $\Pi^Z_h : Z \to Z_h$ with the following properties:

(i) **Divergence preservation of $\Pi^\text{div}_h$ in $Z_h^*$**: It holds for all $w \in X$ and $\eta_h \in Z_h$

$$
\langle \text{div } w, \eta_h \rangle_Z = \langle \text{div } \Pi^\text{div}_h w, \eta_h \rangle_Z.
$$

(ii) **$W^{1,1}$-stability of $\Pi^\text{div}_h$**: There exists a constant $c > 0$, independent of $h > 0$, such that for every $w \in X$ and $K \in \mathcal{T}_h$

$$
\int_K |\Pi^\text{div}_h w| \, dx \leq c \int_K |w| \, dx + c h_K \int_{S_K} |\nabla w| \, dx.
$$

(iii) **$L^1$-stability of $\Pi^Z_h$**: There exists a constant $c > 0$, independent of $h > 0$, such that for every $\eta \in Z$ and $K \in \mathcal{T}_h$

$$
\int_K |\Pi^Z_h \eta| \, dx \leq c \int_{S_K} |\eta| \, dx.
$$

Certainly, the existence of projection operators $\Pi^\text{div}_h$ and $\Pi^Z_h$ satisfying Assumption 1.17 depends on the choice of $X_h$ and $Z_h$. It is shown in [12], [19], [20] that $\Pi^\text{div}_h$ exists for a variety of spaces $X_h$ and $Z_h$, which, e.g., include the Taylor-Hood, the conforming Crouzeix-Raviart, and the MINI element in dimension two and three. Projection operators $\Pi^Z_h$ satisfying Assumption 1.17 (iii) are e.g. the Clément interpolation operator (cf. [14]) and a version of the Scott-Zhang interpolation operator (cf. [29]). The abstract assumptions allow for an easy extension of our results to other choices of $X_h$ and $Z_h$ in future works.

Plan of the paper: In Section 2 we recall some basic definitions and results concerning the theory of pseudo-monotone operators and evolution equations. In Section 3 we introduce the concept of quasi non-conforming approximations. In Section 4 we introduce the quasi non-conforming...
Bochner pseudo-monotonicity, and give sufficient and easily verifiable conditions on families of operators such that the corresponding induced operator satisfies this concept. In Section 5 we recall some basic facts about the Rothe scheme. In Section 6 we formulate the scheme of a fully discrete, quasi non-conforming approximation of an evolution equation, prove that this scheme is well-defined, i.e., the existence of iterates, that the corresponding family of piece-wise constant interpolants satisfies certain a-priori estimates and that there exist a diagonal subsequence which weakly converges to a weak solution of the corresponding evolution equation. In Section 7 we apply this approximation scheme on the unsteady \( p \)-Navier-Stokes equations. In Section 8 we present some numerical experiments.

2 Preliminaries

2.1 Operators

For a Banach space \( X \) with norm \( \| \cdot \|_X \) we denote by \( X^* \) its dual space equipped with the norm \( \| \cdot \|_{X^*} \). The duality pairing is denoted by \( \langle \cdot, \cdot \rangle_X \). All occurring Banach spaces are assumed to be real.

Definition 2.1 Let \( X \) and \( Y \) be Banach spaces. The operator \( A : X \to Y \) is said to be

(i) bounded, if for all bounded subsets \( M \subseteq X \) the image \( A(M) \subseteq Y \) is bounded.

(ii) pseudo-monotone, if \( Y = X^* \), and for \( (x_n)_{n \in \mathbb{N}} \subseteq X \) from \( x_n \rightharpoonup x \) in \( X \) \((n \to \infty)\) and \( \limsup_{n \to \infty} \langle Ax_n, x_n - x \rangle_X \leq 0 \), it follows \( \langle Ax, x-y \rangle_X \leq \liminf_{n \to \infty} \langle Ax_n, x_n - y \rangle_X \) for every \( y \in X \).

(iii) coercive, if \( Y = X^* \) and \( \lim_{\|x\|_X \to \infty} \frac{\langle Ax, x \rangle_X}{\|x\|_X} = \infty \).

Proposition 2.2 If \( X \) is a reflexive Banach space and \( A : X \to X^* \) a bounded, pseudo-monotone and coercive operator, then \( R(A) = X^* \).

Proof See [36, Corollary 32.26].

Lemma 2.3 If \( X \) is a reflexive Banach space and \( A : X \to X^* \) a locally bounded and pseudo-monotone operator, then \( A \) is demi-continuous.

Proof See [36, Proposition 27.7].

2.2 Evolution equations

We call \((V,H,j)\) an evolution triple, if \( V \) is a reflexive Banach space, \( H \) a Hilbert space and \( j : V \to H \) a dense embedding, i.e., \( j \) is a linear, injective and bounded operator with \( j(V) \subseteq H \). Let \( R : H \to H^* \) be the Riesz isomorphism with respect to \( \langle \cdot, \cdot \rangle_H \). As \( j \) is a dense embedding the adjoint \( j^* : H^* \to V^* \) and therefore \( e := j^* R j : V \to V^* \) are embeddings as well. We call \( e \) the canonical embedding of \((V,H,j)\). Note that \( \langle ev, w \rangle_V = \langle jv, jw \rangle_H \) for all \( v, w \in V \).

(2.4)
For an evolution triple \((V, H, j)\), \(I := (0, T)\), \(T < \infty\), and \(1 \leq p \leq q \leq \infty\) we define operators \(j : L^p(I, V) \rightarrow L^p(I, H)\): \(x \rightarrow jx\) and \(j^* : L^q(I, H^*) \rightarrow L^q(I, V^*)\): \(y \rightarrow j^*y\), where \(jx\) and \(j^*y\) are for every \(x \in L^p(I, V)\) and \(y \in L^q(I, H^*)\) given via

\[
(jx)(t) := j(x(t)) \quad \text{in } H \quad \text{for a.e. } t \in I, \\
(j^*y)(t) := j^*(y(t)) \quad \text{in } V^* \quad \text{for a.e. } t \in I.
\]

It is shown in [23, Prop. 2.19] that both \(j\) and \(j^*\) are embeddings, which we call induced embeddings. In particular, note that we will use throughout the entire article bold letters, i.e., \(x\), to indicate that a function is a Bochner-Lebesgue function. Moreover, we define the intersection space

\[
L^p(I, V) \cap_j L^q(I, H) := \{x \in L^p(I, V) \mid jx \in L^q(I, H)\},
\]

which forms a Banach space equipped with the canonical sum norm

\[
\|\cdot\|_{L^p(I, V) \cap_j L^q(I, H)} := \|\cdot\|_{L^p(I, V)} + \|\cdot\|_{L^q(I, H)}.
\]

If \(1 < p \leq q < \infty\), then \(L^p(I, V) \cap_j L^q(I, H)\) is additionally reflexive. Furthermore, for each \(x^* \in (L^p(I, V) \cap_j L^q(I, H))^*\) there exist functions \(g \in L^p(I, V^*)\) and \(h \in L^q(I, H^*)\), such that for every \(x \in L^p(I, V) \cap_j L^q(I, H)\) it holds

\[
\langle x^*, x \rangle_{L^p(I, V) \cap_j L^q(I, H)} = \int_I \langle g(t) + (j^*h)(t), x(t) \rangle_V dt,
\]

and \(\|x^*\|_{(L^p(I, V) \cap_j L^q(I, H))^*} := \|g\|_{L^p(I, V^*)} + \|h\|_{L^q(I, H^*)}\), i.e., \((L^p(I, V) \cap_j L^q(I, H))^*\) is isometrically isomorphic to the sum \(L^p(I, V^*) + j^*(L^q(I, H^*))\) (cf. [18, Kapitel I, Bemerkung 5.13 & Satz 5.13]), which is a Banach space equipped with the norm

\[
\|f\|_{L^p(I, V^*) + j^*(L^q(I, H^*))} := \min_{h \in L^q(I, H^*)} \|g\|_{L^p(I, V^*)} + \|h\|_{L^q(I, H^*)},
\]

for \(f = g + j^*h\).

**Definition 2.6 (Generalized time derivative)** Let \((V, H, j)\) be an evolution triple, \(I := (0, T)\), \(T < \infty\), and \(1 < p \leq q < \infty\). A function \(x \in L^p(I, V) \cap_j L^q(I, H)\) possesses a **generalized derivative with respect to the canonical embedding** \(e\) of \((V, H, j)\) if there exists a function \(x^* \in L^p(I, V^*) + j^*(L^q(I, H^*))\) such that for all \(v \in V\) and \(\varphi \in C^\infty_0(I)\)

\[
- \int_I (j(x(s)), jv)_{H^*} \varphi'(s) ds = \int_I \langle x^*(s), v \rangle_V \varphi(s) ds.
\]

As this function \(x^* \in L^p(I, V^*) + j^*(L^q(I, H^*))\) is unique (cf. [35, Proposition 23.18]), \(\frac{dx}{dt} := x^*\) is well-defined. By

\[
W_{e^{L^p,q}}^1(I, V, H) := \left\{ x \in L^p(I, V) \cap_j L^q(I, H) \mid \exists \frac{dx}{dt} \in L^p(I, V^*) + j^*(L^q(I, H^*)) \right\}
\]

we denote the **Bochner-Sobolev space with respect to** \(e\).

**Proposition 2.7 (Formula of integration by parts)** Let \((V, H, j)\) be an evolution triple, \(I := (0, T)\), \(T < \infty\), and \(1 < p \leq q < \infty\). Then, it holds:
(i) The space $\mathcal{W}_{e}^{1,p,q}(I, V, H)$ forms a Banach space equipped with the norm

$$
\| \| \mathcal{W}_{e}^{1,p,q}(I, V, H) := \| \| L^{p}(I, V) \cap L^{q}(I, H) + \left\| \frac{d}{dt} \right\|_{L^{p}(I, V') + L^{q}(I, H')} .
$$

(ii) Given $x \in \mathcal{W}_{e}^{1,p,q}(I, V, H)$ the function $j_{c}x \in L^{q}(I, H)$ possesses a unique representation $j_{c}x \in C^{0}(\mathcal{T}, H)$, and the resulting mapping $j_{c} : \mathcal{W}_{e}^{1,p,q}(I, V, H) \rightarrow C^{0}(\mathcal{T}, H)$ is an embedding.

(iii) Generalized integration by parts formula: It holds

$$
\int_{t'}^{t} \left\langle \frac{d}{dt}(s), y(s) \right\rangle_{V} \ dt = \left\langle \left( (j_{c}, x)(s), (j_{c}, y)(s) \right)_{H} \right\rangle_{s = t'} - \int_{t'}^{t} \left\langle \frac{d}{dt}(s), x(s) \right\rangle_{V} \ dt
$$

for all $x, y \in \mathcal{W}_{e}^{1,p,q}(I, V, H)$ and $t, t' \in \mathcal{T}$ with $t' \leq t$.

Proof See [18, Kapitel IV, Satz 1.16 & Satz 1.17].

For an evolution triple $(V, H, j)$, $I := (0, T)$, $T < \infty$, and $1 < p \leq q < \infty$ we call an operator $\mathcal{A} : L^{p}(I, V) \cap L^{q}(I, H) \rightarrow (L^{p}(I, V) \cap L^{q}(I, H))^{*}$ induced by a family of operators $A(t) : V \rightarrow V^{*}$, $t \in I$, if for every $x, y \in L^{p}(I, V) \cap L^{q}(I, H)$ it holds

$$
\langle \mathcal{A}x, y \rangle_{L^{p}(I, V) \cap L^{q}(I, H)} = \int_{I} \langle A(t)(x(t)), y(t) \rangle_{V} \ dt.
$$

Remark 2.9 (Need for $L^{p}(I, V) \cap L^{q}(I, H)$) Note that an operator family $A(t) : V \rightarrow V^{*}$, $t \in I$ can define an induced operator in different spaces. In [23], [22] the induced operator $\mathcal{A}$ is considered as an operator from $L^{p}(I, V) \cap L^{q}(I, H)$ into $(L^{p}(I, V))^{*}$. Here, we consider the induced operator $\mathcal{A}$ as an operator from $L^{p}(I, V) \cap L^{q}(I, H)$ into $(L^{p}(I, V) \cap L^{q}(I, H))^{*}$, which is a more general and enables us to consider operator families with significantly worse growth behavior. Here, the so-called Temam modification $\tilde{B} : X \rightarrow X^{*}$, tracing back to [32], [33], of the convective term $B : X \rightarrow X^{*}$ defined in (1.3), defined for $p > \frac{3d + 2}{d + 2}$ and all $u, v \in X$ via

$$
\langle \tilde{B}u, v \rangle_{X} = \frac{1}{2} \int_{Q} u \otimes v : Du \ dx - \frac{1}{2} \int_{Q} u \otimes u : Dv \ dx,
$$

serves as a prototypical example. In fact, following [23, Example 5.1], one can prove that $B : X \rightarrow X^{*}$ satisfies for $d = 3$ and $p \geq \frac{11}{2}$ the estimate

$$
\| Bu \|_{X^{*}} \leq c(1 + \| u \|_{Y})(1 + \| u \|_{X}^{p-1})
$$

for all $u \in X$ and that corresponding induced operator $\mathcal{B}$ is well-defined and bounded as an operator from $L^{p}(I, X) \cap L^{q}(I, Y)$ to $(L^{p}(I, Y))^{*}$. Regrettably, for the remaining term in Temam’s modification, i.e., for the operator $\tilde{B} := \tilde{B} - \frac{1}{2} B : X \rightarrow X^{*}$, we can prove (2.11) for $d = 3$ only for $p > \frac{13}{3}$. In order to reach $p > \frac{11}{2}$ for $d = 3$, one is forced to use a larger target space, i.e., we view the induced operator of $B$ as an operator from $L^{p}(I, X) \cap L^{q}(I, Y)$ to $(L^{p}(I, X) \cap L^{q}(I, Y))^{*}$, where $q \in [p, \infty)$ is specified in Proposition 4.21.
Definition 2.12 (Weak solution) Let \((V, H, j)\) be an evolution triple, \(I := (0, T), T < \infty\), and \(1 < p \leq q < \infty\). Moreover, let \(x_0 \in H\) be an initial value, \(f \in L^p(I, V^*)\) a right-hand side, and \(A : L^p(I, V) \cap_j L^q(I, H) \to (L^p(I, V) \cap_j L^q(I, H))^*\) induced by a family of operators \(A(t) : V \to V^*,\ t \in I\). A function \(x \in \mathcal{W}_I^{1,p,q}(I, V, H)\) is called weak solution of the initial value problem (1.1) if 
\[
(j, x)(0) = x_0 \quad \text{in} \quad H 
\]
and for all \(\phi \in C_c^0(I, V)\) there holds 
\[
\int_I \left\langle \frac{dx}{dt}(t), \phi(t) \right\rangle_V dt + \int_I (A(t)(x(t)), \phi(t))_V dt = \int_I (f(t), \phi(t))_V dt.
\]

Here, the initial condition is well-defined since due to Proposition 2.7 (ii) there exists the unique continuous representation \(j, x \in C^0(T, H)\) of \(x \in \mathcal{W}_I^{1,p,q}(I, V, H)\).

3 Quasi non-conforming approximation

In this section we introduce the concept of quasi non-conforming approximations.

Definition 3.1 (Quasi non-conforming approximation) Let \((V, H, j)\) and \((X, Y, j)\) be evolution triples such that \(V \subseteq X\) with \(\| \cdot \|_V = \| \cdot \|_X\) in \(V\) and \(H \subseteq Y\) with \(\langle \cdot, \cdot \rangle_H = \langle \cdot, \cdot \rangle_Y\) in \(H \times H\). Moreover, let \(I := (0, T), T < \infty\), and let \(1 < p < \infty\). We call a sequence of closed subspaces \((V_n)_{n \in \mathbb{N}}\) of \(X\) a quasi non-conforming approximation of \(V\) in \(X\), if the following properties are satisfied:

(QNC.1) There exists a dense subset \(D \subseteq V\), such that for each \(v \in D\) there exist elements \(v_n \in V_n, \ n \in \mathbb{N}\), such that \(v_n \to v\) in \(X\) \((n \to \infty)\).

(QNC.2) For each sequence \(x_n \in L^p(I, V_{m_n}), \ n \in \mathbb{N}\), where \((m_n)_{n \in \mathbb{N}} \subseteq \mathbb{N}\) with \(m_n \to \infty\) \((n \to \infty)\), from \(x_n \to x\) in \(L^p(I, X)\) \((n \to \infty)\), it follows that \(x \in L^p(I, V)\).

The next proposition shows that our motivating example, namely the approximation of divergence-free Sobolev functions through discretely divergence-free finite element spaces perfectly fits into the framework of quasi non-conforming approximations.

Proposition 3.2 For \(p \geq \frac{2d}{d+2}\), we set as in the introduction \((V, H, \text{id}_V) := (W_0^{1,p}(\Omega), L_0^{2}(\Omega), \text{id})\), \((X, Y, \text{id}_X) := (W_0^{1,p}(\Omega)^d, L^2(\Omega)^d, \text{id})\) and \(Z := L^p(\Omega)\). Moreover, for given \(m, \ell \in \mathbb{N}_0\) and all \(h > 0\) let \(X_h \subseteq P_m(T_h)^d \cap X\) and \(Z_h \subseteq P_{\ell}(T_h) \cap Z\) be finite element spaces satisfying Assumption 1.17. Then, for a null sequence \((h_n)_{n \in \mathbb{N}} \subseteq (0, \infty)\) the sequence \((V_n)_{n \in \mathbb{N}}\), for every \(n \in \mathbb{N}\) given via 
\[
V_n := V_{h_n} := \{v_n \in X_{h_n} \mid \langle \text{div} v_n, \eta_n \rangle_Z = 0 \quad \text{for all} \quad \eta_n \in Z_{h_n}\},
\]
forms a quasi non-conforming approximation of \(V\) in \(X\).

Proof Clearly, \((V, H, \text{id}_V)\) and \((X, Y, \text{id}_X)\) form evolution triples, such that \(V \subseteq X\) with \(\| \cdot \|_V = \| \cdot \|_X\) in \(V\) and \(H \subseteq Y\) with \(\langle \cdot, \cdot \rangle_H = \langle \cdot, \cdot \rangle_Y\) in \(H \times H\). So, let us verify that \((V_n)_{n \in \mathbb{N}}\) satisfies (QNC.1) and (QNC.2):

ad (QNC.1) Due to their finite dimensionality, the spaces \((V_n)_{n \in \mathbb{N}}\) are closed. We set \(D := Y := \{v \in C^{\infty}_0(\Omega)^d \mid \text{div} v = 0\}\). Let \(v \in D\). Then, owning to standard estimates for polynomial projection operators (cf. [34, Lemma 2.25]), the sequence \(v_n := \Pi_{h_n} v \in V_n, n \in \mathbb{N}\), satisfies 
\[
\|v - v_n\|_X \leq c h_n \|v\|_{W^{2,2}(\Omega)^d} \to 0 \quad (n \to \infty).
\]
Ad (QNC.2) Let $x_n \in L^p(I, V_{m_n})$, $n \in \mathbb{N}$, where $(m_n)_{n \in \mathbb{N}} \subseteq \mathbb{N}$ with $m_n \to \infty (n \to \infty)$, be such that $x_n \rightharpoonup x$ in $L^p(I, X)$ ($n \to \infty$). Let $\eta \in C_0^\infty(\Omega)$ and $\varphi \in C_0^\infty(I)$. As in the previous step we infer that the sequence $\eta_n := \Pi_{m_n}^Z \eta \in Z_{p_{m_n}}$, $n \in \mathbb{N}$, satisfies $\eta_n \rightharpoonup \eta$ in $Z$ ($n \to \infty$). On the other hand, since $\langle \text{div} x_n(t), \eta_n \rangle_Z = 0$ for almost every $t \in I$ and $n \in \mathbb{N}$, as $x_n(t) \in V_{m_n}$ for almost every $t \in I$ and all $n \in \mathbb{N}$, there holds for every $n \in \mathbb{N}$

$$\int_I \langle \text{div} x_n(t), \eta_n \rangle_Z \varphi(s) \, ds = 0. \quad (3.3)$$

By passing in (3.3) for $n \to \infty$, we obtain for every $\eta \in C_0^\infty(\Omega)$ and $\varphi \in C_0^\infty(I)$

$$\int_I \langle \text{div} x(s), \eta \rangle_Z \varphi(s) \, ds = 0,$$

i.e., $x \in L^p(I, V)$.

The next proposition shows that the notion of quasi non-conforming approximation is indeed a generalization of the usual notion of conforming approximation.

**Proposition 3.4** Let $(X, Y, j)$ and $(V, H, j)$ be as in Definition 3.1. Then, it holds:

(i) The constant approximation $V_n = V$, $n \in \mathbb{N}$, is a quasi non-conforming approximation of $V$ in $X$.

(ii) If $(V_n)_{n \in \mathbb{N}}$ is a conforming approximation of $V$, i.e., $(V_n)_{n \in \mathbb{N}}$ satisfy (C.1) and (C.2), then $(V_n)_{n \in \mathbb{N}}$ is a quasi non-conforming approximation of $V$ in $X$.

**Proof** Ad (i) Follows right from the definition.

Ad (ii) We set $D := \bigcup_{n \in \mathbb{N}} V_n$. Then, for each $v \in D$ there exists an integer $n_0 \in \mathbb{N}$ such that $v \in V_n$ for every $n \geq n_0$. Therefore, the sequence $v_n \in V_n$, $n \in \mathbb{N}$, given via $v_n := 0$ if $n < n_0$ and $v_n := v$ if $n \geq n_0$, satisfies $v_n \rightharpoonup v$ in $V$ ($n \to \infty$), i.e., $(V_n)_{n \in \mathbb{N}}$ satisfies (QNC.1). Apart from that, $(V_n)_{n \in \mathbb{N}}$ obviously fulfills (QNC.2). □

The following proposition will be crucial in verifying that the induced operator $A$ of a family of operators $(A(t))_{t \in I}$ is quasi non-conforming Bochner pseudo-monotone (cf. Definition 4.1).

**Proposition 3.5** Let $(V, H, j)$ and $(X, Y, j)$ be as in Definition 3.1 and let $(V_n)_{n \in \mathbb{N}}$ be a quasi non-conforming approximation of $V$ in $X$. Then, the following statements hold true:

(i) For a sequence $v_n \in V_{m_n}$, $n \in \mathbb{N}$, where $(m_n)_{n \in \mathbb{N}} \subseteq \mathbb{N}$ with $m_n \to \infty (n \to \infty)$, from $v_n \rightharpoonup v$ in $X$ ($n \to \infty$), it follows that $v \in V$.

(ii) For a sequence $v_n \in V_{m_n}$, $n \in \mathbb{N}$, where $(m_n)_{n \in \mathbb{N}} \subseteq \mathbb{N}$ with $m_n \to \infty (n \to \infty)$, with $\sup_{n \in \mathbb{N}} \|v_n\|_X < \infty$, and $v \in V$ the following statements are equivalent:

\( a \) $v_n \rightharpoonup v$ in $X$ ($n \to \infty$).

\( b \) $P_H jv_n \rightharpoonup jv$ in $H$ ($n \to \infty$), where $P_H : Y \to H$ is the orthogonal projection of $Y$ into $H$.

(iii) For each $h \in H$ there exists a sequence $v_n \in V_{m_n}$, $n \in \mathbb{N}$, where $(m_n)_{n \in \mathbb{N}} \subseteq \mathbb{N}$ with $m_n \to \infty (n \to \infty)$, such that $jv_n \rightharpoonup h$ in $Y$ ($n \to \infty$).

**Proof** Ad (i) Immediate consequence of (QNC.2).

Ad (ii) (a) $\Rightarrow$ (b) Follows from the weak continuity of $j : X \to Y$ and $P_H : Y \to H$.

(b) $\Rightarrow$ (a) From the reflexivity of $X$, we obtain a subsequence $(v_{n_k})_{k \in \mathbb{A}}$, with $\mathbb{A} \subseteq \mathbb{N}$, and an element $\tilde{v} \in X$, such that $v_{n_k} \rightharpoonup \tilde{v}$ in $X$ ($\mathbb{A} \ni n \to \infty$). Due to (i) we infer $\tilde{v} \in V$. From the weak
continuity of \( j : X \to Y \) and \( P_H : Y \to H \) we conclude \( P_H j v_n \to P_H j \tilde{v} = j \tilde{v} \) in \( H \) (\( \Lambda \ni n \to \infty \)).

In consequence, we have \( j \tilde{v} = jv \) in \( H \), which in virtue of the injectivity of \( j : V \to H \) implies that \( \tilde{v} = v \) in \( V \), and therefore

\[
v_n \to v \quad \text{in} \quad X \quad (\Lambda \ni n \to \infty).
\]

(3.6)

Since this argumentation stays valid for each subsequence of \( (v_n)_{n \in \mathbb{N}} \subseteq X \), \( v \in V \) is weak accumulation point of each subsequence of \( (v_n)_{n \in \mathbb{N}} \subseteq X \). Therefore, the standard convergence principle (cf. [18, Kap. I, Lemma 5.4]) guarantees that (3.6) remains true even if \( \Lambda = \mathbb{N} \).

**ad (iii)** Since \((V,H,j)\) is an evolution triple, \( j(V) \) is dense in \( H \). As a result, for fixed \( h \in H \) there exists a sequence \((v_n)_{n \in \mathbb{N}} \subseteq V\), such that \( \|h - jv_n\|_H \leq 2^{-n} \) for all \( n \in \mathbb{N} \). Due to (QNC.1) there exist a sequence \((w_n)_{n \in \mathbb{N}} \subseteq D\), such that \( \|v_n - w_n\|_V \leq 2^{-n-1} \) for all \( n \in \mathbb{N} \) and a double sequence \((v^n_k)_{n,k \in \mathbb{N}} \subseteq X\), with \( v^n_k \in V_k \) for all \( k \), \( n \in \mathbb{N} \), such that \( v^n_k \to w_n \) in \( X \) (\( k \to \infty \)) for all \( n \in \mathbb{N} \). Thus, for each \( n \in \mathbb{N} \) there exists \( m_n \in \mathbb{N} \), such that \( \|w_n - v^n_{k_n}\|_X \leq 2^{-n-1} \) for all \( k \geq m_n \). Then, we have \( v^n_{m_n} \in V_{m_n} \) for all \( n \in \mathbb{N} \) and \( \|h - jv^n_{m_n}\|_Y \leq (1 + c)2^{-n} \) for all \( n \in \mathbb{N} \), where \( c > 0 \) is the embedding constant of \( j \).

\[ \text{\( \blacksquare \)} \]

### 4 Quasi non-conforming Bochner pseudo-monotonicity

In this section we introduce an extended notion of Bochner pseudo-monotonicity (cf. [23], [22]), which incorporates a given quasi non-conforming approximation \((V_n)_{n \in \mathbb{N}}\).

**Definition 4.1** Let \((X,Y,j)\) and \((V,H,j)\) be as in Definition 3.1 and let \((V_n)_{n \in \mathbb{N}}\) be a quasi non-conforming approximation of \( V \) in \( X \), \( I := (0,T) \), with \( 0 < T < \infty \), and \( 1 < p \leq q < \infty \). An operator \( A : L^p(I,X) \cap_j L^q(I,Y) \to (L^p(I,X) \cap_j L^q(I,Y))^* \) is said to be **quasi non-conforming Bochner pseudo-monotone with respect to \((V_n)_{n \in \mathbb{N}}\)** if for a sequence \( x_n \in L^\infty(I,V_{m_n}), n \in \mathbb{N}, \) where \((m_n)_{n \in \mathbb{N}} \subseteq \mathbb{N} \) with \( m_n \to \infty \) (\( n \to \infty \)), from

\[
x_n \to x \quad \text{in} \quad L^p(I,X) \quad (n \to \infty),
\]

\[
jx_n \xrightarrow{*} jx \quad \text{in} \quad L^\infty(I,Y) \quad (n \to \infty),
\]

\[
P_H(jx_n)(t) \to (jx)(t) \quad \text{in} \quad H \quad (n \to \infty) \quad \text{for a.e.} \ t \in I,
\]

and

\[
\limsup_{n \to \infty} \langle Ax_n, x_n - x \rangle_{L^p(I,X) \cap_j L^q(I,Y)} \leq 0
\]

(4.5)

it follows for all \( y \in L^p(I,X) \cap_j L^q(I,Y) \) that

\[
\langle Ax, x - y \rangle_{L^p(I,X) \cap_j L^q(I,Y)} \leq \liminf_{n \to \infty} \langle Ax_n, x_n - y \rangle_{L^p(I,X) \cap_j L^q(I,Y)}.
\]

Note that (4.2) and (4.3) guarantee that \( x \in L^p(I,V) \cap_j L^\infty(I,H) \) due to Definition 3.1.

The basic idea of quasi non-conforming Bochner pseudo-monotonicity, in comparison to the original notion of Bochner pseudo-monotonicity tracing back to [23], consists in incorporating the approximation \((V_n)_{n \in \mathbb{N}}\). We will see in the proof of Theorem 1.7 that (4.2)–(4.5) are natural properties of a sequence \( x_n \in L^p(I,V_{m_n}), n \in \mathbb{N}, \) coming from a quasi non-conforming Rothe-Galerkin approximation of (1.1), if \( A \) satisfies appropriate additional assumptions. In fact, (4.2) usually is a consequence of the coercivity of \( A \), (4.3) stems from the time derivative, while (4.4) and (4.5) follow directly from the approximation scheme.
Proposition 4.6 Let \((X, Y, j)\) and \((V, H, j)\) be as in Definition 3.1 and let \((V_n)_{n \in \mathbb{N}}\) be a quasi
non-conforming approximation of \(V\) in \(X\), \(I := (0, T)\), \(T < \infty\), and \(1 < p, q < \infty\). Moreover, let
\(A(t) : X \to X^*, t \in I\), be a family of operators with the following properties:

(A.1) \(A(t) : X \to X^*\) is pseudo-monotone for almost every \(t \in I\).

(A.2) \(A(\cdot) : I \to X^*\) is Bochner measurable for every \(x \in X\).

(A.3) For some constants \(c_0 > 0\) and \(c_1, c_2 \geq 0\) holds for almost every \(t \in I\) and every \(x \in X\)
\[
\langle A(t)x, x \rangle_X \geq c_0 \|x\|_X^2 - c_1 \|jx\|_{L^p}^2 - c_2.
\]

(A.4) For constants \(\gamma \geq 0\) and \(\lambda \in (0, c_0)\) holds for almost every \(t \in I\) and every \(x, y \in X\)
\[
\|A(t)x, y\|_X \leq \lambda \|x\|_X^p + \gamma [1 + \|jx\|_{L^p}^q + \|jy\|_{L^p}^q + \|y\|_{L^q}^q].
\]

Then, the induced operator \(\mathcal{A} : L^p(I, X) \cap \gamma L^q(I, Y) \to (L^p(I, X) \cap \gamma L^q(I, Y))^*\), given via (2.8), is
well-defined, bounded and quasi-non-conforming Bochner pseudo-monotone with respect to \((V_n)_{n \in \mathbb{N}}\).

Proof 1. Well-definiteness: For \(x_1, x_2 \in L^p(I, X) \cap \gamma L^q(I, Y)\) there exists sequences of simple
functions \((s_n^m)_{n \in \mathbb{N}} \subseteq L^\infty(I, X), m = 1, 2, \ldots\), i.e., \(s_n^m(t) = \sum_{i=1}^{k_n^m} s_n^{i,m} \chi_{E_{n,i}^m}(t)\) for \(t \in I\) and \(m = 1, 2\),
where \(s_n^{i,m} \in X, k_n^m \in \mathbb{N}\) and \(E_{n,i}^m \in L^1(I)\) with \(\bigcup_{i=1}^{k_n^m} E_{n,i}^m = I\) and \(E_{n,i}^m \cap E_{n,j}^m = \emptyset\) for \(i \neq j\), such
that \(s_n^m(t) \to x_m(t)\) in \(X\) for almost every \(t \in I\) and \(m = 1, 2\). Moreover, it follows from Lemma 2.3
that \(A(t) : X \to X^*\) is for almost every \(t \in I\) semi-continuous, since it is for almost every \(t \in I\)
pseudo-monotone (cf. (A.1)) and bounded (cf. (A.4)). This yields for almost every \(t \in I\)
\[
\langle A(t)(s_n^1(t), s_n^2(t))_X, s_n^1(t), s_n^2(t) \rangle_X \to \sum_{i=1}^{k_n^1} \sum_{j=1}^{k_n^2} \langle A(t)s_n^{1,i}, s_n^{2,j} \rangle_X \chi_{E_{n,i}^1 \cap E_{n,j}^2}(t) \to \sum_{i=1}^{k_n^1} \sum_{j=1}^{k_n^2} \langle A(t)x_i(t), x_j(t) \rangle_X.
\]

Thus, since the functions \(t \to \langle A(t)s_n^{i,j} \rangle_X : I \to \mathbb{R}, i = 0, \ldots, k_n^1, j = 1, \ldots, k_n^2, n \in \mathbb{N}\), are
Lebesgue measurable due to (A.2), we conclude from (4.7) that \(t \to \langle A(t)(x_1(t), x_2(t))_X : I \to \mathbb{R}\)
is Lebesgue measurable. In addition, using (A.4), we obtain
\[
\int_I \langle A(t)(x_1(t), x_2(t))_X \rangle dt \leq \lambda \|x_1\|_{L^p(I, X)}^p + \gamma [T + \|jx_1\|_{L^q(I, Y)}^q + \|jx_2\|_{L^q(I, Y)}^q + \|x_2\|_{L^p(I, X)}^p] \leq \lambda \|x_1\|_{L^p(I, X)}^p + \gamma [T + 2],
\]

i.e., \(\mathcal{A} : L^p(I, X) \cap \gamma L^q(I, Y) \to (L^p(I, X) \cap \gamma L^q(I, Y))^*\) is well-defined.

2. Boundedness: As \(\|y\|_{L^p(I, X)} + \|jy\|_{L^q(I, Y)} \leq 1\) implies that \(\|y\|_{L^p(I, X)} + \|jy\|_{L^q(I, Y)} \leq 2\) for
every \(y \in L^p(I, X) \cap \gamma L^q(I, Y)\), we infer from (4.8) for every \(x \in L^p(I, X) \cap \gamma L^q(I, Y)\) that
\[
\|\mathcal{A}x\|_{(L^p(I, X) \cap \gamma L^q(I, Y))^*} \leq \sup \left\{ \langle Ax, y \rangle_{L^p(I, X) \cap \gamma L^q(I, Y)} \right\} \leq \lambda \|x\|_{L^p(I, X)} + \gamma \|jx\|_{L^q(I, Y)} + \gamma [T + 2],
\]
i.e., \(\mathcal{A} : L^p(I, X) \cap \gamma L^q(I, Y) \to (L^p(I, X) \cap \gamma L^q(I, Y))^*\) is bounded.

3. Quasi non-conforming Bochner pseudo-monotonicity: In principle, we proceed analogously to [23, Proposition 3.13]. However, as we have solely almost everywhere weak convergence
of the orthogonal projections available, i.e., (4.4), in the definition of quasi-nonconforming Bochner
pseudo-monotonicity (cf. Definition 4.1), the arguments in [23] ask for some slight modifications. In fact, in this context the properties of the quasi non-conforming approximation \((V_n)_{n \in \mathbb{N}}\) come into play. Especially the role of Proposition 3.5 will be crucial. We split the proof of the quasi non-conforming Bochner pseudo-monotonicity into four steps:

### 3.1. Collecting information:
Let \(x_n \in L^\infty(I, V_{m_n}), n \in \mathbb{N}\), where \((m_n)_{n \in \mathbb{N}} \subseteq \mathbb{N}\) with \(m_n \to \infty\) \((n \to \infty)\), be a sequence satisfying (4.2)–(4.5). We fix an arbitrary \(y \in L^p(I, X) \cap \bigcap_j L^q(I, Y)\), and choose a subsequence \((x_n)_{n \in A}\), with \(A \subseteq \mathbb{N}\), such that

\[
\lim_{n \to \infty} \langle A x_n, x_n - y \rangle_{L^p(I, X) \cap \bigcap_j L^q(I, Y)} = \lim_{n \to \infty} \inf_n \langle A x_n, x_n - y \rangle_{L^p(I, X) \cap \bigcap_j L^q(I, Y)}.
\]

(4.9)

Due to (4.4) there exists a subset \(E \subseteq I\), with \(I \setminus E\) a null set, such that for all \(t \in E\)

\[
P_H(jx_n)(t) \rightharpoonup (jx)(t) \quad \text{in} \quad H \quad (n \to \infty).
\]

(4.10)

Using (A.3) and (A.4), we get for every \(z \in L^p(I, X) \cap \bigcap_j L^q(I, Y)\) and almost every \(t \in I\)

\[
\langle A(t)(x_n(t)), x_n(t) - z(t) \rangle_X \\
\geq c_0 \|x_n(t)\|_X^p - c_1 \|j(x_n(t))\|_Y^p - c_2 - \langle A(t)(x_n(t)), z(t) \rangle_X \\
\geq (c_0 - \lambda) \|x_n(t)\|_X^p - c_1 K^2 - c_2 - \gamma [1 + K^q + \|z(t)\|_Y^p + \|z(t)\|_X^p],
\]

(4.11)

where \(K := \sup_{n \in \mathbb{N}} \|j x_n\|_{L^\infty(I, Y)} < \infty\) (cf. (4.3)). If we set \(\mu_z(t) := -c_1 K^2 - c_2 - \gamma [1 + K^q + \|z(t)\|_Y^p + \|z(t)\|_X^p] \in L^1(I)\), then (4.11) reads

\[
\langle A(t)(x_n(t)), x_n(t) - z(t) \rangle_X \geq (c_0 - \lambda) \|x_n(t)\|_X^p - \mu_z(t) \quad \text{for all} \quad t \in I \quad \text{and all} \quad n \in A.
\]

(4.12)

for almost every \(t \in I\) and all \(n \in A\). Next, we define

\[
S := \{ t \in E \mid A(t) : X \to X^* \text{ is pseudo-monotone, } |\mu_z(t)| < \infty \text{ and } (*)_{z, n, t} \text{ holds for all } n \in A \}.
\]

Apparently, \(I \setminus S\) is a null set.

### 3.2. Intermediate objective:
Our next objective is to verify for all \(t \in S\) that

\[
\liminf_{n \to \infty} \langle A(t)(x_n(t)), x_n(t) - x(t) \rangle_X \geq 0.
\]

(4.13)

To this end, let us fix an arbitrary \(t \in S\) and define

\[
A_t := \{ n \in A \mid \langle A(t)(x_n(t)), x_n(t) - x(t) \rangle_X < 0 \}.
\]

We assume without loss of generality that \(A_t\) is not finite. Otherwise, (4.13) would already hold true for this specific \(t \in S\) and nothing would be left to do. But if \(A_t\) is not finite, then

\[
\limsup_{n \to \infty} \langle A(t)(x_n(t)), x_n(t) - x(t) \rangle_X \leq 0.
\]

(4.12)

The definition of \(A_t\) and \((*)_{z, n, t}\) imply for all \(n \in A_t\)

\[
(c_0 - \lambda) \|x_n(t)\|_X^p \leq \langle A(t)(x_n(t)), x_n(t) - x(t) \rangle_X + |\mu_z(t)| < |\mu_z(t)| < \infty.
\]

(4.13)
This and \( \lambda < c_0 \) yield that the sequence \((x_n(t))_{n \in A_t}\) is bounded in \(X\). In view of (4.10), Proposition 3.5 (ii) implies that

\[
x_n(t) \to x(t) \quad \text{in} \quad X \quad (A_t \ni n \to \infty).
\]

Since \(A(t) : X \to X^*\) is pseudo-monotone, we get from (4.14) and (4.12) that

\[
\liminf_{n \to \infty} (A(t)(x_n(t)), x_n(t) - x(t))_X \geq 0.
\]

Due to \((A(t)(x_n(t)), x_n(t) - x(t))_X \geq 0\) for all \(n \in A \setminus A_t\), \((**)_t\) holds for all \(t \in S\).

3.3. Switching to the image space level: In this passage we verify the existence of a subsequence \((x_n)_{n \in A_0} \subseteq L^p(I, X) \cap_j L^\infty(I, Y)\), with \(A_0 \subseteq A\), such that for almost every \(t \in I\)

\[
x_n(t) \to x(t) \quad \text{in} \quad X \quad (A_0 \ni n \to \infty),
\]

\[
\limsup_{n \to \infty} (A(t)(x_n(t)), x_n(t) - x(t))_X \leq 0.
\]

As a consequence, we are in a position to exploit the almost everywhere pseudo-monotonicity of the operator family. Thanks to \((A(t)(x_n(t)), x_n(t) - x(t))_X \geq -\mu x(t)\) for all \(t \in S\) and \(n \in A\) (cf. \((*)_x,\), Fatou’s lemma (cf. [27, Theorem 1.18]) is applicable. It yields, also using (4.5)

\[
0 \leq \int_I \liminf_{n \to \infty} (A(s)(x_n(s)), x_n(s) - x(s))_X \, ds \quad (4.16)
\]

\[
\leq \liminf_{n \to \infty} \int_I (A(s)(x_n(s)), x_n(s) - x(s))_X \, ds \leq \limsup_{n \to \infty} \langle Ax_n, x_n - x \rangle_{L^p(I, X) \cap_j L^\infty(I, Y)} \leq 0.
\]

Let us define \(h_n(t) := \langle A(t)(x_n(t)), x_n(t) - x(t) \rangle_X\). Then, \((**)_t\) and (4.16) read:

\[
\liminf_{n \to \infty} h_n(t) \geq 0 \quad \text{for all} \quad t \in S.
\]

\[
\lim_{n \to \infty} \int_I h_n(s) \, ds = 0. \quad (4.17)
\]

As \(s \mapsto s^- := \min\{0, s\}\) is continuous and non-decreasing we deduce from (4.17) that

\[
0 \geq \limsup_{n \to \infty} h_n(t)^- \geq \liminf_{n \to \infty} h_n(t)^- \geq \min\left\{0, \liminf_{n \to \infty} h_n(t)\right\} = 0,
\]

i.e., \(h_n(t)^- \to 0\) \((A \ni n \to \infty)\) for all \(t \in S\). Since \(0 \geq h_n(t)^- \geq -\mu x(t)\) for all \(t \in S\) and \(n \in A\), Vitali’s theorem yields \(h_n^- \to 0\) in \(L^1(I)\) \((A \ni n \to \infty)\). From the latter, \(|h_n| = h_n - 2h_n^-\) and (4.18), we conclude that \(h_n \to 0\) in \(L^1(I)\) \((A \ni n \to \infty)\). This provides a further subsequence \((x_n)_{n \in A_0}\) with \(A_0 \subseteq A\) and a subset \(F \subseteq I\) such that \(I \setminus F\) is a null set and for all \(t \in F\)

\[
\lim_{n \to \infty} \langle A(t)(x_n(t)), x_n(t) - x(t) \rangle_X = 0. \quad (4.19)
\]
Let us first consider the case
\[ A \]
Eventually, it can readily be seen by exploiting (S.3) that
\[ S \]
Hence,
\[ S \]
This and (4.13) implies for all \( t \in \mathcal{S} \cap F \)
\[ x_n(t) \rightarrow x(t) \quad \text{in } X \quad (A_0 \ni n \rightarrow \infty). \] (4.20)

The relations (4.19) and (4.20) are just (4.15).

3.4. Switching to the Bochner-Lebesgue level: From the pseudo-monotonicity of the operators \( A(t) : X \rightarrow X^* \) for all \( t \in \mathcal{S} \cap F \) we obtain almost every \( t \in I \)
\[ \langle A(t)(x(t)), x(t) - y(t) \rangle_X \leq \liminf_{n \rightarrow \infty} \langle A(t)(x_n(t)), x_n(t) - y(t) \rangle_X. \]

Due to \((*)_{y,n,t}\), we have \( \langle A(t)(x_n(t)), x_n(t) - y(t) \rangle_X \geq -\mu y(t) \) for almost every \( t \in I \) and all \( n \in A_0 \).
Thus, using the definition of the induced operator (2.8), Fatou’s lemma and (4.9) we deduce
\[ \langle Ax, x - y \rangle_{L^p(I,X) \cap L^q(I,Y)} \leq \int_I \liminf_{n \rightarrow \infty} \langle A(s)(x_n(s)), x_n(s) - y(s) \rangle_X \, ds \]
\[ \leq \liminf_{n \rightarrow \infty} \int_I \langle A(s)(x_n(s)), x_n(s) - y(s) \rangle_X \, ds \]
\[ = \liminf_{n \rightarrow \infty} \langle Ax_n, x_n - y \rangle_{L^p(I,X) \cap L^q(I,Y)} \]
\[ = \liminf_{n \rightarrow \infty} \langle Ax_n, x_n - y \rangle_{L^p(I,X) \cap L^q(I,Y)} \]
\[ \frac{\|x_n - x\|}{\|y\|_X} \leq \frac{\|x_n - x\|}{\|y\|_X} \quad (4.20) \]

As \( y \in L^p(I,X) \cap L^q(I,Y) \) was chosen arbitrary, this completes the proof of Proposition 4.6. \( \square \)

Proposition 4.21 Let \((X,Y,j)\) be as in Proposition 3.2 with \( p > \frac{3d+2}{d+2} \) and \( d \geq 2 \). Moreover, let \( S(t), \hat{B} : X \rightarrow X^*, \ t \in I \), be defined in (1.3) and (2.10), respectively. Then, the operator family \( A(t) := S(t) + \hat{B} : X \rightarrow X^*, \ t \in I \), satisfies (A.1)–(A.4).

Proof Let us first consider \( S(t) : X \rightarrow X^*, \ t \in I \), separately. From (S.1) and (S.2) in conjunction with the standard theory of Nemytskii operators (cf. [27, Theorem 1.43]) we deduce for almost every \( t \in I \) the well-definiteness and continuity of \( S(t) : X \rightarrow X^* \), including the conditions (A.2) and (A.3). (S.4) certainly implies for almost every \( t \in I \) the monotonicity of \( S(t) : X \rightarrow X^* \).
Hence, \( S(t) : X \rightarrow X^* \) is for almost every \( t \in I \) pseudo-monotone, i.e., condition (A.1) is satisfied. Eventually, it can readily be seen by exploiting (S.3) that \( S(t) : X \rightarrow X^*, \ t \in I \), satisfies (A.4).

Next, we treat the more delicate part \( \hat{B} : X \rightarrow X^* \). Here, we limit ourselves to the case \( d \geq 3 \), as the case \( d = 2 \) simplifies due to better Sobolev embeddings. In the same manner, one can verify by the standard theory of Nemytskii operators and Rellich’s compactness theorem that \( \hat{B} : X \rightarrow X^* \) is bounded and pseudo-monotone, i.e., satisfies (A.1) and (A.2). In addition, by Hölder’s inequality there holds for every \( u, v \in X \)
\[ |\langle \hat{B}u, v \rangle_X| \leq \|u\|_{L_2^p(I)^d} \|v\|_X + \|u\|_{L_2^p(I)^d} \|v\|_{L_{2p'}(I)^d} \|u\|_X. \] (4.22)
If \( p \geq d \), then \( p - 1 \geq 2 \), as \( d \geq 3 \), and \( \|u\|_{L^{2p'(t)}(Q)^d} \leq c\|u\|_{X} \) for all \( u \in X \) by means of Sobolev embedding. Thus, using \( a^2 \leq (1 + a)^{p-1} \leq 2^{p-2}(1 + a^{p-1}) \) for all \( a \geq 0 \) and the weighted \( \varepsilon \)-Young inequality with constant \( c_p(\varepsilon) := (p\varepsilon)^{1-p}/p \), we obtain for every \( u, v \in X \) and \( \varepsilon > 0 \)

\[
|\langle \hat{B}u, v \rangle_X | \leq c\|u\|_X^2\|v\|_X \leq c2^{p-2}(1 + \|u\|_X^{p-1})\|v\|_X \leq c\varepsilon 2^{p(p-2)}\|u\|_X^p + (c2^{p-2} + c_p(\varepsilon))\|v\|_X^p,
\]

i.e., \((A.3)\) for \( \varepsilon > 0 \) sufficiently small.

If \( p \in (\frac{3d+2}{d+2}, d) \), then by interpolation with \( \frac{1}{\rho} = \frac{1}{p} + \frac{\theta}{2} \), where \( \rho = p\frac{d+2}{d} \), \( \theta = \frac{2}{d+2} \) and \( p^* = \frac{dp}{d-p} \), we obtain for all \( u \in X \)

\[
\|u\|_{L^\rho(t)} \leq \|u\|_{\frac{2d+2}{d+2}} \|u\|_{\frac{d}{d+2}} \leq c\|u\|_{\frac{2d+2}{d+2}}\|u\|_{\frac{d}{d+2}}^2.
\]

(4.23)

Hence, since also \( p \geq 2p' \), we further conclude from (4.23) in (4.22) that for all \( u, v \in X \)

\[
|\langle \hat{B}u, v \rangle_X | \leq c\|u\|_{\frac{4}{Y^2}}\|u\|_{\frac{2d}{X^2}}\|v\|_X + c\|u\|_{\frac{2}{Y^2}}\|u\|_{\frac{2d}{X^2}}\|v\|_X \|v\|_X \|v\|_X \|v\|_X \|v\|_X
\]

(4.24)

\[
\leq c\|u\|_{\frac{4p'}{Y^2}}\|u\|_{\frac{2d}{X^p}} + c\|u\|_{\frac{2p}{Y^2}}\|u\|_{\frac{2d}{X^p}}\|v\|_X^2 + c\|u\|_{\frac{4p'}{Y^2}}\|u\|_{\frac{2d}{X^p}}\|v\|_X^2 + 2\|u\|_X^p + c\|v\|_X^p,
\]

(4.25)

where we applied Young’s inequality with exponent \( p \) in the first summand in (4.24) and with exponent \( \frac{d+2}{d}p \) in the second summand. Since \( s := \frac{p(2d+2)}{2d+2-d} < p \) and \( r := \frac{2d+2}{d+2}p' < p \) for \( p > \frac{3d+2}{d+2} \), we apply the weighted \( \varepsilon \)-Young inequality in the first two summands in (4.25) with exponent \( \frac{p}{r} > 1 \) and \( \frac{p}{2} \) > 1, respectively. In doing so, we obtain for every \( u, v \in X \) and \( \varepsilon > 0 \) that

\[
|\langle \hat{B}u, v \rangle_X | \leq c [c_\varepsilon(\varepsilon)]\|u\|_{\frac{4p'}{Y^2}}\|u\|_{\frac{4p}{Y^2}} \|v\|_X^2 + c(\varepsilon)\|u\|_{\frac{2p}{Y^2}}\|u\|_{\frac{2d}{X^p}}\|v\|_X^2 + 2\|u\|_X^p + \|v\|_X^p
\]

\[
\leq c[c_\varepsilon(\varepsilon)]\|u\|_{\frac{4p'}{Y^2}}\|u\|_{\frac{4p}{Y^2}} + c(\varepsilon)\|u\|_{\frac{2p}{Y^2}}\|u\|_{\frac{2d}{X^p}}\|v\|_X^2 + 2\|u\|_X^p + \|v\|_X^p,
\]

i.e., \((A.3)\) for \( \varepsilon > 0 \) sufficiently small and

\[
q := \max \left\{ \frac{4p'}{d+2}, \frac{4p}{d+2}, \frac{p}{d-p}, \frac{4p^2}{p-s} \right\}
\]

Altogether, \( \hat{B} : X \to X^* \) satisfies \((A.3)\). As a result, \( A(t) : X \to X^* \), \( t \in I \), satisfies \((A.1)-(A.3)\), and thanks to \( \langle \hat{B}u, u \rangle_X = 0 \) for all \( u \in X \) also \((A.4)\). \( \Box \)

5 Rothe scheme

Let \( X \) be a Banach space, and \( I := (0, T) \), \( T < \infty \). For \( K \in \mathbb{N} \) we set \( \tau := \frac{T}{K} \), \( I_k^\tau := ((k-1)\tau, k\tau) \), \( k = 1, ..., K \), and \( I_\tau := \{ I_k^\tau \}_{k=1}^{K} \). Moreover, we denote by

\[
\mathcal{S}(I_\tau, X) := \{ x : I \to X \mid x(s) = x(t) \text{ in } X \text{ for all } t, s \in I_k^\tau, k = 1, ..., K \} \subset L^\infty(I, X)
\]

the space of piece-wise constant functions with respect to \( I_\tau \). For a given finite sequence \((x^k)_{k=0,...,K} \subset X \) the backward difference quotient operator is defined via

\[
d_{-\tau}x^k := \frac{1}{\tau}(x^k - x^{k-1}) \text{ in } X, \quad k = 1, ..., K.
\]
Furthermore, we denote for a given finite sequence \((x^k)_{k=0,...,K} \subseteq X\) by \(\mathbf{x}^\tau \in S^0(I_\tau, X)\) the piecewise constant interpolant, and by \(\hat{x}^\tau \in W^{1,\infty}(I, X)\) the piecewise affine interpolant, for every \(t \in I^*_K\) and \(k = 1,...,K\) given via

\[
\mathbf{x}^\tau(t) := x^k, \quad \hat{x}^\tau(t) := \left(\frac{t}{\tau} - (k - 1)\right)x^k + \left(k - \frac{t}{\tau}\right)x^{k-1} \quad \text{in } X. \quad (5.1)
\]

In addition, if \((X, Y, J)\) is an evolution triple and \((x^k)_{k=0,...,K} \subseteq X\) a finite sequence, then it holds for \(k, l = 0,...,K\) the discrete integration by parts formula

\[
\int_{I^*_K} \left< \frac{d_x x^\tau}{dt}(t), \mathbf{x}^\tau(t) \right>_X \, dt \geq \frac{1}{2} \|jx^l\|_Y^2 - \frac{1}{2} \|jx^k\|_Y^2, \quad (5.2)
\]

which is an immediate consequence of the identity \(\left< d_x x^k, x^k \right>_X = \frac{1}{2} d_x \|jx^k\|_Y^2 + \frac{1}{2} \|d_x j x^k\|_Y^2\) for every \(k = 1,...,K\).

For the discretization of the right-hand side in (1.1) we use the following construction. Let \(X\) be a Banach space, \(I = (0, T), T < \infty, K \in \mathbb{N}, \tau := \frac{T}{K} > 0\) and \(1 < p < \infty\). The Clemént 0-order quasi-interpolation operator \(\mathcal{J}_\tau : L^p(I, X) \rightarrow S^0(I_\tau, X)\) is defined for every \(x \in L^p(I, X)\) via

\[
\mathcal{J}_\tau[x] := \sum_{k=1}^K [x]_k^\tau \chi_{I_k^\tau} \quad \text{in } S^0(I_\tau, X), \quad [x]_k^\tau := \int_{I_k^\tau} x(s) \, ds \in X.
\]

**Proposition 5.3** For every \(x \in L^p(I, X)\) it holds:

(i) \(\mathcal{J}_\tau[x] \rightarrow x\) in \(L^p(I, X)\) (\(\tau \rightarrow 0\)), i.e., \(\bigcup_{\tau > 0} S^0(I_\tau, X)\) is dense in \(L^p(I, X)\).

(ii) \(\sup_{\tau > 0} \|\mathcal{J}_\tau[x]\|_{L^p(I, X)} \leq \|x\|_{L^p(I, X)}\).

**Proof** See [27, Remark 8.15].

Since we treat non-autonomous evolution equations we also need to discretize the time dependent family of operators in (1.1). This will also be obtained by means of the Clemént 0-order quasi-interpolant. Let \((X, Y, J)\) be an evolution triple, \(I := (0, T), T < \infty, K \in \mathbb{N}, \tau := \frac{T}{K} > 0\) and \(1 < p \leq q < \infty\). Let \(A(t) : X \rightarrow X^*, t \in I\), be a family of operators satisfying the conditions (A.1)–(A.4), and denote by \(A : L^p(I, X) \cap J L^q(I, Y) \rightarrow (L^p(I, X) \cap J L^q(I, Y))^*\) the induced operator (cf. (2.8)). The k-th temporal mean \([A]_k^\tau : X \rightarrow X^*, k = 1,...,K\), of \(A(t) : X \rightarrow X^*, t \in I\), is defined for every \(x \in X\) via

\[
[A]_k^\tau x := \int_{I_k^\tau} A(s)x \, ds \quad \text{in } X^*.
\]

The Clemént 0-order quasi-interpolant \(\mathcal{J}_\tau[A](t) : X \rightarrow X^*, t \in I, \) of \(A(t) : X \rightarrow X^*, t \in I\), is defined for almost every \(t \in I\) and \(x \in X\) via

\[
\mathcal{J}_\tau[A](t)x := \sum_{k=1}^K \chi_{I_k^\tau}(t)[A]_k^\tau x \quad \text{in } X^*.
\]
The **Clement 0-order quasi-interpolant** \( J_\tau[A] : L^p(I, X) \cap_j L^q(I, Y) \to (L^p(I, X) \cap_j L^q(I, Y))^* \), of \( A : L^p(I, X) \cap_j L^q(I, Y) \to (L^p(I, X) \cap_j L^q(I, Y))^* \) is for all \( x, y \in L^p(I, X) \cap_j L^q(I, Y) \) defined via

\[
(J_\tau[A] x, y)_{L^p(I, X) \cap_j L^q(I, Y)} := \int_I (J_\tau[A](x(t)), y(t)) x dt.
\]

Note that \( J_\tau[A] \) is the induced operator of the family of operators \( J_\tau[A](t) : X \to X^*, \ t \in I \).

**Proposition 5.4 (Clement 0-order quasi-interpolant for induced operators)**

With the above notation we have:

(i) \( [A]_k^\tau : X \to X^* \) is well-defined, bounded, pseudo-monotone, and satisfies:

(i.a) \( ([A]_k^\tau x, y)_X \leq \lambda \|x\|_X^p + \gamma [1 + \|x\|_X^{p-1} \|y\|_Y^q + \|y\|_X^q] \) for all \( x, y \in X \).

(ii) \( J_\tau[A](t) : X \to X^*, \ t \in I \), satisfies the conditions (A.1)–(A.4).

(iii) \( J_\tau[A] : L^p(I, X) \cap_j L^q(I, Y) \to (L^p(I, X) \cap_j L^q(I, Y))^* \) is well-defined, bounded and satisfies:

(iii.a) For all \( x_\tau \in S^0(I, X) \), \( y \in L^p(I, X) \cap_j L^q(I, Y) \) holds

\[
(J_\tau[A] x_\tau, y)_{L^p(I, X) \cap_j L^q(I, Y)} = (A x_\tau, J_{\tau}[y])_{L^p(I, X) \cap_j L^q(I, Y)}.
\]

(iii.b) If the functions \( x_\tau \in S^0(I, X), \tau > 0 \), are bounded in \( L^p(I, X) \cap_j L^q(I, Y) \), then

\[
A x_\tau - J_\tau[A] x_\tau \to 0 \quad \text{in } (L^p(I, X) \cap_j L^q(I, Y))^* \quad (\tau \to 0).
\]

(iii.c) If \( x_\tau \in S^0(I, X) \), then \( \|J_\tau[A] x_\tau\|_{(L^p(I, X) \cap_j L^q(I, Y))^*} \leq \|A x_\tau\|_{(L^p(I, X) \cap_j L^q(I, Y))^*} \).

**Proof ad (i)** Let \( x \in X \). Due to (A.2) the function \( A(:) : I \to X^* \) is Bochner measurable. (A.4) guarantees that \( \|A(:)x\|_{X^*} \in L^1(I) \), and thus the Bochner integrability of \( A(:)x : I \to X^* \). As a result, the Bochner integral \( [A]_k^\tau x = \int_I A(s) x ds \) exists, i.e., \( [A]_k^\tau : X \to X^* \) is well-defined. The inequalities (i.a) and (i.b) are obvious. In particular, we gain from inequality (i.a) the boundedness of \( [A]_k^\tau : X \to X^* \). So, it is left to show the pseudo-monotonicity. Therefore, let \( (x_n)_{n \in \mathbb{N}} \subseteq X \) be a sequence such that

\[
x_n \to x \quad \text{in } X \quad (n \to \infty),
\]

\[
\limsup_{n \to \infty} ([A]_k^\tau x_n, x_n - x)_X \leq 0.
\]

If we set \( x_n := x_n \chi_{I_k}^\tau \in L^\infty(I, X), \ n \in \mathbb{N}, \) and \( x := x \chi_{I_k}^\tau \in L^\infty(I, X) \), then (5.5), the Lebesgue theorem on dominated convergence and the properties of the induced embedding \( j \) imply

\[
x_n \to x \quad \text{in } L^p(I, X) \quad (n \to \infty),
\]

\[
\|j x_n - j x\|_{L^\infty(I, Y)} \to 0 \quad (n \to \infty),
\]

\[
J_\tau[A] x_n(t) \xrightarrow{n \to \infty} j x(t) \quad \text{in } Y \quad \text{for a.e. } t \in I.
\]

In addition, from (5.6) we infer

\[
\limsup_{n \to \infty} (A x_n, x_n - x)_{L^p(I, X) \cap_j L^q(I, Y)} = \tau \limsup_{n \to \infty} ([A]_k^\tau x_n, x_n - x)_X \leq 0.
\]
Since \( \mathcal{A} : L^p(I, X) \cap J^q(I, Y) \rightarrow (L^p(I, X) \cap J^q(I, Y))^* \) is quasi non-conforming Bochner pseudo-monotone with respect to the constant approximation \( V_n = X, n \in \mathbb{N} \), by Proposition 4.6, we obtain from (5.7)–(5.10) that for all \( y \in L^p(I, X) \cap J^q(I, Y) \)

\[
\langle Ax, x - y \rangle_{L^p(I, X) \cap J^q(I, Y)} \leq \liminf_{n \to \infty} \langle Ax_n, x_n - y \rangle_{L^p(I, X) \cap J^q(I, Y)}.
\]

(5.11)

If we choose in (5.11) \( y := y \chi_{I_k^*} \in L^\infty(I, X) \) with \( y \in X \) and divide by \( \tau > 0 \), we conclude

\[
\langle [A]_k^x x - y \rangle_X \leq \liminf_{n \to \infty} \langle [A]_k^x x_n - y \rangle_X.
\]

In other words, \( [A]_k^x : X \to X^* \) is pseudo-monotone.

**ad (ii)** The assertion follows immediately from (i) and the definition of \( \mathcal{J}_r[A](t) \), \( t \in I \).

**ad (iii)** Since \( \mathcal{J}_r[A] \) is the induced operator of the family of operators \( \mathcal{J}_r[A](t) \), \( t \in I \), the well-definiteness and boundedness of \( \mathcal{J}_r[A] \) results from (ii) in conjunction with Proposition 4.6.

**ad (iii.a)** Let \( x_t \in \mathcal{S}^0(I, X) \) and \( y \in L^p(I, X) \cap J^q(I, Y) \). Then, using for every \( t, s \in I_k^* \), \( k = 1, \ldots, K \), that \( (A(s)(x_t(s)), y(t))_X = (A(s)(x_t(s)), y(t))_X \) and Fubini’s theorem, we infer

\[
\langle \mathcal{J}_r[A]x_t, y \rangle_{L^p(I, X) \cap J^q(I, Y)} = \int_I \langle \mathcal{J}_r[A](t)(x(t)), y(t) \rangle_X dt
\]

\[
= \sum_{k=1}^K \int_{I_k^*} \left\langle \int_{I_k^*} A(s)(x_t(t)) ds, y(t) \right\rangle_X dt
\]

\[
= \sum_{k=1}^K \int_{I_k^*} \left\langle A(s)(x_t(s)), \int_{I_k^*} y(t) dt \right\rangle_X ds
\]

\[
= \int_I \langle A(s)(x_t(s)), \mathcal{J}_r[y](s) \rangle_X ds
\]

\[
= \langle Ax_t, \mathcal{J}_r[y] \rangle_{L^p(I, X) \cap J^q(I, Y)}.
\]

**ad (iii.b)** Let the family \( x_t \in \mathcal{S}^0(I, X) \), \( \tau > 0 \), be bounded in \( L^p(I, X) \cap J^q(I, Y) \). Then, by the boundedness of \( \mathcal{A} : L^p(I, X) \cap J^q(I, Y) \rightarrow (L^p(I, X) \cap J^q(I, Y))^* \) (cf. Proposition 4.6), the family \( \langle x_t \rangle_{\tau > 0} \subseteq (L^p(I, X) \cap J^q(I, Y))^* \) is bounded as well. Therefore, also using (iii.a), we conclude for every \( y \in L^p(I, X) \cap J^q(I, Y) \) that

\[
\langle Ax_t - \mathcal{J}_r[A]x_t, y \rangle_{L^p(I, X) \cap J^q(I, Y)} = \langle Ax_t, y - \mathcal{J}_r[y] \rangle_{L^p(I, X) \cap J^q(I, Y)} \rightarrow 0 \quad (\tau \to 0),
\]

where we also used Proposition 5.3 (i).

**ad (iii.c)** Using (iii.a) and Proposition 5.3 (ii), we deduce

\[
\| \mathcal{J}_r[A]x_t \|_{(L^p(I, X) \cap J^q(I, Y))^*} = \sup_{\|y\|_{L^p(I, X) \cap J^q(I, Y)} \leq 1} \langle \mathcal{J}_r[A]x_t, y \rangle_{L^p(I, X) \cap J^q(I, Y)}
\]

\[
= \sup_{\|y\|_{L^p(I, X) \cap J^q(I, Y)} \leq 1} \langle Ax_t, \mathcal{J}_r[y] \rangle_{L^p(I, X) \cap J^q(I, Y)}
\]

\[
\leq \|Ax_t\|_{(L^p(I, X) \cap J^q(I, Y))^*}.
\]

\( \square \)
6 Fully discrete, quasi non-conforming approximation

In this section we formulate the exact framework of a quasi non-conforming Rothe-Galerkin approximation, prove its well-posedness, i.e., the existence of iterates, its stability, i.e., the boundedness of the corresponding double sequence of piece-wise constant interpolants, and its weak convergence, i.e., the weak convergence of a diagonal subsequence towards a weak solution of (1.1).

Assumption 6.1 Let $I := (0, T)$, $T < \infty$ and $1 < p \leq q < \infty$. We make the following assumptions:

(i) **Spaces:** $(V, H, j)$ and $(X, Y, j)$ are as in Definition 3.1 and $(V_n)_{n \in \mathbb{N}}$ is a quasi non-conforming approximation of $V$ in $X$.

(ii) **Initial data:** $x_0 \in H$ and there is a sequence $x_n^0 \in V_n$, $n \in \mathbb{N}$, such that $x_n^0 \to x_0$ in $Y$ ($n \to \infty$) and $\sup_{n \in \mathbb{N}} \|j x_n^0\|_Y \leq \|x_0\|_H$.\footnote{For a quasi non-conforming approximation Proposition 3.5 guarantees the existence of such a sequence.}

(iii) **Right-hand side:** $f \in L^p(I, X^*)$.

(iv) **Operators:** $A(t) : X \to X^*$, $t \in I$, is a family of operators satisfying (A.1)–(A.4) and $\mathcal{A} : L^p(I, X) \cap j H \to (L^p(I, X) \cap j L^q(I, Y))^*$ the corresponding induced operator.

Furthermore, we set $H_n := j(V_n) \subseteq Y$ equipped with $(\cdot, \cdot)_Y$, denote by $j_n : V_n \to H_n$ the restriction of $j$ to $V_n$ and by $R_n : H_n \to H_n^*$ the corresponding Riesz isomorphism with respect to $(\cdot, \cdot)_Y$. As $j_n$ is an isomorphism, the triple $(V_n, H_n, j_n)$ is an evolution triple with canonical embedding $\varepsilon_n := j_n^* R_n j_n : V_n \to V_n^*$, which satisfies

$$\langle \varepsilon_n v_n, w_n \rangle_{V_n} = (j_n v_n, j_n^* w_n)_Y \quad \text{for all } v_n, w_n \in V_n. \quad (6.2)$$

Putting all together leads us to the following algorithm:

**Algorithm 6.3 (Quasi non-conforming Rothe-Galerkin scheme)** Let Assumption (6.1) be satisfied. For given $K, n \in \mathbb{N}$ the sequence of iterates $(x_n^k)_{k=0,\ldots,K} \subseteq V_n$ is given via the implicit Rothe-Galerkin scheme for $\tau = \frac{T}{K}$

$$(d_t j_n^k \cdot j_n v_n)_Y + \langle [A]^\tau j_n^k, v_n \rangle_X = (\|f\|_Y^\tau, v_n)_X \quad \text{for all } v_n \in V_n. \quad (6.4)$$

**Remark 6.5** Note that the Rothe-Galerkin scheme (6.4) also covers pure Rothe schemes, i.e., without spatial approximation, and fully discrete conforming approximations:

(i) If $X = V$, $Y = H$, and $V_n = X$, $n \in \mathbb{N}$, then (6.4) forms a pure Rothe scheme.

(ii) If $X = V$, $Y = H$, and the closed subspaces $(V_n)_{n \in \mathbb{N}}$ satisfy (C.1)–(C.2), then (6.4) forms a conforming Rothe-Galerkin scheme.

**Proposition 6.6 (Well-posedness of (6.4))** Let Assumption (6.1) be satisfied and set $\tau_0 := \frac{1}{4c_1}$. Then, for all $K, n \in \mathbb{N}$ with $\tau = \frac{T}{K} < \tau_0$ there exist iterates $(x_n^k)_{k=1,\ldots,K} \subseteq V_n$, solving (6.4).

**Proof** Using (6.2) and the identity mapping $id_{V_n} : V_n \to X$, we see that (6.4) is equivalent to

$$(id_{V_n})^* ([f]^\tau_k) + \frac{1}{\tau} \varepsilon_n x_n^{k-1} \in R\left(\frac{1}{\tau} \varepsilon_n + (id_{V_n})^* \circ [A]^\tau_k \circ id_{V_n}\right), \quad \text{for all } k = 1, \ldots, K, \quad (6.7)$$
We fix an arbitrary \( k = 1, \ldots, K \). Apparently, \( \frac{1}{\tau} e_n : V_n \rightarrow V_n^* \) is linear and continuous. Using (6.2), we infer that \( \frac{1}{\tau} e_n(x, x) x V_n = \frac{1}{\tau}\| j_n x \|_Y^2 \geq 0 \) for all \( x \in V_n \), i.e., \( \frac{1}{\tau} e_n : V_n \rightarrow V_n^* \) is positive definite, and thus monotone. In consequence, \( \frac{1}{\tau} e_n : V_n \rightarrow V_n^* \) is pseudo-monotone. Since the conditions (A.1)–(A.4) are inherited from \( A : X \rightarrow X^* \) to \( (id_{V_n})^* \circ A \circ id_{V_n} : V_n \rightarrow V_n^* \) and since \( (id_{V_n})^* \circ A \circ id_{V_n} = [(id_{V_n})^* \circ A \circ id_{V_n}]_K \), Proposition 5.4 (i) guarantees that the operator \( (id_{V_n})^* \circ A \circ id_{V_n} : V_n \rightarrow V_n^* \) is bounded and pseudo-monotone. Altogether, we conclude that the sum \( \frac{1}{\tau} e_n + (id_{V_n})^* \circ A \circ id_{V_n} : V_n \rightarrow V_n^* \) is bounded and pseudo-monotone. In addition, as \( \tau < \frac{1}{2c_1} \), combining (6.2) and Proposition 5.4 (i,b), provides for all \( x \in V_n \)

\[
\left\langle \left( \frac{1}{\tau} e_n + (id_{V_n})^* \circ A \circ id_{V_n} \right) x, x \right\rangle_{V_n} \geq 3c_1 \| j_n x \|_Y^2 + c_0 \| x \|_{X}^p - c_2,
\]

i.e., \( \frac{1}{\tau} e_n + (id_{V_n})^* \circ A \circ id_{V_n} : V_n \rightarrow V_n^* \) is coercive. Hence, Proposition 2.2 proves (6.7).

**Proposition 6.8 (Stability of (6.4))** Let Assumption (6.1) be satisfied and set \( \tau_0 := \frac{1}{4c_1} \). Then, there exists a constant \( \tau_{\max} > 0 \) (not depending on \( K, n \in \mathbb{N} \)), such that the piece-wise constant interpolants \( \Xi_n \in S^0(\mathcal{Z}_\tau, X) \), \( K, n \in \mathbb{N} \) with \( \tau = \frac{T}{K} \in (0, \tau_0) \), and piece-wise affine interpolants \( \tilde{\Xi}_n \in W^{1,\infty}(I, X) \), \( n \in \mathbb{N} \), \( \tau \in (0, \tau_0) \) (cf. (5.1)) generated by iterates \( (x_n^k)_{k=0,\ldots,K-1} \subseteq V_n \), \( K \in \mathbb{N} \) with \( \tau = \frac{T}{K} \in (0, \tau_0) \), solving (6.4), satisfy the following estimates:

\[
\| \Xi_n^\tau \|_{L^p(I,X)} \leq M, \tag{6.9}
\]

\[
\| j_0 \Xi_n^\tau \|_{L^\infty(I,Y)} \leq M, \tag{6.10}
\]

\[
\| \mathcal{A} \Xi_n^\tau \|_{(L^p(I,Y) \cap L^q(I,Y))^*} \leq M, \tag{6.11}
\]

\[
\| e_n(\Xi_n^\tau - \Xi_k^\tau) \|_{L^p(I,V_n^*)} \leq \tau \| f \|_{L^p(I,X^*)} + M. \tag{6.12}
\]

**Proof** We use \( v_n = x_n^k \in V_n \), \( k = 1, \ldots, l \), for arbitrary \( l = 1, \ldots, K \) in (6.4), multiply by \( \tau \in (0, \tau_0) \), sum with respect to \( k = 1, \ldots, l \), use (5.2) and use \( \sup_{n \in \mathbb{N}}\| j_n x \|_Y \leq \| x_0 \|_H \), to obtain for every \( l = 1, \ldots, K \)

\[
\frac{1}{2} \| j_n^\tau x \|_Y^2 + \sum_{k=1}^l \tau (\langle A_k^{x_n^k}, x_n^k \rangle_X) \leq \frac{1}{2} \| x_0 \|_H^2 + \sum_{k=1}^l \tau (\langle f_k^{x_n^k}, x_n^k \rangle_X). \tag{6.13}
\]

Applying the weighted \( \varepsilon \)-Young inequality with constant \( c(\varepsilon) := (p \varepsilon)^{1-p'}/p' \) for all \( \varepsilon > 0 \), using \( \| f N [f] \|_{L^p(I,X^*)} \leq \| f \|_{L^p(I,X^*)} \) (cf. Proposition 5.3 (ii)), we deduce for every \( l = 1, \ldots, K \)

\[
\sum_{k=1}^l \tau (\langle f_k^{x_n^k}, x_n^k \rangle_X) = \langle f N [f], \Xi_n^\tau \rangle_{L^p(I,X)} \leq c(\varepsilon) \| f \|_{L^p(I,X^*)} + \varepsilon \int_0^{\tau_T} \| \Xi_n^\tau(s) \|_X^p \, ds.
\]

In addition, using Proposition 5.4 (i.b), we obtain for every \( l = 1, \ldots, K \)

\[
\sum_{k=1}^l \tau (\langle A_k^{x_n^k}, x_n^k \rangle_X) \geq c_0 \int_0^{\tau_T} \| \Xi_n^\tau(s) \|_X^p \, ds - \tau c_1 \| j_n^\tau X \|_Y^2 - \sum_{k=1}^{l-1} \tau c_1 \| j_n^\tau X \|_Y^2 - c_2 T. \tag{6.14}
\]
We set \( \varepsilon := \frac{q}{2} \), \( \alpha := \frac{1}{2} \| x_0 \|_Y^2 + c(\varepsilon) \| f^\prime \|^p \| J(\alpha, I, V) \|_p + c_2 T, \) \( \beta := 4 \tau c_1 < 1 \) and \( y^k_n := \frac{1}{4} \| j x_n^k \|_Y^2 \) for \( k = 1, \ldots, K \). Thus, we infer for every \( l = 1, \ldots, K \) from (6.13)–(6.14) that

\[
y^l_n + \frac{c_0}{2} \int_0^t \| \mathbf{p}_n^l(s) \|_Y^2 \, ds \leq \alpha + \beta \sum_{k=1}^{l-1} y^k_n. \tag{6.15}
\]

The discrete Gronwall inequality applied on (6.15) yields

\[
\frac{1}{4} \| j \mathbf{f}_n \|_{L^\infty(I,Y)} + \frac{c_0}{2} \| \mathbf{p}_n \|_{L^p(I,X)} \leq \alpha \exp(K \beta) = \alpha \exp(4Tc_1) =: C_0,
\]

which proves (6.9). From the boundedness of \( \mathcal{A} : L^p(I, X) \cap j L^q(I, Y) \to (L^p(I, X) \cap j L^q(I, Y))^* \) (cf. Proposition 4.6) and (6.9) we infer \( \| \mathcal{A} \mathbf{f}_n \|_{L^p(I,X) \cap j L^q(I,Y))^*} \leq C_1 \) for some \( C_1 > 0 \), i.e., (6.11).

In addition, it holds \( \| j \mathbf{x}_n \|_{L^\infty(I,Y)} \leq \| j \mathbf{f}_n \|_{L^\infty(I,Y)} \leq 4C_0 \) for every \( n \in \mathbb{N} \) and \( \tau \in (0, \tau_0) \), i.e., (6.10). Moreover, since \( e_n (\mathbf{x}_n(t) - \mathbf{f}_n(t)) = (t - k \tau) e_n \mathbf{x}_n(t) = (t - k \tau) \frac{d}{dt} j \mathbf{x}_n(t) \) in \( V_n \) and \( |t - k \tau| \leq \tau \) for every \( t \in I_k \), \( k = 1, \ldots, K \), \( n \in \mathbb{N} \), there holds for every \( n \in \mathbb{N} \) and \( \tau \in (0, \tau_0) \)

\[
\| e_n (\mathbf{x}_n - \mathbf{f}_n) \|_{L^p(I,V_n^*)} \leq \tau \left\| \frac{d}{dt} j \mathbf{x}_n \right\|_{L^p(I,V^*_n)} = \tau \left( \| j \mathbf{f}_n \|_{L^p(I,X^*)} + C_1 \right),
\]

i.e., the estimate (6.12), where we used Proposition 5.3 (ii) and Proposition 5.4 (iii.c).

We can now prove the main abstract convergence results, that is Theorem 1.7.

**Proof (Proof of Theorem 1.7)** We split the proof into four steps:

1. **Convergences:** From the estimates (6.9)–(6.12), the reflexivity of \( L^p(I, X) \cap j L^q(I, Y) \), also using Proposition 5.4 (iii.b), we obtain not relabelled subsequences \( (\mathbf{x}_n)_{n \in \mathbb{N}} \), \( (\mathbf{f}_n)_{n \in \mathbb{N}} \subseteq L^p(I, X) \cap j L^q(I, Y) \), where \( \mathbf{f}_n := \mathbf{f}_n \) for all \( n \in \mathbb{N} \), as well as \( \mathbf{x} \in L^p(I, X) \cap j L^q(I, Y) \), \( j \mathbf{x} \in L^q(I, Y) \) and \( \mathbf{f} \in (L^p(I, X) \cap j L^q(I, Y))^* \) such that

\[
\mathbf{x}_n \to \mathbf{x} \quad \text{in} \quad L^p(I, X) \quad (n \to \infty),
\]

\[
j \mathbf{x}_n \to j \mathbf{x} \quad \text{in} \quad L^q(I, Y) \quad (n \to \infty),
\]

\[
j \mathbf{f}_n \to j \mathbf{f} \quad \text{in} \quad L^q(I, Y) \quad (n \to \infty),
\]

\[
\mathcal{J}_\mathcal{A} \mathbf{x}_n \to \mathbf{f} \quad \text{in} \quad (L^p(I, X) \cap j L^q(I, Y))^* \quad (n \to \infty).
\]

From (QNC 2) we immediately obtain that \( \mathbf{x} \in L^p(I, V) \cap j L^q(I, H) \). In particular, there exists \( g \in L^q(I, V^*) + j^*(L^q(I, H)^*) \) (cf. (2.5)), such that for every \( v \in L^p(I, V) \cap j L^q(I, H) \)

\[
\langle \mathbf{x}, v \rangle_{L^p(I,X) \cap j L^q(I,Y)} = \int_I \langle g(t), v(t) \rangle_Y \, dt.
\]

Due to (6.12) there exists a subset \( E \subset I \), with \( I \setminus E \) a null set, such that for every \( t \in E \)

\[
\| e_{mn} (\mathbf{x}_n(t) - \mathbf{f}_n(t)) \|_{V_n^m} \to 0 \quad (n \to \infty).
\]
Owing to (QNC.1) we can choose for every element $v$ of the dense subset $D \subseteq V$ a sequence $v_n \in V_{m_n}$, $n \in \mathbb{N}$, such that $v_n \to v$ in $X$ ($n \to \infty$). Then, using the definition of $P_H$, (6.2), (6.9), (6.10) and (6.18), we infer for every $t \in E$ that

$$
|\langle P_H [j \hat{x}_n](t) - (j \overline{x}_n)(t), jv \rangle_H| = |\langle (j \hat{x}_n)(t) - (j \overline{x}_n)(t), jv \rangle_Y |
\leq |\langle \epsilon_{m_n} [j \hat{x}_n](t) - (j \overline{x}_n)(t), v_n \rangle_{V_{m_n}}| + |\langle (j \hat{x}_n)(t) - (j \overline{x}_n)(t), jv - jv_n \rangle_Y |
\leq \| \epsilon_{m_n} [\hat{x}_n(t) - \overline{x}_n(t)] \|_{V_{m_n}} \| v_n \|_X + \| (j \hat{x}_n)(t) - (j \overline{x}_n)(t) \|_Y \| jv - jv_n \|_Y
\leq \| \epsilon_{m_n} [\hat{x}_n(t) - \overline{x}_n(t)] \|_{V_{m_n}} \| v_n \|_X + 2M \| jv - jv_n \|_Y \to 0 \quad (n \to \infty).
$$

Since $D$ is dense in $V$ and $j(V)$ is dense in $H$, we conclude from (6.19) for every $t \in E$ that

$$
P_H [j \hat{x}_n](t) - (j \overline{x}_n)(t) \to 0 \quad \text{in $H$} \quad (n \to \infty).
$$

Since the sequences $(P_H j \overline{x}_n)_{n \in \mathbb{N}}$, $(P_H j \hat{x}_n)_{n \in \mathbb{N}} \subseteq L^\infty(I, H)$ are bounded (cf. (6.9) and (6.10)), [23, Prop. 2.15] yields, due to (6.20), that $P_H (j \hat{x}_n - j \overline{x}_n) \to 0$ in $L^q(I, H)$ ($n \to \infty$). From (6.16) we easily deduce that $P_H (j \hat{x}_n - j \overline{x}_n) \to 0$ in $L^q(I, H)$ ($n \to \infty$). Thus, $P_H (j \hat{x}) = P_H (j \overline{x}) = j \overline{x}$ in $L^\infty(I, H)$, where we used that $\overline{x} \in L^p(I, V) \cap \ell L^\infty(I, H)$.

2. Regularity and trace of the weak limit: Let $v \in D$ and $v_n \in V_{m_n}$, $n \in \mathbb{N}$, be a sequence such that $v_n \to v$ in $X$ ($n \to \infty$). Testing (6.4) for $n \in \mathbb{N}$ by $v_n \in V_{m_n}$, multiplication by $\varphi \in C^\infty(\overline{T})$ with $\varphi(T) = 0$, integration over $I$, and integration by parts yields for every $n \in \mathbb{N}$

$$
\langle \mathcal{J}_{m_n} \mathcal{A}[\overline{x}_n, v_n \varphi]_{L^p(I, X)} \cap \ell L^q(I, Y), \mathcal{J}_{m_n} \mathcal{J} \varphi(s) \rangle_{L^p(I, X)} = \int_I ((j \hat{x}_n)(s), jv_n)^\prime \varphi(s) \, ds + (x_{m_n}^0, jv_n)^\prime \varphi(0).
$$

By passing in (6.21) for $n \to \infty$, using (6.16), (6.17), Proposition 5.3 (i), $P_H (j \hat{x}) = j \overline{x}$ in $L^\infty(I, H)$, $x_{m_n}^0 \to x_0$ in $Y$ ($n \to \infty$) and the density of $D$ in $V$, we obtain that for all $v \in V$ and $\varphi \in C^\infty(\overline{T})$ with $\varphi(T) = 0$ there holds

$$
\int_I ((g(s) - f(s), v)^\prime \varphi(s) \, ds = \int_I ((j \hat{x})(s), jv)^\prime \varphi(s) \, ds + (x_0, jv)^\prime \varphi(0)
\leq \int_I ((j \overline{x})(s), jv)^\prime \varphi(s) \, ds + (x_0, jv)^\prime \varphi(0).
$$

In the case $\varphi \in C^\infty_0(I)$ in (6.22), recalling Definition 2.6, we conclude that $\overline{x} \in \mathcal{W}^{1,p,q}_\varepsilon(I, V, H)$ with continuous representation $j \overline{x} \in C^0(\overline{T}, H)$ and

$$
d_{t} \overline{x} = f - g \quad \text{in $L^p(I, V^*) + j^\ast(L^q(I, H^*))$},
$$

Thus, we are able to apply the generalized integration by parts formula in $\mathcal{W}^{1,p,q}_\varepsilon(I, V, H)$ (cf. Proposition 2.7) in (6.22) in the case $\varphi \in C^\infty(\overline{T})$ with $\varphi(T) = 0$ and $\varphi(0) = 1$, which yields for all $v \in V$

$$
((j \overline{x})(0) - x_0, jv)^\prime = 0.
$$

As $j(V)$ is dense in $H$ and $(j \overline{x})(0) \in H$, we deduce from (6.23) that $(j \overline{x})(0) = x_0$ in $H$. 

3. **Pointwise weak convergence:** Next, we show that \( P_H(j \hat{x}_n)(t) \to (j, \varphi)(t) \) in \( H \) (\( n \to \infty \)) for all \( t \in \overline{I} \), which due to (6.20) in turn yields that \( P_H(j \bar{x}_n)(t) \to (j, \varphi)(t) \) in \( H \) (\( n \to \infty \)) for almost every \( t \in \overline{I} \). To this end, let us fix an arbitrary \( t \in I \). From the a-priori estimate \( \| (j \hat{x}_n)(t) \|_Y \leq M \) for all \( t \in \overline{I} \) and \( n \in \mathbb{N} \) (cf. (6.10)) we obtain a subsequence \( ((j \hat{x}_n)(t))_{n \in A_t} \subseteq Y \) with \( A_t \subseteq \mathbb{N} \), initially depending on this fixed \( t \), and an element \( \hat{x}_{A_t} \in Y \) such that

\[
(j \hat{x}_n)(t) \to \hat{x}_{A_t} \quad \text{in} \ Y \quad (A_t \ni n \to \infty).
\] (6.24)

Let \( v \in D \) and \( v_n \in V_{m_n} \), \( n \in \mathbb{N} \), be such that \( v_n \to v \) in \( X \) (\( n \to \infty \)). Then, we test (6.4) for \( n \in A_t \) by \( v_n \in V_{m_n} \), multiply by \( \varphi \in C^\infty(\overline{T}) \) with \( \varphi(0) = 0 \) and \( \varphi(t) = 1 \), integrate over \([0, t]\) and integrate by parts, to obtain for all \( n \in A_t \)

\[
\langle j \hat{x}_n, v_n \hat{x}_n \rangle_{L^p(I,X) \cap L^q(I,Y)} - \int_0^t \langle j \hat{x}_n, v_n \varphi(s) \rangle_{X} \varphi(s) \, ds = - \int_0^t (j \hat{x}_n)(s), jv_n \varphi'(s) \, ds - ((j \hat{x}_n)(t), jv_n)_Y. \tag{6.25}
\]

By passing in (6.25) for \( n \in A_t \) to infinity, using (6.16), (6.17), Proposition 5.3 (i), (6.24) and the density of \( D \) in \( V \), we obtain for all \( v \in V \)

\[
\int_0^t \langle g(s) - f(s), v \rangle_{V} \varphi(s) \, ds = - \int_0^t (j \bar{x}(s), jv)_H \varphi'(s) \, ds - (\hat{x}_{A_t}, jv)_Y. \tag{6.26}
\]

From (6.23), (6.26), the integration by parts formula in \( W^{1,p,q}_e(I, V, H) \) and the properties of \( P_H \) we obtain

\[
0 = ((j, \varphi)(t) - \hat{x}_{A_t}, jv)_Y = ((j, \varphi)(t) - P_H \hat{x}_{A_t}, jv)_H \tag{6.27}
\]

for all \( v \in V \). Thanks to the density of \( j(V) \) in \( H \), (6.27) yields \( (j, \varphi)(t) = P_H \hat{x}_{A_t} \) in \( H \), i.e.,

\[
P_H(j \hat{x}_n)(t) \to (j, \varphi)(t) \quad \text{in} \ H \quad (A_t \ni n \to \infty). \tag{6.28}
\]

As this argumentation stays valid for each subsequence of \( (P_H(j \hat{x}_n)(t))_{n \in \mathbb{N}} \subseteq H \), \( (j, \varphi)(t) \in H \) is a weak accumulation point of each subsequence of \( (P_H(j \hat{x}_n)(t))_{n \in \mathbb{N}} \subseteq H \). The standard convergence principle (cf. [18, Kap. I, Lemma 5.4]) yields \( A_t = \mathbb{N} \) in (6.28). Therefore, using (6.20) and that \( (j, \varphi)(t) = (j \bar{x})(t) \) in \( H \) for almost every \( t \in I \), there holds for almost every \( t \in I \)

\[
P_H(j \bar{x}_n)(t) \to (j \bar{x})(t) \quad \text{in} \ H \quad (n \to \infty). \tag{6.29}
\]

4. **Identification of \( A \bar{x} \) and \( \bar{x} \):** Inequality (6.13) in the case \( \tau = \tau_n \), \( n = m_n \) and \( l = K_n \), using Proposition 5.4 (iii.a), \( (j, \varphi)(0) = x_0 \) in \( H \), \( \| P_H(j \hat{x}_n)(T) \|_H \leq \| (j \hat{x}_n)(T) \|_Y = \| (j \bar{x}_n)(T) \|_Y \)

and \( \langle j \bar{x}_n \rangle_{L^p(I,X) \cap L^q(I,Y)} = \langle f, \bar{x}_n \rangle_{L^p(I,X)} \) for all \( n \in \mathbb{N} \), yields for all \( n \in \mathbb{N} \)

\[
\langle A \bar{x}_n, \bar{x}_n \rangle_{L^p(I,X) \cap L^q(I,Y)} \leq - \frac{1}{2} \| P_H(j \hat{x}_n)(T) \|_H^2 + \frac{1}{2} \| (j \bar{x}_n)(0) \|_H^2 + \| f, \bar{x}_n \|_{L^p(I,X)}. \tag{6.30}
\]
Thus, the limit superior with respect to \( n \in \mathbb{N} \) on both sides in (6.30), (6.16), (6.17), (6.28) with \( A_t = \mathbb{N} \) in the case \( t = T \), the weak lower semi-continuity of \( \| \cdot \|_H \), the integration by parts formula in \( \mathcal{W}^{1,p,q}_{\text{e}}(I, V, H) \) and (6.23) yield

\[
\limsup_{n \to \infty} \langle A\bar{x}_n, \bar{x}_n - \bar{x} \rangle_{L^p(I,X) \cap jL^q(I,Y)} \leq -\frac{1}{2} \| (j, \bar{x})(T) \|_H^2 + \frac{1}{2} \| (j, \bar{x})(0) \|_H^2 + \int_I \langle f(t) - g(t), \bar{x}(t) \rangle_V \ dt \leq -\int_I \langle \frac{d}{dt} \bar{x}(t) + f(s) - g(s), \bar{x}(t) \rangle_V \ dt = 0.
\]

As a result of (6.16), (6.29), (6.31) and the quasi non-conforming Bochner pseudo-monotonicity of \( A : L^p(I, X) \cap jL^q(I, Y) \to (L^p(I, X) \cap jL^q(I, Y))^* \), there holds

\[
\langle A\bar{x}, \bar{x} - y \rangle_{L^p(I,X) \cap jL^q(I,Y)} \leq \liminf_{n \to \infty} \langle A\bar{x}_n, \bar{x}_n - y \rangle_{L^p(I,X) \cap jL^q(I,Y)} \leq \langle \bar{x}^*, \bar{x} - y \rangle_{L^p(I,X) \cap jL^q(I,Y)}
\]

for any \( y \in L^p(I, X) \cap jL^q(I, Y) \), which in turn implies that \( A\bar{x} = \bar{x}^* \) in \( (L^p(I, X) \cap jL^q(I, Y))^* \). This completes the proof of Theorem 1.7. \( \square \)

7 Application: \((p\text{-}\text{Navier-Stokes equations}, \ p > \frac{3d+2}{d+2})\)

We follow the procedure of Section 6 in order to prove the well-posedness, stability and weak convergence of the quasi non-conforming Rothe-Galerkin scheme (1.16) of the to the unsteady \( p \)-Navier-Stokes equations (1.2) corresponding evolution equation (1.4) by means of discretely divergence-free FEM spaces.

**Assumption 7.1** Let \( \Omega \subseteq \mathbb{R}^d, \ d \geq 2, \) a bounded polygonal Lipschitz domain, \( I := (0, T), \ T < \infty, \) and \( p > \frac{3d+2}{d+2} \). We make the following assumptions:

(i) \( (V, H, \text{id}), (X, Y, \text{id}) \) and \( (V_n)_{n \in \mathbb{N}} \) are defined as in Proposition 3.2.

(ii) \( u_0 \in H \) and \( u_0^0 \in V_n, \ n \in \mathbb{N} \), such that \( u_0^0 \to u_0 \) in \( Y \) \( (n \to \infty) \) and \( \sup_{n \in \mathbb{N}} \| u_0^0 \|_Y \leq \| u_0 \|_H \).

(iii) \( f \in L^p(I, X^*). \)

(iv) \( S(t) : X \to X^*, \ t \in I, \) and \( \hat{B} : X \to X^* \) are defined as in Proposition 4.21.

Furthermore, we denote by \( e := (\text{id}_V)^* R_H : V \to V^* \) the canonical embedding with respect to the evolution triple \((V, H, \text{id})\). Let us next recall the quasi non-conforming Rothe-Galerkin scheme we already revealed in the introduction.

**Algorithm 7.2** Let Assumption 7.1 be satisfied. For given \( K, n \in \mathbb{N} \) the sequence of iterates \((u^k_n)_{k=0,\ldots,K} \subseteq V_n\) is given via the implicit Rothe-Galerkin scheme for \( \tau = \frac{T}{K} \)

\[
(d_{\tau} u^k_n, v_n)_Y + \langle [S^*_{\tau}] u^k_n, v_n \rangle_X + \langle \hat{B} u^k_n, v_n \rangle_X = \langle [f]_{\tau}^*, v_n \rangle_X \quad \text{for all} \ v_n \in V_n, \quad (7.3)
\]

By means of Proposition 6.6, Proposition 6.8, Theorem 1.7 and the observation we already made in Proposition 4.21, we can immediately conclude the following results.
Theorem 7.4 (Well-posedness, stability and weak convergence of (7.3))

Let Assumption 7.1 be satisfied. Then, it holds:

(I) **Well-posedness:** For every $K, n \in \mathbb{N}$ there exist iterates $(u^k_n)_{k=0,\ldots,K} \subseteq V_n$, solving (7.3), without any restrictions on the step-size.

(II) **Stability:** The corresponding piece-wise constant interpolants \( \overline{u}_n^k \in S^0(I, X) \), $K, n \in \mathbb{N}$ with $\tau = \frac{T}{K}$, are bounded in $L^p(I, X) \cap L^\infty(I, Y)$.

(III) **Weak convergence:** If $(\overline{u}_n)_{n \in \mathbb{N}} := (\overline{u}_n^m)_{n \in \mathbb{N}}$, where $\tau_n = \frac{T}{K}$ and $K_n, m_n \to \infty (n \to \infty)$, is an arbitrary diagonal sequence of the piece-wise constant interpolants $\overline{u}_n^m \in S^0(I, X)$, $K, n \in \mathbb{N}$ with $\tau = \frac{T}{K}$, then there exists a not relabelled subsequence and a weak limit $\overline{u} \in L^p(I, V) \cap L^\infty(I, H)$ such that

\[
\overline{u}_n^k \to \overline{u} \quad \text{in} \quad L^p(I, X), \\
\overline{u}_n \rightharpoonup \overline{u} \quad \text{in} \quad L^\infty(I, Y), \quad (n \to \infty).
\]

Furthermore, it follows that $\overline{u} \in W^{1,p,p}_e(I, V, H) \cap L^\infty(I, H)$ satisfies $\overline{u}(0) = u_0$ in $H$ and for all $\phi \in L^p(I, V)$

\[
\int_I \left\langle \frac{d}{dt} \overline{u}(t), \phi(t) \right\rangle_V dt + \int_I \left\langle S(t)(\overline{u}(t)) + B(\overline{u}(t)), \phi(t) \right\rangle_X dt = \int_I \left\langle f(t), \phi(t) \right\rangle_X dt.
\]

**Proof ad (I)/(II)** The assertions follow immediately from Proposition 6.6 and Proposition 6.8, since the operator family $A(t) := S(t) + \hat{B} : X \to X^*$, $t \in I$, satisfies (A.1)–(A.4) with $c_1 = 0$ due to Proposition 4.21.

ad (III) The assertions follow from Theorem 1.7. To be more precise, Theorem 1.7 initially yields that $\overline{u} \in W^{1,p,p}_e(I, V, H)$, where $q > 1$ is specified in the proof of Proposition 4.21, satisfies $\overline{u}(0) = u_0$ in $H$ and for all $\phi \in C^1_0(I, V)$

\[
\int_I \left\langle \frac{d}{dt} \overline{u}(t), \phi(t) \right\rangle_V dt + \int_I \left\langle S(t)(\overline{u}(t)) + \hat{B}(\overline{u}(t)), \phi(t) \right\rangle_X dt = \int_I \left\langle f(t), \phi(t) \right\rangle_X dt.
\]

Since $\langle \hat{B}(\overline{u}(t)), \phi(t) \rangle_X = \langle B(\overline{u}(t)), \phi(t) \rangle_X$ for almost every $t \in I$ and all $\phi \in C^1_0(I, V)$ as well as $B(\overline{u}(\cdot)) \in L^p(I, V^*)$ (cf. [23, Example 5.1]), as $\overline{u}(t) \in L^p(I, V) \cap L^\infty(I, H)$, we actually proved that $\overline{u} \in W^{1,p,p}_e(I, V, H) \cap L^\infty(I, H)$, such that for all $\phi \in C^1_0(I, V)$

\[
\int_I \left\langle \frac{d}{dt} \overline{u}(t), \phi(t) \right\rangle_V dt + \int_I \left\langle S(t)(\overline{u}(t)) + B(\overline{u}(t)), \phi(t) \right\rangle_X dt = \int_I \left\langle f(t), \phi(t) \right\rangle_X dt. \quad \square
\]

**Remark 7.5** The results in Theorem 7.4 are among others already contained in [34] (cf. [31]). There a numerical analysis of the unsteady $p$-Navier-Stokes equations using the framework of maximal monotone graphs is performed. The convergence of a conforming implicitly fully discrete Rothe-Galerkin scheme of an evolution problem with Bochner pseudo-monotone operators is proved in [6]. Convergence results with optimal rates for the unsteady $p$-Navier-Stokes equations for $p \leq 2$ and space periodic boundary conditions can be found in [8]. Optimal convergence rates in the case of homogeneous Dirichlet boundary conditions for unsteady $p$-Stokes equations can be found in [16]. First results for optimal convergence rates of a fully space-time discretization for an unsteady problem with $(p, \delta)$-structure depending on the symmetric gradient are contained in [9] (cf. [10] for a setting with variable exponents).
8 Numerical experiments

Eventually, we want to present some numerical experiments with low regularity data which perfectly suit the framework of this article, but are to irregular to apply usual optimal convergence results, see e.g. Remark 7.5 for an overview. The choice of appropriate data for the first experiment is motivated by [13]. However, differing from [13] we will consider constant exponents and a time-dependent viscosity. All numerical experiments were conducted using the finite element software FEniCS [24]. The graphics are generated using the Matplotlib library [21].

Experiment 8.1 Let \( \Omega = (0, 3) \times (0, 1) \subseteq \mathbb{R}^2 \), \( T = 1 \), \( p = 11 \), and \( f = 0 \). The initial data \( u_0 \in H \) is described in Figure 1, in which \( |u_0| = 1 \) in the left vortex, i.e., in \((0, 2) \times (0, 1)\), and \( |u_0| = 10 \) in the right vortex, i.e., in \((2, 3) \times (0, 1)\). The spatial discretization of our domain \( \Omega \) is obtained by a uniform finite element mesh consisting of triangles with straight sides, shown in Figure 2. The mapping \( S : Q_T \times M^{d \times d}_{\text{sym}} \rightarrow M^{d \times d}_{\text{sym}} \) is given via \( \nu(t, x)(\delta + |A|)^{p-2} A \) in \( M^{d \times d}_{\text{sym}} \) for almost every \((t, x) \in Q_T\) and all \( A \in M^{d \times d}_{\text{sym}}\), where \( \nu \in L^\infty(Q_T) \) is given via \( \nu(t, x) = t^2 + \exp(-x_1^2 + x_2^2) \) for almost every \((t, x) \in Q_T\) and \( \delta = 1e^{-2} \). We consider both the MINI element (cf. Figure 3) and the conforming Crouzeix-Raviart element (cf. Figure 4). Moreover, we choose the step-size \( \tau = 1e^{-2} \), i.e., \( K = 100 \). Then, the iterates \((u^k)_{k=0}^K \subseteq V_0\), solving (7.3), are approximated by means of Newton’s iteration, where the linear system emerging in each Newton step is solved using the LU solver of PETSc [2]. Under all these circumstances, we gain the following pictures:

![Fig. 1: Velocity field of initial data \( u_0 \in H \). The length of the vectors is suitable scaled.](image)

![Fig. 2: Uniform mesh consisting of 48 x 16 rectangles, each divided into a pair of triangles, i.e., 1536 triangles.](image)

Leaving the discontinuities of the pressure in the conforming Crouzeix-Raviart element aside, we observe that both elements produce similar pictures. More precisely, the right (faster) vortex...
clearly dominates the left (slower) vortex, in the sense that the right vortex quickly accelerates the left vortex while simultaneously decelerating itself. This acceleration takes place in the discontinuity line \( \{2\} \times (0,1) \) of the velocity field and thus causes a large magnitude of the pressure, when both vortexes clash together with different orientations around the points \((2,0)^T\) and \((2,1)^T\). Still, we observe that this evolution quickly results in an equilibrium state, since we already see at the relatively short time scale \(t = 0.2\) that both vortexes have more or less the same average length of vectors.

In the next experiment we are interested in the convergence speed of the scheme (7.3) in the case of low regularity data, which just falls into the scope of application of the weak convergence result Theorem 7.4. The considered data is mainly motivated by [7].
Experiment 8.2 Let $\Omega = (-1,1)^2 \subseteq \mathbb{R}^2$, $T = \frac{1}{2}$, $Q_T := I \times \Omega$, $\Gamma := I \times \partial \Omega$ and $p = \frac{11}{6}$.
We consider solutions with a point singularity at the origin in the velocity. More precisely, for $\alpha = \frac{6}{5} - \frac{2}{p} \approx 0.291$, let
\[
\mathbf{u}(t,\mathbf{x}) := \left( \frac{x_2}{r^2} \right) + |\mathbf{x}|^{\alpha - 1} \left( \frac{x_2}{-x_1} \right), \quad \mathbf{p}(t,\mathbf{x}) := 0.
\]
The mapping $S : \mathbb{M}^{2 \times 2}_{\text{sym}} \to \mathbb{M}^{2 \times 2}_{\text{sym}}$ is given via $S(A) := (\delta + |A|^2)^{p-2}A$ in $\mathbb{M}^{2 \times 2}_{\text{sym}}$ for all $A \in \mathbb{M}^{2 \times 2}_{\text{sym}}$, where $\delta = 1 \times 4$. Then, choosing $f := \partial_t \mathbf{u} - \text{div} S(D\mathbf{u}) + \text{div}(\mathbf{u} \otimes \mathbf{u}) \in L^p(I, (W^{1,p'}_0(\Omega))^d)$, $u(0) := u_0 \in W^{1,p}(\Omega) := \{v \in W^{1,p}(\Omega)^2 \mid \text{div } v = 0\}$ and $u_D := u|_{\Gamma_T}$, we trivially have
\[
\partial_t \mathbf{u} - \text{div} S(D\mathbf{u}) + \text{div}(\mathbf{u} \otimes \mathbf{u}) + \nabla p = f \quad \text{in } Q_T, \\
\text{div } \mathbf{u} = 0 \quad \text{in } Q_T, \\
\mathbf{u} = u_D \quad \text{in } \Gamma_T, \\
\mathbf{u}(0) = u_0 \quad \text{in } \Omega.
\]
In particular, note that the parameter $\alpha$ is chosen so small that just $\mathbf{u}(t) \in W^{1,p'}(\Omega)^2$ for every $t \in T$, as then $|D\mathbf{u}(t)| \sim |\cdot|^{\alpha - 1} \in L^p(\Omega)$, but $u(t) \notin W^{2,1}(\Omega)^2$ for every $t \in T$. Similar, this choice guarantees $S(D\mathbf{u}(t)) \in L^p(\Omega)^{2 \times 2}$ for every $t \in T$, but neither $S(D\mathbf{u}(t)) \notin W^{1,p'}(\Omega)^{2 \times 2}$, nor $\text{div } S(D\mathbf{u}(t)) \notin L^p(\Omega)^2$ for every $t \in T$. Thus, the right-hand side has just enough regularity, namely $f \in L^p(I, (W^{1,p'}_0(\Omega)^2)^d)$, to fall into the framework of our weak convergence result Theorem 7.4. In consequence, we can expect (weak) convergence of the scheme (7.3) for this choice of $\alpha$. However, the attaching experimental results indicate that we presumably cannot expect convergence with rates.

![Fig. 5: Snapshots of the exact solution $u(t)$ at times $t = 0, 0.25, 0.5$.](image)

The spatial discretization of our domain $\Omega$ is obtained by a sequence of uniform finite element meshes $(\mathcal{T}_{hn})_{n \in \mathbb{N}}$, consisting of triangles with straight sides and diameter $h_n := \frac{h_0}{n}$, $h_0 := 2\sqrt{2}$, for every $n \in \mathbb{N}$. Beginning with $\mathcal{T}_{h_1}$, see e.g. Figure 6 the first and the third picture, and for every $n \in \mathbb{N}$ with $n \geq 2$ the mesh $\mathcal{T}_{hn}$ is a refinement of $\mathcal{T}_{hn-1}$ obtained by dividing each triangle into four, which is based an edge midpoint or regular 1:4 refinement algorithm.

---

6 The exact solution is not zero on the boundary of the computational domain. However, the error is mainly concentrated around the singularity in the origin and thus small inconsistency with the setup of the theory does not have any influence in the results.
Once more we consider both the MINI element (cf. Table 1), the Taylor-Hood element (cf. Table 2) and the conforming Crouzeix-Raviart element (cf. Table 3). Moreover, we choose the step-size \( \tau = 4e^{-3} \), i.e., \( K = 250 \). Then, the iterates \( (u^k_{hn})_{k=0,...,K} \subseteq V_n \), solving (7.3), are again approximated by means of Newton’s iteration. Let the mapping \( F : \mathbb{M}^{2\times 2}_{\text{sym}} \rightarrow \mathbb{M}^{2\times 2}_{\text{sym}} \) be given via \( F(A) := (\delta + |A|)^{p-2}A \) for every \( A \in \mathbb{M}^{2\times 2}_{\text{sym}} \). We are interested in the parabolic errors
\[
e_{F,h_n} := \left( \sum_{k=0}^{K} \tau \|F(Du(t_k)) - F(Du^k_{hn})\|_{L^2(\Omega)^d}^d \right)^{\frac{1}{2}},
\]
\[
e_{L^2,h_n} := \max_{0 \leq k \leq K} \|u(t_k) - u^k_{hn}\|_{L^2(\Omega)^d}, \quad n = 1, ..., 7,
\]
which are approximations of the errors \( \|F(Du) - F(Du^k_{hn})\|_{L^2(Q_T)^d} \) and \( \|u - u^k_{hn}\|_{L^\infty(Q_T)} \). In particular, we are interested in the total parabolic error \( e_{\text{tot},h_n} := e_{F,h_n} + e_{L^2,h_n}, \quad n = 1, ..., 7 \). As an estimation of the convergence rates, we use the experimental order of convergence (EOC):
\[
\text{EOC}(e_{h_n}) := \frac{\log(e_{h_n}/e_{h_{n-1}})}{\log(h_n/h_{n-1})}, \quad n = 2, ..., 7,
\]
where \( e_{h_n}, n = 2, ..., 7 \), either denote \( e_{F,h_n}, e_{L^2,h_n} \) or \( e_{\text{tot},h_n}, n = 2, ..., 7 \), respectively. In order to obtain higher accuracy in the computation of these errors, especially in view of the singularity of the exact solution around the origin, we interpolate both \( u(t_k) \) and \( u^k_{hn} \) into a higher polynomial space with respect to an appropriate refined mesh, namely into the interior of the Taylor-Hood element. Moreover, we choose the step-size \( \tau \) to be such that we are not able to extract experimental convergence rates. This circumstance may be traced back two the low regularity of the data provided by the exact solution controlled by the parameter \( \alpha > 0 \). In fact, for even lower values of \( \alpha > 0 \), e.g., for \( \alpha = \frac{11}{10} - \frac{2}{p} \approx 0.19 \), the authors were no longer able to observe (strong) convergence of the scheme (7.3). In addition, Figure 7 and Figure 8 illustrate the temporal developments of the errors \( \|F(Du(t_k)) - F(Du^k_{hn})\|_{L^2(\Omega)^d} \) and \( \|u(t_k) - u^k_{hn}\|_{L^2(\Omega)^d}, k = 0, ..., 250 \).
Fully discrete, quasi non-conforming approximation of evolution equations

<table>
<thead>
<tr>
<th>$h_n = \frac{h_0}{2^n}$</th>
<th>$\epsilon_{L^2,h_n}$</th>
<th>$\text{EOC}(\epsilon_{L^2,h_n})$</th>
<th>$\epsilon_{F,h_n}$</th>
<th>$\text{EOC}(\epsilon_{F,h_n})$</th>
<th>$\text{EOC}(\epsilon_{\text{tot},h_n})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n = 1$</td>
<td>$\sqrt{2} \approx 1.414$</td>
<td>$4.698e-1$</td>
<td>-</td>
<td>$1.256$</td>
<td>-</td>
</tr>
<tr>
<td>$n = 2$</td>
<td>$\sqrt{3} \approx 1.732$</td>
<td>$2.451e-1$</td>
<td>$0.939$</td>
<td>$1.062$</td>
<td>$0.243$</td>
</tr>
<tr>
<td>$n = 3$</td>
<td>$\sqrt{4} \approx 2$</td>
<td>$1.578e-1$</td>
<td>$0.635$</td>
<td>$9.407e-1$</td>
<td>$0.174$</td>
</tr>
<tr>
<td>$n = 4$</td>
<td>$\sqrt{5} \approx 2.236$</td>
<td>$1.037e-1$</td>
<td>$0.606$</td>
<td>$8.484e-1$</td>
<td>$0.149$</td>
</tr>
<tr>
<td>$n = 5$</td>
<td>$\sqrt{6} \approx 2.449$</td>
<td>$6.940e-2$</td>
<td>$0.579$</td>
<td>$7.602e-1$</td>
<td>$0.158$</td>
</tr>
<tr>
<td>$n = 6$</td>
<td>$\sqrt{7} \approx 2.645$</td>
<td>$5.293e-2$</td>
<td>$0.391$</td>
<td>$6.802e-1$</td>
<td>$0.161$</td>
</tr>
<tr>
<td>$n = 7$</td>
<td>$2.21e-2$</td>
<td>$4.610e-2$</td>
<td>$0.199$</td>
<td>$6.138e-1$</td>
<td>$0.148$</td>
</tr>
</tbody>
</table>

Table 1: Error analysis for the MINI element.

<table>
<thead>
<tr>
<th>$h_n = \frac{h_0}{2^n}$</th>
<th>$\epsilon_{L^2,h_n}$</th>
<th>$\text{EOC}(\epsilon_{L^2,h_n})$</th>
<th>$\epsilon_{F,h_n}$</th>
<th>$\text{EOC}(\epsilon_{F,h_n})$</th>
<th>$\text{EOC}(\epsilon_{\text{tot},h_n})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n = 1$</td>
<td>$\sqrt{2} \approx 1.414$</td>
<td>$1.876e-1$</td>
<td>-</td>
<td>$9.336e-1$</td>
<td>-</td>
</tr>
<tr>
<td>$n = 2$</td>
<td>$\sqrt{3} \approx 1.732$</td>
<td>$1.261e-1$</td>
<td>$0.573$</td>
<td>$8.289e-1$</td>
<td>$0.172$</td>
</tr>
<tr>
<td>$n = 3$</td>
<td>$\sqrt{4} \approx 2$</td>
<td>$1.058e-1$</td>
<td>$0.256$</td>
<td>$7.527e-1$</td>
<td>$0.139$</td>
</tr>
<tr>
<td>$n = 4$</td>
<td>$\sqrt{5} \approx 2.236$</td>
<td>$8.191e-2$</td>
<td>$0.370$</td>
<td>$6.919e-1$</td>
<td>$0.122$</td>
</tr>
<tr>
<td>$n = 5$</td>
<td>$\sqrt{6} \approx 2.449$</td>
<td>$6.206e-2$</td>
<td>$0.400$</td>
<td>$6.295e-1$</td>
<td>$0.136$</td>
</tr>
<tr>
<td>$n = 6$</td>
<td>$\sqrt{7} \approx 2.645$</td>
<td>$5.090e-2$</td>
<td>$0.286$</td>
<td>$5.724e-1$</td>
<td>$0.137$</td>
</tr>
<tr>
<td>$n = 7$</td>
<td>$2.21e-2$</td>
<td>$4.552e-2$</td>
<td>$0.161$</td>
<td>$5.263e-1$</td>
<td>$0.121$</td>
</tr>
</tbody>
</table>

Table 2: Error analysis for the Taylor-Hood element.

<table>
<thead>
<tr>
<th>$h_n = \frac{h_0}{2^n}$</th>
<th>$\epsilon_{L^2,h_n}$</th>
<th>$\text{EOC}(\epsilon_{L^2,h_n})$</th>
<th>$\epsilon_{F,h_n}$</th>
<th>$\text{EOC}(\epsilon_{F,h_n})$</th>
<th>$\text{EOC}(\epsilon_{\text{tot},h_n})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n = 1$</td>
<td>$\sqrt{2} \approx 1.414$</td>
<td>$2.085e-1$</td>
<td>-</td>
<td>$9.627e-1$</td>
<td>-</td>
</tr>
<tr>
<td>$n = 2$</td>
<td>$\sqrt{3} \approx 1.732$</td>
<td>$1.336e-1$</td>
<td>$0.643$</td>
<td>$8.431e-1$</td>
<td>$0.191$</td>
</tr>
<tr>
<td>$n = 3$</td>
<td>$\sqrt{4} \approx 2$</td>
<td>$1.091e-1$</td>
<td>$0.292$</td>
<td>$7.654e-1$</td>
<td>$0.139$</td>
</tr>
<tr>
<td>$n = 4$</td>
<td>$\sqrt{5} \approx 2.236$</td>
<td>$8.400e-2$</td>
<td>$0.388$</td>
<td>$7.041e-1$</td>
<td>$0.121$</td>
</tr>
<tr>
<td>$n = 5$</td>
<td>$\sqrt{6} \approx 2.449$</td>
<td>$6.252e-2$</td>
<td>$0.416$</td>
<td>$6.401e-1$</td>
<td>$0.137$</td>
</tr>
<tr>
<td>$n = 6$</td>
<td>$\sqrt{7} \approx 2.645$</td>
<td>$5.109e-2$</td>
<td>$0.291$</td>
<td>$5.812e-1$</td>
<td>$0.139$</td>
</tr>
<tr>
<td>$n = 7$</td>
<td>$2.21e-2$</td>
<td>$4.567e-2$</td>
<td>$0.161$</td>
<td>$5.334e-1$</td>
<td>$0.124$</td>
</tr>
</tbody>
</table>

Table 3: Error analysis for the conforming Crouzeix-Raviart element.

References

Fig. 7: Temporal development of the errors $\| F(Du(t_k)) - F(Du_k) \|_{L^2(\Omega)^2 \times 2}$, $k = 0, \ldots, 250$.  

Fig. 8: Temporal development of the errors $\| u(t_k) - u_k \|_{L^2(\Omega)^2}$, $k = 0, \ldots, 250$.  