

# Variational Space-Time Elements for Large-Scale Systems

C. Hesch<sup>a\*</sup>, S. Schuß<sup>a</sup>, M. Dittmann<sup>a</sup>, S. R. Eugster<sup>b</sup>,  
M. Favino<sup>c</sup> and R. Krause<sup>c</sup>

<sup>a</sup> Chair of Computational Mechanics, University of Siegen, Siegen, Germany

<sup>b</sup> Institute for Nonlinear Mechanics, University of Stuttgart, Stuttgart, Germany

<sup>c</sup> Institute of Computational Science, Università della Svizzera Italiana, Lugano, Switzerland

## Abstract

In this paper, we introduce a new Galerkin based formulation for transient continuum problems, governed by partial differential equations in space and time. Therefore, we aim at a direct finite element discretisation of the space-time, suitable for massive parallel analysis of the arising large-scale problem. The proposed formulation is applied to thermal, mechanical and fluid systems, as well as to a Kuramoto-Sivashinsky problem, representing the general class of higher-order formulations in material science using NURBS based shape functions. We verify whenever possible the conservation properties of the formulation. Finally, a series of examples demonstrate the applicability to all systems presented in this paper.

**Keywords:** Space-time, large-scale, Galerkin, solids, fluids, higher-order, IGA, NURBS, parallel, multigrid

## 1 Introduction

The most important mathematical framework for the numerical analysis of partial differential equations (PDEs) has been developed by Galerkin using of a weak form on a Hilbert space. On this basis, Courant formulated in 1943 a variational method for the solution of Galerkin type problems [13]. This leads to the first application of finite elements in 1956 by Turner et al. [67]. Driven by statical problems in engineering, finite element solutions of spatial systems have been further improved (see, among many others, Zienkiewicz et al. [69] and Belytschko et al. [62, 63]) and higher continuity basis functions like B-Splines in the context of Isogeometric Analysis (IGA) was introduced by Hughes et al. [40]. Further extensions have been proposed throughout the past two decades for the calculation of multi-physical phenomena, e.g. thermomechanical Systems (see, among many other, Simo & Holzapfel [36], Hesch & Betsch [33]), fluid systems (see Donea & Huerta [18]), material science (see e.g. Anders et al. [1]) and many other formulations.

The fundament remains: variational formulations in space are widely used in structural mechanics, whereas in the temporal regime, both, weak and strong forms have been

---

\*Corresponding author. E-mail address: christian.hesch@uni-siegen.de

proposed, see Hughes et al. [49, 35] for an example of a classical time stepping method. First applications on finite elements in time can be traced back to Bailey [4], dealing with the transient behaviour of beams. Further developments can be found in Gellin & Pitarresi [29], Pitarresi & Manolis [64] and Atilgan et al. [66], see also Bottasso [54]. Space-time finite element methods have also been considered in fluid dynamics, see Hughes and Hulbert [43], Hughes et al. [42] and Shakib and Hughes [60], for two-phase flow in [51], see although Betsch et al. [6, 7, 8] for a series on finite element formulations in time. In the majority of finite element formulations in time, the semidiscrete equations are, however, multiplied by weighting functions in space and time and integrated over time. As already shown in Hughes & Hulbert [43], *this negates one of the most powerful features of the finite element method: unstructured meshes*. More importantly, it restricts adaptive methodologies either to space only or/and to global refinement in time.

Due to restricted computational capacities, and above all due to our own personal experience, all methods have been designed to march forward in time. Even the above mentioned finite element formulations in time have been formulated using a Petrov-Galerkin method with different shape function of the functional solution and test spaces to obtain a time-discontinuous Galerkin formalism. This approach leads to a lower triangular matrix for the total system to be solved in space and time and ensures that information flows always in the direction of positive time, i.e. solutions are said to be *causal* since they depend upon the past but not on the future. And although this is natural in our human perspective, it certainly may not be optimal from a computational point of view. We note, however, that the first time-parallel approach for ordinary differential equations was already presented more than 50 years ago in [55].

As a remedy, we propose a Bubnov-Galerkin approach using continuous finite elements in time and space-time, i.e. we apply the same shape functions for the solution as well as for the test functional space. This approach can be derived in a variationally consistent way, hence we can demonstrate all necessary conservation properties of the different systems under consideration. The application of domain-decomposition methods is formally straight forward, we only have to consider to work in the  $\mathbb{R}^{n+1}$  dimensional space-time. For two-dimensional problems in space, existing solvers for three-dimensional problems can be applied directly to solve the arising, massive large-scale problem. However, certain modifications might be necessary as the differential operator in time behaves differently than in space.

Multigrid methods, which have been developed since the 50's of the last century, allow for the solution of sparse symmetric positive definite linear system with optimal complexity, cf. [32]. Thanks to their optimality and also their parallel scalability, multigrid methods have turned into a standard choice for the solution of large scale systems arising from the discretization of elliptic PDEs. During the last decades, extensions of multigrid have been developed which have also shown to be efficient and robust also in non-linear systems, such as frictional non-linear elasticity, and saddle-point problems. Multigrid methods have also been successfully applied to the linear systems arising from the full space-time discretization of parabolic systems, see, e.g. [37, 26]. The fact that

for parabolic equations, the differential operator in time direction is only of first order and not of second order, as it is in space, requires some modifications of the multigrid method. These modifications are realized usually by special choices of the smoother, i.e. line smoothers, or by special coarsening strategies. In particular the modifications of the smoother are related to smoothing strategies which have been proposed for convection-dominated elliptic problems, such as convection-diffusion. Another possibility is to introduce on the discrete level mesh-dependent diffusion in time, cf. [47]. For general non-linear problems, multigrid methods for non-linear problems can be employed, leading to the PFASST-method (Parallel Full Approximation Scheme in Space and Time), see [19]. Again, the treatment of the time-direction within a multi-level scheme requires particular care, cf., e.g. [61].

In general, for the discretization in time direction for space-time multigrid methods, in the literature the following approaches can be identified:

- to use discontinuous Galerkin in time. This approach is the most widely used, see e.g. [27, 25, 47], based on the formulation originally presented in [45].
- to use tailored test functions, containing a time-derivative of the space-time continuous test functions. This means to replace a standard test function  $\delta\theta$  with  $\delta\theta \Rightarrow \nabla_t \delta\theta$ , or to add this expression to the otherwise unchanged test-function. This formalism has been presented in [3] and has been shown to be equivalent to a Petrov-Galerkin formulation
- to use continuous ansatz-spaces in space and time, see, e.g. [24] and [2]. We note that these works are dealing with the wave equation, where time and space derivative are of the same order.

Space-time multigrid for parabolic problems, using finite differences and a Gauß-Seidel smoother have been used in [31]. Interestingly, if only coarsening in space is applied, standard multigrid performance is achieved. If coarsening in time is done nevertheless, standard multigrid performance is not guaranteed, and the method might even diverge. This behaviour can be traced back to errors which are smooth in space, but non-smooth in time.

In [37], multigrid is applied in space-time for parabolic equations. For the 2-dimensional or 3-dimensional heat equation, good  $V$ -cycle contraction rates are obtained, but mesh independent convergence rates is achieved only for  $F$ -cycles. This is in contrast to multigrid for elliptic problems, which shows already mesh-independence for the  $V$ -cycle. Horton and Vandewalle propose also an adaptive coarsening strategy based on the space and time discretisation parameters  $\Delta x$  and  $\Delta t$ . Defining an inverse numerical diffusion

$$\lambda = \frac{\Delta t}{\Delta x^2}, \quad (1)$$

they showed that standard space-time multigrid provided good convergence rates only for  $\lambda$  larger than a critical value  $\lambda_{\text{crit}}$ , which depends on the accuracy of the time discretisation. For  $\lambda < \lambda_{\text{crit}}$ , only time coarsening was numerically proven to be effective.

Recently, a space-time multigrid based on a discontinuous Galerkin formulation in time and standard finite element in space for parabolic problems has been presented [25], which uses a fully parallel block Jacobi smoother. The MGRIT approach (Multigrid Reduction in Time) can be found in [20], where standard one-step methods are employed as smoothers in a multilevel method in time. For hyperbolic problems, the situation is more complex than for the parabolic case. Space-time discontinuous Galerkin methods for the compressible Navier-Stokes equation have been considered in [46], and multigrid methods for Navier-Stokes with periodic boundary conditions and point-block smoothers can be found in [5].

From a multigrid point of view, the main difficulty for space-time solution is that in time we have to deal with a purely convective problem (after possible reduction to a first order system). The multigrid approach considered here is therefore based on a stabilization in time. Here, we follow the approach of [26, 47] and add diffusion in time. We moreover adapt our coarsening strategy the strategy proposed and analysed in [37, 25].

For the application on continuous Lagrangian finite elements in time and space-time, we construct the coarser approximation spaces using semi-geometric multigrid methods, see [16]. These methods create a nested hierarchy of finite element spaces based on a hierarchy of possibly non-nested meshes. To this end, a discrete (pseudo-)  $L^2$ -projection operator between the finite element spaces related to the non-nested meshes is computed, [48]. The weights of the resulting scaled mass-matrices are then used for defining the coarse level spaces based on linear combinations of the fine level basis functions. This means that we can easily construct the desired multi-level hierarchies by providing appropriate coarser meshes which fulfil the condition (1). We note the one advantage of this approach is the resulting low operator complexity.

We start with a most common description in Section 2 followed by a first and easy to understand application to thermal systems in Section 3. Afterwards, we demonstrate the application to mechanical systems (see Section 4), using a direct discretisation of Hamilton's principle of varying action in the Hamiltonian framework. The following two sections 5 and 6 deal with fluids as well as a Kuramoto-Sivashinsky formulation. The fluid system will demonstrate the general applicability of the proposed formulation to constrained (incompressible) fluids. We notice that finite volume and similar formulations are more common to the fluid community, so we present here a proof of concept for fluid systems within the scope of this paper using variational consistent finite element formulation. The Kuramoto-Sivashinsky model is chosen as a prototypical higher-order formulation in material science using a state of the art Isogeometric Analysis (IGA) approach based on NURBS shape functions in space-time. Eventually, in Section 7, we discuss modifications that can be done in order to achieve multigrid performance in Bubnov-Galerkin formulations of problems which are only first order in time. Finally, a series of examples in Section 8 demonstrates the applicability as well as the accuracy of the proposed formulation along with additional investigations on the multigrid performance.

## 2 Space-time formulation

Let  $\mathcal{B} \subset \mathbb{R}^n$ ,  $n \in \{1, 2, 3\}$ , be a Lipschitz bounded domain and  $I = [t_1, t_2] \subset \mathbb{R}$ ,  $t_1 < t_2$ , a temporal interval. We consider systems of  $m$  partial differential equations<sup>†</sup> in the general form

$$\mathbf{A}(\mathbf{u}(\mathbf{x}, t), \mathbf{x}, t) \nabla_t \mathbf{u}(\mathbf{x}, t) + \mathbf{L}\mathbf{u}(\mathbf{x}, t) = \mathbf{f}(\mathbf{x}, t) \quad \forall (\mathbf{x}, t) \in \mathcal{B} \times I, \quad (2)$$

along with initial- and boundary-conditions

$$\begin{aligned} \mathbf{u}(\mathbf{x}, t) &= \mathbf{u}_0(\mathbf{x}) & \text{on } (\mathbf{x}, t) \in \mathcal{B} \times \{t_1\}, \\ \mathbf{u}(\mathbf{x}, t) &= \mathbf{g}(\mathbf{x}, t) & \text{on } (\mathbf{x}, t) \in \partial\mathcal{B}^u \times I, \\ (\mathbf{R}\mathbf{u})(\mathbf{x}, t) &= \mathbf{h}(\mathbf{x}, t) & \text{on } (\mathbf{x}, t) \in \partial\mathcal{B}^\sigma \times I. \end{aligned} \quad (3)$$

In keeping with common nomenclature, we will denote the Dirichlet boundary conditions  $\partial\mathcal{B}^u$  and Neumann conditions  $\partial\mathcal{B}^\sigma$ , respectively, both satisfying  $\partial\mathcal{B} = \partial\mathcal{B}^u \cup \partial\mathcal{B}^\sigma$  and  $\partial\mathcal{B}^u \cap \partial\mathcal{B}^\sigma = \emptyset$  throughout the time interval  $I$ . Moreover, we assume that  $\mathbf{A} = \text{diag}(a_1, \dots, a_m): \mathbb{R}^m \times \mathcal{B} \times I \rightarrow \mathbb{R}^{m \times m}$  and  $\mathbf{f}: \mathcal{B} \times I \rightarrow \mathbb{R}^m$  are sufficiently smooth functions,  $\nabla_t$  denotes the partial derivative with respect to  $t$  and  $\mathbf{L} = (\mathbf{L}_1, \dots, \mathbf{L}_m)$ ,  $\mathbf{R} = (\mathbf{R}_1, \dots, \mathbf{R}_m)$  are component-wise defined differential operators,

$$\mathbf{L}\mathbf{u} := (\mathbf{L}_1\mathbf{u}_1, \dots, \mathbf{L}_m\mathbf{u}_m), \quad \mathbf{R}\mathbf{u} := (\mathbf{R}_1\mathbf{u}_1, \dots, \mathbf{R}_m\mathbf{u}_m), \quad (4)$$

acting on the spatial coordinates of  $\mathbf{u}$ . A commonly used approach to deal with such problems in the context of the finite element method (FEM) is to discretise the temporal part of (2) separately using methods for solving initial value problems such as Runge-Kutta methods to obtain a semi-discrete boundary value problem. This results in a series of boundary value problems which are solved successively with finite elements. In the present paper we avoid separate treatment of spatial and temporal derivations by treating the temporal part as well as the spatial part of (2) within the scope of the FEM.

For the space-time formulation proposed here, we redefine and reinterpret the initial boundary value problem (2),(3) as a pure boundary value problem in the space-time domain  $\tilde{\mathcal{B}} := \mathcal{B} \times I \subset \mathbb{R}^{n+1}$  as follows: find  $\mathbf{u}: \tilde{\mathcal{B}} \rightarrow \mathbb{R}^m$  such that

$$\mathbf{A}(\mathbf{u}(\mathbf{y}), \mathbf{y}) \nabla_t \mathbf{u}(\mathbf{y}) + \mathbf{L}\mathbf{u}(\mathbf{y}) = \mathbf{f}(\mathbf{y}) \quad \forall \mathbf{y} \in \tilde{\mathcal{B}}, \quad (5)$$

supplemented by

$$\begin{aligned} \mathbf{u}(\mathbf{y}) &= \tilde{\mathbf{g}}(\mathbf{y}) & \text{on } \mathbf{y} \in \partial\tilde{\mathcal{B}}^u, \\ (\mathbf{R}\mathbf{u})(\mathbf{y}) &= \tilde{\mathbf{h}}(\mathbf{y}) & \text{on } \mathbf{y} \in \partial\tilde{\mathcal{B}}^\sigma, \end{aligned} \quad (6)$$

where  $\mathbf{y} = (\mathbf{x}, t)$  represents an event in the space-time. Moreover, the partial derivatives are split using the notion  $\nabla_{\mathbf{y}} = [\nabla_{\mathbf{x}}, \nabla_t]$ . The outer boundaries of the space-time is given

---

<sup>†</sup>For a thermal system,  $m$  is equal one, for a mechanical system,  $m$  is equal  $n$ , for a thermomechanical system,  $m$  is equal  $n + 1$ , etc.

by  $\partial\tilde{\mathcal{B}}^u = (\partial\mathcal{B}^u \times I) \cup (\mathcal{B} \times \{t_1\})$ ,  $\partial\tilde{\mathcal{B}}^\sigma = \partial\mathcal{B}^\sigma \times I$  and

$$\tilde{\mathbf{g}}(\mathbf{y}) := \begin{cases} \mathbf{u}_0(\mathbf{y}) & \text{if } \mathbf{y} \in \mathcal{B} \times \{t_1\} \\ \mathbf{g}(\mathbf{y}) & \text{if } \mathbf{y} \in \partial\mathcal{B}^u \times I. \end{cases} \quad (7)$$

Note that we apply Dirichlet conditions in  $\mathcal{B} \times \{t_1\}$ , i.e. we fixate the initial conditions. Within the proposed framework, we can alternatively constrain the final configuration and move backwards in time. The application of the FEM to (5),(6) is now straight forward: define appropriate trial and weighting spaces  $\mathcal{S}$ ,  $\mathcal{V}$ , derive a variational statement of (5) and solve the weak formulation using finite elements. This will be demonstrated in detail on different initial boundary value problems in the subsequent sections.

### 3 Thermal conduction

In a first step, we investigate the heat equation to describe the evolution of the absolute temperature distribution  $\theta: \mathcal{B}_0 \times I \rightarrow \mathbb{R}$  in the context of the proposed space-time formulation. Assuming the absence of mechanical contributions, the initial mesh remains static and thus, the nominal heat flux vector  $\mathbf{Q}$  is defined by

$$\mathbf{Q} = -\mathbf{K}(\theta)\nabla_x(\theta), \quad (8)$$

where  $\mathbf{K}$  is a thermal conductivity tensor which must be positive semi-definite. Note that this constitutive (Fourier-) law is thermodynamically consistent in the sense that it has to be in accordance to the second law of thermodynamics. The Lagrangian form of the local energy balance equation now reads

$$\rho c_p \nabla_t(\theta) = \nabla_x \cdot \mathbf{Q} + R, \quad \forall \mathbf{y} \in \tilde{\mathcal{B}}, \quad (9)$$

which has to be valid in the space-time domain  $\tilde{\mathcal{B}}$ . Here,  $c_p := c_p(\mathbf{y}, \theta)$  can be interpreted as specific heat capacity excluding the mechanical load (e.g. constant pressure),  $\rho := \rho(\mathbf{x})^\ddagger$  denotes the mass density and  $R := R(\mathbf{y})$  represents the heat supply per unit volume. To complete this space-time boundary value problem, appropriate Dirichlet and Neumann conditions are specified by

$$\begin{aligned} \theta(\mathbf{y}) &= \tilde{\theta}(\mathbf{y}) & \text{on } \mathbf{y} \in \mathcal{B} \times t_1 &= \partial\tilde{\mathcal{B}}^{t_1}, \\ \theta(\mathbf{y}) &= \bar{\theta}(\mathbf{y}) & \text{on } \mathbf{y} \in \partial\mathcal{B}^\theta \times I &= \partial\tilde{\mathcal{B}}^\theta, \\ \mathbf{Q} \cdot \mathbf{N} &= -\bar{Q} & \text{on } \mathbf{y} \in \partial\mathcal{B}^Q \times I &= \partial\tilde{\mathcal{B}}^Q. \end{aligned} \quad (10)$$

Here,  $\partial\mathcal{B}^\theta$  and  $\partial\mathcal{B}^Q$  are subsets of the domain boundary  $\partial\mathcal{B}$ , satisfying  $\partial\mathcal{B}^\theta \cup \partial\mathcal{B}^Q = \partial\mathcal{B}$  and  $\partial\mathcal{B}^\theta \cap \partial\mathcal{B}^Q = \emptyset$ . Moreover,  $\mathbf{N}$  denotes the unit outward normal field and  $\tilde{\theta}$  can be interpreted as prescribed initial temperature distribution.

---

<sup>‡</sup>Here, we assume that the mass density does not change in time.

### 3.1 Weak formulation

For finite element analysis we reformulate (9) along with (10). Therefore the collection of solution and weighting functions are defined as

$$\mathcal{S} = \left\{ \theta \in H^1(\tilde{\mathcal{B}}) \mid \left( \theta = \bar{\theta} \text{ on } \partial\tilde{\mathcal{B}}^\theta \right) \wedge \left( \theta = \tilde{\theta} \text{ on } \partial\tilde{\mathcal{B}}^{t_1} \right) \right\}, \quad (11)$$

$$\mathcal{V} = \left\{ \delta\theta \in H^1(\tilde{\mathcal{B}}) \mid \left( \delta\theta = 0 \text{ on } \partial\tilde{\mathcal{B}}^\theta \right) \wedge \left( \delta\theta = 0 \text{ on } \partial\tilde{\mathcal{B}}^{t_1} \right) \right\}, \quad (12)$$

where  $H^1(\tilde{\mathcal{B}})$  is the Sobolev space of square integrable functions  $\theta: \tilde{\mathcal{B}} \rightarrow \mathbb{R}$  with square integrable weak derivatives of first order. Note that we often suppress for the remainder of the article the dependence on  $(\mathbf{x}, t)$  and on  $\mathbf{y}$ , respectively, for the sake of brevity. The weak problem corresponding to (9) reads: find  $\theta \in \mathcal{S}$  such that

$$B(\theta, \delta\theta) = L(\delta\theta) \quad \forall \delta\theta \in \mathcal{V}, \quad (13)$$

where

$$B(\theta, \delta\theta) := \int_{\tilde{\mathcal{B}}} \delta\theta \rho c_p(\theta) \nabla_t(\theta) \, dW + \int_{\tilde{\mathcal{B}}} \nabla_{\mathbf{x}}(\delta\theta) \cdot \mathbf{K}(\theta) \nabla_{\mathbf{x}}(\theta) \, dW, \quad (14)$$

$$L(\delta\theta) := \int_{\tilde{\mathcal{B}}} \delta\theta R \, dW + \int_{\partial\tilde{\mathcal{B}}^Q} \delta\theta \bar{Q} \, d\Gamma. \quad (15)$$

Here,  $dW = dV \, dt$  and  $d\Gamma = dA \, dt$ . Moreover, we have made use of the divergence theorem in the spatial domain and the symmetry of the conductivity tensor.

### 3.2 First law of thermodynamics

The first law of thermodynamics states that energy can neither be created nor destroyed, i.e. that the total inner energy is a conserved quantity throughout time. Testing (13) using  $\delta\theta = \mu$ , where  $\mu \in \mathbb{R}$  is arbitrary and constant, and assuming the absence of Dirichlet boundary conditions  $\partial\tilde{\mathcal{B}}^\theta$  yields

$$\int_{\tilde{\mathcal{B}}} \rho c_p(\theta) \nabla_t(\theta) \, dW = \int_{\tilde{\mathcal{B}}} R \, dW + \int_{\partial\tilde{\mathcal{B}}^Q} \bar{Q} \, d\Gamma \equiv \mathcal{Q}, \quad (16)$$

where  $\mathcal{Q}$  represents the total net heating of the continuum body within the time interval  $I$ . Introducing the specific inner energy  $e := e(\theta(\mathbf{y}))$ , where  $\partial e / \partial \theta = \rho c_p$  and defined such that the left hand side of (16) can be rewritten as follows

$$\int_{\tilde{\mathcal{B}}} \rho c_p(\theta) \nabla_t(\theta) \, dW = \int_{\tilde{\mathcal{B}}} \frac{\partial e}{\partial \theta} \nabla_t(\theta) \, dW = \int_{\tilde{\mathcal{B}}} \nabla_t(e) \, dW. \quad (17)$$

Summarised, (16) reads

$$\int_{\tilde{\mathcal{B}}} \nabla_t(e) dW = \mathcal{Q}, \quad (18)$$

$$\int_{\mathcal{B}} e(t = t_2) dV - \int_{\mathcal{B}} e(t = t_1) dV = \mathcal{Q}, \quad (19)$$

where we have made use of the fundamental theorem of calculus, which relates the definite integral with the antiderivative. The last statement represents the space-time version of the first law of thermodynamics.

### 3.3 Discretisation

To achieve a numerical solution for a thermal problem, we apply a finite element framework to the thermal field. In particular, we introduce a finite dimensional approximation of  $\theta$  and  $\delta\theta$  so that

$$\theta^h = \sum_{A \in \omega} N^A(\mathbf{y}) \theta_A \quad \text{and} \quad \delta\theta^h = \sum_{A \in \omega} N^A(\mathbf{y}) \delta\theta_A. \quad (20)$$

Here,  $A \in \omega = \{1, \dots, n_{\text{node}}\}$  are nodes in the space-time and  $N^A: \tilde{\mathcal{B}} \rightarrow \mathbb{R}$  are the corresponding global shape functions within the space-time. The spaces  $\mathcal{S}$ ,  $\mathcal{V}$  are approximated by finite-dimensional spaces

$$\mathcal{S}^h = \left\{ \theta^h \in H^1(\tilde{\mathcal{B}}^h) \mid \left( \theta^h = \bar{\theta} \text{ on } \partial\tilde{\mathcal{B}}^{\theta,h} \right) \wedge \left( \theta^h = \tilde{\theta} \text{ on } \partial\tilde{\mathcal{B}}^{t_1,h} \right) \right\} \quad (21)$$

$$\mathcal{V}^h = \left\{ \delta\theta^h \in H^1(\tilde{\mathcal{B}}^h) \mid \left( \delta\theta^h = 0 \text{ on } \partial\tilde{\mathcal{B}}^{\theta,h} \right) \wedge \left( \delta\theta^h = 0 \text{ on } \partial\tilde{\mathcal{B}}^{t_1,h} \right) \right\}. \quad (22)$$

Taking the linearity of  $B$  in its second argument and the linearity of  $L$  into account, the resulting discrete problem reads: find  $\theta^h \in \mathcal{S}^h$  such that

$$B(\theta^h, N^A) = L(N^A) \quad \forall A \in \omega, \quad (23)$$

with

$$B(\theta^h, N^A) = \int_{\tilde{\mathcal{B}}^h} N^A \rho c_p(\theta^h) \nabla_t(N^B) dW \theta_B + \int_{\tilde{\mathcal{B}}^h} \nabla_x(N^A) \cdot \mathbf{K}(\theta^h) \nabla_x(N^B) dW \theta_B, \quad (24)$$

$$L(N^A) = \int_{\tilde{\mathcal{B}}^h} N^A R dW + \int_{\partial\tilde{\mathcal{B}}^{\mathcal{Q},h}} N^A \bar{Q} d\Gamma, \quad (25)$$

is valid. Note that the Einstein summation convention is employed for repeated indices throughout the article.

### 3.4 Balance of energy

To verify the algorithmic conservation of energy as stated by the first law of thermodynamics, we set  $\delta\theta_A = \mu$  and obtain

$$B(\theta^h, N^A \mu) = \int_{\tilde{\mathcal{B}}^h} \left( \sum_{A \in \omega} N^A \right) \mu \rho c_p(\theta) \nabla_t(N^B) dW \theta_B + \int_{\tilde{\mathcal{B}}^h} \left( \nabla_x \left( \sum_{A \in \omega} N^A \right) \right) \mu \cdot \mathbf{K}(\theta^h) \nabla_x(N^B) dW \theta_B, \quad (26)$$

$$L(N^A \mu) = \int_{\tilde{\mathcal{B}}^h} \left( \sum_{A \in \omega} N^A \right) \mu R dW + \int_{\partial \tilde{\mathcal{B}}^h} \left( \sum_{A \in \omega} N^A \right) \mu \bar{Q} d\Gamma = \mathcal{Q}^h, \quad (27)$$

Next, we take the properties  $\sum_{A \in \omega} N^A = 1$  and  $\nabla_x \left( \sum_{A \in \omega} N^A \right) = 0$  into account. Moreover, a subsequent division on both sides with  $\mu$  and introducing the discrete inner energy  $e^h$  yields

$$\int_{\tilde{\mathcal{B}}^h} \rho c_p(\theta) \nabla_t(N^B) dW \theta_B = \int_{\tilde{\mathcal{B}}^h} \nabla_t(e^h) dW = \mathcal{Q}^h. \quad (28)$$

The last equation confirms algorithmic energy consistency of the proposed space-time formulation. Hence, the balance law for energy is correctly reproduced in the discrete space-time setting.

## 4 Non-linear elasticity

The most challenging system under investigation deals with general non-linear elasticity. Again, a Lipschitz bounded domain is defined in  $\mathcal{B}_0 \subset \mathbb{R}^n$ ,  $n \in \{1, 2, 3\}$  its reference or unstressed configuration. The current configuration  $\mathcal{B}_t \subset \mathbb{R}^n$  at time  $t \in I = [t_1, t_2]$  is related to the reference configuration via the deformation mapping  $\varphi : \tilde{\mathcal{B}}_0 \rightarrow \mathbb{R}^n$ , where  $\tilde{\mathcal{B}}_0 := \mathcal{B}_0 \times I \subset \mathbb{R}^{n+1}$ , such that  $\mathcal{B}_t = \varphi(\mathcal{B}_0, t)$ . Material points are labelled by  $\mathbf{X} \in \mathcal{B}_0$  with corresponding actual or spatial positions defined by  $\mathbf{x}(t) = \varphi(\mathbf{X}, t)$ . The dynamics of an elastic body caused by conservative external forces is described by its Lagrangian

$$\mathcal{L}(\varphi, \nabla_{\mathbf{X}}(\varphi), \nabla_t(\varphi)) = T(\nabla_t(\varphi)) - V(\varphi, \nabla_{\mathbf{X}}(\varphi)), \quad (29)$$

where the kinetic energy  $T$  and the potential energy  $V$  are given by the functions

$$T(\nabla_t(\varphi)) = \frac{1}{2} \int_{\mathcal{B}_0} \rho_0 \|\nabla_t(\varphi)\|^2 dV = \int_{\mathcal{B}_0} \mathcal{T} dV, \quad (30)$$

$$V(\varphi, \nabla_{\mathbf{X}}(\varphi)) = \int_{\mathcal{B}_0} [\Psi - \rho_0 \varphi \cdot \mathbf{B}] dV - \int_{\partial \mathcal{B}_0^g} \varphi \cdot \bar{\mathbf{T}} dA.$$

The contribution of the external forces to the potential energy is due to the dead loads  $\mathbf{B} = \mathbf{B}(\mathbf{X})$  and  $\bar{\mathbf{T}} = \bar{\mathbf{T}}(\mathbf{X})$  which are distributed over the volume of the continuum body and its von Neumann boundary, respectively. Moreover,  $\mathcal{T} = \mathcal{T}(\mathbf{X}, \nabla_t(\boldsymbol{\varphi})) = \frac{1}{2}\rho_0\|\nabla_t(\boldsymbol{\varphi})\|^2$  and  $\Psi = \Psi(\mathbf{X}, \nabla_{\mathbf{X}}(\boldsymbol{\varphi}))$  represent the kinetic energy and the strain energy function per unit reference volume.

As in elasto-dynamics the main interest lies in initial value problems, the classical form of Hamilton's principle with prescribed initial and end configurations is not a suitable starting point. A somehow weakened formulation is obtained by postulating Hamilton's principle of varying action

$$\delta \int_{t_1}^{t_2} \mathcal{L}(\boldsymbol{\varphi}, \nabla_{\mathbf{X}}(\boldsymbol{\varphi}), \nabla_t(\boldsymbol{\varphi})) dt - \delta \boldsymbol{\varphi} \cdot \frac{\partial \mathcal{T}}{\partial(\nabla_t(\boldsymbol{\varphi}))} \Big|_{t_1}^{t_2} = 0 \quad \forall \delta \boldsymbol{\varphi}, \quad (31)$$

which can be found for instance in Izadpanah [44] or Borri [11].

Inserting the Lagrangian defined in (29) and (30) into Hamilton's principle of varying action (31), the dynamics of an elastic body subjected to external dead loads is described by the variational expression

$$\int_{t_1}^{t_2} \left\{ \int_{\mathcal{B}_0} \delta \mathcal{T}(\nabla_t(\boldsymbol{\varphi})) + \rho_0 \delta \boldsymbol{\varphi} \cdot \mathbf{B} - \delta \Psi(\nabla_{\mathbf{X}}(\boldsymbol{\varphi})) dV + \int_{\partial \mathcal{B}_0^c} \delta \boldsymbol{\varphi} \cdot \bar{\mathbf{T}} dA \right\} dt - \int_{\mathcal{B}_0} \delta \boldsymbol{\varphi} \cdot \frac{\partial \mathcal{T}}{\partial(\nabla_t(\boldsymbol{\varphi}))} dV \Big|_{t_1}^{t_2} = 0 \quad \forall \delta \boldsymbol{\varphi}. \quad (32)$$

Carrying out the variations followed by suitable integration by parts, the variational statement (32) leads directly to the local form of the balance of linear momentum together with suitable boundary conditions.

What follows next has to be understood point wise for every material point  $\mathbf{X}$  for every instant of time  $t$ , i.e. for every event in space-time. The density of the kinetic energy  $\mathcal{T}(\nabla_t(\boldsymbol{\varphi}))$  is a convex and differentiable function with respect to  $\nabla_t(\boldsymbol{\varphi})$ . The conjugate function  $\mathcal{T}^* = \mathcal{T}^*(\mathbf{p})$  is defined as

$$\mathcal{T}^*(\mathbf{p}) = \sup_{\nabla_t(\boldsymbol{\varphi}) \in \mathbb{R}^n} \{ \mathbf{p} \cdot \nabla_t(\boldsymbol{\varphi}) - \mathcal{T}(\nabla_t(\boldsymbol{\varphi})) \}. \quad (33)$$

Denoting the variation with respect to  $\nabla_t(\boldsymbol{\varphi})$  by  $\delta_{\nabla_t(\boldsymbol{\varphi})}$ , the necessary and sufficient condition for the supremum in (33) can be written as

$$\delta_{\nabla_t(\boldsymbol{\varphi})} (\nabla_t(\boldsymbol{\varphi}) \cdot \mathbf{p} - \mathcal{T}(\nabla_t(\boldsymbol{\varphi}))) = \delta(\nabla_t(\boldsymbol{\varphi})) \cdot \left( \mathbf{p} - \frac{\partial \mathcal{T}}{\partial(\nabla_t(\boldsymbol{\varphi}))} \right) = 0 \quad \forall \delta(\nabla_t(\boldsymbol{\varphi})). \quad (34)$$

Using  $\mathcal{T} = \frac{1}{2}\rho_0\|\nabla_t(\boldsymbol{\varphi})\|^2$ , the condition (34) can also be expressed by the relation  $\nabla_t(\boldsymbol{\varphi}) = \rho_0^{-1}\mathbf{p}$ , which then inserted into (33) leads to the conjugate kinetic energy

density  $\mathcal{T}^*$  and the conjugate kinetic energy  $T^*$

$$\mathcal{T}^*(\mathbf{p}) = \frac{1}{2}\rho_0^{-1}\|\mathbf{p}\|^2, \quad T^*(\mathbf{p}) = \frac{1}{2}\int_{\mathcal{B}_0}\rho_0^{-1}\|\mathbf{p}\|^2 dV. \quad (35)$$

Since the kinetic energy is a quadratic form, we can write the conjugate kinetic energy also as  $T^*(\mathbf{p}) = T(\rho_0^{-1}\mathbf{p})$ . Moreover, due to the differentiability of  $\mathcal{T}$  the biconjugate function  $\mathcal{T}^{**} = (\mathcal{T}^*)^*$  is equal to the function itself, i.e.  $\mathcal{T}^{**} = \mathcal{T}$ . This allows us to rewrite the kinetic energy density as

$$\mathcal{T}(\nabla_t(\boldsymbol{\varphi})) = \sup_{\mathbf{p} \in \mathbb{R}^n} \{\nabla_t(\boldsymbol{\varphi}) \cdot \mathbf{p} - \mathcal{T}^*(\mathbf{p})\}. \quad (36)$$

The supremum in Equation (36) can be dropped as long as the necessary and sufficient condition for the supremum is satisfied. Similar to (34), this condition can be written as  $\delta_{\mathbf{p}}(\mathbf{p} \cdot \nabla_t \boldsymbol{\varphi} - \mathcal{T}^*(\mathbf{p})) = 0$ , where we denoted by  $\delta_{\mathbf{p}}$  the variation with respect to  $\mathbf{p}$ . Consequently, we can rewrite Hamilton's principle of varying action (32) as

$$\begin{aligned} & \int_{t_1}^{t_2} \left\{ \int_{\mathcal{B}_0} \delta_{\nabla_t(\boldsymbol{\varphi})}(\nabla_t(\boldsymbol{\varphi}) \cdot \mathbf{p} - \mathcal{T}^*(\mathbf{p})) + \delta_{\mathbf{p}}(\nabla_t(\boldsymbol{\varphi}) \cdot \mathbf{p} - \mathcal{T}^*(\mathbf{p})) + \rho_0 \delta \boldsymbol{\varphi} \cdot \mathbf{B} \right. \\ & \left. - \delta \Psi(\nabla_{\mathbf{X}}(\boldsymbol{\varphi})) dV + \int_{\partial \mathcal{B}_0^g} \delta \boldsymbol{\varphi} \cdot \bar{\mathbf{T}} dA \right\} dt - \int_{\mathcal{B}_0} \delta \boldsymbol{\varphi} \cdot \frac{\partial \mathcal{T}}{\partial(\nabla_t(\boldsymbol{\varphi}))} dV \Big|_{t_1}^{t_2} = 0 \quad \forall \delta \boldsymbol{\varphi}, \delta \mathbf{p} \end{aligned} \quad (37)$$

in which the first term comes from the dynamic principle, while the second term is the condition allowing us to drop the supremum of (36). It is easy to see that the first two variations can be written together by demanding a variation with respect to all variables  $\delta$ . Furthermore, the derivative of the kinetic energy on the time boundary corresponds to the momentum  $\mathbf{p}$  and so

$$\begin{aligned} & \int_{t_1}^{t_2} \left\{ \int_{\mathcal{B}_0} \delta(\nabla_t(\boldsymbol{\varphi}) \cdot \mathbf{p} - \mathcal{T}^*(\mathbf{p})) + \rho_0 \delta \boldsymbol{\varphi} \cdot \mathbf{B} - \delta \Psi(\nabla_{\mathbf{X}}(\boldsymbol{\varphi})) dV + \int_{\partial \mathcal{B}_0^g} \delta \boldsymbol{\varphi} \cdot \bar{\mathbf{T}} dA \right\} dt \\ & - \int_{\mathcal{B}_0} \delta \boldsymbol{\varphi} \cdot \mathbf{p} dV \Big|_{t_1}^{t_2} = 0 \quad \forall \delta \boldsymbol{\varphi}, \delta \mathbf{p} \end{aligned} \quad (38)$$

Introducing the Hamiltonian

$$\mathcal{H}(\boldsymbol{\varphi}, \nabla_{\mathbf{X}}(\boldsymbol{\varphi}), \mathbf{p}) = T^*(\mathbf{p}) + V(\boldsymbol{\varphi}, \nabla_{\mathbf{X}}(\boldsymbol{\varphi})), \quad (39)$$

Equation (38) writes as

$$\delta \int_{t_1}^{t_2} \left\{ \int_{\mathcal{B}_0} (\nabla_t(\boldsymbol{\varphi}) \cdot \mathbf{p} dV - \mathcal{H}(\boldsymbol{\varphi}, \nabla_{\mathbf{X}}(\boldsymbol{\varphi}), \mathbf{p})) \right\} dt - \int_{\mathcal{B}_0} \delta \boldsymbol{\varphi} \cdot \mathbf{p} dV \Big|_{t_1}^{t_2} = 0 \quad \forall \delta \boldsymbol{\varphi}, \delta \mathbf{p} \quad (40)$$

which, except from the boundary terms, corresponds with the classical expression of a Hamiltonian field theory. Carrying out the variations in (38) and making explicit use of the conjugate kinetic energy density (35), after integrating by parts with respect to time the variational expression (38) takes the form

$$\begin{aligned}
& - \int_{\tilde{\mathcal{B}}_0} \delta \mathbf{p} \cdot (\nabla_t(\boldsymbol{\varphi}) - \rho_0^{-1} \mathbf{p}) \, dW + \\
& \int_{\tilde{\mathcal{B}}_0} \delta \boldsymbol{\varphi} \cdot \nabla_t(\mathbf{p}) + \nabla_{\mathbf{X}}(\delta \boldsymbol{\varphi}) : \mathbf{P} \, dW = \int_{\tilde{\mathcal{B}}_0} \delta \boldsymbol{\varphi} \cdot \rho_0 \mathbf{B} \, dW + \int_{\partial \tilde{\mathcal{B}}_0^\sigma} \delta \boldsymbol{\varphi} \cdot \bar{\mathbf{T}} \, d\Gamma \quad \forall \delta \boldsymbol{\varphi}, \delta \mathbf{p},
\end{aligned} \tag{41}$$

where we can now introduce the functional spaces of trial functions in space-time

$$\begin{aligned}
\mathcal{V}^p &= \left\{ \delta \mathbf{p} \in \mathbf{H}^1(\tilde{\mathcal{B}}_0) \mid (\delta \mathbf{p} = \mathbf{0} \text{ on } \partial \tilde{\mathcal{B}}_0^{t_1}) \right\}, \\
\mathcal{V}^\varphi &= \left\{ \delta \boldsymbol{\varphi} \in \mathbf{H}^1(\tilde{\mathcal{B}}_0) \mid (\delta \boldsymbol{\varphi} = \mathbf{0} \text{ on } \partial \tilde{\mathcal{B}}_0^{t_1}) \wedge (\delta \boldsymbol{\varphi} = \mathbf{0} \text{ on } \partial \tilde{\mathcal{B}}_0^\sigma) \right\}.
\end{aligned} \tag{42}$$

with boundary conditions, summarised as follows

$$\begin{aligned}
\boldsymbol{\varphi}(\mathbf{Y}) &= \tilde{\boldsymbol{\varphi}}(\mathbf{Y}) \quad \text{on} \quad \mathbf{Y} \in \mathcal{B}_0 \times t_1 = \partial \tilde{\mathcal{B}}_0^{t_1}, \\
\mathbf{p}(\mathbf{Y}) &= \tilde{\mathbf{p}}(\mathbf{Y}) \quad \text{on} \quad \mathbf{Y} \in \mathcal{B}_0 \times t_1 = \partial \tilde{\mathcal{B}}_0^{t_1}, \\
\boldsymbol{\varphi}(\mathbf{Y}) &= \bar{\boldsymbol{\varphi}}(\mathbf{Y}) \quad \text{on} \quad \mathbf{Y} \in \partial \mathcal{B}_0^\sigma \times I = \partial \tilde{\mathcal{B}}_0^\sigma, \\
\mathbf{P} \mathbf{N} &= \bar{\mathbf{T}}(\mathbf{Y}) \quad \text{on} \quad \mathbf{Y} \in \partial \mathcal{B}_0^\sigma \times I = \partial \tilde{\mathcal{B}}_0^\sigma.
\end{aligned} \tag{43}$$

This specific functional spaces are further restricted insofar, that the variations in direction of time are constrained to be equal zero, since stretching and compression in temporal direction is only possible in relativistic mechanics<sup>§</sup>. Furthermore,  $\mathbf{P} := \partial \Psi / \partial \mathbf{F}$  denotes the first Piola-Kirchhoff stress tensor and  $\mathbf{F} : \tilde{\mathcal{B}}_0 \rightarrow \mathbb{R}^{n \times n}$ ,  $\mathbf{F} = \nabla_{\mathbf{X}}(\boldsymbol{\varphi})$  is the deformation gradient tensor. Eventually, the right Cauchy-Green tensor is  $\mathbf{C} : \tilde{\mathcal{B}}_0 \rightarrow \mathbb{R}^{n \times n}$ ,  $\mathbf{C} = \mathbf{F}^T \mathbf{F}$ . In order to guarantee objectivity of the material law, it is convenient to formulate the strain energy density of a certain material in terms of the right Cauchy-Green tensor, i.e.  $\Psi(\mathbf{X}, \mathbf{F}) = W(\mathbf{X}, \mathbf{C}(\mathbf{F}))$ . Introducing the second Piola-Kirchhoff stress tensor  $\boldsymbol{\Sigma} := 2\partial W / \partial \mathbf{C}$ , the variation of the strain energy function of  $\Psi$  required in (41) can then be reformulated as

$$\delta \Psi(\mathbf{F}, \mathbf{X}) = \frac{\partial W}{\partial \mathbf{C}} : \frac{\partial \mathbf{C}}{\partial \mathbf{F}} : \delta \mathbf{F} = \boldsymbol{\Sigma} : (\mathbf{F}^T \delta \mathbf{F}) = (\mathbf{F} \boldsymbol{\Sigma}) : \delta \mathbf{F} \tag{44}$$

In the discretization we will make use of this expression in terms of deformation functions only, i.e.  $(\mathbf{F} \boldsymbol{\Sigma}) : \delta \mathbf{F} = (\nabla_{\mathbf{X}}(\boldsymbol{\varphi}) \boldsymbol{\Sigma}) : \nabla_{\mathbf{X}}(\delta \boldsymbol{\varphi})$ .

---

*Remark 1.* We have converted the set of originally hyperbolic equations first to a set of  $2n$  parabolic equations and will discretise this set of equations directly in the space-time in the subsequent section. This procedure ensures stability in temporal direction;

---

<sup>§</sup>Note that this would require to introduce the Minkowski space  $\mathbb{M}^4$ .

otherwise higher continuity for the solution would be required (at least to be in  $H^2$ ) with additional challenges to embed the initial velocity field. Integration by parts in temporal direction to reduce the continuity requirement does not work in the discrete space, since we can not enforce continuity of the velocity across the element boundaries, a fundamental prerequisite for the conservation of linear momentum.

---

## 4.1 First integrals

A first integral, constant of motion or conserved quantity of a general differential equation is a function  $G : \mathbb{R}^{2n} \rightarrow \mathbb{R}$  which is constant along all solution curves of the system, see Leimkuhler & Reich [52] for further details. This concept will now be adapted to the space-time domain.

**Balance of linear momentum** To elaborate the connection between specific symmetry properties and associated momentum maps, we focus first on translational invariance of the Hamiltonian  $\mathcal{H}$  as defined in 39, which implies conservation of total linear momentum. In particular, we introduce a momentum map, defined by

$$L_\xi = \mathbf{p} \cdot \boldsymbol{\xi}, \quad (45)$$

where  $\boldsymbol{\xi} \in \mathbb{R}^n$  is arbitrary and constant and  $L_\xi$  is to be identified as the linear momentum map. Now, we obtain by replacing  $\delta\boldsymbol{\varphi} := \boldsymbol{\xi}$  and  $\delta\mathbf{p} := \boldsymbol{\xi}$  in (41) and noting that replacing the variation  $\delta\mathbf{p}$  immediately shows that the first term in (41), stating that  $\nabla_t(\boldsymbol{\varphi}) = \rho_0^{-1}\mathbf{p}$ , is fulfilled and we obtain

$$\begin{aligned} \int_{\tilde{\mathcal{B}}_0} \boldsymbol{\xi} \cdot \nabla_t(\mathbf{p}) + \underbrace{\nabla_{\mathbf{X}}(\boldsymbol{\xi}) : \mathbf{P}}_{=0} dW &= \int_{\tilde{\mathcal{B}}_0} \nabla_t(L_\xi(\mathbf{Y})) dW, \\ &= \int_{\mathcal{B}_0} L_\xi(\mathbf{Y}(t = t_2)) dV - \int_{\mathcal{B}_0} L_\xi(\mathbf{Y}(t = t_1)) dV, \\ &= \int_{\tilde{\mathcal{B}}_0} \boldsymbol{\xi} \cdot \rho_0 \mathbf{B} dW + \int_{\partial\tilde{\mathcal{B}}_0^\sigma} \boldsymbol{\xi} \cdot \bar{\mathbf{T}} d\Gamma. \end{aligned} \quad (46)$$

Hence, translational invariance of the Hamiltonian implies conservation of the total linear momentum in the sense that the change of total linear momentum is equal to the total external contributions.

**Balance of angular momentum** Next, we elaborate the connection between rotational invariance and conservation of the total angular momentum. Assuming that the Hamiltonian is invariant under the action of a proper orthogonal matrix  $\mathbf{Q} \in SO(n)$ , we can state

$$\mathcal{H}(\mathbf{Q}\boldsymbol{\varphi}, \mathbf{Q}\mathbf{p}) \equiv \mathcal{H}(\boldsymbol{\varphi}, \mathbf{p}). \quad (47)$$

A one-parameter group of rotation matrices can be written in the form  $\mathbf{Q}_\varepsilon = \exp(\varepsilon \hat{\boldsymbol{\xi}}) \in SO(n)$  where the Lie algebra  $\hat{\boldsymbol{\xi}} \in \mathfrak{so}(n)$  is a skew-symmetric matrix. Next, the gradient of the Hamiltonian reads

$$0 = \left. \frac{d}{d\varepsilon} \right|_{\varepsilon=0} \mathcal{H}(\exp(\varepsilon \hat{\boldsymbol{\xi}}) \circ \boldsymbol{\varphi}, \exp(\varepsilon \hat{\boldsymbol{\xi}}) \circ \mathbf{p}) = [\nabla_{\boldsymbol{\varphi}} \mathcal{H}(\boldsymbol{\varphi}, \mathbf{p}), \nabla_{\mathbf{p}} \mathcal{H}(\boldsymbol{\varphi}, \mathbf{p})] \cdot \boldsymbol{\xi}_P(\boldsymbol{\varphi}, \mathbf{p}), \quad (48)$$

where  $\boldsymbol{\xi}_P$  denotes an infinitesimal generator, see Marsden & Ratiu [53], Chapter 11.4 for details. Introducing  $\boldsymbol{\xi} \in \mathbb{R}^3$  as the axial vector of  $\hat{\boldsymbol{\xi}}$  (i.e.  $\hat{\boldsymbol{\xi}}\mathbf{a} = \boldsymbol{\xi} \times \mathbf{a}$  for all  $\mathbf{a} \in \mathbb{R}^3$ ) and  $\mathbf{J} \in \mathbb{R}^3$  can be identified as total angular momentum given by

$$\mathbf{J}(\mathbf{Y}) = \boldsymbol{\varphi} \times \mathbf{p}, \quad J_\xi(\mathbf{Y}) = \mathbf{J}(\mathbf{Y}) \cdot \boldsymbol{\xi}, \quad (49)$$

associated with the previously introduced infinitesimal generator. Now we can replace  $\delta\boldsymbol{\varphi} := \boldsymbol{\xi} \times \boldsymbol{\varphi}$  and obtain

$$\int_{\tilde{\mathcal{B}}_0} \boldsymbol{\xi} \cdot \boldsymbol{\varphi} \times \nabla_t(\mathbf{p}) + \underbrace{\mathbf{F}^T \nabla_X(\boldsymbol{\xi} \times \boldsymbol{\varphi}) : \boldsymbol{\Sigma}}_{=0} dW = \int_{\tilde{\mathcal{B}}_0} \nabla_t(\boldsymbol{\xi} \cdot \boldsymbol{\varphi} \times \mathbf{p}) - \boldsymbol{\xi} \cdot \nabla_t(\boldsymbol{\varphi}) \times \mathbf{p} dW \quad (50)$$

where we have made use of the skew-symmetry of  $\mathbf{F}^T \nabla_X(\boldsymbol{\xi} \times \boldsymbol{\varphi})$  and the symmetry of  $\boldsymbol{\Sigma}$ . Integration by parts in temporal direction of the first term on the right hand side and noting that  $\nabla_t(\boldsymbol{\varphi}) \times \mathbf{p} = \mathbf{p} \times \mathbf{p} = \mathbf{0}$  yields immediately

$$\begin{aligned} \int_{\tilde{\mathcal{B}}_0} \nabla_t(\boldsymbol{\xi} \cdot \boldsymbol{\varphi} \times \mathbf{p}) dW &= \int_{\tilde{\mathcal{B}}_0} \nabla_t(J_\xi(\mathbf{Y})) dW, \\ &= \int_{\tilde{\mathcal{B}}_0} J_\xi(\mathbf{Y}(t = t_2)) dV - \int_{\tilde{\mathcal{B}}_0} J_\xi(\mathbf{Y}(t = t_1)) dV, \\ &= \boldsymbol{\xi} \cdot \left\{ \int_{\tilde{\mathcal{B}}_0} \boldsymbol{\varphi} \times \rho_0 \mathbf{B} dW + \int_{\partial \tilde{\mathcal{B}}_0^g} \boldsymbol{\varphi} \times \bar{\mathbf{T}} d\Gamma \right\}. \end{aligned} \quad (51)$$

Thus, the total change in angular momentum corresponds to the external contributions.

**Balance of energy** Next, a translation of the Hamiltonian itself in time reveals, that replacing  $\delta\boldsymbol{\varphi} := \nabla_t \boldsymbol{\varphi}$  will serve us as connection with the conservation of energy. Ap-

proprate substitution yields

$$\begin{aligned}
\int_{\tilde{\mathcal{B}}_0} \nabla_t(\boldsymbol{\varphi}) \cdot \nabla_t(\mathbf{p}) + \nabla_X(\nabla_t(\boldsymbol{\varphi})) : \mathbf{P} \, dW &= \int_{\tilde{\mathcal{B}}_0} \underbrace{\rho_0^{-1} \mathbf{p} \nabla_t(\mathbf{p})}_{=\nabla_t(1/2\rho_0^{-1}\mathbf{p}\cdot\mathbf{p})} + \nabla_t(\Psi) \, dW, \\
&= \int_{\tilde{\mathcal{B}}_0} \nabla_t \mathcal{H} \, dW, \\
&= \int_{\tilde{\mathcal{B}}_0} \mathcal{H}(t = t_2) \, dV - \int_{\tilde{\mathcal{B}}_0} \mathcal{H}(t = t_1) \, dV, \\
&= 0
\end{aligned} \tag{52}$$

hence, the total energy  $\int_{\mathcal{B}_0} \mathcal{H} \, dV = T + V$  is a conserved quantity. Note that we have committed the external contributions for the sake of a clear presentation. Independent of the presents of these contributions, the last statement holds, since the external contributions are included in  $\mathcal{H}$ .

## 4.2 Discretisation

Next, we introduce finite dimensional approximation of  $\boldsymbol{\varphi}$  and  $\delta\boldsymbol{\varphi}$  as well as for  $\mathbf{p}$  and  $\delta\mathbf{p}$  as follows

$$\begin{aligned}
\boldsymbol{\varphi}^h &= \sum_{A \in \omega} N^A(\mathbf{Y}) \mathbf{q}_A \quad \text{and} \quad \delta\boldsymbol{\varphi}^h = \sum_{A \in \omega} N^A(\mathbf{Y}) \delta\mathbf{q}_A, \\
\mathbf{p}^h &= \sum_{A \in \omega} N^A(\mathbf{Y}) \mathbf{p}_A \quad \text{and} \quad \delta\mathbf{p}^h = \sum_{A \in \omega} N^A(\mathbf{Y}) \delta\mathbf{p}_A.
\end{aligned} \tag{53}$$

Following the arguments presented in 3.3,  $A \in \omega = \{1, \dots, n_{\text{node}}\}$  are once again nodes in the space-time and  $N^A: \tilde{\mathcal{B}} \rightarrow \mathbb{R}$  are the corresponding global shape functions in the reference configuration within the space-time. The discrete functional space of test or trial functions reads

$$\begin{aligned}
\mathcal{V}^{p,h} &= \left\{ \delta\mathbf{p}^h \in H^1(\tilde{\mathcal{B}}_0^h) \mid (\delta\mathbf{p}^h = \mathbf{0} \text{ on } \partial\tilde{\mathcal{B}}_0^{t_1,h}) \right\}, \\
\mathcal{V}^{\varphi,h} &= \left\{ \delta\boldsymbol{\varphi}^h \in H^1(\tilde{\mathcal{B}}_0^h) \mid (\delta\boldsymbol{\varphi}^h = \mathbf{0} \text{ on } \partial\tilde{\mathcal{B}}_0^{t_1,h}) \wedge (\delta\boldsymbol{\varphi}^h = \mathbf{0} \text{ on } \partial\tilde{\mathcal{B}}_0^{\varphi,h}) \right\}.
\end{aligned} \tag{54}$$

Both,  $\mathbf{q}_A$  as well as  $\delta\mathbf{q}_A$  are  $\in \mathbb{R}^{n+1}$ . As stated before, we do not allow stretching or compression in time. Therefore, we place an additional constraint on the  $(n+1)$ th dimension, which can be considered as additional Dirichlet constraint on each node of the system. The discrete set of equations of motion follows immediately from (41) as follows

$$\delta\mathbf{p}_A \cdot \int_{\tilde{\mathcal{B}}_0^h} N^A \nabla_t(N^B) \mathbf{q}_B - N^A \rho_0^{-1} N^B \mathbf{p}_B \, dW = 0, \quad \forall \delta\mathbf{p}_A \tag{55}$$

$$\delta\mathbf{q}_A \cdot \int_{\tilde{\mathcal{B}}_0^h} N^A \nabla_t(N^B) \mathbf{p}_B + \nabla_X(N^A) \cdot \boldsymbol{\Sigma}^h \nabla_X(N^B) \mathbf{q}_B \, dW = \delta\mathbf{q}_A \cdot \mathbf{P}^{A,ext}, \quad \forall \delta\mathbf{q}_A. \tag{56}$$

Here, we have made use of discrete variation of  $\mathbf{C}^h = \mathbf{q}_A \cdot \mathbf{q}_B \nabla_{\mathbf{X}}(N^A) \otimes \nabla_{\mathbf{X}}(N^B)$ . Moreover,  $\Sigma^h := \partial\Psi/\partial\mathbf{C}^h$  represents the discrete second Piola-Kirchhof stress tensor and

$$\mathbf{P}^{A,ext} = \int_{\tilde{\mathcal{B}}_0^h} N^A \rho_0 \mathbf{B} \, dW + \int_{\partial\tilde{\mathcal{B}}_0^{\sigma,h}} N^A \bar{\mathbf{T}} \, d\Gamma, \quad (57)$$

are the discrete external contributions.

### 4.3 Conservation properties

**Linear momentum** To verify conservation of linear momentum, we postulate again the absence of Dirichlet boundary conditions. Insertion of  $\delta\mathbf{q}_A = \boldsymbol{\mu}$  into (56) yields

$$\int_{\tilde{\mathcal{B}}_0} \boldsymbol{\mu} \cdot \sum_A (N^A) \nabla_t(N^B) \mathbf{p}_B + \boldsymbol{\mu} \cdot \nabla_{\mathbf{X}} \left( \sum_A (N^A) \right) \cdot \Sigma^h \nabla_{\mathbf{X}}(N^B) \mathbf{q}_B \, dW = \boldsymbol{\mu} \cdot \sum_A \mathbf{P}^{A,ext}. \quad (58)$$

As before, we take the properties  $\sum_{A \in \omega} N^A = 1$  and  $\nabla_{\mathbf{x}} \left( \sum_{A \in \omega} N^A \right) = 0$  into account and obtain

$$\int_{\tilde{\mathcal{B}}_0} \nabla_t(N^B) \, dW \mathbf{p}_B = \sum_A \mathbf{P}^{A,ext}. \quad (59)$$

Using again the fundamental theorem of calculus in temporal direction, we obtain

$$\int_{\tilde{\mathcal{B}}_0} N^B(\mathbf{X}, t = t_2) \, dV \mathbf{p}_B - \int_{\tilde{\mathcal{B}}_0} N^B(\mathbf{X}, t = t_1) \, dV \mathbf{p}_B = \sum_A \mathbf{P}^{A,ext}, \quad (60)$$

and thus, linear momentum is algorithmically conserved.

**Angular momentum** Next, we substitute  $\delta\mathbf{q}_A = \boldsymbol{\mu} \times \mathbf{q}_A$  into (56) and obtain

$$\int_{\tilde{\mathcal{B}}_0} \boldsymbol{\mu} \cdot \mathbf{q}_A \times N^A \nabla_t(N^B) \mathbf{p}_B + \boldsymbol{\mu} \cdot \mathbf{q}_A \times \underbrace{\nabla_{\mathbf{X}}(N^A) \cdot \Sigma^h \nabla_{\mathbf{X}}(N^B)}_{=S^{AB}} \mathbf{q}_B \, dW = \boldsymbol{\mu} \cdot \mathbf{q}_A \times \mathbf{P}^{A,ext}, \quad (61)$$

where we have made use of the identity  $(\mathbf{a} \times \mathbf{b}) \cdot \mathbf{c} = \mathbf{a} \cdot (\mathbf{b} \times \mathbf{c})$ . With regard to the skew-symmetry of  $\mathbf{q}_A \times \mathbf{q}_B$  and the symmetry of  $S^{AB} = \nabla_{\mathbf{X}}(N^A) \cdot \Sigma^h \nabla_{\mathbf{X}}(N^B)$ , the second term on the left hand side of (61) vanishes and we obtain

$$\int_{\tilde{\mathcal{B}}_0} \boldsymbol{\mu} \cdot \mathbf{q}_A \times N^A \nabla_t(N^B) \mathbf{p}_B \, dW = \boldsymbol{\mu} \cdot \mathbf{q}_A \times \mathbf{P}^{A,ext}. \quad (62)$$

This can be easily rewritten as follows

$$\int_{\tilde{\mathcal{B}}_0} \boldsymbol{\mu} \cdot \left[ \mathbf{q}_A \times \nabla_t(N^A N^B) \mathbf{p}_B - \mathbf{q}_A \times \nabla_t(N^A) N^B \mathbf{p}_B \right] \, dW = \boldsymbol{\mu} \cdot \mathbf{q}_A \times \mathbf{P}^{A,ext}. \quad (63)$$

With regard to (55), we can state that

$$\mathbf{q}_A \nabla_t(N^A) \approx \rho_0^{-1} N^A \mathbf{p}_A, \quad (64)$$

where the error depends on the element size, i.e. we fulfil this requirement only in an integral sense throughout the element. Assuming that the last statement is locally fulfilled, the second term on the left hand side vanishes, we obtain immediately

$$\boldsymbol{\mu} \cdot \mathbf{q}_A \times \int_{\tilde{\mathcal{B}}_0} \nabla_t(N^A N^B) dW \mathbf{p}_B = \boldsymbol{\mu} \cdot \mathbf{q}_A \times \mathbf{P}^{A,ext}, \quad (65)$$

such that

$$\begin{aligned} \boldsymbol{\mu} \cdot \mathbf{q}_A \times \int_{\tilde{\mathcal{B}}_0} (N^A(\mathbf{X}, t = t_2) N^B(\mathbf{X}, t = t_2)) dW \mathbf{p}_B - \\ \boldsymbol{\mu} \cdot \mathbf{q}_A \times \int_{\tilde{\mathcal{B}}_0} (N^A(\mathbf{X}, t = t_1) N^B(\mathbf{X}, t = t_1)) dW \mathbf{p}_B = \boldsymbol{\mu} \cdot \mathbf{q}_A \times \mathbf{P}^{A,ext}. \end{aligned} \quad (66)$$

Thus, angular momentum is conserved in an approximative sense depending on the local accuracy of (55).

**Total energy** Next, we investigate the impact of space-time elements on the energy conservation. Following the arguments outlined in (52), we set

$$\int_{\tilde{\mathcal{B}}_0} \mathbf{q}_A \cdot \nabla_t(N^A) \nabla_t(N^B) \mathbf{p}_B dW + \int_{\tilde{\mathcal{B}}_0} \mathbf{q}_A \cdot \nabla_t(\nabla_{\mathbf{X}}(N^A)) \cdot \boldsymbol{\Sigma}^h \nabla_{\mathbf{X}}(N^B) \mathbf{q}_B dW, \quad (67)$$

for the left hand side of (41). Assuming again that

$$\mathbf{q}_A \nabla_t(N^A) \approx \rho_0^{-1} N^A \mathbf{p}_A, \quad (68)$$

holds, the antiderivative of the first term on the right hand side reads

$$\begin{aligned} \int_{\tilde{\mathcal{B}}_0} \mathbf{p}_A \cdot N^A \rho_0^{-1} \nabla_t(N^B) \mathbf{p}_B dW &= \frac{1}{2} \mathbf{p}_A \cdot \int_{\tilde{\mathcal{B}}_0} \nabla_t(N^A \rho_0^{-1} N^B) dW \mathbf{p}_B, \\ &= \frac{1}{2} \mathbf{p}_A \cdot \left[ \int_{\tilde{\mathcal{B}}_0} N^A(\mathbf{X}, t = t_2) \rho_0^{-1} N^B(\mathbf{X}, t = t_2) dV \right. \\ &\quad \left. - \int_{\tilde{\mathcal{B}}_0} N^A(\mathbf{X}, t = t_1) \rho_0^{-1} N^B(\mathbf{X}, t = t_1) dV \right] \mathbf{p}_B, \\ &= T_{t_2} - T_{t_1}, \end{aligned} \quad (69)$$

i.e. we obtain the change in the kinetic energy. To calculate the change in the strain energy function, we use

$$\begin{aligned} \int_{\tilde{\mathcal{B}}_0} \mathbf{q}_A \cdot \nabla_t(\nabla_{\mathbf{X}}(N^A)) \cdot \Sigma^h \nabla_{\mathbf{X}}(N^B) \mathbf{q}_B \, dW &= \int_{\tilde{\mathcal{B}}_0} \underbrace{\frac{1}{2} \mathbf{q}_A \cdot \mathbf{q}_B \nabla_t(\nabla_{\mathbf{X}}(N^A) \otimes \nabla_{\mathbf{X}}(N^B))}_{\nabla_t \mathbf{C}^h} : \Sigma^h \, dW, \\ &= \int_{\tilde{\mathcal{B}}_0} \nabla_t \mathbf{C}^h : \nabla_{\mathbf{C}} \Psi(\mathbf{C}^h) \, dW. \end{aligned} \quad (70)$$

Assuming again for the sake of a clear presentation, that external contributions are absent

$$\begin{aligned} \int_{\tilde{\mathcal{B}}_0} \nabla_t \mathbf{C}^h : \nabla_{\mathbf{C}} \Psi(\mathbf{C}^h) \, dW &= \int_{\tilde{\mathcal{B}}_0} \nabla_t \Psi(\mathbf{C}^h) \, dW, \\ &= \int_{\tilde{\mathcal{B}}_0} \nabla_t \Psi(\mathbf{C}(\mathbf{X}, t = t_2)) \, dV - \int_{\tilde{\mathcal{B}}_0} \nabla_t \Psi(\mathbf{C}(\mathbf{X}, t = t_1)) \, dV, \\ &= V_{t_2} - V_{t_1}. \end{aligned} \quad (71)$$

In the presents of external contributions, additional terms can be added in a straightforward manner to the last term using (57). Eventually, we can state

$$T_{t_2} - T_{t_1} + V_{t_2} - V_{t_1} = 0, \quad (72)$$

which proofs conservation of total energy in an approximative sense. For this proof, the second derivative of the shape functions has been used. To avoid this construction, we assume that partial derivatives are exchangeable and obtain by using (64)

$$\begin{aligned} \int_{\tilde{\mathcal{B}}_0} \mathbf{q}_A \cdot \nabla_t(\nabla_{\mathbf{X}}(N^A)) \cdot \Sigma^h \nabla_{\mathbf{X}}(N^B) \mathbf{q}_B \, dW &\approx \int_{\tilde{\mathcal{B}}} \rho_0^{-1} \mathbf{p}_A \cdot \mathbf{q}_B \nabla_{\mathbf{X}}(N^A) \otimes \nabla_{\mathbf{X}}(N^B) : \Sigma^h \, dW, \\ &\approx \int_{\tilde{\mathcal{B}}_0} \nabla_t \mathbf{C}^h : \nabla_{\mathbf{C}} \Psi(\mathbf{C}^h) \, dW = V_{t_2} - V_{t_1}. \end{aligned} \quad (73)$$

which holds again in an approximative sense, analogues to the angular momentum.

## 5 Fluids

As usual, we write the fluid system in terms of an Eulerian description using the inverse mapping  $\mathbf{Y} = \boldsymbol{\varphi}^{-1}(\mathbf{y})$ . For the time differential of a physical quantity  $f(\mathbf{y})$ , it follows immediately that

$$\dot{f} = \nabla_t f + \mathbf{v} \cdot \nabla_{\mathbf{x}} f, \quad (74)$$

where  $\mathbf{v}(\mathbf{y})$  denotes the velocity at a specific event. Without loss of generality we restrict ourselves to the incompressible case and obtain for the conservation of mass equation

$$\nabla_x \cdot \mathbf{v} = 0. \quad (75)$$

For a classical Newtonian viscous fluid, the Cauchy stress tensor  $\boldsymbol{\sigma} : \tilde{\mathcal{B}} \rightarrow \mathbb{R}^{n \times n}$  is defined by

$$\boldsymbol{\sigma} = -p\mathbf{I} + \lambda\mathbf{I}\nabla_x \cdot \mathbf{v} + \mu(\nabla_x \mathbf{v} + \nabla_x \mathbf{v}^T), \quad (76)$$

where  $\mu$  denotes the dynamic viscosity and  $\lambda$  the second coefficient of viscosity. Here, the pressure  $p : \tilde{\mathcal{B}} \rightarrow \mathbb{R}$  is a sufficiently smooth function and can be regarded as a Lagrange multiplier introduced to enforce condition (75). Note that for the case of an incompressible fluid the second term on the right hand side vanishes. The non-conservative Eulerian form of the balance of linear momentum reads

$$\rho\dot{\mathbf{v}} = \nabla_x \cdot \boldsymbol{\sigma} + \rho\mathbf{g}, \quad (77)$$

where  $\rho$  denotes the density and  $\mathbf{g}$  a prescribed body force per unit of mass. To complete the strong form of the space-time boundary value problem, suitable boundary conditions need to be introduced as follows

$$\begin{aligned} \mathbf{v}(\mathbf{y}) &= \tilde{\mathbf{v}}(\mathbf{y}) \quad \text{on} \quad \mathbf{y} \in \mathcal{B} \times t_1 = \partial\tilde{\mathcal{B}}^{t_1}, \\ \mathbf{v}(\mathbf{y}) &= \bar{\mathbf{v}}(\mathbf{y}) \quad \text{on} \quad \mathbf{y} \in \partial\mathcal{B}^v \times I = \partial\tilde{\mathcal{B}}^v, \\ \boldsymbol{\sigma} \cdot \mathbf{n} &= \mathbf{h}(\mathbf{y}) \quad \text{on} \quad \mathbf{y} \in \partial\mathcal{B}^h \times I = \partial\tilde{\mathcal{B}}^h. \end{aligned} \quad (78)$$

Here, we assume that  $\partial\mathcal{B}^v \cup \partial\mathcal{B}^h = \partial\mathcal{B}$  and  $\partial\mathcal{B}^v \cap \partial\mathcal{B}^h = \emptyset$ . Moreover,  $\mathbf{n}$  denotes the unit outward normal field and  $\tilde{\mathbf{v}}$  can be interpreted as prescribed initial velocity field.

## 5.1 Weak formulation

Suitable functional spaces of test functions for both the velocity and pressure fields are introduced as follows,

$$\begin{aligned} \mathcal{V}^v &=: \left\{ \delta\mathbf{v} \in \mathbf{H}^1(\tilde{\mathcal{B}}) \mid (\delta\mathbf{v} = \mathbf{0} \text{ on } \partial\tilde{\mathcal{B}}^v) \wedge (\delta\mathbf{v} = \mathbf{0} \text{ on } \partial\tilde{\mathcal{B}}^{t_1}) \right\}, \\ \mathcal{V}^p &=: \left\{ \delta p \in L_2(\tilde{\mathcal{B}}) \right\}. \end{aligned} \quad (79)$$

Next, we rewrite the balance of linear momentum using Gauss divergence theorem along with an integration by parts and obtain

$$\int_{\tilde{\mathcal{B}}} \delta\mathbf{v} \cdot \rho(\dot{\mathbf{v}} - \mathbf{g}) \, dW + \int_{\tilde{\mathcal{B}}} \nabla_x(\delta\mathbf{v}) : \boldsymbol{\sigma} \, dW - \int_{\partial\tilde{\mathcal{B}}^h} \delta\mathbf{v} \cdot \mathbf{h} \, d\Gamma = 0, \quad (80)$$

supplemented by the kinematic constraint

$$\int_{\tilde{\mathcal{B}}} \delta p \nabla_x \cdot \mathbf{v} \, dW = 0. \quad (81)$$

Both equations (80) and (81) have to hold for all  $\delta\mathbf{v} \in \mathcal{V}^v$  and  $\delta p \in \mathcal{V}^p$ .

## 5.2 Discretisation

Without prejudice to the generality of the formulation, a Q2Q1 Lagrangian finite element discretisation is chosen, widely known as Taylor-Hood element. Therefore we consider a finite element discretisation of the space-time

$$\begin{aligned} \mathbf{v}^h &= \sum_{A \in \omega} N^A(\mathbf{y}) \mathbf{v}_A; & \delta \mathbf{v}^h &= \sum_{A \in \omega} N^A(\mathbf{y}) \delta \mathbf{v}_A \\ p^h &= \sum_{B \in \bar{\omega}} M^B(\mathbf{y}) p_B; & \delta p^h &= \sum_{B \in \bar{\omega}} M^B(\mathbf{y}) \delta p_B \end{aligned} \quad (82)$$

where  $N^A(\mathbf{y}) : \tilde{\mathcal{B}} \rightarrow \mathbb{R}$  are quadratic shape functions associated with nodes  $A \in \omega = \{1, \dots, n\}$  and  $M^B(\mathbf{x}) : \tilde{\mathcal{B}} \rightarrow \mathbb{R}$  are linear shape functions associated with nodes  $B \in \bar{\omega} = \{1, \dots, m\}$ . This element is known to satisfy the LBB condition and provides optimal quadratic convergence of the velocity field, see Donea & Huerta [18], chapter 6.5. Hence, the Taylor-Hood element allows to circumvent stabilisation techniques (cf. Tezduyar [65]) using extended test function spaces and we can define in a standard fashion the functional spaces

$$\begin{aligned} \mathcal{V}^{v,h} &=: \left\{ \delta \mathbf{v}^h \in H^1(\tilde{\mathcal{B}}^h) \mid \left( \delta \mathbf{v}^h = \mathbf{0} \text{ on } \partial \tilde{\mathcal{B}}^{v,h} \right) \wedge \left( \delta \mathbf{v}^h = \mathbf{0} \text{ on } \partial \tilde{\mathcal{B}}^{t_1,h} \right) \right\}, \\ \mathcal{V}^{p,h} &=: \left\{ \delta p^h \in L_2(\tilde{\mathcal{B}}^h) \right\}. \end{aligned} \quad (83)$$

The full-discrete balance of momentum reads

$$\delta \mathbf{v}_A \cdot \int_{\tilde{\mathcal{B}}} N^A \rho (\dot{\mathbf{v}}^h - \mathbf{g}) dW + \delta \mathbf{v}_A \cdot \int_{\tilde{\mathcal{B}}} \nabla_x(N^A) : \boldsymbol{\sigma}^h dW - \delta \mathbf{v}_A \cdot \int_{\partial \tilde{\mathcal{B}}^h} N^A \mathbf{h}^h d\Gamma = 0, \quad (84)$$

whereas the incompressibility constraint is evaluated as

$$\delta p_A \int_{\tilde{\mathcal{B}}} M^B \nabla_x(N^B) dW \cdot \mathbf{v}_B = 0. \quad (85)$$

Eventually, we note that

$$\dot{\mathbf{v}}^h = \nabla_t(N^A) \mathbf{v}_A + \left( \sum_B \mathbf{v}_B \otimes \nabla_x(N^B) \right) \sum_C N^C \mathbf{v}_C, \quad (86)$$

represents the acceleration term.

## 6 Kuramoto-Sivashinsky equation

In this section we consider a representative example for higher order diffusion problems. In particular, the Kuramoto-Sivashinsky model is used for the simulation of different thermodynamical systems far away from equilibrium state, e.g. fluctuations in fluid

films and instabilities in laminar flame fronts, see Paniconi & Elder [57] among many others. This transient evolution equation of fourth-order accounts for the physics of arising chaotic instabilities in the spatial as well as the temporal domain. The numerical resolution of the fourth-order operator is traditionally realized using finite difference schemes. In the context of FEA for irreducible systems, global  $C^1$ -continuity is required for spatial discretization, see Gomez & Paris [30] and Hesch et al. [34] for an application of NURBS basis functions.

Assume a Lipschitz bounded domain  $\mathcal{B} \subseteq \mathbb{R}^n$ ,  $n \in \{2, 3\}$ , in the time interval  $I = [0, T]$ . Then the Kuramoto-Sivashinsky problem is stated as

$$\nabla_t(u) + \nu \Delta_x^2(u) + \Delta_x(u) + \alpha u - \|\nabla_x(u)\|^2 = 0, \quad (87)$$

where the constitutive parameters  $\nu$  and  $\alpha$  are used to model diffusion induced chaos and dissipative structures in reaction diffusion systems. Analogues to the thermal system, we do not distinguish between the reference and the actual configuration, i.e.  $\mathbf{y} = \mathbf{Y}$ . This space-time boundary value problem is supplemented by appropriate boundary conditions given as

$$\begin{aligned} u(\mathbf{y}) &= \tilde{u}(\mathbf{y}) \quad \text{on} \quad \mathbf{y} \in \mathcal{B} \times t_1 = \partial\tilde{\mathcal{B}}^{t_1}, \\ \nabla_x(u(\mathbf{y})) \cdot \mathbf{n} &= 0 \quad \text{on} \quad \mathbf{y} \in \partial\mathcal{B} \times I = \partial\tilde{\mathcal{B}}^u, \\ \nabla_x(u(\mathbf{y}) + \Delta_x(u(\mathbf{y}))) \cdot \mathbf{n} &= 0 \quad \text{on} \quad \mathbf{y} \in \partial\mathcal{B} \times I = \partial\tilde{\mathcal{B}}^u. \end{aligned} \quad (88)$$

Note, that (88)<sub>2</sub> and (88)<sub>3</sub> are often replaced by periodic boundary conditions, see e.g. Gomez & Paris [30].

## 6.1 Weak formulation

Next we introduce the functional space of solution and weighting functions

$$\begin{aligned} \mathcal{S} &= \{u \in H^2(\tilde{\mathcal{B}}) \mid u = \tilde{u} \text{ on } \partial\tilde{\mathcal{B}}^{t_1}\}, \\ \mathcal{V} &= \{\delta u \in H^2(\tilde{\mathcal{B}}) \mid \delta u = 0 \text{ on } \partial\tilde{\mathcal{B}}^{t_1}\}, \end{aligned} \quad (89)$$

where  $H^2(\tilde{\mathcal{B}})$  is the Sobolev space of square integrable functions  $u : \tilde{\mathcal{B}} \rightarrow \mathbb{R}$  with square integrable derivatives of first and second order. The weak form of the problem reads

$$\int_{\tilde{\mathcal{B}}} \delta u (\nabla_t(u) + \alpha u - \|\nabla_x(u)\|^2) + \nu \Delta_x(\delta u) \Delta_x(u) - \nabla_x(\delta u) \cdot \nabla_x(u) \, dW = 0, \quad (90)$$

for all  $\delta u \in \mathcal{V}$ .

## 6.2 Discretisation

For the finite element approximation of (90) we employ quadratic B-spline based shape functions which fulfil the continuity requirements specified in (89). Accordingly, the discrete functional spaces are defined as

$$\begin{aligned}\mathcal{S}^h &= \left\{ u^h \in \mathbb{H}^2(\tilde{\mathcal{B}}^h) \mid u^h = \tilde{u} \text{ on } \partial\tilde{\mathcal{B}}^{t_1,h} \right\}, \\ \mathcal{V}^h &= \left\{ \delta u^h \in \mathbb{H}^2(\tilde{\mathcal{B}}^h) \mid \delta u^h = 0 \text{ on } \partial\tilde{\mathcal{B}}^{t_1,h} \right\},\end{aligned}\tag{91}$$

and the approximations of  $u$  and  $\delta u$  read

$$u^h = \sum_{A \in \omega} R^A u_A, \quad \delta u^h = \sum_{A \in \omega} R^A \delta u_A,\tag{92}$$

where  $R^A : \tilde{\mathcal{B}} \rightarrow \mathbb{R}$  are the global shape functions associated with a net of control points  $\mathbf{y}_A$ ,  $A \in \omega = \{1, \dots, n_{\text{node}}\}$  within the space-time domain. The multivariate B-spline basis functions are defined within the  $n + 1$  dimensional parametric domain  $\hat{\mathcal{B}}$  as dyadic product of univariate B-splines

$$R^A = \prod_{l=1}^{n+1} R_{i_l,p}(\xi^l).\tag{93}$$

A univariate B-spline of the  $l$ -th parametric direction and polynomial order  $p$  is recursively defined via a sequence of knots  $[\xi_1^l \leq \xi_2^l \leq \dots \leq \xi_{n_l+p+1}^l]$  as follows

$$R_{i,p}(\xi^l) = \frac{\xi^l - \xi_{i-1}^l}{\xi_i^l - \xi_{i-1}^l} R_{i,p-1}(\xi^l) + \frac{\xi_{i+1}^l - \xi^l}{\xi_{i+1}^l - \xi_i^l} R_{i+1,p-1}(\xi^l),\tag{94}$$

starting with piecewise constant functions

$$R_{i,0}(\xi^l) = \begin{cases} 1 & \text{if } \xi_{i-1}^l \leq \xi^l < \xi_i^l \\ 0 & \text{otherwise} \end{cases},\tag{95}$$

see Dittmann [17] and Hesch et al. [34] for further details. A geometrical map  $\mathfrak{F} : \hat{\mathcal{B}} \rightarrow \tilde{\mathcal{B}}$  is introduced to link the parametric domain and the space-time domain

$$\mathbf{y}^h := \mathfrak{F}(\boldsymbol{\xi}) = R^A(\boldsymbol{\xi}) \mathbf{y}_A.\tag{96}$$

Eventually, the fully discrete setting of (90) reads

$$\begin{aligned}\delta u_A \int_{\tilde{\mathcal{B}}} R^A (\nabla_t(R^B) u_B + \alpha R^B u_B - \|\nabla_{\mathbf{x}}(R^B) u_B\|^2) \\ + \nu \Delta_{\mathbf{x}}(R^A) \Delta_{\mathbf{x}}(R^B) u_B - \nabla_{\mathbf{x}}(R^A) \cdot \nabla_{\mathbf{x}}(R^B) u_B \, dW = 0,\end{aligned}\tag{97}$$

representing the discrete space-time version of the Kuramoto-Sivashinsky model.

## 7 Semi-geometric multigrid methods for a continuous space-time approach

Based on the fact that classical iterative methods quickly smooth the high-frequency error, multigrid solvers combine two different methods on two different finite element spaces.

On a coarse space  $\mathcal{S}^H$  a correction is easily computed, usually by means of a direct solver, resolving the low frequencies of the error. Then the correction is prolonged on a fine space  $\mathcal{S}^h$  and few iterations of Jacobi or Gauß-Seidel methods are employed in order to remove the high-frequency components of the error. In order to solve the coarse level problem, the fine-level residual has to be projected on the coarse space.

The smoother and restriction operator are two main ingredients which affect the multigrid performance. In the purely geometric approach, usually two nested meshes are employed in order to solve for the correction on the coarse level and smoothing the current iteration of the solution vector on the fine level. The restriction operator used to project the residual from the fine to the coarse mesh is usually taken as the transpose of the interpolation operator between the coarse and the fine mesh. Convergence and optimality proofs of this approach have been shown for elliptic problems but it has been successfully employed also in saddle-point problems. In the algebraic multigrid, projection and coarse level operators are constructed only by means of aggregation of the entries of the fine level operator. While this approach gives good convergence properties for positive definite matrices, optimality in general can be shown for discretizations which results in a stiffness matrix which enjoys the M-matrix property.

The semi-geometric approach allows to extend the good convergence properties of geometric multigrid to non-nested couple of meshes and keeping a purely algebraic construction of the coarse level operators [15].

In this section, we describe the semi-geometric multigrid for space-time discretization, considering specifically the thermal conduction problem (9). Let  $\tau_h$  and  $\tau_H$  denote two different triangulations of  $\tilde{\mathcal{B}}$ , where the subscripts  $h$  and  $H$  denote their mesh-size parameter. We assume that  $h < H$ , so that  $\tau_h$  is finer than  $\tau_H$ . As described in Section 3, we denote by  $\mathcal{S}^h$  and  $\mathcal{S}^H$  the two interpolation spaces, whose basis function are denoted by as  $N_h^A$  and  $N_H^B$ . Here,  $A \in \omega_h = \{1, \dots, n_{\text{node}}^h\}$  and  $B \in \omega_H = \{1, \dots, n_{\text{node}}^H\}$  are the space-time indices of the nodes on each of the meshes.

The algebraic formulation of the finite element problem (23) reads

$$\mathbf{B}^h \underline{\theta}^h = \underline{\mathbf{f}}^h.$$

The stiffness matrix  $\mathbf{B}^h \in \mathbb{R}^{n_{\text{node}}^h \times n_{\text{node}}^h}$  and the right-hand-side vector  $\underline{\mathbf{f}}^h \in \mathbb{R}^{n_{\text{node}}^h}$  are defined as

$$\mathbf{B}_{i,j}^h = B(N_h^i, N_j^A) \quad \text{with} \quad i, j \in \omega_h,$$

and

$$\underline{\mathbf{f}}_i^h = l(N_h^i) \quad \text{with} \quad i \in \omega_h.$$

The vector  $\underline{\theta}^h$  is the vector having as components the unknown coefficients of the finite element function  $\theta^h$ .

The prolongation operator  $\mathbf{P}_{H \rightarrow h}$  is defined as the  $L^2$  projection from  $\mathcal{V}^H$  into  $\mathcal{V}^h$  [14]. Observe that the coarse-level correction is defined on the test function space. Given a function  $c^H \in \mathcal{V}^H$ , its projection can be computed by solving

$$\mathbf{M}_h \underline{c}^h = \mathbf{T}_{H \rightarrow h} c^H,$$

where  $c^h \in \mathbb{R}^{n_{\text{node}}^h}$  and  $\underline{c}^h \in \mathbb{R}^{n_{\text{node}}^h}$  denote the nodal vectors of the finite element functions  $c^h$  and  $c^H$ , respectively. The matrix  $\mathbf{M}_h \in \mathbb{R}^{n_{\text{node}}^h \times n_{\text{node}}^h}$  is a mass matrix like structure on the fine finite element space, whose entries are defined as

$$(\mathbf{M}_h)_{i,j} = \int_{\tilde{\mathcal{B}}} \theta^j(\mathbf{y}) \theta^i(\mathbf{y}) \, d\mathbf{y} \quad \text{with} \quad i, j \in \omega_h.$$

The entries of the matrix  $\mathbf{T}_{H \rightarrow h} \in \mathbb{R}^{n_{\text{node}}^h \times n_{\text{node}}^H}$  are instead defined as

$$(\mathbf{T}_{H \rightarrow h})_{i,k} = \int_{\tilde{\mathcal{B}}} \theta^k(\mathbf{y}) \theta^j(\mathbf{y}) \, d\mathbf{y} \quad \text{with} \quad i \in \omega_h, k \in \omega_H.$$

The restriction operator is simply defined as  $\mathbf{R}_{h \rightarrow H} = \mathbf{P}_{H \rightarrow h}^T$ .

As in the algebraic approach, once prolongation and restriction operators are computed, the coarse level matrix  $\mathbf{B}^H$  can be computed by means of the Galerkin assembling procedure

$$\mathbf{B}^H = \mathbf{R}_{h \rightarrow H} \mathbf{B}^h \mathbf{P}_{H \rightarrow h}. \tag{98}$$

Obviously this approach is equivalent to the geometric multigrid in case the two grids are nested. The computation of the prolongation operator  $\mathbf{P}_{H \rightarrow h} = \mathbf{M}_h^{-1} \mathbf{T}_{H \rightarrow h}$  requires formally the inversion of a mass matrix on the fine level. The use of bi-orthogonal basis functions allows to lump this matrix and hence directly store its approximation  $\tilde{\mathbf{P}}_{H \rightarrow h}$  [48].

## 7.1 Time stabilization

Multigrid theory is well-established for elliptic problems. Space-time problems are only first order in time direction and they can be considered as the limit of singularly perturbed diffusion-reaction problems. For this class of operators, multigrid solvers are known not to perform efficiently [68, 10, 9, 56]. The main problem is that standard finite element methods provide matrices that are not M-matrices, and hence do not preserve the positivity and monotonicity properties of the solution.

To understand this problem, we consider a low-order discretization of the diffusion-convection problem

$$Au = -\mu \frac{d^2u}{dx^2} + b \frac{du}{dx}.$$

The discretization with linear finite elements gives a local stiffness matrix that reads

$$\mathbf{A}_h = \frac{\mu}{h} \begin{vmatrix} 1 & -1 \\ -1 & 1 \end{vmatrix} + \frac{b}{2} \begin{vmatrix} -1 & 1 \\ -1 & 1 \end{vmatrix}, \quad (99)$$

where the first term at the right-hand-side comes from the discretization of the diffusive term and the second from the discretization of convective term. The parameter  $h$  is the element size.

In order to ensure that the local stiffness matrix is an M-matrix, the entry  $(\mathbf{A}_h)_{1,2}$  has to be smaller than zero. This leads to the condition on the Péclet number [58]

$$Pe = \frac{bh}{2\mu} < 1.$$

When the ratio  $b/\mu \gg 1$ , the mesh has to be fine enough in order to ensure that the local matrix is an M-matrix. When it is not possible, two possible strategies are available

- Numerical diffusion: the diffusion parameter  $\mu$  is replaced by  $\mu_h = \mu(1 + \phi(Pe))$ , where  $\phi$  is a function which tends to zero when  $h \rightarrow 0$ , such that the discretization is consistent. The choice  $(\phi(x) = x)$  leads to upwind schemes [38] but more complicated choices are possible, such as the Scharfetter-Gummel scheme [59].
- Residual-based stabilizations: the space of the test functions is enriched with local basis functions [39]. Since these basis functions are local, they can be eliminated leading to the so-called unusual-stabilization schemes [23].

In order to render the problem elliptic in the space-time formulation, we modify Eq. 16 adding a stabilisation of the form

$$c_{\Delta t}(\theta, \delta\theta) = \int_{\tilde{\mathcal{B}}} \gamma \Delta t \nabla_t \theta \nabla_t \delta\theta \, dW, \quad (100)$$

where  $\gamma$  is a positive parameter to be accurately chosen depending on the physical properties of the space-time operator. A similar approach has been used in [47], in order to stabilize mortar finite element discretization in a space-time approach. Our approach can be classified as an upwind scheme for which the analysis is simpler [58]. As in upwinded schemes, the parameter  $\Delta t \gamma$  plays the role of artificial diffusion and the  $\Delta t$  factor renders the stabilization consistent.

To study the optimal choice of the parameter  $\gamma$ , we consider the ordinary differential equation

$$\begin{cases} b \frac{du}{dt} = -u \\ u(0) = 1. \end{cases} \quad (101)$$

Introducing the time stabilization (100), the stabilized differential operator reads

$$-\alpha \Delta t \frac{d^2 u}{dt^2} + b \frac{du}{dt},$$

and the stabilized stiffness matrix is

$$\mathbf{A}_h = \alpha \begin{vmatrix} 1 & -1 \\ -1 & 1 \end{vmatrix} + \frac{b}{2} \begin{vmatrix} -1 & 1 \\ -1 & 1 \end{vmatrix}. \quad (102)$$

The study of the entry  $(\mathbf{A}_h)_{1,2}$  leads to the condition

$$b/2 < \alpha. \quad (103)$$

Since the condition (103) does not depend on  $\Delta t$ , the coarse level operators obtained with the Galerkin assembly procedure (98) will be M-matrices and hence they do not require a level-dependent  $\alpha$ .

In order to quantify the error introduced by the upwind stabilization, we perform convergence tests for the equation (101). Results for different values of  $\alpha$  are reported in Figure 1. We observe that, as expected, the upwinded schemes have only first order accuracy and it holds true both for monotone and non-monotone schemes.

## 8 Examples

In this section we investigate the numerical performance of the proposed finite element formulation. We consider a series of planar benchmark tests for all systems shown in the theoretical part. To demonstrate the accuracy, we compare the results with traditional time stepping schemes and verify different conservation properties whenever possible.

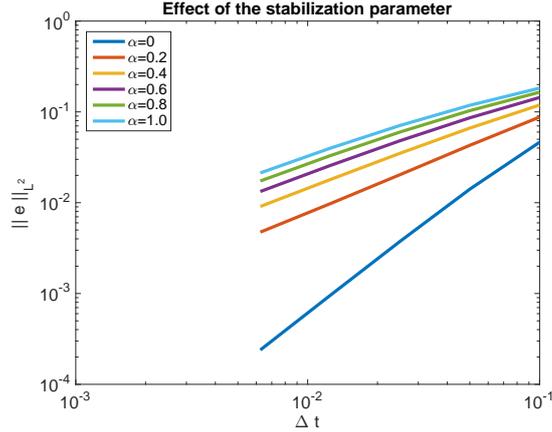


Figure 1: Convergence plot for different values of the upwind stabilization parameter.

An incremental iterative Newton-Raphson solution procedure has been implemented to solve monolithically all arising non-linear system of equations.

We present the convergence properties of a space-time multigrid solver applied to the thermal conduction and the Navier-Stokes problems presented in Section 3 and Section 5, respectively. The first one represent a standard test-case for parabolic problems and the second an interesting non-linear problem with saddle-point structure. Numerical tests of multigrid solver has been carried out in the application HART [50], previously EWE, developed inside the finite element framework MOOSE. In particular, on the side of the smoothers, we use time-line smoothers for the heat-equation and standard point block-smoothers for Navier-Stokes problem. The former show a huge improvement in terms of convergence speed, which can be expected, taking into account the convective nature of the time direction.

## 8.1 Thermal conduction

As a first example we consider a planar conduction problem with an initial temperature distribution as depicted in Figure 2. The geometry has a size of  $[-2.5, 2.5] \times [-2.5, 2.5]$  in [cm], the temporal domain has a size of  $[0, 4]$  and the initial temperature is given by  $\theta(x_1, x_2, 0) = 0.04 \cdot (x_1 + 2.5) \cdot (x_2 + 2.5) + 293$ . The Neumann boundaries are equal zero and except of the initial conditions no further Dirichlet boundary conditions are given, i.e. the system is adiabatically isolated. The number of elements is given by  $[32, 32, 64]$  and the material parameters for copper at room temperature are used: conductivity  $\mathbf{K} = 400\mathbf{I}$ , heat capacity  $c_p = 385$  and density  $\rho = 8920$ . Note that unless otherwise stated, all values are given in SI units.

Results are displayed in Figure 3, the z-axis represents the temporal domain. For comparison, Figure 3 also presents the results of a classical implicit Euler scheme using

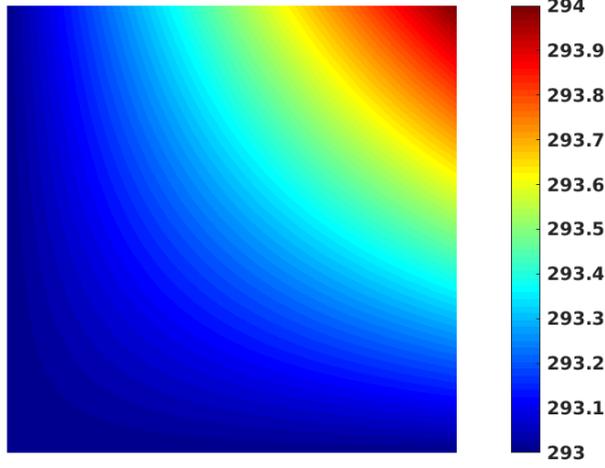


Figure 2: Initial temperature distribution.

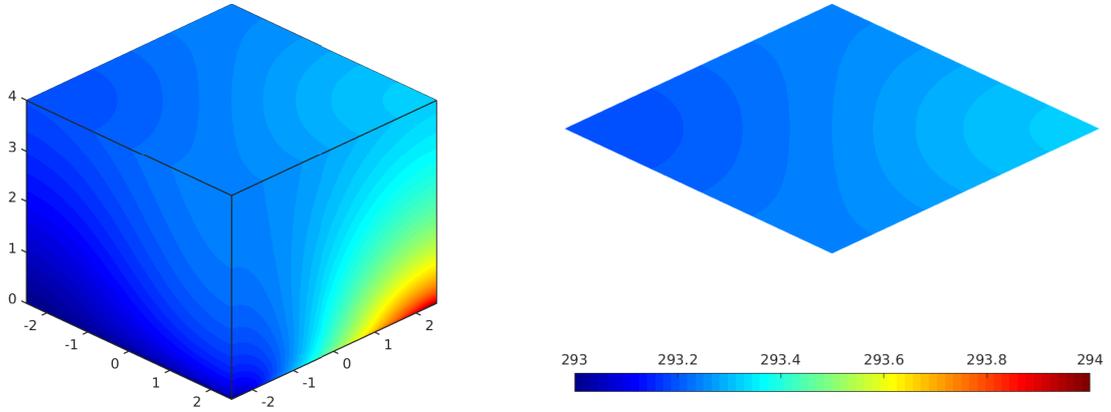


Figure 3: Results of the conductivity problem. Left: space-time cube; right: classical time stepping scheme.

2048 time-steps. The change in total energy throughout time for this adiabatic isolated system is given by

$$\int_{\tilde{\mathcal{B}}^h} \nabla_t(e^h) dW = 1.5353 \times 10^{-8}, \quad (104)$$

which is within the range of round-off errors.

The eigenvalue distribution of the stiffness matrix of this specific system with 69696 degrees of freedom is plotted in Figure 4. Note that the formulation is non-symmetric in general, i.e. we obtain complex eigenvalues. Moreover Figure 4 shows a convergence plot for the present thermal conduction problem. Therefore, we calculated an overkill solution using 100000 time-steps using a classical implicit Euler scheme and use

$$error = \left\| [\theta_1, \dots, \theta_{n_{\text{node}}}] - [\theta_1^{\text{overkill}}, \dots, \theta_{n_{\text{node}}}^{\text{overkill}}] \right\|, \quad (105)$$

as measure of the error.

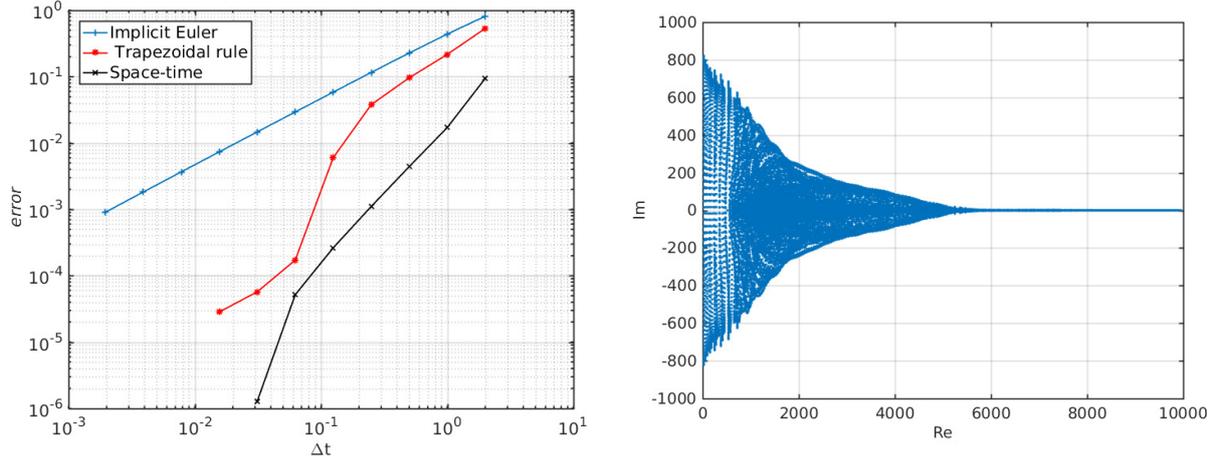


Figure 4: Left: convergence plot; right: eigenvalues of the discrete thermal system.

### 8.1.1 Multigrid performance for thermal systems

We first consider the direct application of multigrid solvers to the space-time discretisation of heat equation using linear finite elements on space-time hexahedral meshes. The space and time discretisation parameters are denoted by  $dt$  and  $dx$ , respectively. Again, we obtain a large linear system of equations in the space-time domain, even for this spatially 2-dimensional system. The applied multigrid method is characterised by the following components: 3 pre- and post-smoothing steps of Gauss-Seidel, two level V-cycle, and nested meshes. Assembly of coarse levels is realised by means of a Galerkin approach, which is also suitable for non-nested meshes. Convergence is related to the  $L^2$  norm of the linear residual  $\|r_m\|$  at the  $m$ -th iteration and the convergence rate is defined as  $\rho_m = \|r_m\|/\|r_{m-1}\|$ . V-cycle iterations are performed until the norm of the residual is reduced by a factor of  $10^{12}$ . The reported convergence rates are computed as the mean convergence rate over the different cycles.

**A naive approach to heat equation** Figure 6 shows the convergence rate of the for different space-time meshes. It can be clearly seen that, differently from the application of elliptic problems, the method is not robust with respect to space-time refinements. In particular, we observe a decrease of multigrid performances with respect to mesh refinement. As discussed in Subsection 7.1, the non-optimal behaviour of this approach can be understood from parabolic structure of the equations. The equation in time can be seen as a degenerate convection-diffusion equation. For these kind of equations, the coarse grid approximation may deteriorate the convergence since convection is responsible of grid-scale effects: oscillations present in the coarse scale are not smoothed and smoothers may diverge for convection-dominated problems [9].

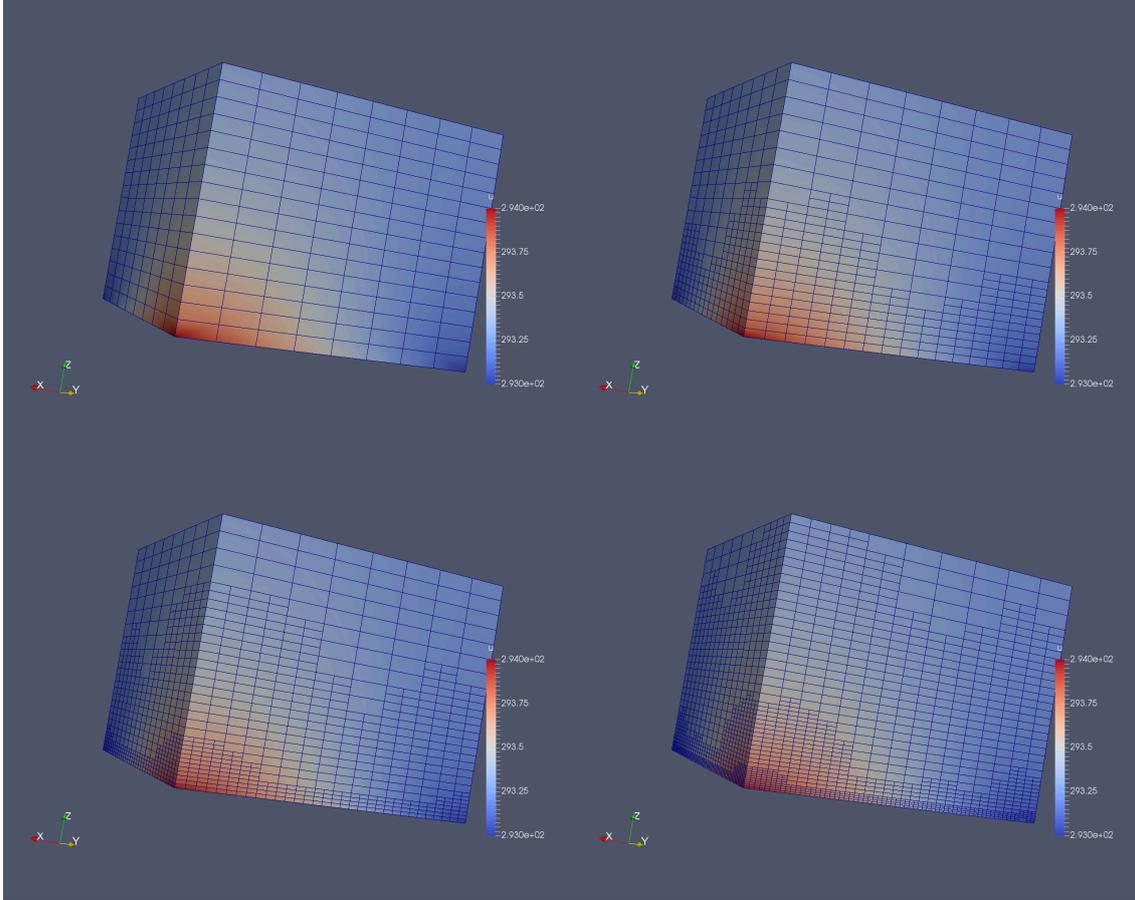


Figure 5: Space-time adaptive approach based on a gradient error indicator.

**Diffusion-in-time stabilisation** Following the approach presented in [47], in order to render the problem elliptic in the space-time formulation, we modify Eq. 16 in a consistent way adding a stabilisation term of the form (100). This new term renders the space-time heat equation elliptic for each  $\Delta t$  and but reduces the convergence order of the discretization scheme. In the one-dimensional case an optimal value for the parameter can be found studying the monotonicity property of the stiffness matrix. In these (2+1)-dimensional case, a similar study is more complicated. Here we want to show by means of a numerical study, how a threshold for  $\gamma$  exists in order to gain optimality of multigrid solvers.

In Table 1, we report the convergence properties of a multigrid solver applied to the modified system. It can be observed that, if  $\gamma$  is smaller than a certain threshold, multigrid performances are not achieved and are almost  $\gamma$ -independent. On the other hand, for  $\gamma > 0.128$ , the solver has an optimal behaviour showing mesh-independency. In particular for larger values of the artificial diffusion, a better behaviour of multigrid can be observed.

This is also confirmed by the results reported in Figure 7: an increase in the ellipticity

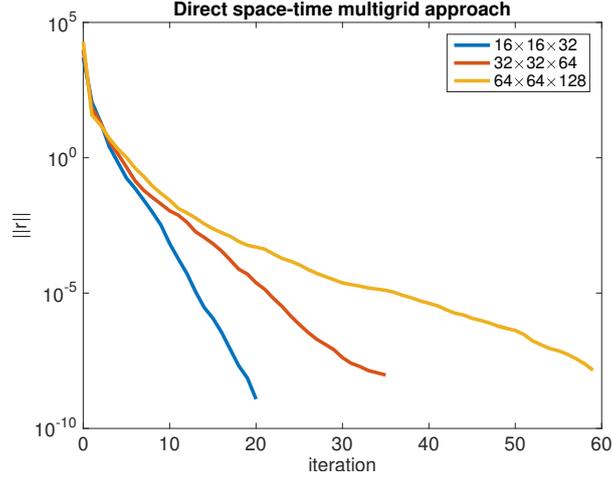


Figure 6: Convergence of standard multigrid in the space-time domain.

Fine mesh $\gamma$	$16 \times 16 \times 32$	$32 \times 32 \times 64$	$64 \times 64 \times 128$
	0.004	0.264	0.519
0.008	0.264	0.519	0.692
0.016	0.264	0.519	0.692
0.032	0.264	0.519	0.692
0.064	0.264	0.519	0.692
0.128	0.014	0.020	0.065
0.256	0.008	0.008	0.020
0.512	0.005	0.003	0.005

Table 1: Two-level convergence factor for different stabilization parameters.

of the problem provide a faster convergence rate. The numerical simulations also confirm that the stabilization in Eq. (100) is only problem-dependent. The choice of  $\gamma$  is not related to the coarsest or finest levels employed, rendering our semi-geometric approach possible since it does not require further or different corrections on the coarsest levels. and the Galerkin assembly procedure gives naturally the correct coarse level operators.

**Three-level method** Similar results can be observed from the convergence rates reported in Table 2 where the convergence rate for a three-level method is reported. Again, we can observe the existence of threshold value for  $\gamma$  for which optimal convergence is achieved. This value seems not to depend on the time-space mesh-size.

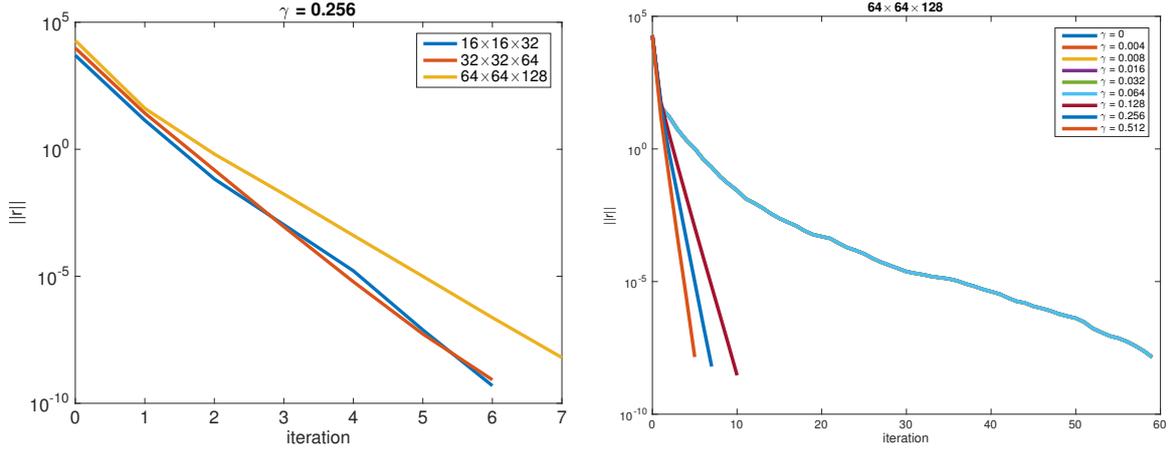


Figure 7: Left: convergence of stabilized MG for a given  $\gamma$ . Right: convergence behaviour of MG with respect to the stabilisation parameter.

$\gamma$	Fine mesh		
	$32 \times 32 \times 64$	$64 \times 64 \times 128$	$128 \times 128 \times 256$
0.004	0.598	0.788	0.939
0.008	0.598	0.788	0.939
0.016	0.598	0.788	0.939
0.032	0.598	0.788	0.939
0.064	0.598	0.788	0.939
0.128	0.006	0.007	0.008
0.256	0.003	0.003	0.004
0.512	0.003	0.003	0.003

Table 2: Three-level convergence factor for different stabilization parameters.

**A new time-line smoothing strategy** Line-smoothers have been developed for convection dominated problems [56, 68, 10] and are based on the idea of re-ordering strategy for grid-points following the flow direction and, possible, processing nodes in forward and backward directions. Again, we follow the similarities between time-flow and convection. Being the convective term only along the time co-ordinate axis, we extend this idea by smoothing the spatial nodes following the time direction.

In Table 3, we can observe how the use of this additional information is responsible for a huge reduction of the overall cost of the solution procedure. In a parallel approach, this approach requires a careful distribution of the nodes keeping all spatial nodes at different times on the same processor. An alternative variant for the parallelisation of this inherent sequential smoother has been presented in [56].

**A note on space-time adaptivity** The continuous-in-time space-time approach allows for a simple and quick re-use of spatially adaptive finite element softwares. The ellipticity of the problem allows to use error indicator which are usually used in time sequential codes. In Figure 8 we report a comparison of the  $H^1$ -error between three uniform mesh refinements and four adaptive mesh refinement steps based on gradient-jump indicator. We can observe how, also in the space-time approach, a lower error can be obtained with less degrees of freedom. The resulting meshes of the four space-time adaptivity steps for the thermal conduction problem considered in this section are reported in Figure 5.

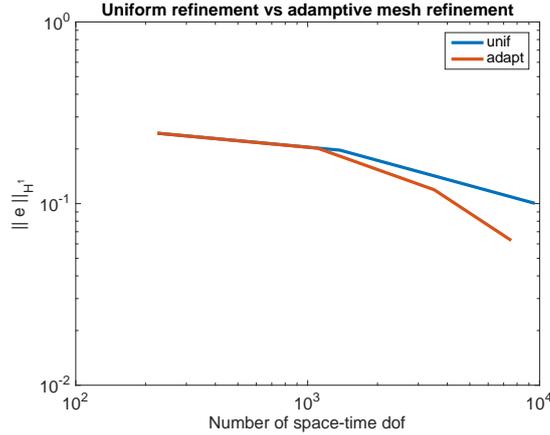


Figure 8: Comparison between uniform mesh refinement and adaptive mesh refinement based on a gradient indicator.

	$8 \times 8 \times 16$	$16 \times 16 \times 32$	$32 \times 32 \times 64$
Gauss-Seidel	0.010 (7)	0.014 (7)	0.007 (6)
Time-line smoother	3.01e-04 (4)	4.92e-04 (6)	6.50e-05 (3)

Table 3: Two-level convergence factor and, in bracket, the number of V-cycle iterations for a standard Gauss-Seidel smoother and a time-line smoother.

## 8.2 Non-linear elasticity

This example deals with a continuum body which assumes the form of an L-shaped block in the stress-free initial configuration (see Figure 9, left). This planar example consists of 36 elements in the spatial direction. Moreover, the interval in temporal direction  $T = [0, 1.5]$  is divided in 30 slices. Hence, we have to solve for 8835 unknowns. Hyperelastic constitutive behaviour is assumed to be governed by a compressible Neo-Hooke material with associated stored energy function

$$W(\mathbf{C}) = \frac{\mu}{2} [\text{tr}(\mathbf{C}) - n] + \frac{\lambda}{2} (\ln(J))^2 - \mu \ln(J). \quad (106)$$

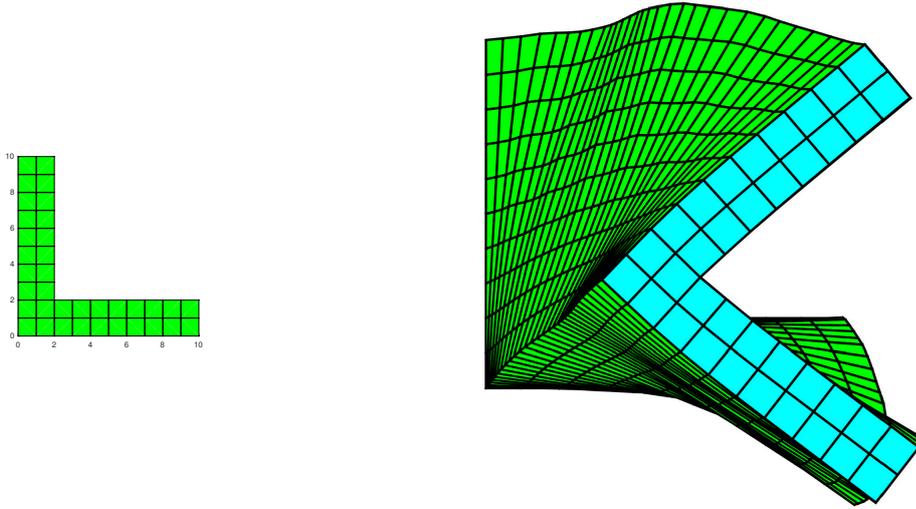


Figure 9: Non-linear elasticity. Left: initial configuration, right: comparison of results

The corresponding material constants are assumed to take the values  $\lambda = 1298.100$ ,  $\mu = 865.3846$  and  $\rho = 0.1$ , where  $\lambda$  and  $\mu$  denote Lamé parameters. An initial velocity distribution is predefined such that the  $x$ -component  $v_x$  is equal the  $y$ -component of the reference configuration, i.e. zero at  $y = 0$  and 10 at the upper end of the L-shape. Thus, we obtain an inhomogeneous initial velocity distribution of this L-shape, leading to a deformation in time. Further Dirichlet and Neumann boundary conditions are not given, i.e. the L-shape can move free in space.

In Figure 9, right, results for the space-time element (green) and a classical implicit Euler scheme (cyan) are displayed. About 2500 timesteps were necessary to obtain a similar accurate solution compared to the space-time results, demonstrating the superiority of the proposed formulation. In Figure 10, the corresponding von Mises stress distribution for both, the space-time as well as the classical time stepping scheme is shown.

The change of linear and angular momentum over all time-steps is given as follows

$$\int_{\tilde{B}_0} \nabla_t(L_\xi(\mathbf{Y})) dW = 0.3137 \times 10^{-13}, \quad \int_{\tilde{B}_0} \nabla_t(J_\xi(\mathbf{Y})) dW = 0.0888. \quad (107)$$

Reducing gradually the element size in temporal direction yields for the total change in angular momentum and total energy

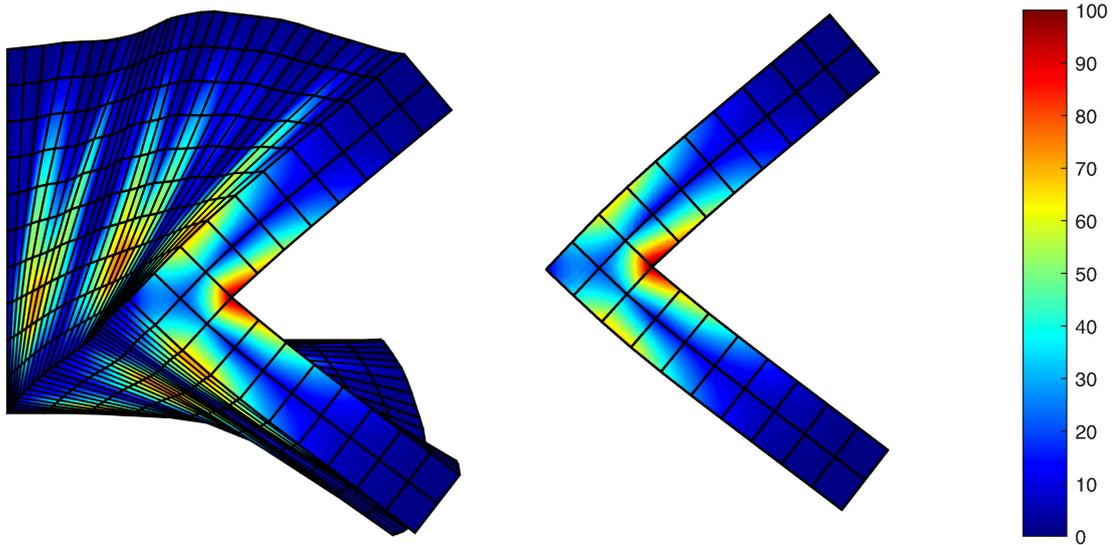


Figure 10: Results of the elasticity problem, colors indicate von Mises stress distribution. Left: space-time formulation, right: classical time stepping scheme.

time slices	30	90	150	210
change in AM	0.088	0.009	0.002	$6.192 \times 10^{-04}$
change in energy	0.7496	0.6134	0.3649	0.0773

i.e. linear momentum is algorithmically conserved and we obtain a clear convergence of angular momentum and total energy. Note that the displayed values represent the integral over space and time and thus, the global change for the whole time period. Figure 11, left, shows a convergence plot for this transient L-shape problem in comparison to an implicit Euler formulation and a midpoint rule. Therefore, we calculated an overkill solution using 25000 time-steps of an implicit Euler scheme and use the norm

$$error = \|[q_1, \dots, q_{n_{node}}] - [q_1^{\text{overkill}}, \dots, q_{n_{node}}^{\text{overkill}}]\|, \quad (108)$$

as error indicator. Note that a specific numerical error in the overkill solution remains independent of the time-step size, since the stop criterion of the Newton-Raphson iteration is  $1 \times 10^{-7}$  and the total error accumulates over all time steps. Both, the implicit time stepping scheme as well as the space-time formulation clearly converge to the same solution. Moreover, it is easy to see that, as expected, the implicit Euler is of first order accuracy, whereas the linear shape function in the temporal direction of the space-time formulation provide a quadratic term and hence, we obtain second order accuracy. Moreover, the absolute error is tremendously lower compared to traditional time-stepping methods. Overall, the superiority of the space-time formulation is obvious.

Eventually, in Figure 11, right, the complex eigenvalues of the mechanical system using 30 time slices are plotted.

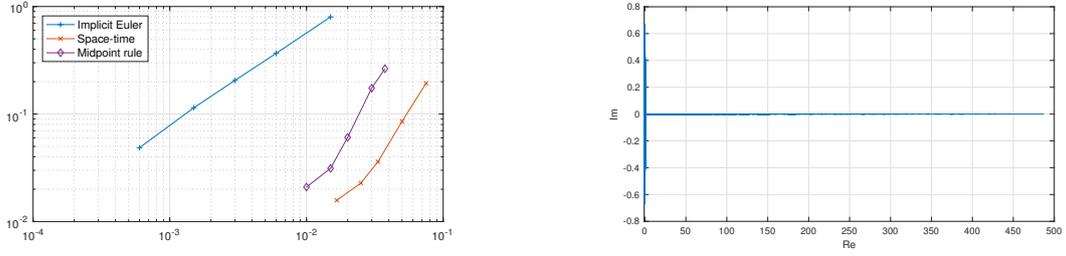


Figure 11: Left: convergence plot, *error* is plotted over  $\Delta t$ ; right: eigenvalues of the discrete mechanical system.

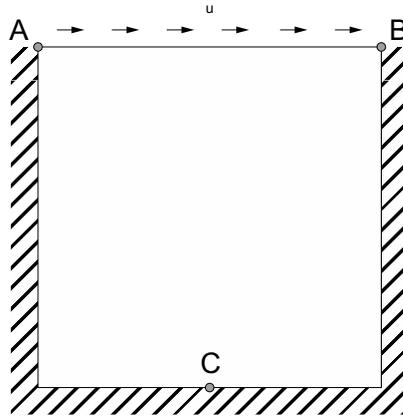


Figure 12: Problem set up of lid driven cavity flow.

### 8.3 Fluids

In this section we consider a common benchmark in the context of fluid problems. In order to demonstrate the proposed space-time formulation, the lid driven cavity flow is simulated using a Taylor-Hood Q2Q1 velocity-pressure space-time elements. The geometry of size  $[0, 1] \times [0, 1]$  and  $[0, 10]$  in temporal direction as well as the boundary

conditions of the problem are depicted in Figure 12. The initial velocity is set to zero on the boundary  $\partial\tilde{\mathcal{B}}^{t_1}$  except at the lid, where the horizontal velocity is prescribed as  $\mathbf{v}(\mathbf{y}) = [1, 0] \in \partial\tilde{\mathcal{B}}^v$  for all time  $t \in [0, 10]$ , including the edge nodes A and B („leaky cavity“). On all remaining boundaries the velocity is set to zero throughout time. The width and the height of the cavity are  $b = h = 1$  and the density of the fluid is set to  $\rho = 1$ . The fluid viscosity is set to  $\mu = 0.01$ , resulting in a global Reynold number of  $Re = 100$ .

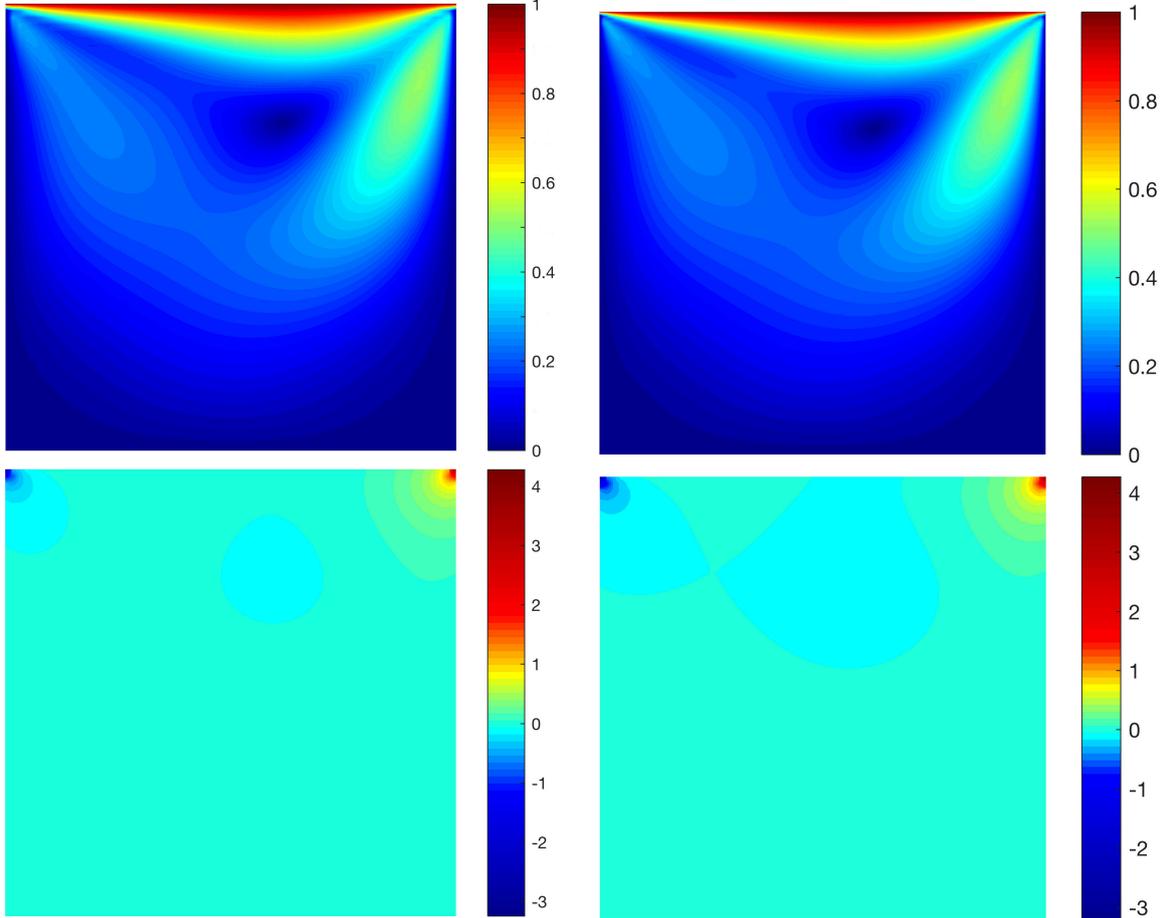


Figure 13: Results of the flow problem. Up: Norm of the velocity, below: pressure field. Left: space-time formulation, right: classical time stepping scheme.

Figure 13 displays the results of the space-time formulation and the conventional time marching scheme, respectively. In particular, the upper row shows the norm of the velocity, whereas the lower row shows the pressure field. All examples have been calculated using  $64 \times 64$  elements in spatial direction. In temporal direction, only four element rows for the space-time formulation have been used, whereas 100 time-steps were used to calculate the results using an implicit Euler scheme. As can be seen, the results match nearly perfectly. In Figure 14, the flow field at time  $t = [0.25, 5, 10]$  is plotted for the space time formulation.

Overall, we can state that the space-time formulation is suitable for the calculation of fluid systems using finite elements. The usage of finite volume or similar discretisation techniques may be considered for space-time as well, a proof of concept for the more complex finite element formulation is given here.

### 8.3.1 Multigrid performance for fluid system

The application of multigrid solvers to the Navier-Stokes problem allows to test their performances in case of non-linear saddle-point problems. For this latter, convergence theory is still a challenge. One of the standard choices in order to improve the convergence rate for vector and saddle-point problems is to employ point-block smoothers [28], which are based on the idea of performing a Gauß-Seidel or Jacobi step together for the degrees of freedom associated to a mesh node. To have the same number of degree of freedom per node, we employ a P1-P1 discretisation stabilised with pressure space Laplacian of the form

$$-\int_{\tilde{\mathcal{B}}} \alpha \Delta x^2 \nabla_x p \cdot \nabla_x \delta p \, d\mathbf{y}.$$

This stabilization presented in [12] is consistent and allows to remove spurious modes. It does not affect the convergence order both when linear finite element or Taylor-Hood schemes are employed. For Stokes problem, an optimal value of  $\alpha$  is the inverse of the Reynolds number [41]. Differently from the artificial diffusion term, this stabilization depending on  $\Delta x^2$  is not mesh-independent. Similar observation has been done in the study of similar saddle-point problems, such as poroelasticity [21, 22].

In the solution of the non-linear problem, we employ an inexact Newton's method, for which each non-linear step is solved with an relative tolerance of  $10^{-5}$ . The global Newton's procedure is solved with a relative tolerance of  $10^{-12}$ . For the solution of the linear solver, we employ a GMRES method, preconditioned with a semi-geometric multigrid for which 3 pre- and 3 post-smoothing point-block Gauß-Seidel steps are performed.

The role of the presented simulations is to study the effect of the different stabilizations on the convergence of multigrid. While the role of  $\gamma$  is well-understood, the choice of  $\alpha$  depends on the coarsest mesh. This latter also influence the properties of the point-block matrices, hence affecting the smoothing properties of iterative solvers.

Convergence studies of semi-geometric multigrid is reported in Tables 4 and 5 for two different meshes. It can be observed that for  $\alpha$  larger than a certain value, the overall number of Newton iterations and V-cycle is constant. As for the thermal conduction problem, the number of multigrid steps in order to converge decreases increasing the time-stabilization parameter.

In Figure 15, convergence tests are reported for different mesh sizes. In the first step, where a pure Stokes problem is solved, optimality of multigrid solver can be observed. From the second iteration, we can observe that for finer meshes a better convergence rate is obtained.

$\gamma \backslash \alpha$	1.6	3.2	6.4	12.8	25.6	51.2	102.4
0.02	N.C.	23 (5)	22 (5)	22 (5)	22 (5)	22 (5)	22 (5)
0.04	N.C.	23 (5)	22 (5)	22 (5)	22 (5)	22 (5)	22 (5)
0.08	N.C.	23 (5)	22 (5)	22 (5)	22 (5)	22 (5)	22 (5)
0.16	N.C.	23 (5)	22 (5)	22 (5)	22 (5)	22 (5)	22 (5)
0.32	N.C.	23 (5)	22 (5)	22 (5)	22 (5)	22 (5)	22 (5)
0.64	N.C.	23 (5)	22 (5)	22 (5)	22 (5)	22 (5)	22 (5)
1.28	N.C.	22 (5)	21 (5)	20 (5)	20 (5)	20 (5)	20 (5)
2.56	N.C.	20 (5)	20 (5)	20 (5)	20 (5)	16 (4)	16 (4)
5.12	N.C.	16 (4)	16 (4)	16 (4)	16 (4)	16 (4)	16 (4)

Table 4: Convergence rate for different time and space stabilisation parameters for a  $32 \times 32 \times 50$  mesh.

$\gamma \backslash \alpha$	1.6	3.2	6.4	12.8	25.6	51.2	102.4
0.02	30 (5)	36 (5)	40 (5)	44 (5)	46 (5)	47 (5)	48 (5)
0.04	30 (5)	36 (5)	40 (5)	44 (5)	46 (5)	47 (5)	48 (5)
0.08	30 (5)	36 (5)	40 (5)	44 (5)	46 (5)	47 (5)	48 (5)
0.16	30 (5)	36 (5)	40 (5)	44 (5)	46 (5)	47 (5)	48 (5)
0.32	30 (5)	36 (5)	40 (5)	44 (5)	46 (5)	47 (5)	48 (5)
0.64	30 (5)	36 (5)	40 (5)	44 (5)	46 (5)	47 (5)	48 (5)
1.28	28 (5)	30 (5)	33 (5)	34 (5)	34 (5)	34 (5)	34 (5)
2.56	27 (5)	29 (5)	31 (5)	31 (5)	25 (4)	25 (4)	25 (4)
5.12	20 (4)	22 (4)	23 (4)	23 (4)	23 (4)	23 (4)	22 (4)

Table 5: Convergence rate for different time and space stabilisation parameters for a  $64 \times 64 \times 100$  mesh.

## 8.4 Kuramoto Sivashinsky problem

In this last section we investigate the (2+1) dimensional version of the Kuramoto-Sivashinsky problem on a geometry of size  $[0, 100]^2$  and a temporal domain  $[0, 115]$ . The computational mesh consists of  $64 \times 64 \times 200$  quadratic B-spline elements. Following the investigations in Gomez and Paris [30] periodic conditions along with a uniform random perturbation between  $[-0.05, 0.05]$  of the homogeneous field  $u = 0$  as initial configuration (see Figure 16) are applied to the boundaries. Moreover, the constitutive parameter setting is given by  $\nu = 1$  and  $\alpha = 0.195$  which corresponds to the so called disordered state of the problem and shows a chaotic behavior of the solution, see Paniconi and Elder [57].

The solutions obtained via the novel space-time approach and a classical midpoint time integration scheme are compared in Figure 17. Although the system is chaotic, we expect the same results for an identical initial random field. Obviously, this is given in Figure 17. Note that we used quadratic NURBS functions in space and time and thus, this example demonstrates additionally the application of higher-order shape functions in time as well.

## 9 Conclusions

We have introduced a new framework for the numerical analysis of general transient continuum problems. The proposed formulation leads to a massive large-scale system with all degrees of freedom over all time-steps to be solved simultaneously. In contrast to classical time-stepping schemes, which decompose the temporal direction artificially in time steps, a more efficient domain-decomposition can be applied by modern multigrid solver, which can now deploy their full potential and capabilities.

The proposed formulation conserves the necessary constants of motion and provide a high level of accuracy, as shown by several examples. The applicability to the presented wide range of examples demonstrates also the generality of the proposed formulation. Moreover, this approach allows for refinement simultaneously in space and time. This leads to new opportunities for research, e.g. for contact problems, where we will be able to resolve not only the contact area, we can directly refine the space-time in a contact event. Further investigations on phase-field methods to brittle and ductile fracture can be considered as well, where we will be able to refine the fracturing event directly.

## Acknowledgements

Support for this research was provided by the Deutsche Forschungsgemeinschaft (DFG) under grant HE5943/6-1 and HE5943/8-1. The authors C. Hesch and R. Krause grate-

fully acknowledge support by the DFG in the framework of the Priority Programme „Reliable Simulation Techniques in Solid Mechanics. Development of Non-standard Discretisation Methods, Mechanical and Mathematical Analysis“ (SPP 1748) under projects HE5943/5-1 and 200021E-154090 (subproject Large-scale simulation of pneumatic and hydraulic fracture with a phase-field approach). The authors Rolf Krause and Marco Favino acknowledge the support by the Swiss Commission for Technology and Innovation via the SCCER-SoE and by the Swiss National Science Foundation, via the projects 200021E-145271 ExaSolvers - Extreme Scale Solvers for Coupled Systems and 200021L-144298 Parallel multilevel solvers for coupled interface problems. The authors also thank Alena Kopanicakova for her help on the simulations.

## References

- [1] D. Anders, C. Hesch, and K. Weinberg. Computational modeling of phase separation and coarsening in solder alloys. *Int. J. Solids Structures*, 49:1557–1572, 2012.
- [2] M. Anderson and J.-H. Kimn. A numerical approach to space-time finite elements for the wave equation. *J. Comput. Phys.*, 226:466–4767, 2007.
- [3] A.K. Aziz and P. Monk. Continuous finite elements in space and time for the heat equation. *Math. Comput.*, 52:255–274, 1989.
- [4] C. Bailey. Hamilton, Ritz, and elastodynamics. *Journal of Applied Mechanics*, 43:684–688, 1976.
- [5] Pietro Benedusi, Daniel Hupp, Peter Arbenz, and Rolf Krause. A Parallel Multigrid Solver for Time-Periodic Incompressible Navier-Stokes Equations in 3d. In *Numerical Mathematics and Advanced Applications - ENUMATH 2015*, Ankara, Turkey, 2016. Springer. bibtex: benedusi\_parallel\_2016.
- [6] P. Betsch and P. Steinmann. Inherently energy conserving time finite elements for classical mechanics. *Journal of Computational Physics*, 160:88–116, 2000.
- [7] P. Betsch and P. Steinmann. Conserving properties of a time FE method - Part II: Time-stepping schemes for non-linear elastodynamics. *Int. J. Numer. Methods Eng.*, 50:1931–1955, 2001.
- [8] P. Betsch and P. Steinmann. Conservation Properties of a Time FE Method. Part III: Mechanical systems with holonomic constraints. *Int. J. Numer. Methods Eng.*, 53:2271–2304, 2002.
- [9] Jürgen Bey, Arnold Reusken, and Stephan Dahlke. On the convergence of basic iterative methods for convection-diffusion equations. *Numerical Linear Algebra with Applications*, 1(1):1–7, 1993.

- [10] Jürgen Bey and Gabriel Wittum. Downwind numbering: Robust multigrid for convection-diffusion problems. *Applied Numerical Mathematics*, 23(1):177–192, 1997.
- [11] M. Borri. Helicopter rotor dynamics by finite element time approximation. *Computers & mathematics with applications*, 12:149–160, 1986.
- [12] Franco Brezzi and Juhani Pitkäranta. On the stabilization of finite element approximations of the stokes equations. In *Efficient solutions of elliptic systems*, pages 11–19. Springer, 1984.
- [13] R. Courant. Variational methods for the solution of problems of equilibrium and vibration. *American Math Society*, 49:1–61, 1943.
- [14] T. Dickopf and R. Krause. A pseudo-L2-projection for multilevel methods based on non-nested meshes. *ICS Preprint*, 2009.
- [15] T. Dickopf and R. Krause. A study of prolongation operators between non-nested meshes. *ICS Preprint*, 2009.
- [16] Thomas Dickopf and Rolf Krause. Numerical Study of the Almost Nested Case in a Multilevel Method Based on Non-nested Meshes. In Randolph Bank, Michael Holst, Olof Widlund, and Jinchao Xu, editors, *Domain Decomposition Methods in Science and Engineering XX*, volume 91, pages 551–558, Berlin, Heidelberg, 2013. Springer Berlin Heidelberg. 00000.
- [17] M. Dittmann. *Isogeometric analysis and hierarchical refinement for multi-field contact problems*. PhD thesis, University of Siegen, 2017.
- [18] J. Donea and A. Huerta. *Finite element methods for flow problems*. John Wiley & Sons, 2003.
- [19] Matthew Emmett and Michael L. Minion. Toward an Efficient Parallel in Time Method for Partial Differential Equations. *Communications in Applied Mathematics and Computational Science*, 7:105–132, 2012.
- [20] R. D. Falgout, S. Friedhoff, Tz. V. Kolev, Scott P. MacLachlan, and Jacob B. Schroder. Parallel time integration with multigrid. *SIAM Journal on Scientific Computing*, 36(6):C635–C661, 2014. bibtex: FalgoutEtAl2014\_MGRIT.
- [21] Marco Favino, Alfio Grillo, and Rolf Krause. A stability condition for the numerical simulation of poroelastic systems. In *Poromechanics V: Proceedings of the Fifth Biot Conference on Poromechanics*, pages 919–928, 2013.
- [22] Marco Favino, Jürg Hunziker, Klaus Holliger, and Rolf Krause. An accuracy condition for the finite element discretization of biot’s equations on triangular meshes. In *Poromechanics VI*, pages 172–181. 2017.

- [23] Leopoldo P Franca and Charbel Farhat. Bubble functions prompt unusual stabilized finite element methods. *Computer Methods in Applied Mechanics and Engineering*, 123(1-4):299–308, 1995.
- [24] D.A. French and T.E. Peterson. A continuous space-time finite element method for the wave equation. *Math. Comput.*, 65:491–506, 1996.
- [25] Martin J. Gander and Martin Neumüller. Analysis of a New Space-Time Parallel Multigrid Algorithm for Parabolic Problems. *SIAM Journal on Scientific Computing*, 38(4):A2173–A2208, January 2016. 00008.
- [26] Martin J. Gander and Martin Neumüller. Analysis of a new space-time parallel multigrid algorithm for parabolic problems. *SIAM Journal on Scientific Computing*, 38(4):A2173–A2208, 2016.
- [27] M.J. Gander. 50 years of Time Parallel Time Integration. In *Multiple Shooting and Time Domain Decomposition*. Springer, 2015.
- [28] Francisco José Gaspar, Francisco Javier Lisbona, Cornelis W Oosterlee, and Roman Wienands. A systematic comparison of coupled and distributive smoothing in multigrid for the poroelasticity system. *Numerical linear algebra with applications*, 11(2-3):93–113, 2004.
- [29] S. Gellin and J.M. Pitarresi. Nonlinear analysis using temporal finite elements. *Engineering analysis*, 5:126–132, 1988.
- [30] H. Gomez and J. Paris. Numerical Simulation of Asymptotic States of the Damped Kuramoto-Sivashinsky Equation. *Phys. Rev. E*, 93:046702, 2011.
- [31] W. Hackbusch. Parabolic multi-grid methods. *Computing Methods in Applied Sciences and Engineering*, VI, pages 189–197, 1984.
- [32] W. Hackbusch. *Multi-Grid Methods and Applications*. Springer-Verlag, 1985.
- [33] C. Hesch and P. Betsch. Energy-momentum consistent algorithms for dynamic thermomechanical problems - Application to mortar domain decomposition problems. *Int. J. Numer. Methods Eng.*, 86:1277–1302, 2011.
- [34] C. Hesch, S. Schuß, M. Dittmann, M. Franke, and K. Weinberg. Isogeometric analysis and hierarchical refinement for higher-order phase-field models. *Computer Methods in Applied Mechanics and Engineering*, 303:185–207, 2016.
- [35] H.M. Hilbert, T.J.R. Hughes, and R.L. Taylor. Improved Numerical Dissipation for Time Integration Algorithms in Structural Dynamics. *Earthquake Engineering and Structural Dynamics*, pages 282–292, 1977.
- [36] G.A. Holzapfel and J.C. Simo. Entropy elasticity of isotropic rubber-like solids at finite strains. *Comput. Methods Appl. Mech. Engrg.*, 132:17–44, 1996.
- [37] G. Horton and S. Vandewalle. A space time multigrid method for parabolic partial differential equations. *SIAM J. Sci. Comput.*, 16:848–864, 1995.

- [38] Thomas JR Hughes. A simple scheme for developing upwind finite elements. *International journal for numerical methods in engineering*, 12(9):1359–1365, 1978.
- [39] Thomas JR Hughes. Multiscale phenomena: Green’s functions, the dirichlet-to-neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods. *Computer methods in applied mechanics and engineering*, 127(1-4):387–401, 1995.
- [40] T.J.R. Hughes, J.A. Cottrell, and Y. Bazilevs. Isogeometric analysis: CAD, finite elements, NURBS, exact geometry and mesh refinement. *Comput. Methods Appl. Mech. Engrg.*, 194:4135–4195, 2005.
- [41] T.J.R. Hughes, L.P. Franca, and M. Balestra. A new finite element formulation for computational fluid dynamics: V. Circumventing the Babuška-Brezzi condition: A stable Petrov-Galerkin formulation of the Stokes problem accommodating equal-order interpolations. *Comput. Methods Appl. Mech. Engrg.*, 59:85–99, 1986.
- [42] T.J.R. Hughes, L.P. Franca, and G.M. Hulbert. A new finite element formulation for computational fluid dynamics: VIII. The Galerkin/least-squares method for advective-diffusive equations. *Comput. Methods Appl. Mech. Engrg.*, 73:173–189, 1089.
- [43] T.J.R. Hughes and G. M. Hulbert. Space-time finite element methods for elastodynamics: Formulations and error estimates. *Comput. Methods Appl. Mech. Engrg.*, 66:339–363, 1988.
- [44] A.P. Izadpanah. *p-version finite elements for the space-time domain with application to Floquet theory*. PhD thesis, Georgia Institute of Technology, 1986.
- [45] P. Jamet. Galerkin-Type Approximations which are Discontinuous in Time for Parabolic Equations in a Variable Domain. *SIAM J. Numer. Anal.*, 15:912–928, 1978.
- [46] C.M. Klaij, J.J.W. van der Vegt, and H. van der Ven. Space-time discontinuous galerkin method for the compressible navier-stokes equations. *Journal of Computational Physics*, 217(2):589 – 611, 2006.
- [47] D. Krause and R. Krause. Enabling local time stepping in the parallel implicit solution of reaction-diffusion equations via space-time finite elements on shallow tree meshes. *Applied Mathematics and Computation*, 277, March 2016.
- [48] R. Krause and P. Zulian. A Parallel Approach to the Variational Transfer of Discrete Fields between Arbitrarily Distributed Unstructured Finite Element Meshes. *SIAM Journal on Scientific Computing*, 38(3):C307–C333, 2016. 00000.
- [49] O.A. Ladyzhenskaya, V.A. Solonnikov, and N.N. Uraltseva. *Linear and quasilinear equations of parabolic type*. American Mathematical Society, Providence, R.I., 1968.

- [50] Sander Land, Viatcheslav Gurev, Sander Arens, Christoph M. Augustin, Lukas Baron, Robert Blake, Chris Bradley, Sebastian Castro, Andrew Crozier, Marco Favino, Thomas E. Fastl, Thomas Fritz, Hao Gao, Alessio Gizzi, Boyce E. Griffith, Daniel E. Hurtado, Rolf Krause, Xiaoyu Luo, Martyn P. Nash, Simone Pezzuto, Gernot Plank, Simone Rossi, Daniel Ruprecht, Gunnar Seemann, Nicolas P. Smith, Joakim Sundnes, J. Jeremy Rice, Natalia Trayanova, Dafang Wang, Zhinuo Jenny Wang, and Steven A. Niederer. Verification of cardiac mechanics software: benchmark problems and solutions for testing active and passive material behaviour. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 471(2184), 2015.
- [51] Christoph Lehrenfeld and Arnold Reusken. Finite Element Techniques for the Numerical Simulation of Two-Phase Flows with Mass Transport. In Mohammad Rahni, Mohsen Karbaschi, and Reinhard Miller, editors, *Computational Methods for Complex Liquid-Fluid Interfaces*, Progress in Colloid and Interface Science, pages 353–370. CRC Press, December 2015. DOI: 10.1201/b19337.
- [52] B.J. Leimkuhler and S. Reich. *Simulating Hamiltonian Dynamics*. Cambridge University Press, 2004.
- [53] A. Lew, J.E. Marsden, M. Ortiz, and M. West. Variational time integrators. *Int. J. Numer. Methods Eng.*, 60:153–212, 2004.
- [54] A new look at finite elements in time: a variational interpretation of Runge-Kutta methods. Bottasso, C.L. *Applied Numerical Mathematics*, 25:355–368, 1997.
- [55] J. Nievergelt. Parallel methods for integrating ordinary differential equations. *Commun. ACM*, 7(12):731–733, 1964.
- [56] CW Oosterlee, FJ Gaspar, T Washio, and R Wienands. Fast multigrid solvers for higher order upwind discretizations of convection-dominated problems. In *Multigrid Methods V*, pages 212–224. Springer, 1998.
- [57] M. Paniconi and K.R. Elder. Stationary, dynamical, and chaotic states of the two-dimensional damped Kuramoto-Sivashinsky equation. *Phys. Rev. E*, 56:2713–2721, 1997.
- [58] Alfio Quarteroni. *Numerical models for differential problems*, volume 2. Springer Science & Business Media, 2010.
- [59] Donald L Scharfetter and Hermann K Gummel. Large-signal analysis of a silicon read diode oscillator. *IEEE Transactions on electron devices*, 16(1):64–77, 1969.
- [60] F. Shakib and T.J.R. Hughes. A new finite element formulation for computational fluid dynamics: IX. Fourier analysis of space-time Galerkin/least-squares algorithms. *Comput. Methods Appl. Mech. Engrg.*, 87:35–58, 1991.

- [61] Robert Speck, Daniel Ruprecht, Matthew Emmett, Michael Minion, Matthias Bolten, and Rolf Krause. A multi-level spectral deferred correction method. *BIT. Numerical Mathematics*, 55(3):843–867, 2015. bibtex: Speck:2015 bibtex[mrclass=65M06 (65Y05);mrreviewer=BÄijlent KarasÄűzen].
- [62] H. Stolarski and T. Belytschko. On the equivalence of mode decomposition and mixed finite elements based on the Hellinger-Reissner principle. Part I: Theory. *Comput. Methods Appl. Mech. Engrg.*, 58:249–263, 1986.
- [63] H. Stolarski and T. Belytschko. On the equivalence of mode decomposition and mixed finite elements based on the Hellinger-Reissner principle. Part II: Applications. *Comput. Methods Appl. Mech. Engrg.*, 58:265–284, 1986.
- [64] The temporal finite element method in structural dynamics. Pitarresi, J. and Manolis, G. *Computers & structures*, 41:647–655, 1991.
- [65] T.E. Tezduyar. Stabilized Finite Element Formulations for Incompressible Flow Computations. *Advances in applied mechanics*, 28:1–44, 1992.
- [66] Space time mixed finite elements for rods. Atilgan, A. and Hodges, D. and Özbek, M. and Zhou, W. *Journal of sound and vibration*, 192:731–739, 1996.
- [67] M.J. Turner, R.W. Clough, H.C. Martin, and L.J. Topp. Stiffness and Deflection Analysis of Complex Structures. *Journal of the Aeronautical Sciences*, 23:805–823, 1956.
- [68] Jinchao Xu and Ludmil Zikatanov. A monotone finite element scheme for convection-diffusion equations. *Mathematics of Computation of the American Mathematical Society*, 68(228):1429–1446, 1999.
- [69] O.C. Zienkiewicz, S. Qu, R.L. Taylor, and S. Nakazawa. The patch test for mixed formulations. *Int. J. Numer. Methods Eng.*, 23:1873–1883, 1986.

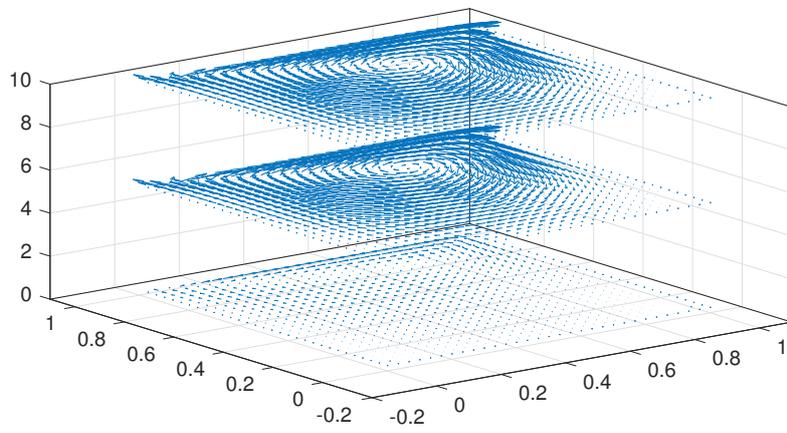


Figure 14: Flow field at time  $t = [0.25, 5, 10]$ .

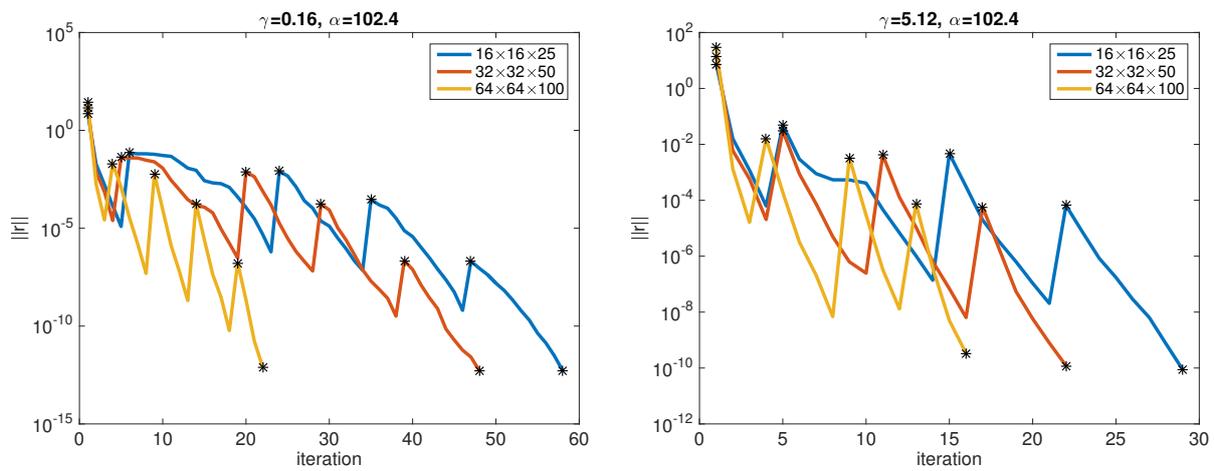


Figure 15: Convergence for different time-stabilizations.

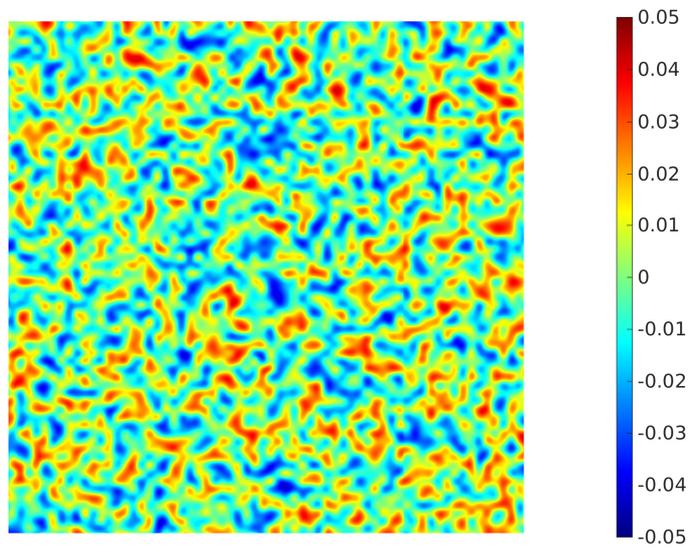


Figure 16: Initial configuration of the Kuramoto-Sivashinsky problem.

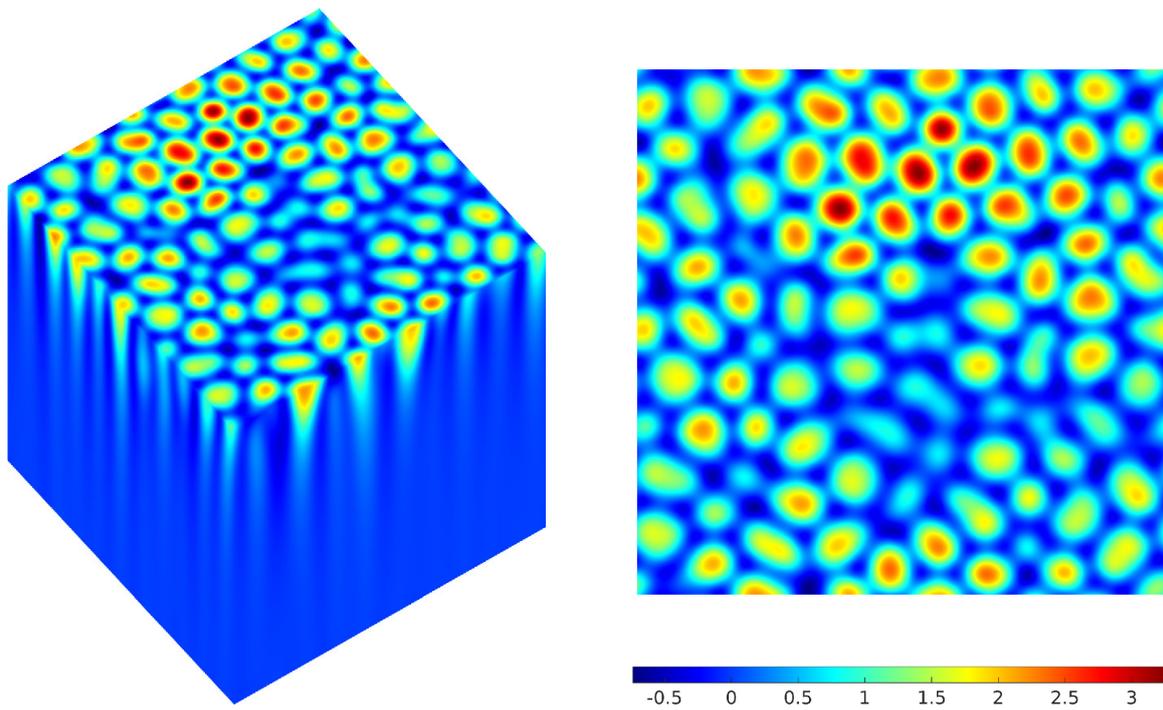


Figure 17: Solution of the Kuramoto-Sivashinsky problem. Left: result of the space-time approach, right: result of a classical time stepping scheme.