Simulation and Control of Two-Phase Flow Using Diffuse-Interface Models

Dissertation
zur Erlangung des Doktorgrades
der Fakultät für Mathematik, Informatik
und Naturwissenschaften
der Universität Hamburg

vorgelegt
im
Fachbereich Mathematik

von
Christian Kahle
aus Winsen/Luhe

Hamburg
2014

Als Disseration angenommen vom Fachbereich
Mathematik der Universität Hamburg

auf Grund der Gutachten von	Prof. Dr. Michael Hinze
und	Prof. Dr. L'ubomír Baňas

Hamburg, den 29.10.2014

Prof. Dr. Michael Hinze
Leiter des Fachbereichs Mathematik

Ich möchte mich bei Herrn Prof. Dr. Michael Hinze für die Betreuung und seine umfangreiche Unterstützung bei der Anfertigung dieser Dissertation bedanken, so wie für die langjährige Förderung und Zusammenarbeit.

Bei meinen Kollegen aus der Arbeitsgruppe von Professor Hinze und dem Fachbereich Mathematik der Universität Hamburg möchte ich mich für die sehr angenehme Arbeitsatmosphäre bedanken.

Des Weiteren bedanke ich mich für die Unterstützung des Fachbereichs Mathematik der Universität Hamburg und der Lothar–Collatz School for Computing in Science.

Liebe Julia – Danke.

Contents

Introduction

The present work concerns simulation and closed-loop control of two-phase flows.

The simulation of two-phase fluids and multi-phase fluids has attained a growing interest in the last decades. The treatment of multi-phase flows is involved for several reasons. The phases are separated by fluid-fluid interfaces, whose evolution has to be tracked. Furthermore, processes on the interface, which depend on the fluids involved, influence the evolution of the free boundaries, i.e. the fluid-fluid interfaces. Here by a phase we mean a distinct fluid that is not mixed with other distinct fluids inside the fluid domain, and in the present work we deal with the case of two phase flow.

There exist several approaches for modelling two-phase flow. In the so called 'sharp-interface' approach the interface between two phases is modelled as a lower dimensional manifold. Thus the thickness of the interface is assumed to vanish. A commonly used model is presented in [GR11, 1.1.2]. It consists of two Navier–Stokes equations in the two domains of pure fluid and containes coupling conditions at the interface between the two fluid domains.

Approaches for the numerical solution of these equations differ in the way how the unknown interface is represented. In interface tracking approaches the unknown interface is explicitly discretized and its evolution is tracked through the simulation. In volume tracking approaches the interface is implicitly described by the zero-level-set of an appropriate level-set-function and an additional equation for the evolution of this level-set-function is derived. There exist many well developed codes implementing these approaches, see e.g. [HTK$^+$09] and [GR11].

On the other hand, topology changes like breakup or coalescence of interfaces have to be carefully captured by the model. For these phenomena phase field models allow for a natural description of the topology changes in the model. Such models assume that the fluid-fluid interface has a small positive thickness and that the two phases are mixed inside this region. This is also called 'diffuse-interface' approach.

One of the first models using this approach is the so-called Cahn–Hilliard model ([CH58]) that describes phase separation of a binary fluid. It contains a free energy that yields the separating effect. For a specific choice of free energy, the 'double-obstacle' free energy, the model contains a variational inequality. This free energy is first analytically investigated by Blowey and Elliot in [BE91] and the existence of solutions is shown. We also note the review [Ell89] for a deviation of the model and an overview on analytical results and generalisations with a more general free energy. Concerning approaches for the numerical solution of the Cahn–Hilliard system with double-obstacle free energy we note the publications [BE92, BBG99, GK07, BN09, BBG11, HHT11]. In [BE92] a discretization scheme for the solution of the Cahn–Hilliard system is provided, that preserves the variational inequality in the fully discrete setting. The authors provide stability, convergence and error bounds for the fully discrete scheme. The fully discrete variational inquality is treated iteratively by an active-set strategy. In [BBG99] the case of degenerate mobility is investigated. Here degenerate mobility means, that some diffusion is restricted to the interface. The authors ensure properties of the solution arising from the degenerate mobility by a variational inequality, and they present a discrete scheme for which they provide well-posedness and stability results. They show numerical simulations based on a splitting scheme for the variational inequality and a discrete cosine transformation on homogeneous meshes. In [GK07] an Uzawa iteration for a fully discretized Cahn–Hilliard system together with a multigrid preconditioner on the interface is proposed. The variational inequality is treated by monotone multigrid. In [BN09] an Uzawa multigrid iteration is used for the numerical solution of the Cahn–Hilliard system and reliable a posteriori error estimation is performed. In [BBG11] the authors reformulate the Cahn–Hilliard equation as an optimization problem and use a primal dual active set method for the solution of the resulting optimality system in the discrete setting. In [HHT11] the variational inequality is relaxed using Moreau–Yosida regularization and a semi-smooth Newton method in function space is applied for the numerical treatment of the relaxed Cahn–Hilliard system. That paper also contains a convergence analysis of the relaxed solutions to the solution of the variational inequality, as well as a reliable and efficient error estimator for the error in the discrete solution.

The Cahn–Hilliard model only encorporates diffusion for the transport of the particles. If additional advection occures, extensions of the model have to be used. We refer to the review [AMW98] for an overview of available diffuse interface models for hydrodynamics. In the case of fluids with the same density a commonly used model is the so called model 'H' ([HH77]). It couples the Cahn–Hilliard system for the description of the two-phase structure to a Navier–Stokes system for the description of the velocity field. It fulfills energy inequalities and thus is thermodynamically consistent, see e.g. [Abe07]. Concerning results on the existence of solutions we refer to [Boy99, Abe07]. For a convergent finite element scheme for this model we refer to [KSW08]. The authors solve linear systems arising through the simulation by combining the multigrid solver for the Cahn–Hilliard system developed in [KW06] and the preconditioning technique for Navier–Stokes equations presented in [KLW02].

The model 'H' copes only with fluids of the same density. To overcome this restriction several new models are developed in the last decades. We note the models presented in [LT98, Boy02, DSS07, AGG12]. The model presented in [LT98] is termodynamically consistent but leads to a velocity field, that is not solenoidal and the model contains a strong coupling of the Cahn–Hilliard and the Navier–Stokes part. In [KKL04] a multigrid solver for this model is proposed. The model presented in [Boy02] contains a solenoidal velocity field, but is not consistent with thermodynamics, while the same holds for the model presented in [DSS07]. For the latter model this can be overcome by redefining the kinetic energy, see [SY10]. Recently, in [AGG12], a thermodynamically consistent model is proposed that contains a solenoidal velocity field. Most couplings between the Cahn–Hilliard and the Navier–Stokes part of the model are well known from the model 'H'. Concerning existence of solutions for the model presented in [AGG12] with different assumptions on the free energy and data we refer to [ADG13a, ADG13b, Grü13]. A stable discretization concept is proposed for the numerical simulation of this model in [GK14]. In [AV12] the models [Boy02, DSS07, AGG12] are numerically compared by running the rising bubble benchmark [HTK$^+$09]. The results of the simulations are also compared to sharp-interface numeric.

In Part A of this thesis we extend the results from [HHT11] to model 'H'. We present a time discretization that is similiar to the scheme proposed in [KSW08] and that yields a sequential coupling of the Cahn–Hilliard and the Navier–Stokes system. We use Moreau–Yosida relaxation to treat the variational inequality introduced through the double-obstacle free energy and propose a semi-smooth Newton solver in function space. Reliable and efficient error estimation is presented both for the Cahn–Hilliard and the Navier–Stokes part, where the spatial discretization of the Cahn–Hilliard equation and the Navier–Stokes equation are nearly independent.

We further apply this approach to the model [AGG12] to pass the rising bubble benchmark [HTK$^+$09] and we present a stable discretization scheme for model [AGG12] that preserves the consistence with thermodynamics in the fully discrete setting.

In the last decades there was notable progress on the mathematical theory of control of one-phase flows, see [Bar11, NMT11]. Here we consider the case of feedback control or closed-loop control, and in particular we are concerned with the concept of 'model predictive control', see [GP11], which is also known as receding horizon control. Here the control feedback is obtained from open-loop control over short time horizons. This open-loop control is based on an appropriate model of the underlying process. In [BMT01] and [Pro02] this concept is used to control 2D channel flow, and in [HM07] it is applied to the control of the Boussinesq approximation of the Navier–Stokes equation. In the latter publication also the 'instantaneous control' concept is investigated. The instantantaneous control approach is a variant of the model predictive control approach. Here the open-loop optimal control problems are only solved approximately, resulting in a faster evaluation of the feedback control law. For the control of flows this concept is applied in [Cho95, CHK99]. In [Hin05a] it is shown that instantaneous control is able to steer a velocity field towards a desired configuration exponentially fast.

In Part B of this thesis we apply model predictive feedback control to two-phase flow. We apply the instantaneous control concept to the system with different densities investigated in Part A and give a short outlook to the general model predictive control concept applied to the same model.

Notation

For convenience, we start with some notation we use throughout this work.

Let $\Omega \subset \mathbb{R}^d$, $d \in \{2, 3\}$, be an open bounded domain. Its boundary is denoted by $\partial\Omega$ and is assumed to be sufficiently smooth. The outer normal vector of unit length is denoted by ν_Ω. $I = (0, T]$ denotes a time interval.

By $L^p(\Omega)^d$ we denote the space of all measurable functions on Ω, whose modulus is Lebesgue-integrable up to the power of $p \geq 1$ with values in $\mathbb{R}^d$.

We use the conventional notation for Sobolev spaces. Especially we denote by $W^{k,p}(\Omega)^d$ the Sobolev space of all functions possessing weak derivatives up to order k in $L^p(\Omega)$. For $p = 2$ and $k \geq 1$ we write $H^k(\Omega)^d$ instead of $W^{k,2}(\Omega)^d$. For a subset $D \subset \Omega$ and $f, g \in L^2(\Omega)$ we define $(f, g)_D = \int_D f g \, dx$ with corresponding norm $\|f\|_D^2 = (f, f)_D$. For $f \in L^\infty(\Omega)$ we further define $\|f\|_{\infty,D} = \|f_{|D}\|_{L^\infty(\Omega)}$.

We frequently need the following subspaces

$$\begin{aligned} L^2_{(0)}(\Omega) &:= \{v \in L^2(\Omega) \,|\, \int_\Omega v \, dx = 0\}, \\ H^1_0(\Omega) &:= \{v \in H^1(\Omega) \,|\, v_{|\partial\Omega} = 0 \text{ in the sense of traces}\}, \\ H(\operatorname{div}, \Omega) &:= \{v \in H^1_0(\Omega)^d \,|\, (\operatorname{div}(v), q) = 0 \, \forall q \in L^2_{(0)}(\Omega)\}. \end{aligned}$$

The abbreviation ‘a.e.’ stands for ‘almost everywhere’.
As norm in $H^1_0(\Omega)$ we use $\|v\|^2_{H^1_0(\Omega)} = \|v\|^2_{L^2(\Omega)} + \|\nabla v\|^2_{L^2(\Omega)^d}$ and we recall the equivalence of this norm and the norm $\|v\|_* = \|\nabla v\|_{L^2(\Omega)^d}$ on $H^1_0(\Omega)$.

For a normed vector space V by V^* we denote its dual space, which is the set of all linear and continous functionals $A : V \to \mathbb{R}$.

For more details on Lebesgue and Sobolev Spaces and also their dual spaces we refer to [AF03].

Part A
Simulation of the Cahn–Hilliard Navier–Stokes system

In this main part of our work, we start with a description of the process we consider in Section 1. We briefly introduce the Cahn–Hilliard equation in Section 2. Thereafter we state the Cahn–Hilliard Navier–Stokes system with equal densities and double-obstacle free energy in Section 3. This part is devoted to the numerical treatment of this system. For this in Section 4 we state a time discretization scheme, that results in a sequential coupling of the Cahn–Hilliard system and an Oseen system, and show existence of time discrete solutions. Thereafter, in Section 5, we apply a relaxation method for the numerical treatment of the variational inequality involved, and show convergence of the solutions to the relaxed problem to the solutions of the original problem for vanishing relaxation. In Section 6 we show that the relaxed systems can be solved using semi-smooth Newton's method in function space, and in Section 7 we provide a fully discrete concept for the numerical treatment using the finite element method. The construction of spatial meshes is investigated in Section 8, where we develop reliable and efficient error estimators for the adaptation of the spatial meshes. In Section 9 the all-over concept is numerically investigated. In [HHT11] the presented concept is applied to the Cahn–Hilliard system without transport and especially a residual-based adaptive concept is developed. Here we extend these results to the case of the Cahn–Hilliard Navier–Stokes system, where we adapt results obtained for the Oseen system in [Jus11].

In Section 10 we extend our approach to the case of two-phase flow with different densities. In particular we develop a time discretization scheme preserving the sequential coupling exploited in the equal densitiy case. In Section 11 we present a new time discretization scheme giving rise to a discrete in time energy inequality that can also be conserved in the fully discrete setting, provided the adaptive concept is modified accordingly.

1 Problem description and solution concept

We consider a domain $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$, filled with a fluid consisting of two immiscible components A and B. Examples for such configurations can be given by butan or toluene in water ([GR11]) or air in glycerol ([ABH$^+$13]). Thus when investigating the flow inside this domain one in particular has to take care of the two-phase structure of the fluid. If the fluids are separated by a sharp interface, the numerical treatment of this situation has to cope equations for the fluid flow in both fluid domaines. Additionaly one also has to track the evolution of the interface which is transported by the flow and on the other hand also influences the flowfield.

A standard model for this situation for example is given in [GR11, 1.1.2]. It consist of two Navier–Stokes equations in the respective phases together with conditions on the stress tensor and the velocity across the interface. This results in a Navier–Stokes system with discontinuous coefficients. Numerical techniques for the treatment of two-phase flow based on a sharp-interface formulation, such as front tracking approaches and level set methods are described in e.g. [GR11, 6.2].

In the present work we use the diffuse interface approach, where the interface is assumed to be smeared out over a small region of width $0 < \gamma \ll 1$. This technique is already proposed in the 19th century by Rayleigh and van der Waals, see [Ray92, vanon]. The diffuse interface is described by an order parameter c which satisfies $c \equiv 1$ in the pure phase of fluid A and $c \equiv -1$ in the pure phase of fluid B. A transition layer between these so-called bulk phases is called diffuse interface. Instead of *order-parameter* we also use the termini *phase field* or *concentration.*

In Sections 3–9 of Part A we consider the solution of the diffuse interface model governed by a coupled Cahn–Hilliard Navier–Stokes system assuming equal densities (model ‘H’ [HH77]). In Sections 10 and 11 we then investigate a new Cahn–Hilliard Navier–Stokes model proposed in [AGG12] treating the case of different densities. This model we control in Part B of this work.

2 Brief description of the Cahn–Hilliard system

In [CH58] J. W. Cahn and J. E. Hilliard develope a model for describing spinodal decomposition of a binary alloy (see e.g. [FM08, Sig79]) using a diffuse interface approach. Spinodal decomposition is observed if e.g. the temperature of a homogeneous alloy of two metals with specific properties (e.g. Ag and Au in [EAK$^+$01]) is rapidly decreased. The mixture gets unstable and demixes into its two components.

We denote the two fluids involved by A and B and denote the corresponding concentrations by c_A and c_B. For describing the spatial distribution of the two phases with only one variable we introduce an order parameter c as

$$c = \frac{c_A - c_B}{c_A + c_B}.$$

This order parameter c fulfills $|c| \leq 1$, $c \equiv 1$ in the pure A-phase and $c \equiv -1$ in the pure B-phase. The transition region between these two phases is called diffuse interface and separates the two phases.

The Cahn–Hilliard equation is a fourth order partial differential equation describing the evolution of c in space and time, starting from some initial distribution c_0. It is convenient to split this equation into two equations of second order in space introducing a chemical potential w. The Cahn–Hilliard system for c and w then can be written as

$$\begin{aligned} c_t - \operatorname{div}(m(c)\nabla w) &= 0, \\ -\gamma^2 \Delta c - w + \Psi'(c) &= 0, \end{aligned}$$

with $c(0) = c_0$ and $\nabla c \cdot \nu_\Omega = \nabla w \cdot \nu_\Omega = 0$. The diffusion coefficient function $m(c) \geq 0$ is the so called mobility. It is typically set to a constant positive value, while the case of degenerate mobility, i.e. $m(\pm 1) \equiv 0$, for example was investigated in [EG96]. We also note a numerical comparison of simulations of phase separation with degenerate and non-degenerate mobility in [BNN13].

The free energy Ψ is a of double-well type, i.e. it has exactly two minima at $c = -1$ and $c = +1$. There are three common choices for Ψ. In the original work [CH58]

$$\Psi^{log}(c) = \frac{\theta}{2}\left((1+c)\log(1+c) + (1-c)\log(1-c)\right) - \frac{\theta_0}{2}c^2$$

is chosen, where log denotes the natural logarithm. This energy has two minima if and only if $\theta < \theta_0$ holds and one minimum if $\theta \geq \theta_0$ holds. Due to the obstacle structure of the logarithmic functions involved this potential prevents the concentration from reaching the pure states $c = \pm 1$. Frequently, the free energy Ψ^{log} is approximated by a polynomial of fourth order in the form

$$\Psi^{poly}(c) = \frac{1}{4}(1-c^2)^2.$$

The use of this free energy simplifies the Cahn–Hilliard system, but does not force the concentration to stay within its bounds, i.e. $|c| \leq 1$ can in general not be achieved.

A third choice of the free energy can be obtained by taking the limit $\theta \to 0$ in Ψ^{log} (see e.g. [Abe07, BE91]). This yields the so called double-obstacle free energy

$$\Psi^{obst}(c) = \begin{cases} \frac{1}{2}(1-c^2) & \text{if } c \in [-1,1], \\ \infty & \text{else.} \end{cases}$$

This free energy is proposed in [OP88] and is first investigated analytically in [BE91]. Due to its non-differentiability the Cahn–Hilliard system becomes

$$\begin{aligned} c_t - \operatorname{div}(m(c)\nabla w) &= 0, \\ -(\gamma^2 \nabla c, \nabla(v-c)) - (w, v-c) - (c, v-c) &\geq 0 \qquad \forall v \in H^1(\Omega),\ |v| \leq 1. \end{aligned}$$

In the following we use the double-obstacle free energy. Concerning the existence of solutions of the related Cahn–Hilliard system we refer to [BE91].

We note, that the Cahn–Hilliard system can be derived as a mass conserving gradient flow for minimizing the Ginzburg–Landau energy given by

$$E = \int_\Omega \frac{\gamma^2}{2}|\nabla c|^2 + \Psi(c)\,dx, \tag{2.1}$$

see [Ell89] and [BBG11].

3 The Cahn–Hilliard Navier–Stokes system

The model we use in the first part of this work for the simulation of two-phase flow is the so called model 'H' in the nomenclatura of Hohenberg and Holperin ([HH77]). It is able to cope with the complex interaction between the interface and the flow field. The interface influences the flow field through capillary forces while the flow field transports the interface. It might be regarded as a drawback of this model, that it assumes equal densities for both fluid components. In Sections 10 and 11 we consider a new model which is able to handle the case of different densities. However, since we intend to derive a numerical approach for the simulation of the two-phase fluid structure, we stick to the model 'H' to develop the numerical scheme and the solver components. The analysis presented in the following sections carries over to the model with different densities in a natural way.

The form of the model we consider here is taken from [KSW08]. In strong form it reads:

$$
\begin{aligned}
y_t - \eta\Delta y + (y\nabla)y + \nabla p &= -Kc\nabla w && \text{in } \Omega \times I, && (3.1)\\
\operatorname{div} y &= 0 && \text{in } \Omega \times I, && (3.2)\\
c_t - \frac{1}{\mathrm{Pe}}\Delta w + y\nabla c &= 0 && \text{in } \Omega \times I, && (3.3)\\
-\gamma^2\Delta c + \Psi'(c) &= w && \text{in } \Omega \times I, && (3.4)\\
\nabla c \cdot \nu_\Omega &= 0 && \text{on } \partial\Omega \times I, && (3.5)\\
\nabla w \cdot \nu_\Omega &= 0 && \text{on } \partial\Omega \times I, && (3.6)\\
y &= g && \text{on } \partial\Omega \times I, && (3.7)\\
y(0,x) &= y_0(x) && \text{in } \Omega, && (3.8)\\
c(0,x) &= c_0(x) && \text{in } \Omega. && (3.9)
\end{aligned}
$$

The flowfield is denoted by y, and p is the corresponding pressure. The phase field is denoted by c, and the chemical potential is denoted by w. The viscosity of the fluid is denoted by $\eta = 1/\mathrm{Re}$, where Re denotes the Reynold number, and the capillarity is given by $K := 1$. By Pe we denote the Péclet number of the fluid. It can be regarded as a mobility of the two-phase structure. The diffuse interface covers a region of width $0 < \gamma \ll 1$. For simplicity we use $g \equiv 0$. The initial concentration c_0 is chosen to satisfy $(c_0, 1) = 0$.

The function Ψ denotes the free energy of the system. Here we use the double-obstacle free energy as described in Section 2, and thus (3.4) abbreviates a variational inequality.

In [Abe07, Ch. 6.5] Abels shows the existence of unique weak solutions to (3.1)–(3.9) (at least for short time intervals) in two and three space dimensions. Both the case of the logarithmic and the double-obstacle free energy are investigated. Weak solutions in the sense of [Abe07] especially fulfill

$$\begin{aligned} y &\in BC_\omega\left(I, L^2_\sigma(\Omega)^d\right) \cap L^2\left(I, H^1_0(\Omega)^d \cap L^2_\sigma(\Omega)^d\right), \\ c &\in L^2_{loc}(I, H^2(\Omega)), \\ \nabla\mu &\in L^2(I\times\Omega). \end{aligned}$$

Here $BC_\omega(I, L^2_\sigma(\Omega)^d)$ denotes the topological vector space of bounded and weakly continuous functions from I with values in $L^2_\sigma(\Omega)^d$. It further holds

$$\begin{aligned} L^2_\sigma(\Omega)^d &:= \{f\in L^2(\Omega)^d \,|\, \mathrm{div}\, f = 0, \nu_\Omega\cdot f|_{\partial\Omega} = 0\}, \\ L^2_{loc}(I, H^2(\Omega)) &:= \{f\,|\, f\in L^2(I\cap B, H^2(\Omega)) \text{ for all balls } B \text{ with } \overline{I\cap B}\subset I\}. \end{aligned}$$

4 Time-discrete Cahn–Hilliard Navier–Stokes system

Our simulation technique is based on a semi-implicite time discretization that we describe in this section. The discretization of the Cahn–Hilliard part is performed following [Eyr98]. A comparison of different time discretizations for a smooth free energy is carried out in [GT13]. The discretization of the coupling between the Cahn–Hilliard and the Navier–Stokes system is performed similiar to [KSW08] and yields a sequential coupling of the systems.

We define

$$\mathcal{K} := \{v\in H^1(\Omega)\,|\,|v|\le 1\,\text{a.e.}\}.$$

Let $\tau>0$ be the time step size. Then the values of c and y at $t_{old}\in[0,T-\tau]$ are denoted by $c_{old}\in\mathcal{K}$, and $y_{old}\in H^1_0(\Omega)^d$. The values at time $t = t_{old}+\tau$ are written as c^τ, w^τ, y^τ and p^τ.
Given (y_{old}, c_{old}), the tupel $(c^\tau, w^\tau, y^\tau, p^\tau)$ solves the problem:
Find $c^\tau\in\mathcal{K}$, $w^\tau\in H^1(\Omega)$, $y^\tau\in H^1_0(\Omega)^d$, and $p^\tau\in L^2_{(0)}(\Omega)$ such that

$$\begin{aligned} \xi(y^\tau - y_{old}, v) + \eta(\nabla y^\tau : \nabla v) + a^t(y_{old}, y^\tau, v)& \\ -(p^\tau, \mathrm{div}\, v) + (c^\tau\nabla w^\tau, v) &= 0 \quad \forall v\in H^1_0(\Omega)^d, \quad (4.1)\\ (-\mathrm{div}\, y^\tau, q) &= 0 \quad \forall q\in L^2_{(0)}(\Omega), \quad (4.2)\\ (c^\tau - c_{old}, v) + \frac{\tau}{\mathrm{Pe}}(\nabla w^\tau, \nabla v) - \tau(c^\tau y_{old}, \nabla v) &= 0 \quad \forall v\in H^1(\Omega), \quad (4.3)\\ \gamma^2(\nabla c^\tau, \nabla(v-c^\tau)) - (w^\tau, v-c^\tau) - (c_{old}, v-c^\tau) &\ge 0 \quad \forall v\ \in\mathcal{K}, \quad (4.4) \end{aligned}$$

with $\xi = 1/\tau$. In order to simplify the notation, from now on we write c, w, y, p instead of $c^\tau, w^\tau, y^\tau, p^\tau$.

For $u\in L^q(\Omega)^d, q>d$ and $v, w\in H^1_0(\Omega)^d$ we introduce

$$a^t(u,v,w) := \frac{1}{2}\left(\int_\Omega ((u\nabla)v)w\,dx - \int_\Omega ((u\nabla)w)v\,dx\right). \quad (4.5)$$

We note that for all $u \in H_0^1(\Omega)^d$ with div $u = 0$ there holds

$$a^t(u, v, w) = \int_\Omega ((u\nabla)v)w\, dx. \tag{4.5'}$$

Whenever div $u = 0$ holds we use $a^t(u, v, w)$ as in (4.5'), while (4.5) is used if preservation of anti-symmetry with respect to the last two arguments is needed in the discrete setting, i.e. $a^t(u, v, w) = -a^t(u, w, v)$. The anti-symmetry directly implies $a^t(u, v, v) = 0$. Due to [Tem77, Lem. II.1.1] there holds

$$a^t(u, v, w) \leq C(d)\|\nabla u\|\|\nabla v\|\|\nabla w\|. \tag{4.6}$$

For matrices $A, B \in L^2(\Omega)^{d\times d}$ we use the notation

$$(A : B) := \int_\Omega A : B\, dx = \int_\Omega \sum_{i,j=1}^{d} (A)_{i,j}(B)_{i,j}\, dx.$$

Note that by using $v \equiv 1$ as test function in (4.3), we obtain $(c, 1) = (c_{old}, 1)$ and thus mass conservation. To achieve this, it is essential to have the gradient on the test function in the term arising from transport and that the boundary integrals arising from integration by parts vanishes. The latter is encorporated by prescribing $y = 0$ on the boundary.

Our time discretization sequentially couples the Cahn–Hilliard and the Navier–Stokes system. Thus, for given y_{old} we first solve (4.3)–(4.4) to obtain c and w, and then solve (4.1)–(4.2) with c and w at hand.

Due to this sequential coupling we investigate (4.3)–(4.4) and (4.1)–(4.2) independently and start with (4.3)–(4.4).

Analysis of (4.3)–(4.4)

In this section we show the existence of a unique solution to (4.3)–(4.4) exploiting results from [HHT11].

To prove existence and uniqueness of a solution to (4.3)–(4.4) it is convenient, to introduce the following optimization problem

$$\begin{aligned} \min_{(c,w)\in\mathcal{K}\times V_0} J(c, w) &:= \frac{\gamma^2}{2}\|\nabla c\|^2 + \frac{\tau}{2\mathrm{Pe}}\|\nabla w\|^2 - (c_{old}, c) \\ \text{subject to} &\quad (4.3) \end{aligned} \tag{$\mathcal{P}$}$$

and to interpret the system (4.3)–(4.4) as the first order optimality system for $(\mathcal{P})$, see e.g. [Gar07, GK07]. Here $V_0 = \{v \in H^1(\Omega) \,|\, (v, 1) = 0\}$. Note that, if we interpret $c \in \mathcal{K}$ as control and $w \in V_0$ as an according state, then $(\mathcal{P})$ has the flavour of a linear-quadratic elliptic optimal control problem with box constraints on the control.

We start our investigation by some results concerning J.

Lemma 4.1 ([HHT11, Lem. 3.1]). *Let $\mathcal{F}$ denote the feasible set of $(\mathcal{P})$. Then the following properties hold true:*

(i) $\mathcal{F} \neq \emptyset$ and $\mathcal{F} \subset V_0 \times V_0$.

(ii) $\mathcal{F}$ is a closed and convex subset of $H^1(\Omega) \times H^1(\Omega)$.

(iii) J is strictly convex on $\mathcal{F}$.

(iv) For every sequence $(c_n, w_n)_{n\in\mathbb{N}} \subset \mathcal{F}$ such that

$$\|c_n\|_{H^1(\Omega)} \overset{n\to\infty}{\longrightarrow} +\infty \textit{ or } \|w_n\|_{H^1(\Omega)} \overset{n\to\infty}{\longrightarrow} +\infty$$

we have $\lim_{n\to\infty} J(c_n, w_n) = +\infty$.

Proof.

(i) We have $\mathcal{F} \neq \emptyset$ since, due to Lax–Milgram's theorem (A1), for choosing $\tilde{c} \in \mathcal{K}$ arbitrary with $(\tilde{c}, 1) = (c_{old}, 1)$ there exists a unique $\tilde{w} \in V_0$ such that $(\tilde{c}, \tilde{w})$ is a solution to (4.3) since $(c_{old} - \tilde{c}, \cdot) + \tau(\tilde{c}y_{old}, \nabla\cdot) \in V_0^*(\Omega)$. Taking $v \equiv 1$ as test function in (4.3) we obtain $(c, 1) = (c_{old}, 1) = 0$ and thus $c \in V_0$. Since $w \in V_0$ by construction we have $\mathcal{F} \subset V_0 \times V_0$.

(ii)–(iv) See [HHT11, Lem. 3.1].

□

Theorem 4.2 ([HHT11, Th. 3.2]). *The problem $(\mathcal{P})$ has a unique solution $(c^\star, w^\star)$. Moreover, there exists a Lagrange multiplier $p^\star \in H^1(\Omega)$ such that $w^\star = p^\star - (p^\star, 1)$ and $(c^\star, p^\star)$ is a solution of (4.3)–(4.4). Conversely, if $(c^\star, p^\star)$ is a solution to (4.3)–(4.4), then $(c^\star, w^\star)$ with $w^\star = p^\star - (p^\star, 1)$ is the unique solution of $(\mathcal{P})$.*

Proof. For convenience we here repeat the proof from [HHT11] with slight modification.

The existence and uniqueness of the solution of $(\mathcal{P})$ are immediate consequences of the previous lemma. The existence of a Lagrange multiplier $p^\star$ follows from mathematical programming in Banach space, see [ZK79]. The main result of [ZK79] concerning the existence of a Lagrange multiplier is given in the Appendix (Theorem A2). Here we check that the constraint qualification (a1) is satisfied. For a given $f \in (H^1(\Omega))^*$ in our context it consists in finding $(c, w) \in \mathcal{K} \times V_0$ and $\xi \geq 0$ such that

$$\frac{\tau}{\mathrm{Pe}}(\nabla w, \nabla v) = \langle f, v\rangle - \xi(c - c^\star, v) =: \langle g, v\rangle \quad \forall v \in H^1(\Omega). \tag{4.7}$$

Let $c \in \mathcal{K}$ chosen such that $(c, 1) \neq 0$ and $\xi = \langle f, 1\rangle/(c - c^\star, 1) \geq 0$. Its existence is guaranteed since $\mathcal{K}$ is symmetric with respect to the origin. Note that the right hand side $g \in (H^1(\Omega))^*$ in (4.7) satisfies the compatibility condition

$\langle g, 1\rangle = 0$. Hence, by the Lax–Milgram theorem there exists a unique w such that (4.7) is fullfilled.

Now Theorem (A2) yields the existence of an adjoint state (or Lagrange multiplier associated with (4.3)) $p^\star \in H^1(\Omega)$ such that

$$(c^\star, v) + \frac{\tau}{\mathrm{Pe}}(\nabla w^\star, \nabla v) = \tau(c^\star y_{old}, \nabla v) + (c_{old}, v) \quad \forall v \in H^1(\Omega), \tag{4.8}$$

$$\gamma^2 \left(\nabla c^\star, \nabla(v - c^\star)\right) - (p^\star, v - c^\star) \geq (c_{old}, v - c^\star) \quad \forall v \in \mathcal{K}, \tag{4.9}$$

$$(\nabla p^\star, \nabla v) = (\nabla w^\star, \nabla v) \quad \forall v \in H^1(\Omega). \tag{4.10}$$

Consequently, $(c^\star, p^\star)$ is a solution of (4.3)–(4.4).

We next show the uniqueness of $p^\star$. For this let us assume there exist two multipliers $p_1^\star$ and $p_2^\star$ with $p_1^\star \neq p_2^\star$. From the uniqueness of $w^\star$ it follows

$$w^\star = p_1^\star - (p_1^\star, 1) = p_2^\star - (p_2^\star, 1),$$

and thus

$$p_1^\star - p_2^\star = (p_1^\star, 1) - (p_2^\star, 1) = \kappa \in \mathbb{R}.$$

Since $|(c^\star, 1)| = |(c_0, 1)| < |\Omega|$ holds, the set $\Omega^\star = \{x \in \Omega \,|\, |c^\star| < 1\}$ is of positive measure. Thus, it holds $1 - c^\star \geq 0$ and $-1 - c^\star \leq 0$ as well as $1 - c^\star \not\equiv 0$ and $-1 - c^\star \not\equiv 0$

Inserting $p_1^\star$ and $p_2^\star$ into (4.9) and substracting the resulting equations we obtain

$$(p_1^\star - p_2^\star, v - c^\star) \geq 0 \quad \forall v \in \mathcal{K}.$$

By choosing $v = 1$ and $v = -1$ we obtain

$$\begin{aligned} (p_1^\star - p_2^\star, 1 - c^\star) &= \kappa(1, 1 - c^\star) \geq 0, \\ (p_1^\star - p_2^\star, -1 - c^\star) &= \kappa(1, -1 - c^\star) \geq 0. \end{aligned}$$

Since $1 - c^\star \geq 0$ we have $(1, 1 - c^\star) \geq 0$ and thus $\kappa \geq 0$ from the first inequality. Since $-1 - c^\star \leq 0$ we have $(1, -1 - c^\star) \leq 0$ and thus $\kappa \leq 0$ from the second inequality. Thus $\kappa = 0$ holds and $p_1^\star \equiv p_2^\star$. Thus $p^\star$ is unique.

For the reverse implication it is clear that if $(c^\star, p^\star)$ is a solution of (4.3)–(4.4), then $(c^\star, w^\star, p^\star)$ with $w^\star = p^\star - (p^\star, 1)$ is a solution of the optimality system (4.8)–(4.10). Since $(\mathcal{P})$ is a convex problem, any stationary point of $(\mathcal{P})$, i.e. a solution of (4.8)–(4.10), is also a global solution of $(\mathcal{P})$. Thus, $(c^\star, w^\star)$ is the unique solution of $(\mathcal{P})$. □

Analysis of (4.1)–(4.2)

In this section we show the existence of a unique solution to (4.1)–(4.2) using a general existence result for saddle point problems. We further give a reformulation of (4.1)–(4.2) following [Jus11].

Theorem 4.3 ([Jus11, Satz 3.2]). *There exists a unique solution $(y,p) \in H_0^1(\Omega)^d \times L^2_{(0)}(\Omega)$ to (4.1)–(4.2).*

Proof. We use the result [GR79, Th. I.4.1] for a general class of saddle point problems that we for convenience recall in the Appendix (Theorem A6).

In our case we have $X = H_0^1(\Omega)^d$ and $M = L^2_{(0)}(\Omega)$. For $y,v \in H_0^1(\Omega)^d$ and $p,q \in L^2_{(0)}(\Omega)$ we have $a(y,v) = \xi(y,v) + \eta(\nabla y, \nabla v) + a^t(y_{old}, y, v)$ and $b(y,q) = -(\operatorname{div} y, q)$. We further have $l = \xi y_{old} - c\nabla w \in X^\star$ and $\chi = 0 \in M^\star$.

For applying Theorem A6 we have to show that a is continuous and coercive on $H_0^1(\Omega)^d$ and that b is continuous on $H_0^1(\Omega)^d \times L^2_{(0)}(\Omega)$ and fulfills the $inf - sup$ condition. Moreover we have to show $l \in X^\star$.

We start with the continuity of a. Using Poincaré's inequality we have together with the continuity of a^t (4.6)

$$\begin{aligned}|a(y,v)| &\le \xi\|y\|\|v\| + \eta\|\nabla y\|\|\nabla v\| + c(d)\|\nabla y_{old}\|\|\nabla y\|\nabla v\|\\ &\le C(\Omega)(\xi + \eta + C(d)\|\nabla y_{old}\|)\|y\|_{H_0^1(\Omega)^d}\|v\|_{H_0^1(\Omega)^d}.\end{aligned}$$

Furthermore

$$b(y,q) \le \|\operatorname{div} y\|\|q\| \le \sqrt{d}\|q\|\|\nabla y\|.$$

Concerning the coercivity we have for arbitrary $y \in H_0^1(\Omega)^d$

$$a(y,y) = \xi\|y\|^2 + \eta\|\nabla y\|^2 + \underbrace{a^t(y_{old}, y, y)}_{=0} \ge C(\Omega)\min\{\xi,\eta\}\|y\|^2_{H_0^1(\Omega)^d}.$$

The $inf - sup$ condition for b can be found in e.g. [GR79, Th. I.3.7].

The requirement $\xi y_{old} - c\nabla w \in \left(H_0^1(\Omega)^d\right)^*$ follows from Hölder's inequality together with embedding theory (see [AF03]) for Sobolev spaces. □

For constructing an a posteriori error estimator in Section 8.2 we now introduce a bilinear form B to rewrite (4.1)–(4.2) in a compact form. In addition we introduce a norm on $H_0^1(\Omega)^d \times L^2_{(0)}(\Omega)$ which we later use to measure the local error contributions. This idea and the following proof are taken from [Jus11, Satz 3.4].

Equations (4.1)–(4.2) are equivalent to:
Find $y,p \in H_0^1(\Omega)^d \times L^2_{(0)}(\Omega)$ such that

$$B((y,p),(v,q)) = L((v,q)) \quad \forall v \in H_0^1(\Omega)^d, q \in L^2_{(0)}(\Omega) \tag{4.11}$$

holds. Here

$$\begin{aligned}B((y,p),(v,q)) &= \xi(y,v) + \eta(\nabla y : \nabla v) + a^t(y_{old}, y, v) - (p, \operatorname{div} v) + (q, \operatorname{div} y),\\ L((v,q)) &= (\xi y_{old} - c\nabla w, v).\end{aligned}$$

On $H_0^1(\Omega)^d \times L^2_{(0)}(\Omega)$ we define the following norm

$$|||(y,p)||| := \left\{\eta\|\nabla y\|^2 + \xi\|y\|^2 + \frac{1}{\eta}\|p\|^2\right\}^{1/2}.$$

Then there holds:

Theorem 4.4 ([Jus11, Satz 3.4])**.**

1. *The bilinear form B is continuous, i.e. there exists a constant $c_S = c_S(\eta, y_{old}, d)$ such that*

$$B((y,p),(v,q)) \leq c_S \; ||| \; (u,p) \; ||| \; |||(v,q)|||$$

holds for all $(y,p),(v,q) \in H^1_0(\Omega)^d \times L^2_{(0)}(\Omega)$.

2. *The bilinear form B fulfills a modified $inf - sup$ condition, i.e. there exists a constant $\beta^* = \beta^*(\Omega, \eta, y_{old}, \xi, \beta) > 0$ such that*

$$\inf_{(y,p)\in\left(H^1_0(\Omega)^d\times L^2_{(0)}(\Omega)\right)\setminus\{0\}} \; \sup_{(v,q)\in\left(H^1_0(\Omega)^d\times L^2_{(0)}(\Omega)\right)\setminus\{0\}} \frac{B((y,p),(v,q))}{|||(y,p) \; ||| \; |||(v,q)|||} \geq \beta^*,$$

where $\beta > 0$ is the $inf - sup$ constant for the bilinear form b.

Proof. Using Hölder's inequality and the continuity of a^t (4.6) together with $\|\mathrm{div} y\| \leq \sqrt{d}\|\nabla y\|$ we have

$$\begin{aligned} B((y,p),(v,q)) \leq & \xi\|y\|\|v\| + \eta\|\nabla y\|\|\nabla v\| + C(d)\|\nabla y_{old}\|\|\nabla y\|\|\nabla v\| \\ & + \|p\|\|\mathrm{div} v\| + \|q\|\|\mathrm{div} y\| \\ \leq & \xi\|y\|\|v\| + (\eta + C(d)\|\nabla y_{old}\|)\,\|\nabla y\|\|\nabla v\| \\ & + \sqrt{d}\|p\|\|\nabla v\| + \sqrt{d}\|q\|\|\nabla y\|. \end{aligned}$$

Using Cauchy–Schwarz's inequality we proceed

$$\begin{aligned} & B((y,p),(v,q)) \\ & \leq \left\{\xi\|y\|^2 + (\eta + C(d)\|\nabla y_{old}\|)\,\|\nabla y\|^2 + \eta\sqrt{d}\|\nabla y\|^2 + \frac{1}{\eta}\|p\|^2\right\}^{1/2} \\ & \cdot \left\{\xi\|v\|^2 + (\eta + C(d)\|\nabla y_{old}\|)\,\|\nabla v\|^2 + \eta\sqrt{d}\|\nabla v\|^2 + \frac{1}{\eta}\|q\|^2\right\}^{1/2} \\ & = \left\{\xi\|y\|^2 + \eta\|\nabla y\|^2 + \left(C(d)\|\nabla y_{old}\| + \eta\sqrt{d}\right)\|\nabla y\|^2 + \frac{1}{\eta}\|p\|^2\right\}^{1/2} \\ & \cdot \left\{\xi\|v\|^2 + \eta\|\nabla v\|^2 + \left(C(d)\|\nabla y_{old}\| + \eta\sqrt{d}\right)\|\nabla v\|^2 + \frac{1}{\eta}\|q\|^2\right\}^{1/2} \\ & \leq \left(1 + \left(\frac{C(d)\|\nabla y_{old}\| + \eta\sqrt{d}}{\eta}\right)^{1/2}\right)^2 \; ||| \; (y,p) \; ||| \; |||(v,q) \; ||| \, . \end{aligned}$$

Thus B is continous on $H^1_0(\Omega)^d \times L^2_{(0)}(\Omega)$ with continuity constant

$$c_S = \left(1 + \left(\frac{C(d)\|\nabla y_{old}\| + \eta\sqrt{d}}{\eta}\right)^{1/2}\right)^2.$$

Next we show the $inf - sup$ condition for B. Using (y, p) as test function in B and using the antisymmetry of a^t, there holds

$$B((y,p),(y,p)) = \xi\|y\|^2 + \eta\|\nabla y\|^2.$$

From [GR79, Th. I.3.7] we have, that for each $p \in L^2_{(0)}(\Omega)$ there exists a unique $y_p \in H^1_0(\Omega)^d$ such that $p = -\mathrm{div} y_p$ and $\|\nabla y_p\| \leq \frac{1}{\beta}\|p\|$ holds, where β denotes the $inf - sup$ constant for b. With Poincaré's inequality it further follows that $\|y_p\| \leq C_p\|\nabla y_p\|$. Furthermore

$$\begin{aligned}
B((y,p),&(y_p,0))\\
&=\xi(y,y_p) + \eta(y,y_p) + a^t(y_{old},y,y_p) - (p,\operatorname{div} y_p)\\
&\geq -\xi\|y\|\|y_p\| - \eta\|\nabla y\|\|\nabla y_p\| - C(d)\|\nabla y_{old}\|\|\nabla y\|\|\nabla y_p\| + \|p\|^2\\
&\geq -\frac{\xi C_p}{\beta}\|y\|\|p\| - \frac{\eta}{\beta}\|\nabla y\|\|p\| - \frac{C(d)}{\beta}\|\nabla y_{old}\|\|\nabla y\|\|p\| + \|p\|^2\\
&\geq \frac{1}{4\eta}\|p\|^2 - \frac{\xi^2 C_p^2\eta}{\beta^2}\|y\|^2 - \eta\left(\frac{\eta^2 + C(d)^2\|\nabla y_{old}\|^2}{\beta^2}\right)\|\nabla y\|^2
\end{aligned}$$

holds. Now, for $\kappa \in (0,1)$ we have

$$\begin{aligned}
&B((y,p),(1-\kappa)(y,p)+\kappa(y_p,0)) = (1-\kappa)B((y,p),(y,p)) + \kappa B((y,p),(y_p,0))\\
&\geq(1-\kappa)\left(\xi\|y\|^2 + \eta\|\nabla y\|^2\right)\\
&\quad+\kappa\left(\frac{1}{4\eta}\|p\|^2 - \frac{\xi^2 C_p^2\eta}{\beta^2}\|y\|^2 - \eta\left(\frac{\eta^2 + C(d)^2\|\nabla y_{old}\|^2}{\beta^2}\right)\|\nabla y\|^2\right)\\
&\geq\eta\|\nabla y\|^2\left(1-\kappa-\kappa\left(\frac{\eta^2 + C(d)^2\|\nabla y_{old}\|^2}{\beta^2}\right)\right)\\
&\quad+\xi\|y\|^2\left(1-\kappa-\kappa\frac{\eta\xi C_p^2}{\beta^2}\right) + \kappa\frac{1}{4\eta}\|p\|^2\\
&\geq\left\{\eta\|\nabla y\|^2 + \xi\|y\|^2\right\}\left[1-\kappa-\kappa\left(\frac{\eta^2 + C(d)^2\|\nabla y_{old}\|^2 + \eta\xi C_p^2}{\beta^2}\right)\right] + \kappa\frac{1}{4\eta}\|p\|^2.
\end{aligned}$$

We choose κ in a way such that

$$\left[(1-\kappa-\kappa\left(\frac{\eta^2 + C(d)^2\|\nabla y_{old}\|^2 + \eta\xi C_p^2}{\beta^2}\right)\right] = \frac{1}{4}\kappa$$

holds, i.e.

$$\kappa = \frac{\beta^2}{\beta^2 + \eta^2 + C(d)^2\|\nabla y_{old}\|^2 + \eta\xi C_p^2 + \frac{\beta^2}{4}} \in (0,1).$$

Combination of the previous estimates yields

$$B((y,p),(1-\kappa)(y,p)+\kappa(y_p,0)) \geq \frac{1}{4}\kappa \left|\!\left|\!\left| (y,p) \right|\!\right|\!\right|^2.$$

Since

$$\begin{aligned}|||(y_p,0)||| &= \left(\eta\|\nabla y_p\|^2+\xi\|y_p\|^2\right)^{1/2}\\ &\leq \left(\frac{\eta}{\beta^2}\|p\|^2+\frac{\xi C_p^2}{\beta^2}\|p\|^2\right)^{1/2} = \frac{1}{\beta}\left(\eta^2+\eta\xi C_p^2\right)^{1/2}\frac{1}{\sqrt{\eta}}\|p\|,\end{aligned}$$

the triangular inequality gives

$$|||(1-\kappa)(y,p)+\kappa(y_p,0)||| \leq \left(1-\kappa+\frac{\kappa}{\beta}\left(\eta^2+\eta\xi C_p^2\right)^{1/2}\right)|||\,(y,p)\,|||\,.$$

Thus, we end up with

$$\begin{aligned}\sup_{(v,q)\in(H_0^1(\Omega)^d\times L^2_{(0)}(\Omega))\setminus\{0\}}\frac{B((y,p),(v,q))}{|||(v,q)|||} &\geq \frac{B((y,p),(1-\kappa)(y,p)+\kappa(y_p,0))}{|||(1-\kappa)(y,p)+\kappa(y_p,0)|||}\\ &\geq \frac{\frac{1}{4}\kappa\,|||\,(y,p)|||^2}{\left(1-\kappa+\frac{\kappa}{\beta}\left(\eta^2+\eta\xi C_p^2\right)^{1/2}\right)|||\,(y,p)|||}.\end{aligned}$$

Since (y,p) is chosen arbitraryly, the second statement follows with

$$\beta^* = \frac{\kappa}{4\left(1-\kappa+\frac{\kappa}{\beta}\left(\eta^2+\eta\xi C_p^2\right)^{1/2}\right)} > 0.$$

□

In Section 8.2 we use Theorem 4.4 to show the equivalence of the norm of the actual discretization error and the norm of the residual during the construction of an a-posteriori residual based error estimator for the numerical solution of (4.1)–(4.2).

We finish this section with stating higher regularity for the velocity field assuming higher regularity for the phase field c that we show in Lemma 5.8.

Theorem 4.5. *Assume $\partial\Omega$ of class C^3, $y_{old}\in H^2(\Omega)^d\cap H_0^1(\Omega)^d$, and let $(y,p)\in H_0^1(\Omega)^d\times L^2_{(0)}(\Omega)$ be the solution of (4.1)–(4.2). Assume $c\in H^2(\Omega)$. Then there holds $y\in H^2(\Omega)^d$ and $p\in H^1(\Omega)$.*

Proof. Using regularity results for the Stokes equation ([CF88, Th. 3.7]) we obtain

$$\|y\|_{H^2(\Omega)^d}+\|p\|_{H^1(\Omega)} \leq C\left(\|f\|_{L^2(\Omega)^d}+\|y\|_{H^1(\Omega)^d}+\|p\|_{L^2(\Omega)}\right)$$

where $f := \xi(y_{old}-y)-y_{old}\nabla y - c\nabla w$. We have to show $f\in L^2(\Omega)$. This directly follows from $y_{old}\in H^2(\Omega)^d \hookrightarrow L^\infty(\Omega)^d$ and $c\in H^2(\Omega)$. □

5 Moreau–Yosida relaxation

The optimization problem $(\mathcal{P})$ related to the variational inequality (4.4) could be treated by introducing a Lagrange multiplier $\lambda^\star$ for the constraint $c \in \mathcal{K}$. But since $c \in H^1(\Omega)$ we would have $\lambda^\star \in (H^1(\Omega))^*$ thus not allowing a pointwise interpretation.[1] For better regularity results we replace the problem $(\mathcal{P})$ by its Moreau–Yosida relaxed version according to [HHT11]. We then show existence and uniqueness of a solution to the resulting optimization problem along the lines of [HHT11].

The Moreau–Yosida relaxed problem is defined by:

$$\begin{aligned} \min_{(c,w)\in H^1(\Omega)\times V_0} \quad & J^s(c,w) \\ \text{subject to} \quad & (4.3), \end{aligned} \tag{$\mathcal{P}^s$}$$

where J^s is given by

$$J^s(c,w) = J(c,w) + \frac{s}{2}\|\max(0,c-1)\|^2 + \frac{s}{2}\|\min(0,c+1)\|^2.$$

Here $s > 0$ denotes the associated relaxation or penalization parameter, and the max and min operators form a regularization of the indicator function of $\mathcal{K}$, and are understood pointwise.

In the following we show that for $(\mathcal{P}^s)$ there exists a unique solution $(c^s, w^s) \in H^1(\Omega) \times V_0$. The sequence of solutions converges strongly to the unique solution $(c^\star, w^\star) \in \mathcal{K} \times V_0$ of $(\mathcal{P})$ as $s \to \infty$.

Lemma 5.1. *Let $\mathcal{F}_s$ denote the feasible set for $(\mathcal{P}^s)$. Then the following holds*

(i) *$\mathcal{F}_s \neq \emptyset$ and $\mathcal{F}_s \subset V_0 \times V_0$.*

(ii) *$\mathcal{F}_s$ is a closed and convex subset of $H^1(\Omega) \times H^1(\Omega)$.*

(iii) *J^s is strictly convex on $\mathcal{F}_s$.*

(iv) *For each sequence $(c_n, w_n)_{n\in\mathbb{N}} \subset \mathcal{F}_s$ such that*

$$\|c_n\|_{H^1(\Omega)} \overset{n\to\infty}{\longrightarrow} +\infty \text{ or } \|w_n\|_{H^1(\Omega)} \overset{n\to\infty}{\longrightarrow} +\infty$$

we have $\lim_{n\to\infty} J^s(c_n, w_n) = +\infty$.

This lemma can be proven as Lemma 4.1. Note that the functionals

$$c \to \|\max(0,c-1)\|^2, \quad c \to \|\min(0,c+1)\|^2$$

are convex and Fréchet differentiable on $H^1(\Omega)$.

[1] In [BBG11] it is shown that under certain regularity assumptions indeed $\lambda^\star \in L^2(\Omega)$ holds. The authors of [BBG11] exploit this fact to formulate their numerical scheme.

Theorem 5.2 ([HHT11, Th. 4.1]). *The problem $(\mathcal{P}^s)$ has a unique solution (c_s, w_s). Moreover, there exists a unique $p_s \in H^1(\Omega)$ such that*

$$\begin{aligned}(\nabla p_s, \nabla v) =& (\nabla w_s, \nabla v) \qquad \forall v \in H^1(\Omega), \qquad (5.1)\\ \frac{\tau}{Pe}(\nabla p_s, \nabla v) + (c_s, v) - \tau(cy_{old}, \nabla v) =& (c_{old}, v) \qquad \forall v \in H^1(\Omega), \qquad (5.2)\\ \gamma^2(\nabla c_s,\ \nabla v) + (\lambda_s(c_s), v) - (p_s, v) =& (c_{old}, v) \qquad \forall v \in H^1(\Omega), \qquad (5.3)\end{aligned}$$

where $\lambda_s(c_s) = \lambda_s^+(c_s) + \lambda_s^-(c_s)$ with

$$\lambda_s^+(c_s) := s\max(0, c_s - 1) \quad \textit{and} \quad \lambda_s^-(c_s) := s\min(0, c_s + 1).$$

Conversely, if (c_s, p_s) is a solution of (5.2)–(5.3), then (c_s, w_s) with $w_s = p_s - (p_s, 1)$ is the unique solution of $(\mathcal{P}^s)$.

Proof. Due to Lemma 5.1 problem $(\mathcal{P}^s)$ is a convex problem whose cost function is radially unbounded and strictly convex. This yields existence and uniqueness of (c_s, w_s). Similarly, as in the proof of Theorem 4.2, mathematical programming theory in Banach space guarantees the existence of an adjoint state $p_s \in H^1(\Omega)$ satisfying the following first-order optimality system of $(\mathcal{P}^s)$:

$$\begin{aligned}(c_s, v) + \frac{\tau}{\mathrm{Pe}}(\nabla w_s, \nabla v) - \tau(cy_{old}, \nabla v) &= (c_{old}, v) \qquad \forall v \in H^1(\Omega), \qquad (5.4)\\ \gamma^2(\nabla c_s,\ \nabla v) + (\lambda_s(c_s), v) - (p_s, v) &= (c_{old}, v) \qquad \forall v \in H^1(\Omega), \qquad (5.5)\\ (\nabla p_s, \nabla v) &= (\nabla w_s, \nabla v) \qquad \forall v \in H^1(\Omega). \qquad (5.6)\end{aligned}$$

The uniqueness of p_s follows from the uniqueness of (c_s, w_s) of $(\mathcal{P}^s)$ and (5.5). □

Lemma 5.3 ([HHT11, Prop. 4.2]). *Let c_s, w_s denote the solution to $\mathcal{P}^s$. Then there exists $C > 0$, independend of s, such that*

$$\begin{aligned}\|c_s\|_{H^1(\Omega)} \le C, &\quad \sqrt{s}\|\max(0, c_s - 1)\| \le C\\ \|w_s\|_{H^1(\Omega)} \le C, &\quad \sqrt{s}\|\min(0, c_s + 1)\| \le C.\end{aligned}$$

Proof. Let $c^\star, w^\star$ denote the solution to problem $\mathcal{P}$. By the properties of the respective solutions we have

$$J(c_s, w_s) \le J_s(c_s, w_s) \le J_s(c^\star, w^\star) = J(c^\star, w^\star). \qquad (5.7)$$

Thus, there exists a constant $\beta > 0$, independend of s such that

$$\frac{\gamma^2}{2}\|\nabla c_s\|^2 + \frac{\tau}{2Pe}\|\nabla w_s\|^2 - (c_{old}, c) + \frac{s}{2}\|\max(0, c_s - 1)\|^2 + \frac{s}{2}\|\min(0, c_s + 1)\|^2 \le \beta.$$

Since $(c_s, 1) = (w_s, 1) = 0$ by Young's inequality together with the Poincaré–Friedrichs inequality we get the stated results. □

Theorem 5.4 ([HHT11, Prop. 4.2]). *Let $\{(c_s, w_s)\}_{s>0}$ be a sequence of solutions of $(\mathcal{P}^s)$. Then there exists a subsequence, still denoted by $\{(c_s, w_s)\}_{s>0}$, such that*

$$(c_s, w_s) \to (c^\star, w^\star) \text{ in } H^1(\Omega) \times H^1(\Omega) \tag{5.8}$$

as $s \to +\infty$, where $(c^\star, w^\star)$ denotes the unique solution of $(\mathcal{P})$. In particular, $c^\star$ is the order parameter corresponding to the solution of (4.3)–(4.4).

Proof. From Lemma 5.3 we have the existence of $(c_*, w_*) \in H^1(\Omega) \times H^1(\Omega)$ and a subsequence still denoted by $\{(c_s, w_s)\}_{s>0}$ such that

$$(c_s, w_s) \to (c_*, w_*) \text{ in } L^2(\Omega) \quad \text{and} \quad (c_s, w_s) \rightharpoonup (c_*, w_*) \text{ in } H^1(\Omega) \tag{5.9}$$

as $s \to +\infty$ since $L^2(\Omega) \hookrightarrow H^1(\Omega)$ compactly. Moreover, passing to the limit in the state equation of $(\mathcal{P}^s)$, we obtain

$$(c_*, v) + \frac{\tau}{\mathrm{Pe}}(\nabla w_*, \nabla v) = \tau(c_*, y_{old}\nabla v) + (c_{old}, v) \quad \forall v \in H^1(\Omega). \tag{5.10}$$

Thus c_*, w_* is a solution to the state equation.

On the other hand, from (5.9) we infer

$$\begin{aligned} \max(0, c_s - 1) &\to \max(0, c_* - 1) \text{ in } L^2(\Omega), \\ \min(0, c_s + 1) &\to \min(0, c_* + 1) \text{ in } L^2(\Omega). \end{aligned}$$

This together with Lemma 5.3 yields

$$-1 \le c_* \le 1 \quad \text{a.e. in } \Omega. \tag{5.11}$$

From (5.10) and (5.11) we deduce that $(c, w) \in \mathcal{F}$. Moreover, from (5.7) and the lower semi-continuity of semi-norms in $H^1(\Omega)$ we infer

$$J(c_*, w_*) = J_s(c_*, w_*) \le \liminf_{s\to\infty} J_s(c_s, w_s) \le J_s(c_s, w_s) \le J(c^\star, w^\star). \tag{5.12}$$

The uniqueness of the solution of $(\mathcal{P})$ implies $(c_*, w_*) = (c^\star, w^\star)$.

Finally, we establish the strong convergence result in $H^1(\Omega)$. From above we have

$$J(c^\star, w^\star) \le \liminf_{s\to\infty} J_s(c_s, w_s) \le \limsup_{s\to\infty} J_s(c_s, w_s) \le J(c^\star, w^\star)$$

and thus

$$\lim_{s\to\infty} \|\nabla c_s\| = \|\nabla c^\star\| \quad \text{as well as} \quad \lim_{s\to\infty} \|\nabla w_s\| = \|\nabla w^\star\|.$$

Now, the weak and norm convergence yield the strong convergence result (5.8). □

For studying the limit of the optimality system (5.1)–(5.3) we first establish some useful results.

Lemma 5.5 ([HHT11, Lem. 4.3])**.** *There exist constants $\beta_p > 0$ and $\beta_\lambda > 0$ independent of s, such that*

$$|(p_s, 1)| \le \beta_p, \tag{5.13}$$

$$\|\lambda_s(c_s)\| \le \|\lambda_s^+(c_s)\| + \|\lambda_s^-(c_s)\| \le \beta_\lambda, \tag{5.14}$$

for all $s > 0$.

Proof. See [HHT11, Lem. 4.3]. □

This allows us to study the limit of (5.1)–(5.3) for $s \to \infty$.

Theorem 5.6 ([HHT11, Thm. 4.4])**.** *Let $(p_s)_{s>0}$ denote the sequence of functions from Theorem 5.2. Then $p_s \rightharpoonup p^\star$ for $s \to \infty$ in $H^1(\Omega)$. Moreover, together with $(c^\star, w^\star)$ of Theorem 5.4 the function $p^\star$ satisfies the first order optimality system* (4.8)*–*(4.10)*.*

Proof. The weak convergence of a subsequence of $\{p_s\}_{s>0}$ in $H^1(\Omega)$ with the limit $p^\star$ follows from the uniform boundedness of $\{w_s\}_{s>0}$ in $H^1(\Omega)$, $p_s = w_s + (p_s, 1)$, and the uniform boundedness of $\{|(p_s, 1)|\}_{s>0}$ according to Lemma 5.5.

Concerning the first order optimality system (4.8)–(4.10) we note that (4.10) follows immediately from (5.6) and the boundedness of $\{w_s\}_{s>0}$ and $\{p_s\}_{s>0}$ in $H^1(\Omega)$. Equation (4.8) has already been established in the proof of Theorem 5.4. It remains to study (4.9). For this purpose we observe that for arbitrary but fixed $v \in \mathcal{K}$

$$\begin{aligned}(\lambda_s(c_s), v - c_s) &= s(\max(c_s - 1, 0), v - c_s) + s(\min(c_s + 1, 0), v - c_s)\\ &= s(\max(c_s - 1, 0), v - 1) + s(\max(c_s - 1, 0), 1 - c_s)\\ &\quad + s(\min(c_s + 1, 0), v + 1), + s(\min(c_s + 1, 0), -1 - c_s)\\ &\le 0,\end{aligned}$$

holds, where we use that $-1 \le v \le 1$ holds a.e. in Ω. Hence, we have

$$\lim_{s\to\infty} (\lambda_s(c_s), v - c_s) \le 0. \tag{5.15}$$

Let $v \in \mathcal{K}$ arbitrarily, then $v - c_s \in H^1(\Omega)$ is a valid test function in (5.3) and we replace (5.3) by

$$\begin{aligned}\gamma^2(\nabla c_s, \nabla(v - c_s)) &+ (\lambda_s(c_s), v - c_s)\\ &- (p_s, v - c_s) - (c_{old}, v - c_s) = 0 \quad \forall v \in \mathcal{K}.\end{aligned} \tag{5.3'}$$

Next, we recall that due to Theorem 5.4 we have the strong convergence of $\{c_s\}_{s>0}$ in $H^1(\Omega)$. Since by Lemma 5.5 the sequence $\{\lambda_s(c_s)\}$ is uniformly bounded in $L^2(\Omega)$ there exists a weakly convergent subsequence. Thus, passing to the limit in (5.3'), together with (5.15) we obtain

$$\gamma^2(\nabla c^\star, \nabla(v - c^\star)) - (p^\star, v - c^\star) \ge (c_{old}, v - c^\star) \quad \forall v \in \mathcal{K},$$

which establishes (4.9). □

Remark 5.7. Solving (5.1)–(5.3) for a sequence $(s_k)_{k\in\mathbb{N}}$ with $s_k \to +\infty$ as $k \to \infty$ establishes an iterative way for solving (4.3)–(4.4), where y_{old} is fixed. With c_s, w_s obtained in this way we then solve the discrete Navier–Stokes equation with forcing term $c_s \nabla w_s$.

We now prove higher regularity of c_s, w_s, and p_s.

Lemma 5.8. *Assume that the boundary is sufficiently smooth (e.g. of class $\mathcal{C}^2$, [EG04, Th. 3.10, Def. 1.46]), then there holds $c_s, w_s, p_s \in H^2(\Omega)$ and there exists a constant $C > 0$ independend of s such that*

$$\|c_s\|_{H^2(\Omega)} + \|w_s\|_{H^2(\Omega)} + \|p_s\|_{H^2(\Omega)} \leq C$$

holds.

Proof. Using regularity results for the Laplace problem (see e.g. [EG04, Th. 3.10]) we obtain $c_s \in H^2(\Omega)$. Now [Tay96, Prop. 5.7.4] yields

$$\|c_s\|^2_{H^2(\Omega)} \leq C \left(\| -\lambda_s(c_s) + c_{old} + p_s\|^2_{L^2(\Omega)} + \|c_s\|^2_{H^1(\Omega)} \right),$$

which due to Lemma 5.3 and Lemma 5.5 is bounded independently of s. Analogously we now obtain similar estimates for w_s and p_s. □

For convenience of the reader we here state the complete weak form of the time-discrete and Moreau–Yosida relaxed Cahn–Hilliard Navier–Stokes system.

Find $(y, p, c_s, w_s) \in H^1_0(\Omega)^d \times L^2_{(0)}(\Omega) \times H^1(\Omega) \times H^1(\Omega)$ such that there holds

$$\begin{aligned}
\xi(y - y_{old}, v) + \eta(\nabla y : \nabla v) + a^t(y_{old}, y, v) & \\
-(p, \operatorname{div} v) + (c_s \nabla w_s, v) &= 0 \quad \forall v \in H^1_0(\Omega)^d, && (5.16)\\
(-\operatorname{div} y, v) &= 0 \quad \forall v \in L^2_{(0)}(\Omega), && (5.17)\\
(c_s, v) + \frac{\tau}{\mathrm{Pe}}(\nabla w_s, \nabla v) - (c_{old}, v) - \tau(c_s y_{old}, \nabla v) &= 0 \quad \forall v \in H^1(\Omega), && (5.18)\\
\gamma^2(\nabla c_s, \nabla v) - (w_s, v) + (\lambda_s(c_s), v) - (c_{old}, v) &= 0 \quad \forall v \in H^1(\Omega). && (5.19)
\end{aligned}$$

6 Semi-smooth Newton method in function space

Remark 5.7 motivates our function space algorithm to solve (4.1)–(4.4). We specify a sequence $s \to \infty$ and solve the system (5.18)–(5.19) for c_s and w_s. We next show that for fixed but arbitrarily s this can be done by using Newton's method in function space obtaining superlinear convergence at least for small time steps and in a neigbourhood of the unique solution (c_s, w_s) of (5.18)–(5.19). We again proceed along the lines of [HHT11].

We write (5.18)–(5.19) in the form

$$F_s(c_s, w_s) = \big(F_s^{(1)}(c_s, w_s), F_s^{(2)}(c_s, w_s)\big) = 0 \tag{6.1}$$

with

$$\big\langle F_s^{(1)}(c_s, w_s), v\big\rangle = \frac{\tau}{Pe}(\nabla w_s, \nabla v) + (c_s, v) - (c_{old}, v) - \tau(c_s y_{old}, \nabla v), \tag{6.2}$$

$$\big\langle F_s^{(2)}(c_s, w_s), v\big\rangle = \gamma^2(\nabla c_s, \nabla v) + (\lambda_s(c_s), v) - (w_s, v) - (c_{old}, v), \tag{6.3}$$

where c_s, w_s and v are elements of $H^1(\Omega)$.

Note that since $\lambda_s(\cdot)$ is only Lipschitz continuous F_s is not Fréchet differentiable. But it fulfills a weaker form of differentiability called Newton differentiability or slant-differentiability, see Definition A3 and Theorem A4 in the appendix for a convergence result.

We apply Theorem A4 to the mapping $F_s : H^1(\Omega) \times H^1(\Omega) \to H^1(\Omega)^* \times H^1(\Omega)^*$. We first show the Newton differentiability of F_s.

Theorem 6.1 ([HHT11, Lem. 5.3])**.** *The mapping $F_s : H^1(\Omega) \times H^1(\Omega) \to H^1(\Omega)^* \times H^1(\Omega)^*$ is Newton differentiable. A Newton derivative is given by the operator $G_s(c_s, w_s)$ defined by*

$$\langle G_s(c_s, w_s)(\delta c, \delta w), (v_1, v_2)\rangle := \begin{pmatrix} \frac{\tau}{Pe}(\nabla \delta w, \nabla v_1) + (\delta c, v_1) - \tau(\delta c y_{old}, \nabla v_1) \\ \gamma^2(\nabla \delta c, \nabla v_2) + (\lambda_s'(c)\delta c, v_2) - (\delta w, v_2) \end{pmatrix},$$

where $\lambda_s'(c_s)$ is defined as

$$\lambda_s'(c_s) = \begin{cases} 0 & \textit{if } |c_s| \le 1, \\ s & \textit{if } |c_s| > 1. \end{cases}$$

Proof. Follows from [HIK03] and Sobolev embedding, see [HHT11, Lem. 5.3]. □

For applying Theorem A4 we further need that G_s is invertible. This we establish next.

Lemma 6.2. *For given $c_s \in H^1(\Omega), y_{old} \in H_0^1(\Omega)^d$ and $(y_1, y_2) \in H^1(\Omega)^* \times H^1(\Omega)^*$ the optimization problem*

$$\begin{aligned} \min_{(\delta c,\delta p)\in H^1(\Omega)\times V_0} \ & \frac{\gamma^2}{2}\|\nabla \delta c\|^2 + \frac{\tau}{2Pe}\|\nabla \delta p\|^2 + (\lambda_s'(c_s)\delta c, \delta c) - \langle y_2, \delta c\rangle \\ \textit{s.t.} \ & (\delta c, \nabla v) + \frac{\tau}{Pe}(\nabla \delta p, \nabla v) - \tau(\delta c y_{old}, \nabla v) = \langle y_1, v\rangle \quad \forall v \in H^1(\Omega) \end{aligned} \tag{$\mathcal{P}_{G_s}$}$$

admits a unique solution. Moreover, there exists a unique $\delta w \in H^1(\Omega)$ such that

$$\frac{\tau}{Pe}(\nabla \delta w, \nabla v) + (\delta c, v) - \tau(\delta c y_{old}, \nabla v) = \langle y_1, v\rangle\,, \tag{6.4}$$

$$\gamma^2(\nabla \delta c, \nabla v) + (\lambda_s'(c_s)\delta c, v) - (\delta w, v) = \langle y_2, v\rangle \tag{6.5}$$

for all $v \in H^1(\Omega)$.

If $(\delta c, \delta w)$ *is a solution of* (6.4)–(6.5)*, then* $(\delta c, \delta p)$ *with* $\delta p = \delta w - (\delta w, 1)$ *is the unique solution of* $(\mathcal{P}_{G_s})$*.*

Proof. One proceeds as in the proofs of Theorem 4.2 and Theorem 5.2. □

Theorem 6.3 ([HHT11, Prop. 5.5])**.** *The Newton iteration*

$$(c^{k+1}, w^{k+1}) = (c^k, w^k) - G_s(c^k, w^k)^{-1} F_s(c^k, w^k)$$

converges superlinearly to the solution (c_s, w_s) *of* (5.2)–(5.3)*, provided that the initial value* (c^0, w^0) *is sufficiently close to* (c_s, w_s)*, the time step size* τ *is sufficiently small, and* $y_{old} \in H^2(\Omega)^d \subset L^\infty(\Omega)$*.*

Proof. From Lemma 6.2 we deduce that G_s is invertible. This means that for given $(y_1, y_2) \in H^1(\Omega)^* \times H^1(\Omega)^*$, there exists a unique pair $(\delta c, \delta w) \in H^1(\Omega) \times H^1(\Omega)$ such that (6.4)–(6.5) holds.

We now show the boundedness of $\|G_s(c_s, w_s)^{-1}\|_{\mathcal{L}((H^1(\Omega)^2)^*, H^1(\Omega)^2)}$ independently of c_s and w_s to apply Theorem A4.

We take δw as test function in (6.4), δc as test function in (6.5) and add the resulting equations to obtain

$$\begin{aligned}\frac{\tau}{\mathrm{Pe}}\|\nabla \delta w\|^2 + \gamma^2 \|\nabla \delta c\|^2 + (\lambda_s'(c)\delta c, \delta c)& \\ =& \tau(\delta c y_{old}, \nabla \delta w) + \langle y_1, \delta w\rangle + \langle y_2, \delta c\rangle \\ &\le \tau \|\delta c\| \|y_{old}\|_{L^\infty(\Omega)} \|\nabla \delta w\| \\ &+ \|y_1\|_{H^1(\Omega)^*} \|\delta w\|_{H^1(\Omega)} + \|y_2\|_{H^1(\Omega)^*} \|\delta c\|_{H^1(\Omega)}.\end{aligned}$$

By Poincaré–Friedrichs inequality we have

$$\begin{aligned}\|\delta c\|_{H^1(\Omega)} &\le (1 + C_p)\|\nabla \delta c\| + C_p(\delta c, 1), \\ \|\delta w\|_{H^1(\Omega)} &\le (1 + C_p)\|\nabla \delta w\| + C_p(\delta w, 1),\end{aligned}$$

and by using $v \equiv 1$ in (6.4) and (6.5) we obtain

$$\begin{aligned}(\delta c, 1) =& \langle y_1, 1\rangle, \\ (\delta w, 1) =& (\lambda_s'(c)\delta c, 1) - \langle y_2, 1\rangle \le s|\langle y_1, 1\rangle| + |\langle y_2, 1\rangle|.\end{aligned}$$

Young's inequality now implies

$$\begin{aligned}\frac{\tau}{4\mathrm{Pe}}\|\nabla \delta w\|^2 + &\frac{1}{2}\left(\gamma^2 - \tau \mathrm{Pe} C_p^2 \|y_{old}\|^2_{L^\infty(\Omega)}\right)\|\nabla \delta c\|^2 \\ &\le \frac{\mathrm{Pe}}{\tau}(1 + C_p)^2 \|y_1\|^2_{(H^1(\Omega))^*} \\ &\quad + C_p \|y_1\|^2_{(H^1(\Omega))^*}(s|\langle y_1, 1\rangle| + |\langle y_2, 1\rangle|) \\ &\quad + \frac{1}{2\gamma^2}(1 + C_p)^2 \|y_2\|^2_{(H^1(\Omega))^*} \\ &\quad + C_p \|y_2\|^2_{(H^1(\Omega))^*}\langle y_1, 1\rangle.\end{aligned}$$

By using Hölder's inequality we end up with

$$\|(\delta c, \delta w)\|_{H^1(\Omega)\times H^1(\Omega)} \leq C \left(\|y_1\|_{H^1(\Omega)^*} + \|y_2\|_{H^1(\Omega)^*}\right).$$

Thus we obtain boundness of $\|G_s(c_s, w_s)^{-1}\|_{\mathcal{L}((H^1(\Omega)^2)^\star, H^1(\Omega)^2)}$ uniform in c_s and w_s. Now F_s together with its Newton derivative G_s fulfills Theorem A4 yielding superlinear convergence of the Newton sequence $\{c^k, w^k\}$. Note that we here require $\gamma^2 - \tau \mathrm{Pe} C_p^2 \|y_{old}\|^2_{L^\infty(\Omega)} > 0$, implying a restriction on the size of the time step τ. □

Remark 6.4. The restriction on τ can be interpreted in the sense that it restricts the distance that a particle can move during one time step to a value smaller then the thickness of the interface, which is of order $\mathcal{O}(\gamma)$. This restriction will later be reflected in our adaptive scheme. Note that this restriction does not appear if one replaces $\nabla c y_{old}$ by $\nabla c_{old} y_{old}$ in our scheme (5.18)–(5.19).

7 Finite element discretization

For the purpose of a numerical simulation of the Cahn–Hilliard Navier–Stokes system we next discretize (5.16)–(5.19) using the finite element method. Therefore we introduce shape regular simplicial meshes $\mathcal{T}^{cw}$ and $\mathcal{T}^{yp}$ such that $\overline{\Omega} = \bigcup_{T\in\mathcal{T}^{cw}} T$ and $\overline{\Omega} = \bigcup_{T\in\mathcal{T}^{yp}} T$. Here T are closed triangles or simplices. By $\mathcal{E}^{cw}$ and $\mathcal{E}^{yp}$ we denote the sets of faces associated with $\mathcal{T}^{cw}$ and $\mathcal{T}^{yp}$, respectively. For a triangle $T \in \mathcal{T}^{cw}$ we denote by h_T the diameter of T and by $|T|$ its area. For a face $E \in \mathcal{E}^{cw}$ we denote by h_E its length. We define $h = \max_{T\in\mathcal{T}^{cw}} h_T$.

The phase-field c and the potential w are discretized with piecewise linear, continuous finite elements, i.e. their Ansatz space is given by

$$\mathcal{V}^{cw} = \{v \in C^0(\overline{\Omega}) \,:\, v|_T \in P_1(T), \forall T \in \mathcal{T}^{cw}\} =: \mathrm{span}\{\phi_1^{cw}, \ldots, \phi_{N_{cw}}^{cw}\}.$$

The velocity y of the fluid and the pressure p are approximated by the LBB-stable Taylor–Hood finite element defined on $\mathcal{T}^{yp}$, i.e. we set

$$\mathcal{V}^{y} = \{v \in C^0(\overline{\Omega}) \,:\, v|_T \in P_2(T), \forall T \in \mathcal{T}^{yp}, v|_{\partial\Omega} = 0\} =: \mathrm{span}\{\phi_1^{y}, \ldots, \phi_{N_y}^{y}\},$$

and

$$\mathcal{V}^{p} = \{v \in C^0(\overline{\Omega}) \,:\, v|_T \in P_1(T), \forall T \in \mathcal{T}^{yp}\} =: \mathrm{span}\{\phi_1^{p}, \ldots, \phi_{N_p}^{p}\},$$

see [Ver84]. Here $P_k(T)$ stands for the space of polynomials up to degree k defined on T.

The spatially discretized version of (5.16)–(5.19) then consists of finding $(c_s^h, w_s^h) \in \mathcal{V}^{cw} \times \mathcal{V}^{cw}$ and $(y^h, p^h) \in \mathcal{V}^y \times \mathcal{V}^p$ such that the following system is satisfied:

$$B((y^h, p^h), (v, q)) = L^h((v, q)) \qquad \forall (v, q) \in \mathcal{V}^y \times \mathcal{V}^p, \tag{7.1}$$

$$\left\langle F^{(1)}(c_s^h, w_s^h), v \right\rangle = 0 \qquad \forall v \in \mathcal{V}^{cw}, \tag{7.2}$$

$$\left\langle F^{(2)}(c_s^h, w_s^h), v \right\rangle = 0 \qquad \forall v \in \mathcal{V}^{cw}. \tag{7.3}$$

Here B for $(v,q) \in H_0^1(\Omega)^d \times L^2_{(0)}(\Omega)$ is defined by

$$\begin{aligned} B((y^h,p^h),(v,q)) :=& \xi(y^h,v) + \eta(\nabla y^h, \nabla v) + a^t(y_{old}, y^h, v) \\ & - (\operatorname{div} v, p^h) + (\operatorname{div} y^h, q), \end{aligned}$$

while L^h is given by

$$L^h((v,q)) := \xi(y_{old}, v) - K(c^h \nabla w^h, v).$$

For $v \in H^1(\Omega)$ we have

$$\left\langle F^{(1)}(c_s^h, w_s^h), v \right\rangle := \frac{\tau}{\mathrm{Pe}}(\nabla w_s^h, \nabla v) + (c_s^h - c_{old}, v) - \tau(c_s^h y_{old}, \nabla v), \tag{7.4}$$

$$\left\langle F^{(2)}(c_s^h, w_s^h), v \right\rangle := \gamma^2(\nabla c_s^h,\, \nabla v) + (\lambda_s(c_s^h), v) - (w_s^h, v) - (c_{old}, v). \tag{7.5}$$

Every step of the semi-smooth Newton method for solving

$$\left(F^{(1)}(c_s^h, w_s^h), F^{(2)}(c_s^h, w_s^h)\right)^t = 0$$

then requires to solve the following system for given c_s^h, w_s^h:

$$\begin{aligned} \frac{\tau}{\mathrm{Pe}}(\nabla \delta w^h, \nabla v) + (\delta c^h, v) - \tau(\delta c^h y_{old}, \nabla v) &= -\left\langle F^{(1)}(c_s^h, w_s^h), v \right\rangle, \\ \gamma^2(\nabla \delta c^h, \nabla v) + (\lambda_s'(c_s^h)\delta c^h, v) - (\delta w^h, v) &= -\left\langle F^{(2)}(c_s^h, w_s^h), v \right\rangle. \end{aligned}$$

Using matrix notation this reads

$$\begin{pmatrix} A & -M \\ M - \tau T & D \end{pmatrix} \begin{pmatrix} \delta w \\ \delta c \end{pmatrix} = \begin{pmatrix} B_2 \\ B_1 \end{pmatrix}. \tag{7.6}$$

Here $\delta w, \delta w \in \mathbb{R}^N$ are the node vectors of $\delta w^h, \delta c^h \in \mathcal{V}^{cw}$ and the matrices are given by

$$\begin{aligned} A &= \gamma^2 K + \Lambda(c_s^h), & D &= \frac{\tau}{\mathrm{Pe}} K, \\ K &= (\nabla \phi_i^{cw}, \nabla \phi_j^{cw})_{i,j=1}^{N_{cw}}, & \Lambda(c_s^h) &= (\lambda_s'(c_s^h)\phi_j^{cw}, \phi_i^{cw})_{i,j=1}^{N_{cw}}, \\ M &= (\phi_i^{cw}, \phi_j^{cw})_{i,j=1}^{N_{cw}}, & T &= (\phi_i^{cw} y_{old}, \nabla \phi_j^{cw})_{i,j=1}^{N_{cw}}, \end{aligned}$$

while the right hand side is given by

$$B_2 = -\left\langle F^{(2)}(c_s^h, w_s^h), \phi_j^{cw} \right\rangle_{j=1}^{N_{cw}}, \quad B_1 = -\left\langle F^{(1)}(c_s^h, w_s^h), \phi_j^{cw} \right\rangle_{j=1}^{N_{cw}}.$$

Note that $\Lambda(c_s^h)$ is evaluated exactly, and is symmetric and positive semi definite.

We next show the feasibility of the semi smooth Newton method for solving the time and space discrete system (7.2)–(7.3).

Theorem 7.1 ([HHT11, Prop. 6.1]). *Let $\tau > 0$ be sufficiently small. Then the system* (7.6) *admits a unique solution, i.e. the system matrix* (7.6) *is regular.*

Proof. Since the mass matrix M is regular, and symmetric and positiv definite, one readily finds that (7.6) is equivalent to

$$\begin{pmatrix} -M & A \\ 0 & M - \tau T + DM^{-1}A \end{pmatrix} \begin{pmatrix} \delta c \\ \delta w \end{pmatrix} = \begin{pmatrix} B_2 \\ B_1 - DM^{-1}B_2 \end{pmatrix}.$$

It now is sufficient to show that $S = M - \tau T + DM^{-1}A$ is regular. We use the fact that the product UV of two symmetric matrices U and V with all Eigenvalues in $[u_1, u_2]$ and $[v_1, v_2]$ with $0 \leq u_1 \leq u_2$ and $0 \leq v_1 \leq v_2$, respectively, has all its Eigenvalues in $[u_1v_1, u_2v_2]$. From this we obtain that $M^{-1}DM^{-1}A$ is positive semi definite and thus

$$R = M + DM^{-1}A = M(I + M^{-1}DM^{-1}A)$$

is positive definite, where I denotes the identity matrix of size $N_{cw} \times N_{cw}$. Since the set of regular matrices is open and T is singular, since $(1, \ldots, 1)^t \in \mathbb{R}^{N_{cw}} \in \ker(T)$, also $R - \tau T$ is regular, for τ small enough □

As the solution of the continuous problem also the solution of the discrete problem is bounded in $H^1(\Omega) \times H^1(\Omega)$ independently of s .

Theorem 7.2 ([HHT11, Prop. 6.2])**.** *Let $(c_s^h, w_s^h)_{s>0}$ be a sequence of solutions to (7.2)–(7.3) for $s \to \infty$. Then there exists a constant $C > 0$ independent of s and h such that*

$$\|c_s^h\|_{H^1(\Omega)} \leq C, \tag{7.7}$$

$$\|w_s^h\|_{H^1(\Omega)} \leq C, \tag{7.8}$$

$$\|\lambda_s(c_s^h)\|_{L^2(\Omega)} \leq C \tag{7.9}$$

holds.

Proof. The proof is analogeous to the one of Lemma 5.3 and Lemma 5.5. Since in [HHT11] a different finite element formulation of the term involving λ_s is used, for the proof we do not follow [HHT11, Prop. 6.2].

We introduce the minimization problem

$$\min_{(c_s^h, p_s^h) \in \mathcal{V}^{cw} \times \mathcal{V}^{cw} \cap V_0} J^s(c_s^h, p_s^h) \text{ s.t. } (7.2). \tag{7.10}$$

This is the finite dimensional analogue to problem $(\mathcal{P}^s)$.

By the same arguments as in Theorem 5.2 we obtain that the unique solution to (7.10) is $(c_s^h, w_s^h - (w_s^h, 1)) = (c_s^h, p_s^h)$.

Let (c^*, p^*) denote the solution to $(\mathcal{P})$. By $P_h c^*$ we denote the H^1 orthogonal projection of c^* onto the non-empty, closed and convex subset $\mathcal{K} \cap \mathcal{V}^{cw} \subset H^1(\Omega)$. Thus there holds

$$(c^* - P_h c^*, v - P_h c^*)_{H^1} \leq 0 \quad \forall v \in \mathcal{V}^{cw} \cap \mathcal{K}.$$

By $Q_h p^*$ we denote the H^1 orthogonal projection of p^* onto $\mathcal{V}^{cw}$. Note that there holds $0 = (p^*, 1) = (Q_h p^*, 1)$. Since both are orthogonal projections we have the stability properties $\|P_h c^*\|_{H^1(\Omega)} \leq \|c^*\|_{H^1(\Omega)}$ and $\|Q_h w^*\|_{H^1(\Omega)} \leq \|Q_h w^*\|_{H^1(\Omega)}$, see e.g. [EG04, Lem. 1.131]. Inserting $P_h c^*$ and $P_h p^*$ in J_s we get

$$\begin{aligned}
J_s(c_s^h, p_s^h) \leq & J_s(P_h c^\star, Q_h p^\star) = J(P_h c^\star, Q_h p^\star)\\
=& \frac{\gamma^2}{2}\|\nabla P_h c^*\|^2 + \frac{\tau}{2\mathrm{Pe}}\|\nabla Q_h p^*\|^2 - (P_h c^*, c_{old})\\
\leq & \frac{\gamma^2}{2}\|\nabla P_h c^*\|^2 + \frac{\tau}{2\mathrm{Pe}}\|\nabla Q_h p^*\|^2 + \frac{\gamma^2}{2}\|P_h c^*\|^2 + \frac{1}{2\gamma^2}\|c_{old}\|^2\\
\leq & \frac{\gamma^2}{2}\|c^*\|^2_{H^1(\Omega)} + \frac{\tau}{2\mathrm{Pe}}\|p^*\|^2_{H^1(\Omega)} + \frac{1}{2\gamma^2}\|c_{old}\|^2\\
\leq & C.
\end{aligned}$$

Since $(c_s^h, 1) = (p_s^h, 1) = 0$, we obtain (7.7) and also $\|p_s^h\|_{H^1(\Omega)} \leq C$. As in Lemma 5.5 we get $(w_s^h, 1) \leq C$ and thus there follows (7.8).

Now we show the boundedness with respect to s of $\lambda_s(c_s^h)$ in $L^2(\Omega)$. For this we test (7.3) with $v = \lambda_s^+(c_s^h) = s\max(0, c_s^h - 1)$ and $v = \lambda_s^-(c_s^h) = s\min(0, c_s^h + 1)$ and obtain

$$\begin{aligned}
\|\lambda_s^+(c_s^h)\|^2 + \gamma^2 s^{-1}\|\nabla\lambda_s^+(c_s^h)\|^2 &= (w_s^h, \lambda_s^+(c_s^h)) + (c_{old}, \lambda_s^+(c_s^h)),\\
\|\lambda_s^-(c_s^h)\|^2 + \gamma^2 s^{-1}\|\nabla\lambda_s^-(c_s^h)\|^2 &= (w_s^h, \lambda_s^-(c_s^h)) + (c_{old}, \lambda_s^-(c_s^h)).
\end{aligned}$$

This yields

$$\begin{aligned}
\|\lambda_s^+(c_s^h)\| &\leq \|w_s^h\| + \|c_{old}\|,\\
\|\lambda_s^-(c_s^h)\| &\leq \|w_s^h\| + \|c_{old}\|.
\end{aligned}$$

From this the estimate

$$\|\lambda_s(c_s^h)\| \leq \|\lambda_s^+(c_s^h)\| + \|\lambda_s^-(c_s^h)\| \leq 2(\|w_s^h\| + \|c_{old}\|) \leq C$$

follows. □

The discretization of the Navier–Stokes part (7.1) gives rise to a saddle point problem often considered in literature, see e.g. [Ver84, BGL05, DGSW10], namely

$$\begin{pmatrix} A & B^t \\ B & 0 \end{pmatrix}\begin{pmatrix} y \\ p \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix}. \tag{7.11}$$

The matrices are given as

$$\begin{aligned}
A &= \begin{pmatrix} A_{11} & 0 \\ 0 & A_{22} \end{pmatrix}, \quad A_{11} = A_{22} = (a_{ij})_{i,j=1\ldots N_y},\\
B &= \begin{pmatrix} B_1 & B_2 \end{pmatrix}, \quad B_1 = ((b_1)_{ij})_{i=1,\ldots,N_p}^{j=1,\ldots,N_y}, \quad B_2 = ((b_2)_{ij})_{i=1,\ldots,N_p}^{j=1,\ldots,N_y},
\end{aligned}$$

with

$$\begin{aligned} a_{ij} =& \xi(\phi_i^y, \phi_j^y) + \eta(\nabla \phi_j^y, \nabla \phi_i^y) + a^t(y_{old}, \phi_i^y, \phi_j^y), \\ (b_1)_{ij} =& -(\partial_x \phi_j^y, \phi_i^p), \\ (b_2)_{ij} =& -(\partial_y \phi_j^y, \phi_i^p). \end{aligned}$$

Here y denotes the node vector for the velocity field y^h and p denotes the node vector for the pressure field p^h. The right hand side is given by

$$\begin{aligned} (f_1)_i &= \xi((y_1)_{old}, \phi_i^y) - K(c_s^h \partial_x w_s^h, \phi_i^y), \\ (f_2)_i &= \xi((y_2)_{old}, \phi_i^y) - K(c_s^h \partial_y w_s^h, \phi_i^y). \end{aligned}$$

Here $(y_1)_{old}$, resp. $(y_2)_{old}$, refers to the first, resp. second, component of the vector field y_{old}.

We show the existence of a unique solution to (7.1) following the proof of [Jus11, Satz 3.8].

Theorem 7.3 ([Jus11, Satz 3.8]). *There exists a unique solution* $(y^h, p^h) \in V^y \times V^p$ *to* (7.1) *with* $\int_\Omega p_h \, dx = 0$.

Proof. Since we have a finite dimensional linear equation, it is sufficient to show that the homogenous equation only has the trivial solution. Thus we show that

$$B((y^h, p^h), (v^h, q^h)) = 0 \qquad \forall (v^h, q^h) \in \mathcal{V}^y \times \mathcal{V}^p \tag{7.12}$$

only has the unique solution $(y^h, p^h) = 0$.
Testing with $(v^h, q^h) \equiv (y^h, p^h)$ we obtain

$$0 = B((y^h, p^h), (y^h, p^h)) = \xi \|y^h\|^2 + \eta \|\nabla y^h\|^2,$$

and thus $y^h = 0$ in $H_0^1(\Omega)^d$. Now it immediately follows that

$$0 = (\operatorname{div} v^h, p^h) \, \forall v^h \in \mathcal{V}_h^y. \tag{7.13}$$

Since we use LBB-stable spaces there exists $\beta^h > 0$ such that

$$\sup_{v^h \in \mathcal{V}^y} \frac{(\operatorname{div} v^h, p^h)}{\|\nabla v^h\|} \geq \beta^h \|p^h\|$$

holds, see [Ver10]. The constant β^h is independent of h. Equation (7.13) also holds if we divide by $\|\nabla v^h\|$ and it also holds for the supremum over all v^h:

$$0 = \sup_{v^h \in \mathcal{V}^y} \frac{(\operatorname{div} v^h, p^h)}{\|\nabla v^h\|} \geq \beta^h \|p^h\|.$$

Thus $\|p^h\| = 0$ and we obtain $p^h = 0$ in $L^2(\Omega)$. □

8 The adaptive concept

After having a fully discretized scheme for the solution of (4.1)–(4.4) at hand we can solve (7.1)–(7.3) at each time step in a time marching scheme.

Since the phase field is known to be nearly constant in the pure phases and that it describes the interface of width $\mathcal{O}(\gamma)$ between the two fluid phases one expects large gradients of the phase field in the neighborhood of the interface.

Therefore one should carefully select an appropriate mesh for the simulation of the phase field. In Section 8.1 we desribe a reliable and efficient a posteriori error estimator for the simulation of the phase field, based on the system (7.2)–(7.3). It is constructed along the lines of [HHT11, Sec. 7].

On the other hand a mesh taylored to the numerical simulation of the phase field need not be a good choice for the simulation of the flow field. In Section 8.2 we briefly describe how the results from [Ver10] concerning an a posteriori error estimator for the time-dependend Navier–Stokes system carry over to our case, where we again follow [Jus11].

In Section 8.3 we describe the adaptive cycles which we use together with the error estimators obtained in Section 8.1 and Section 8.2. We finish this section with describing aspects of the implementation arising from discretizing the flow field and the concentration field on different meshes in Section 8.4.

For ease of notation and since all references should be clear, in the following we suppress the index $_s$ for denoting the solution of the relaxed system. So, in what follows, we write (y, p, c, w) as solution of the time-discrete and Moreau–Yosida relaxed system (5.16)–(5.19) and by (y^h, p^h, c^h, w^h) we denote the solution of the discrete in space system (7.1)–(7.3).

8.1 Adaptive concept for the Cahn–Hilliard part

In the present section we derive a reliable and efficient error estimator for the Cahn–Hilliard system (7.2)–(7.3). Here we extend results of [HHT11, Sec. 7] to the case of the Cahn–Hilliard system with transport.

Let us briefly comment on further available concepts for the spatial discretization of the Cahn–Hilliard system in the literature.

In [BN09] a reliable estimator for the Cahn–Hilliard equation with double-obstacle free energy is derived. Besides residual based estimators, taylored heuristical approaches are commonly used to resolve the interface numerically. These approaches exploit the fact, that the location of the interface is known and adapt the mesh accordingly. The approaches distinguish in the way they localize the interface. In [KSW08, BBG11, AV12] a triangle T is refined if it is located in the diffuse interface, i.e. if $\max_{x\in T} |c(x)| \leq 1 - \delta$ with a small $\delta > 0$ holds.

On the other hand, since the interface can be characterized by its large concentration gradients, in [GK14] heuristic error estimation based on the norm of the gradient of c is used. We will give a short comparison to heuristic error estimation in the numerical part in Section 9.1.

We define the following errors

$$e_c := c^h - c, \qquad e_w := w^h - w, \qquad e_{\lambda_s} := \lambda_s(c^h) - \lambda_s(c),$$

residuals

$$\begin{aligned} r^{(1)} &:= c - c_{old} + \tau \nabla c y_{old}, & r^{(2)} &:= \lambda_s(c) - w - c_{old}, \\ r_h^{(1)} &:= c^h - c_{old} + \tau \nabla c^h y_{old}, & r_h^{(2)} &:= \lambda_s(c^h) - w^h - c_{old}, \end{aligned}$$

element indicators

$$\eta_T^{(1)} = h_T \| r_h^{(1)} \|_T \qquad \text{for all } T \in \mathcal{T}^{cw}, \tag{8.1}$$

$$\eta_T^{(2)} = h_T \| r_h^{(2)} \|_T \qquad \text{for all } T \in \mathcal{T}^{cw}, \tag{8.2}$$

and edge indicators

$$\begin{aligned} \eta_E^{(1)} &= h_E^{1/2} \| [\nabla w^h]_E \cdot \nu_E \|_E \qquad \text{for all } E \in \mathcal{E}^{cw}, \\ \eta_E^{(2)} &= h_E^{1/2} \| [\nabla c^h]_E \cdot \nu_E \|_E \qquad \text{for all } E \in \mathcal{E}^{cw}, \end{aligned}$$

where ν_E, for all $E \in \mathcal{E}^{cw}$, denotes the outer unit normal on the edge E, pointing from the triangle with lower global number to the triangle with higher global number. If E is a boundary edge, then ν_E coincides with the outer normal ν_Ω. With $[\cdot]_E$ we denote the jump of the respective function across the edge E.

Further, to each function $f \in L^1(\Omega)$ we assign a piecewise constant function $\overline{f}$ defined by

$$\overline{f}_{|T} = \frac{1}{|T|} \int_T f \, dx \qquad \text{for } T \in \mathcal{T}^{cw}. \tag{8.3}$$

The local as well as the 'regional' data oscillations associated with a function f are defined as

$$\begin{aligned} \operatorname{osc}_h(f, T) &= \| h_T (f - \overline{f}) \|_{L^2(T)} && \text{for } T \in \mathcal{T}^{cw}, \\ \operatorname{osc}_h(f, D) &= \left(\sum_{T \in D} \operatorname{osc}_h(f, T)^2 \right)^{1/2} && \text{for } D \subset \mathcal{T}^{cw}. \end{aligned}$$

By $\Pi_h : H^1(\Omega) \to \mathcal{V}^{cw}$, we denote Clément's interpolation operator ([Clé75, EG04]), which satisfies for each $T \in \mathcal{T}^{cw}$ and $E \in \mathcal{E}^{cw}$

$$\| v - \Pi_h v \|_T \leq C h_T \| \nabla v \|_{\omega_T} \qquad \forall v \in H^1(\Omega), \tag{8.4}$$

$$\| v - \Pi_h v \|_T \leq C h_E^{1/2} \| \nabla v \|_{\omega_E} \qquad \forall v \in H^1(\Omega). \tag{8.5}$$

Here, the domains ω_T and ω_E are given by

$$\omega_T := \{ T' \in \mathcal{T}^{cw} : T \cap T' \neq \emptyset \} \text{ and } \omega_E := \{ T \in \mathcal{T}^{cw} : E \subset T \}.$$

8.1.1 Reliability of the estimator – a posteriori upper bound

For all v in $H^1(\Omega)$, we have

$$\left\langle F_s^{(1)}(c,w),v\right\rangle = \left\langle F_s^{(2)}(c,w),v\right\rangle = 0.$$

This yields

$$\left\langle F_s^{(1)}(c^h,w^h),e_w\right\rangle = \left\langle F_s^{(1)}(c^h,w^h) - F_s^{(1)}(c,w),e_w\right\rangle, \tag{8.6}$$

$$\left\langle F_s^{(2)}(c^h,w^h),e_c\right\rangle = \left\langle F_s^{(2)}(c^h,w^h) - F_s^{(2)}(c,w),e_c\right\rangle, \tag{8.7}$$

which implies

$$\left\langle F_s^{(1)}(c^h,w^h),e_w\right\rangle = \frac{\tau}{Pe}(\nabla e_w,\nabla e_w) + (e_c,e_w) - \tau(e_c y_{old},\nabla e_w), \tag{8.8}$$

$$\left\langle F_s^{(2)}(c^h,w^h),e_c\right\rangle = \gamma^2(\nabla e_c,\nabla e_c) - (e_w,e_c) + (\lambda_s(c^h)-\lambda_s(c),e_c). \tag{8.9}$$

From

$$(\max(0,a)-\max(0,b))(a-b) \geq (\max(0,a)-\max(0,b))^2,$$
$$(\min(0,a)-\min(0,b))(a-b) \geq (\min(0,a)-\min(0,b))^2,$$

for all $a,b\in\mathbb{R}$ we have

$$(\lambda_s(c^h)-\lambda_s(c),e_c) \geq s^{-1}\|e_{\lambda_s}\|^2. \tag{8.10}$$

Hence, adding (8.8) and (8.9), and using (8.10), we obtain

$$\mathcal{E} \leq \mathcal{E}_1 + \mathcal{E}_2 + \mathcal{E}_3, \tag{8.11}$$

with

$$\begin{aligned}
\mathcal{E} &:= s^{-1}\|e_{\lambda_s}\|^2 + \frac{\tau}{Pe}\|\nabla e_w\|^2 + \gamma^2\|\nabla e_c\|^2,\\
\mathcal{E}_1 &:= \left\langle F_s^{(1)}(c^h,w^h),e_w\right\rangle,\\
\mathcal{E}_2 &:= \left\langle F_s^{(2)}(c^h,w^h),e_c\right\rangle,\\
\mathcal{E}_3 &:= \tau(e_c y_{old},\nabla e_w).
\end{aligned}$$

We further estimate $\mathcal{E}_i$, $i=1,2$. For this purpose, we recall that for all v^h in $\mathcal{V}^{cw}$

$$\left\langle F_s^{(1)}(c^h,w^h),v^h\right\rangle = \left\langle F_s^{(2)}(c^h,w^h),v^h\right\rangle = 0,$$

holds, which implies

$$\begin{aligned}
\mathcal{E}_1 &= \left\langle F_s^{(1)}(c^h,w^h),e_w\right\rangle = \left\langle F_s^{(1)}(c^h,w^h),e_w-\Pi_h e_w\right\rangle,\\
\mathcal{E}_2 &= \left\langle F_s^{(2)}(c^h,w^h),e_c\right\rangle = \left\langle F_s^{(2)}(c^h,w^h),e_c-\Pi_h e_c\right\rangle.
\end{aligned}$$

Let

$$\mathcal{E}_1 = \mathcal{E}_1^a + \mathcal{E}_1^b, \quad \mathcal{E}_2 = \mathcal{E}_2^a + \mathcal{E}_2^b,$$

where

$$\begin{aligned} \mathcal{E}_1^a &:= \frac{\tau}{Pe}\left(\nabla w_h, \nabla(e_w - \Pi_h e_w)\right), & \mathcal{E}_2^a &:= \gamma^2\left(\nabla c_h, \nabla(e_c - \Pi_h e_c)\right), \\ \mathcal{E}_1^b &:= (r_h^{(1)}, e_w - \Pi_h e_w), & \mathcal{E}_2^b &:= (r_h^{(2)}, e_c - \Pi_h e_c). \end{aligned}$$

Using integration by parts, (8.4)–(8.5), and the discrete Cauchy–Schwarz inequality, it thus follows that

$$\begin{aligned} \mathcal{E}_1^a &= \sum_{E\in\mathcal{E}^{cw}} \frac{\tau}{\mathrm{Pe}}([\nabla w^h]_E \cdot \nu_E, e_w - \Pi_h e_w)_E \\ &\le C\left(\left(\frac{\tau}{\mathrm{Pe}}\right)^2 \sum_{E\in\mathcal{E}^{cw}} (\eta_E^{(1)})^2\right)^{1/2} \|\nabla e_w\|, \\ \mathcal{E}_1^b &= \sum_{T\in\mathcal{T}^{cw}} (r_h^{(1)}, e_w - \Pi_h e_w)_T \le C\left(\sum_{T\in\mathcal{T}^{cw}} (\eta_T^{(1)})^2\right)^{1/2} \|\nabla e_w\|. \end{aligned}$$

Consequently, we infer

$$\mathcal{E}_1 := \mathcal{E}_1^a + \mathcal{E}_1^b \le C\left(\sum_{T\in\mathcal{T}^{cw}} (\eta_T^{(1)})^2 + \left(\frac{\tau}{Pe}\right)^2 \sum_{E\in\mathcal{E}^{cw}} (\eta_E^{(1)})^2\right)^{1/2} \|\nabla e_w\|. \quad (8.12)$$

In the same way, we find

$$\mathcal{E}_2 := \mathcal{E}_2^a + \mathcal{E}_2^b \le C\left(\sum_{T\in\mathcal{T}^{cw}} (\eta_T^{(2)})^2 + \gamma^4 \sum_{E\in\mathcal{E}^{cw}} (\eta_E^{(2)})^2\right)^{1/2} \|\nabla e_c\|. \quad (8.13)$$

Next we establish reliability of our a posteriori error estimator η_Ω which is defined below. We note that this result is similar to the one for a Cahn–Hilliard system obtained in [HHT11, Prop. 7.1]. Let us first note that testing (8.6) with $1 \in \mathcal{V}^{cw}$ yields $(e_c, 1) = 0$. Thus by Poincaré–Friedrichs inequality there holds $\|e_c\| \le C_p\|\nabla e_c\|$.

Theorem 8.1. *Let $c, w \in H^1(\Omega)$ denote the solution of system* (5.18)–(5.19), *and $c^h, w^h \in \mathcal{V}^{cw}$ denote the solution of system* (7.2)–(7.3). *Assume that $y_{old} \in L^\infty(\Omega)$ holds. Then the following holds:*

There exists a positive constant C, depending only on the domain Ω and the smallest angle of the mesh $\mathcal{T}^{cw}$, such that

$$s^{-1}\|e_{\lambda_s}\|^2 + \frac{\tau}{4Pe}\|\nabla e_w\|^2 + \frac{1}{2}\left(\gamma^2 - \tau Pe C_p^2 \|y_{old}\|^2_{L^\infty(\Omega)}\right)\|\nabla e_c\|^2 \le C\eta_\Omega^2, \quad (8.14)$$

holds, where

$$\begin{aligned}\eta_\Omega^2 = &\left(\frac{\tau}{Pe}\right)^{-1} \sum_{T\in\mathcal{T}^{cw}} \left(\eta_T^{(1)}\right)^2 + \gamma^{-2} \sum_{T\in\mathcal{T}^{cw}} \left(\eta_T^{(2)}\right)^2 \\ &+ \frac{\tau}{Pe} \sum_{E\in\mathcal{E}^{cw}} \left(\eta_E^{(1)}\right)^2 + \gamma^2 \sum_{E\in\mathcal{E}^{cw}} \left(\eta_E^{(2)}\right)^2 .\end{aligned} \tag{8.15}$$

Moreover

$$\eta_\Omega \leq \beta$$

with a constant β independent of s and h.

Thus, provided that

$$\gamma^2 > \tau PeC_p^2 \|y_{old}\|^2_{L^\infty(\Omega)}$$

holds, the estimator is reliable.

Proof. For proving (8.14) it remains to treat $\mathcal{E}_3$.
Straight forward estimation yields

$$\begin{aligned}\mathcal{E}_3 &= \tau(e_c y_{old}, \nabla e_w) \\ &\leq \tau\|e_c\|\|y_{old}\|_{L^\infty(\Omega)}\|\nabla e_w\| \\ &\leq \frac{1}{2}\tau\mathrm{Pe}\|y_{old}\|^2_{L^\infty(\Omega)}\|e_c\|^2 + \frac{\tau}{2\mathrm{Pe}}\|\nabla e_w\|^2 \\ &\leq \frac{1}{2}\tau\mathrm{Pe}C_p^2\|y_{old}\|^2_{L^\infty(\Omega)}\|\nabla e_c\|^2 + \frac{\tau}{2\mathrm{Pe}}\|\nabla e_w\|^2.\end{aligned}$$

This together with the estimates (8.12) and (8.13) gives

$$\begin{aligned}\mathcal{E} \leq & C_1 \left(\sum_{T\in\mathcal{T}^{cw}} \left(\eta_T^{(2)}\right)^2 + \gamma^4 \sum_{E\in\mathcal{E}^{cw}} \left(\eta_E^{(2)}\right)^2 \right)^{1/2} \|\nabla e_c\| \\ &+ C_2 \left(\sum_{T\in\mathcal{T}^{cw}} \left(\eta_T^{(1)}\right)^2 + \left(\frac{\tau}{Pe}\right)^2 \sum_{E\in\mathcal{E}^{cw}} \left(\eta_E^{(1)}\right)^2 \right)^{1/2} \|\nabla e_w\| \\ &+ \frac{1}{2}\tau\mathrm{Pe}C_p^2\|y_{old}\|^2_{L^\infty(\Omega)}\|\nabla e_c\|^2 + \frac{\tau}{2\mathrm{Pe}}\|\nabla e_w\|^2 \\ \leq & C_1\gamma^{-2} \left(\sum_{T\in\mathcal{T}^{cw}} \left(\eta_T^{(2)}\right)^2 + \gamma^4 \sum_{E\in\mathcal{E}^{cw}} \left(\eta_E^{(2)}\right)^2 \right) + \frac{\gamma^2}{2}\|\nabla e_c\|^2 \\ &+ C_2 \left(\frac{\tau}{\mathrm{Pe}}\right)^{-1} \left(\sum_{T\in\mathcal{T}^{cw}} \left(\eta_T^{(1)}\right)^2 + \left(\frac{\tau}{Pe}\right)^2 \sum_{E\in\mathcal{E}^{cw}} \left(\eta_E^{(1)}\right)^2 \right) + \frac{\tau}{4\mathrm{Pe}}\|\nabla e_w\|^2 \\ &+ \frac{1}{2}\tau\mathrm{Pe}C_p^2\|y_{old}\|^2_{L^\infty(\Omega)}\|\nabla e_c\|^2 + \frac{\tau}{2\mathrm{Pe}}\|\nabla e_w\|^2,\end{aligned}$$

where $C_1 > 0$, $C_2 > 0$ denote appropriate constants. This implies

$$s^{-1}\|e_{\lambda_s}\|^2 + \frac{\tau}{4\mathrm{Pe}}\|\nabla e_w\|^2 + \frac{1}{2}\left(\gamma^2 - \tau\mathrm{Pe}C_p^2\|y_{old}\|^2_{L^\infty(\Omega)}\right)\|\nabla e_c\|^2 \leq C\eta_\Omega^2$$

which is the claimed result.

Furthermore

$$\eta_\Omega \le \beta \left(\|c^h\|_{H^1(\Omega)} + \|w^h\|_{H^1(\Omega)} + \|\lambda_s(c^h)\|_{L^2(\Omega)} + 1 \right),$$

follows from the definition of η_Ω, where $\beta > 0$ is an appropriate constant. □

Remark 8.2.

- In [HHT11, HHK13] mass lumping is used for the numerical realization of (7.2)–(7.3). The inner product $(f,g) = \int_\Omega fg\,dx$ then is not evaluated exactly, but by the numerical quadrature rule

  $$(f,g)^h := \int_\Omega \pi^h(f(x)g(x))\,dx = \sum_{i=1}^{N_{cw}} (1,\phi_i^{cw}) f(x_i) g(x_i) \quad \forall f,g \in C(\overline{\Omega}),$$

  where $\pi^h : C(\overline{\Omega}) \to \mathcal{V}^{cw}$ denotes the Lagrange interpolation operator. This gives rise to a further error contribution $\eta_3 = \|\pi^h(\lambda_s(c^h)) - \lambda_s(c^h)\|$, to η_Ω, while the term $\eta_T^{(2)}$ modifies to $\tilde{\eta}_T^{(2)} = h_T\|\pi^h(\lambda_s(c^h)) - w^h - c_{old}\|$. This reduces the numerical effort for solving the corresponding equation systems, but introduces a new error contribution.

- For integrating $\lambda_s(c^h)$ exactly further effort only has to be spent on triangles with a discrete active-inactive interface, i.e. on

  $$\mathcal{I}(c^h) := \{T \in \mathcal{T}^{cw} \mid \max_{x_i \in T} |c^h(x_i)| > 1 \wedge \min_{x_i \in T} |c^h(x_i)| < 1\}.$$

  In our numerics this extra effort results in meshes with less degrees of freedom, compared to meshes obtained with lumping of $\lambda_s(c^h)$. The numerical speed up from this smaller meshes even compensates the increased numerical effort of exact evaluation. The influence of lumping is investigated in Section 9.1.

- The condition $\gamma^2 > \tau \mathrm{Pe} C_p^2 \|y_{old}\|^2_{L^\infty(\Omega)}$ can be interpreted as a restriction on the time step size τ. It has to be choosen so small that a particle can not cross the interface of width $\mathcal{O}(\gamma)$ in just one time step.

- If Ω is a convex domain $C_p \le \mathrm{diam}(\Omega)\pi^{-1}$ holds ([PW60]).

- A similiar error estimator can be derived for other free energies, assuming that the free energy is discretized according to [Eyr98], where an implicit discretization in time for the convex part and an explicit in time discretization of the concave part is proposed, see Section 11.3.

8.1.2 Efficiency of the estimator – a posteriori lower bound

For showing the efficiency of the estimator we use the bubble function technique as proposed in e.g. [AO00], to establish a lower bound on the error representation given in Theorem 8.1. By λ_T we denote the canonical bubble function of $T \in \mathcal{T}^{cw}$, which is the product of the barycentric coordinates of T. By λ_E we denote the bubble function corresponing to $E \in \mathcal{E}^{cw}$. We introduce the mapping

$$\widetilde{\ }: L^2(E) \longrightarrow L^2(\omega_E), \quad \widetilde{\Phi}(x) := \Phi(x_E) \quad x \in T,$$

which extends any function defined on an edge E to the pair of neighboring elements (T^+, T^-) with common edge E. Here $\omega_E := T^+ \cup T^-$. We have $T \in \{T^+, T^-\}$, and $x_E \in E$ is such that $x - x_E$ is parallel to a fixed $E' \in T\backslash\{E\}$.

Referring to [AO00], for all polynomial functions $\Phi_T \in P_k(T)$ and $\Phi_E \in P_k(E)$, $k \in \mathbb{N}$, the following estimates are valid:

$$\|\Phi_T\|_T^2 \leq C(\Phi_T, \Phi_T\lambda_T)_T \qquad \forall T \in \mathcal{T}^{cw}, \tag{8.16}$$
$$\|\Phi_T\lambda_T\|_T \leq \|\Phi_T\|_T \qquad \forall T \in \mathcal{T}^{cw}, \tag{8.17}$$
$$\|\nabla(\Phi_T\lambda_T)\|_T \leq Ch_T^{-1}\|\Phi_T\|_T \qquad \forall T \in \mathcal{T}^{cw}, \tag{8.18}$$
$$\|\Phi_E\|_E^2 \leq C(\Phi_E, \Phi_E\lambda_E)_E \qquad \forall E \in \mathcal{E}^{cw}, \tag{8.19}$$
$$\|\Phi_E\lambda_E\|_E \leq C\|\Phi_E\|_E \qquad \forall E \in \mathcal{E}^{cw}. \tag{8.20}$$

Furthermore, we have

$$\|\widetilde{\Phi_E\lambda_E}\|_{\omega_E} \leq Ch_E^{1/2}\|\Phi_E\|_E \qquad \forall E \in \mathcal{E}^{cw}, \tag{8.21}$$
$$\|\nabla(\widetilde{\Phi_E\lambda_E})\|_{\omega_E} \leq Ch_E^{-1/2}\|\Phi_E\|_E \qquad \forall E \in \mathcal{E}^{cw}. \tag{8.22}$$

We start with two auxiliary results.

Lemma 8.3. *For every $T \in \mathcal{T}^{cw}$ the following estimates hold*

$$\left(\frac{\tau}{Pe}\right)^{-1}(\eta_T^{(1)})^2 \leq C\left(\tau Pe h_T^2\|y_{old}\|_{\infty,T}^2\|\nabla e_c\|_T^2 + \left(\frac{\tau}{Pe}\right)^{-1} osc_h^2(r_h^{(1)}, T)\right.$$
$$\left. + \frac{\tau}{Pe}\|\nabla e_w\|_T^2 + \left(\frac{\tau}{Pe}\right)^{-1}\|h_T e_c\|_T^2\right) \tag{8.23}$$

and

$$\gamma^{-2}(\eta_T^{(2)})^2 \leq C\left(\gamma^2\|\nabla e_c\|_T^2 + \gamma^{-2}\|h_T e_w\|_T^2\right. \tag{8.24}$$
$$\left. + \gamma^{-2}\|h_T e_{\lambda_s}\|_T^2 + \gamma^{-2} osc_h^2(r_h^{(2)}, T)\right).$$

Proof. We have

$$(\eta_T^{(1)})^2 = \|h_T r_h^{(1)}\|_T^2 \leq 2h_T^2\|\overline{r}_h^{(1)}\|_T^2 + 2\mathrm{osc}_h^2(r_h^{(1)}, T), \tag{8.25}$$

with $\overline{r}_h^{(1)} := \overline{c} - \overline{c}_{old} + \tau\overline{y_{old}\nabla c}$. Note that while $r_h^{(1)}$ due to the appearance of y_{old} in general is not a piecewise polynomial on $\mathcal{V}^{cw}$, $\overline{r}_h^{(1)}$ is piecewise constant. We set $\psi_T := \overline{r}_h^{(1)}|_T\lambda_T$ and obtain with the help of (8.16)

$$\begin{aligned}\|\overline{r}_h^{(1)}\|_T^2 &\leq C(\overline{r}_h^{(1)}, \psi_T)_T \\ &\leq C(r_h^{(1)}, \psi_T)_T + Ch_T^{-1}\text{osc}_h(r_h^{(1)}, T)\|\psi_T\|_T. \qquad (8.26)\end{aligned}$$

Using $\Delta w^h|_T = 0$ and $c - c_{old} - \frac{\tau}{Pe}\Delta w + \tau\nabla c y_{old} = 0$ we proceed

$$\begin{aligned}(r_h^{(1)}, \psi_T)_T &= (c^h - c_{old} + \tau\nabla c^h y_{old}, \psi_T)_T \\ &= (c^h - c, \psi_T)_T - \frac{\tau}{Pe}(\Delta w^h - \Delta w, \psi_T)_T \\ &+ \tau(\nabla c^h y_{old} - \nabla c y_{old}, \psi_T)_T \\ &= (e_c, \psi_T)_T + \frac{\tau}{Pe}(\nabla e_w, \nabla\psi_T)_T + \tau(\nabla e_c y_{old}, \psi_T)_T. \qquad (8.27)\end{aligned}$$

Using (8.26) and (8.27) we have

$$\begin{aligned}\|\overline{r}_h^{(1)}\|_T^2 \leq & C\Big(\|e_c\|_T\|\psi_T\|_T + \frac{\tau}{Pe}\|\nabla e_w\|_T\|\nabla\psi_T\|_T + \tau(\nabla e_c y_{old}, \psi_T)_T \\ & + h_T^{-1}\text{osc}_h(r_h^{(1)}, T)\|\psi_T\|_T\Big) \\ \leq & C\Big(\|e_c\|_T\|\overline{r}_h^{(1)}\|_T + \frac{\tau}{Pe}\|\nabla e_w\|_T h_T^{-1}\|\overline{r}_h^{(1)}\|_T \\ & + \tau\|\nabla e_c y_{old}\|\|\overline{r}_h^{(1)}\|_T + h_T^{-1}\text{osc}_h(r_h^{(1)}, T)\|\overline{r}_h^{(1)}\|_T\Big)\end{aligned}$$

from which we conclude

$$\begin{aligned}\|\overline{r}_h^{(1)}\|_T \leq & C\Big(\|e_c\|_T + \frac{\tau}{\text{Pe}}h_T^{-1}\|\nabla e_w\|_T \\ & + \tau\|\nabla e_c\|_T\|y_{old}\|_{\infty,T} + h_T^{-1}\text{osc}_h(r_h^{(1)}, T)\Big). \qquad (8.28)\end{aligned}$$

Estimate (8.23) now follows from (8.25) and (8.28) using Young's inequality.

To achieve (8.24) we proceed similarly with

$$(\eta_T^{(2)})^2 = \|h_T r_h^{(2)}\|_T^2 \leq 2h_T^2\|\overline{r}_h^{(2)}\|_T^2 + 2\text{osc}_h^2(r_h^{(2)}, T), \qquad (8.29)$$

where $\overline{r}_h^{(2)} := \overline{\lambda_s(c^h)} - \overline{w}^h - \overline{c}_{old}$. With $\psi_T := \overline{r}_h^{(2)}|_T\lambda_T$ and (8.16) we get

$$\|\overline{r}_h^{(2)}\|_T^2 \leq C(\overline{r}_h^{(2)}, \psi_T)_T.$$

Furthermore

$$\|\overline{r}_h^{(2)}\|_T^2 \leq C(r_h^{(2)}, \psi_T)_T + Ch_T^{-1}\text{osc}_h(r_h^{(2)}, T)\|\psi_T\|_T. \qquad (8.30)$$

Since $\Delta c^h|_T = 0$ and $-\gamma^2 \Delta c + \lambda_s(c) - w - c_{old} = 0$, we have

$$(r_h^{(2)}, \psi_T)_T = \gamma^2 (\nabla e_c, \nabla \psi_T)_T + (e_{\lambda_s}, \psi_T)_T - (e_w, \psi_T)_T. \tag{8.31}$$

From (8.30)–(8.31) it follows that

$$\begin{aligned} \|\bar{r}_h^{(2)}\|_T^2 \leq C \big(&\gamma^2 \|\nabla e_c\|_T \|\nabla \psi_T\|_T \\ &+ \left(\|e_w\|_T + \|e_{\lambda_s}\|_T + h_T^{-1} \mathrm{osc}_h(r_h^{(2)}, T)\right) \|\psi_T\|_T \Big), \end{aligned}$$

and using (8.17) and (8.18) we obtain

$$\begin{aligned} \|\bar{r}_h^{(2)}\|_T \leq C \Big(&\gamma^2 h_T^{-1} \|\nabla e_c\|_T + \|e_w\|_T + \|e_{\lambda_s}\|_T \\ &+ h_T^{-1} \mathrm{osc}_h(r_h^{(2)}, T)\Big). \end{aligned} \tag{8.32}$$

Estimate (8.24) now follows from (8.29) and (8.32). □

Lemma 8.4. *For every $E \in \mathcal{E}^{cw}$ the following estimates hold*

$$\begin{aligned} \frac{\tau}{Pe} (\eta_E^{(1)})^2 \leq & C \left(\frac{\tau}{Pe} \|\nabla e_w\|_{\omega_E}^2 + \left(\frac{\tau}{Pe}\right)^{-1} \|h_T e_c\|_{\omega_E}^2 + \left(\frac{\tau}{Pe}\right)^{-1} osc_h^2(r_h^{(1)}, \omega_E) \right. \\ & \left. + \tau Pe \|h_T y_{old}\|_{\infty,\omega_E}^2 \|\nabla e_c\|_{\omega_E}^2 \right) \end{aligned} \tag{8.33}$$

and

$$\begin{aligned} \gamma^2 (\eta_E^{(2)})^2 \leq C \big(&\gamma^2 \|\nabla e_c\|_{\omega_E}^2 + \gamma^{-2} \|h_E e_w\|_{\omega_E}^2 + \gamma^{-2} \|h_E e_{\lambda_s}\|_{\omega_E}^2 \\ &+ \gamma^{-2} osc_h^2(r_h^{(2)}, \omega_E)\big). \end{aligned} \tag{8.34}$$

Proof. Let E be an arbitrary edge in $\mathcal{E}^{cw}$ and define $\psi_E := \widetilde{\Phi_E} \lambda_E$. For the proof of (8.33) we use $\Phi_E := [\nabla w^h]_E \cdot \nu_E$.
Due to (8.19) we have

$$\left(\eta_E^{(1)}\right)^2 = h_E \|[\nabla w^h]_E \cdot \nu_E\|_E^2 \leq C h_E ([\nabla w^h]_E \cdot \nu_E, \psi_E)_E. \tag{8.35}$$

Using Green's formula and $\Delta w^h|_T = 0$ we get

$$([\nabla w^h]_E \cdot \nu, \psi_E)_E = \sum_{T \subset \omega_E} (\nabla w^h, \nabla \psi_E)_T = (\nabla w^h, \nabla \psi_E)_{\omega_E}.$$

Since $\left(\frac{\tau}{\mathrm{Pe}}\right)^{-1}(c - c_{old} + \tau\nabla c y_{old}) - \Delta w = 0$ there holds

$$\begin{aligned}
&(\nabla w^h, \nabla\psi_E)_{\omega_E}\\
=&(\nabla w^h, \nabla\psi_E)_{\omega_E} + (\Delta w, \psi_E)_{\omega_E} - \left(\frac{\tau}{\mathrm{Pe}}\right)^{-1}(c - c_{old} + \tau\nabla c y_{old}, \psi_E)_{\omega_E}\\
=&(\nabla e_w, \nabla\psi_E)_{\omega_E} - \left(\frac{\tau}{\mathrm{Pe}}\right)^{-1}(r_h^{(1)}, \psi_E)_{\omega_E}\\
&\quad + \tau(\nabla e_c y_{old}, \psi_E)_{\omega_E} + \left(\frac{\tau}{\mathrm{Pe}}\right)^{-1}(e_c, \psi_E)_{\omega_E}\\
\le&\|\nabla e_w\|_{\omega_E}\|\nabla\psi_E\|_{\omega_E} + \left(\frac{\tau}{\mathrm{Pe}}\right)^{-1}\|r_h^{(1)}\|_{\omega_E}\|\psi_E\|_{\omega_E}\\
&+ \tau\|\nabla e_c\|_{\omega_E}\|y_{old}\|_{\infty,\omega_E}\|\psi_E\| + \left(\frac{\tau}{\mathrm{Pe}}\right)^{-1}\|e_c\|_{\omega_E}\|\psi_E\|_{\omega_E}\\
\le&C\|[\nabla w^h]_E\cdot\nu_E\|_{\omega_E}\Bigg(h_E^{-1/2}\|\nabla e_w\|_{\omega_E}\\
&\quad + \left(\frac{\tau}{\mathrm{Pe}}\right)^{-1}h_E^{1/2}\left(\|r_h^{(1)}\|_{\omega_E} + \|e_c\|_{\omega_E} + \tau\|\nabla e_c\|_{\omega_E}\|y_{old}\|_{\infty,\omega_E}\right)\Bigg).
\end{aligned}$$

Thus we conclude

$$\begin{aligned}
&\|[\nabla w^h]_E\cdot\nu_E\|_E\\
&\le C\Bigg(h_E^{-1/2}\|\nabla e_w\|_{\omega_E}\\
&\quad + \left(\frac{\tau}{\mathrm{Pe}}\right)^{-1}h_E^{1/2}\left(\|r_h^{(1)}\|_{\omega_E} + \|e_c\|_{\omega_E} + \tau\|\nabla e_c\|_{\omega_E}\|y_{old}\|_{\infty,\omega_E}\right)\Bigg).
\end{aligned}$$

Using this estimate in (8.35) we see that

$$\begin{aligned}
\frac{\tau}{\mathrm{Pe}}\left(\eta_E^{(1)}\right)^2 =&\frac{\tau}{\mathrm{Pe}}h_E\|[\nabla w^h]_E\cdot\nu_E\|_E^2\\
\le&C\Bigg(\frac{\tau}{\mathrm{Pe}}\|\nabla e_w\|_{\omega_E}^2 + \left(\frac{\tau}{\mathrm{Pe}}\right)^{-1}\|h_E r_h^{(1)}\|_{\omega_E}^2 + \left(\frac{\tau}{\mathrm{Pe}}\right)^{-1}\|h_E e_c\|_{\omega_E}^2\\
&\quad +h_E^2\tau\mathrm{Pe}\|\nabla e_c\|_{\omega_E}\|y_{old}\|_{\infty,\omega_E}^2\Bigg).
\end{aligned}$$

Using the regularity of the mesh, i.e. $\mathcal{O}(h_E/h_T) = 1$, we have

$$\left(\frac{\tau}{\mathrm{Pe}}\right)^{-1}\|h_E r_h^{(1)}\|_{\omega_E}^2 \le C\sum_{T\subset\omega_E}\left(\frac{\tau}{\mathrm{Pe}}\right)^{-1}(\eta_T^{(1)})^2.$$

Now we use the estimate for $\left(\frac{\tau}{\mathrm{Pe}}\right)^{-1}(\eta_T^{(1)})^2$ from (8.23) and proceed

$$\begin{aligned}
&\frac{\tau}{\mathrm{Pe}}\left(\eta_E^{(1)}\right)^2\\
\leq& C\left(\frac{\tau}{\mathrm{Pe}}\|\nabla e_w\|^2_{\omega_E}+\left(\frac{\tau}{\mathrm{Pe}}\right)^{-1}\|h_E e_c\|^2_{\omega_E}+h_E^2\tau\mathrm{Pe}\|\nabla e_c\|^2_{\omega_E}\|y_{old}\|^2_{\infty,\omega_E}\right.\\
&+C_2\sum_{T\subset\omega_E}\left[\frac{\tau}{\mathrm{Pe}}\|\nabla e_w\|^2_T+\left(\frac{\tau}{\mathrm{Pe}}\right)^{-1}\|h_T e_c\|^2_T+\tau\mathrm{Pe}h_T^2\|y_{old}\|^2_{\infty,T}\|\nabla e_c\|^2_T\right.\\
&\left.\left.+\left(\frac{\tau}{\mathrm{Pe}}\right)^{-1}\mathrm{osc}^2_h(r_h^{(1)},T)\right]\right)\\
\leq& C\left(\frac{\tau}{\mathrm{Pe}}\|\nabla e_w\|^2_{\omega_E}+\left(\frac{\tau}{\mathrm{Pe}}\right)^{-1}\|h_T e_c\|^2_{\omega_E}+\left(\frac{\tau}{\mathrm{Pe}}\right)^{-1}\mathrm{osc}^2_h(r_h^{(1)},\omega_E)\right.\\
&\left.+\tau\mathrm{Pe}\|h_T y_{old}\|^2_{\infty,\omega_E}\|\nabla e_c\|^2_{\omega_E}\right)
\end{aligned}$$

which completes this estimate.

For the proof of (8.34) we use another Φ_E namely $\Phi_E := [\nabla c^h]_E\cdot\nu_E$. Due to (8.19) we have

$$(\eta_E^{(2)})^2 := h_E\|[\nabla c^h]_E\cdot\nu_E\|^2_E\leq Ch_E([\nabla c^h]_E\cdot\nu_E,\psi_E)_E.$$

Green's formula and $\Delta c^h|_T=0$ yield

$$([\nabla c^h]_E\cdot\nu_E,\psi_E)_E=\sum_{T\subset\omega_E}(\nabla c^h,\nabla\psi_E)_T=(\nabla c^h,\nabla\psi_E)_{\omega_E}.$$

Using $-\gamma^2\Delta c+\lambda_s(c)-w-c_{old}=0$ we get

$$\begin{aligned}
&([\nabla c^h]_E\cdot\nu_E,\psi_E)_E\\
&\quad=(\nabla e_c,\nabla\psi_E)_{\omega_E}-\gamma^{-2}(e_w,\psi_E)_{\omega_E}\\
&\qquad+\gamma^{-2}(e_{\lambda_s},\psi_E)_{\omega_E}-\gamma^{-2}(r_h^{(2)}\psi_E)_{\omega_E}.
\end{aligned}$$

Consequently, we obtain

$$\begin{aligned}
([\nabla c^h]_E\cdot\nu_E,\psi_E)_E\leq&\|\nabla e_c\|_{\omega_E}\|\nabla\psi_E\|_{\omega_E}+\gamma^{-2}\|e_w\|_{\omega_E}\|\psi_E\|_{\omega_E}\\
&+\gamma^{-2}\|e_{\lambda_s}\|_{\omega_E}\|\psi_E\|_{\omega_E}+\gamma^{-2}\|r_h^{(2)}\|_{\omega_E}\|\psi_E\|_{\omega_E}.
\end{aligned}$$

Using (8.19), (8.21) and (8.22), it follows that

$$\|[\nabla c^h]_E\cdot\nu_E\|^2_E\leq C([\nabla c^h]_E\cdot\nu_E,\psi_E)_E,$$

and

$$\begin{aligned}
\|[\nabla c^h]_E\cdot\nu_E\|_E\leq C&\left(h_E^{-1/2}\|\nabla e_c\|_{\omega_E}+\gamma^{-2}h_E^{1/2}\|e_w\|_{\omega_E}\right.\\
&\left.+\gamma^{-2}h_E^{1/2}\|e_{\lambda_s}\|_{\omega_E}+\gamma^{-2}h_E^{1/2}\|r_h^{(2)}\|_{\omega_E}\right).
\end{aligned}$$

Therefore, we have

$$\begin{aligned}\gamma^2(\eta_E^{(2)})^2 &:= \gamma^2 h_E \|[\nabla c^h]_E \cdot \nu_E\|_E^2 \\ &\leq C\left(\gamma^2\|\nabla e_c\|^2_{\omega_E} + \gamma^{-2}\|h_E e_w\|^2_{\omega_E}\right. \\ &\left.+\gamma^{-2}\|h_E e_{\lambda_s}\|^2_{\omega_E} + \gamma^{-2}\|h_E r_h^{(2)}\|_{\omega_E}\right).\end{aligned} \tag{8.36}$$

Observe that due to (8.2)

$$\gamma^{-2}\|h_E r_h^{(2)}\|^2_{\omega_E} \leq C \sum_{T\in\omega_E} \gamma^{-2}(\eta_T^{(2)})^2 \tag{8.37}$$

holds, where again the regularity of the mesh, is used. Consequently, by combining (8.24), (8.36) and (8.37) we obtain (8.34). □

Combining Lemma 8.3 and Lemma 8.4 we can prove the efficiency of the error estimator η_Ω.

Theorem 8.5. *There exists a constant β depending on s^{-1}, γ, τ, Pe, Ω, $\|y_{old}\|_\infty$ and the smallest angle of the mesh $\mathcal{T}^{cw}$ such that*

$$s^{-1}\|e_{\lambda_s}\|^2 + \frac{\tau}{Pe}\|\nabla e_w\|^2 + \gamma^2\|\nabla e_c\|^2 \geq \beta\eta_\Omega^2 - osc_h(r_h^{(1)},\Omega)^2 - osc_h(r_h^{(2)},\Omega)^2. \tag{8.38}$$

Proof. Using the estimates of Lemma 8.3 and Lemma 8.4 we obtain for η_Ω

$$\begin{aligned}{\eta_\Omega}^2 \leq{}& C\sum_{T\in\mathcal{T}^{cw}}\left(\left(\frac{\tau}{\mathrm{Pe}}\right)^{-1}\|h_T e_c\|_T^2 + \frac{\tau}{\mathrm{Pe}}\|\nabla e_w\|_T^2 + \left(\frac{\tau}{\mathrm{Pe}}\right)^{-1}\mathrm{osc}_h^2(r_h^{(1)},T)\right. \\ &+\tau\mathrm{Pe}\|h_T y_{old}\|^2_{\infty,T}\|\nabla e_c\|_T^2 + \gamma^2\|\nabla e_c\|_T^2 + \gamma^{-2}\|h_T e_w\|_T^2 \\ &\left.+\gamma^{-2}\|h_T e_{\lambda_s}\|_T^2 + \gamma^{-2}\mathrm{osc}_h^2(r_h^{(2)},T)\right) \\ \leq{}& C\left(\left(\frac{\tau}{\mathrm{Pe}}\right)^{-1}\|he_c\|^2 + \frac{\tau}{\mathrm{Pe}}\|\nabla e_w\|^2 + \gamma^2\|\nabla e_c\|^2 + \gamma^{-2}\|he_w\|^2 + \gamma^{-2}\|he_{\lambda_s}\|^2\right. \\ &\left.+\tau\mathrm{Pe}\|y_{old}\|^2_{L^\infty(\Omega)}\|\nabla e_c\|^2 + \left(\frac{\tau}{\mathrm{Pe}}\right)^{-1}\mathrm{osc}_h^2(r_h^{(1)},\Omega) + \gamma^{-2}\mathrm{osc}_h^2(r_h^{(2)},\Omega)\right),\end{aligned}$$

where $C > 0$ denotes a generic constant. Using $v = 1$ as test function in (5.18) and in (7.2) we obtain $(c_{old}, 1) = (c^h, 1) = (c, 1)$, and thus $(e_c, 1) = 0$. By Poincaré's inequality it follows that $\|e_c\| \leq C_p\|\nabla e_c\|$.

Using $v = 1$ as test function in (5.19) and in (7.3) we obtain

$$(e_w, 1) = (\lambda_s(c^h) - \lambda_s(c), 1) \leq |\Omega|\|e_{\lambda_s}\|.$$

Now the stated result follows from Poincaré–Friedrichs inequality for h small enough. □

Remark 8.6. If lumping is used for $(\lambda_s(c^h), v)$, the resulting estimator is only efficient up to terms arising from this lumping, see [HHK13, HHT11].

8.2 Adaptive concept for the Navier–Stokes part

In this section we derive a residual based a-posteriori error estimator for the Navier–Stokes part of system (5.16)–(5.19). The estimator is derived following the construction in [Jus11]. For the estimator we show reliability and efficiency up to higher order terms.

For a-posteriori error estimation including the solution of local linear problems for the error indicators we refer to [AO00]. Residual based estimators for the Stokes problem are described in [Ver89] for the stationary case and in [Ver10] for the instationary case.

We define the following errors

$$e_y = y - y^h, \quad e_p = p - p^h, \tag{8.39}$$

and define the residual as a linear functional on $H_0^1(\Omega)^d \times L^2_{(0)}(\Omega)$ given by

$$R(v,q) := L((v,q)) - B((y^h,p^h),(v,q)). \tag{8.40}$$

Thus there holds

$$\begin{aligned} R((v,q)) :=& (\xi y_{old} - Kc\nabla w, v) - \xi(y^h, v) - \eta(\nabla y^h, \nabla v) \\ & - \frac{1}{2}(y_{old}\nabla y^h, v) + \frac{1}{2}(y_{old}\nabla v, y^h) \\ & + (\mathrm{div} v, p^h) - (\mathrm{div} y^h, q). \end{aligned}$$

Since

$$L((v,q)) - B((y,p),(v,q)) = 0 \qquad \forall (v,q) \in H_0^1(\Omega)^d \times L^2_{(0)}(\Omega),$$

we have

$$\begin{aligned} R((v,q)) &= L((v,q)) - L((v,q)) - B((y^h - y, p^h - p),(v,q)) \\ &= B((y-y^h, p-p^h),(v,q)). \end{aligned}$$

Using Theorem 4.4 we get

$$\beta^* \;||| \,(e_y,e_p)||| \le \|R\|_* \le c_S \;|||\,(e_y,e_p)|||$$

where $\|R\|_*$ denotes the operator norm of R defined by

$$\|R\|_* := \sup_{(v,q)\in H_0^1(\Omega)^d \times L^2_{(0)}(\Omega)} \frac{|R((v,q))|}{|||(v,q)|||}.$$

Thus any upper bound of the operator norm of the residuum R is a reliable estimator for the error $|||(e_y,e_p)|||$.

8.2.1 Reliability of the estimator – a posteriori upper bound

We proceed with estimating $\|R\|_*$. Integration by parts yields

$$\begin{aligned} R((v,q)) = & \sum_{T\in T^{yp}} (\xi y_{old} - Kc\nabla w + \eta\Delta y^h - y_{old}\nabla y^h - \nabla p^h, v)_T \\ & - \sum_{T\in T^{yp}} (\text{div } y^h, q)_T - \sum_{E\in\mathcal{E}^{yp}} \left(\left[\nu_E \cdot \left(\eta\nabla y^h - p^h I\right)\right]_E, v\right)_E . \end{aligned}$$

Let us introduce

$$\begin{aligned} r_T &:= \xi y_{old} - Kc^h\nabla w^h - \xi y^h + \eta\Delta y^h - y_{old}\nabla y^h - \nabla p^h, \\ f_T &:= K(c^h\nabla w^h - c\nabla w), \\ d_T &:= -\text{div } y^h, \\ j_E &:= \begin{cases} -\left[\nu_E \cdot \left(\eta\nabla y^h - p^h I\right)\right]_E & \text{if } E \notin \partial\Omega, \\ 0 & \text{if } E \in \partial\Omega. \end{cases} \end{aligned}$$

Then

$$R(v,q) = \sum_{T\in T^{yp}} \{(r_T, v) + (d_T, q) + (f_T, v)\} + \sum_{E\in\mathcal{E}^{yp}} (j_E, v). \tag{8.41}$$

From $L^h(v^h, q^h) - B((y^h, p^h), (v^h, q^h)) = 0$ for all $v^h \in \mathcal{V}^y, q^h \in \mathcal{V}^p$ we further infer

$$R((v,q)) = \sum_{T\in\mathcal{T}^{yp}} \left\{(r_T, v - v^h) + (d_T, q - q^h) + (f_T, v)\right\} + \sum_{E\in\mathcal{E}^{yp}} (j_E, v - v^h). \tag{8.42}$$

Let $\Pi_h v \in \mathcal{V}^y$ denote the Clément interpolation of v with $\Pi_h v|_{\partial\Omega} = 0$ (see [EG04, Rem. 1.129]). We test (8.42) with $v^h = \Pi_h v$ and $q^h = 0$ and obtain

$$\begin{aligned}
&R((v,q))\\
&\quad\le \sum_{T\in\mathcal{T}^{yp}} \left\{ \|r_T\|_T \|v-\Pi_h v\|_T + \|d_T\|_T\|q\|_T + \|f_T\|_T\|v\|_T \right\}\\
&\qquad + \sum_{E\in\mathcal{E}^{yp}} \|j_E\|_E \|v-\Pi_h v\|_E,\\
&\quad\le \sum_{T\in\mathcal{T}^{yp}} \left\{ c_1 h_T \|r_T\|_T \|\nabla v\|_{\omega_T} + \|d_T\|_T\|q\|_T + \|f_T\|_T\|v\|_T \right\}\\
&\qquad + \sum_{E\in\mathcal{E}^{yp}} c_2 h_E^{1/2} \|j_E\|_E \|\nabla v\|_{\omega_E}\\
&\quad\le C \Bigg\{ \sqrt{\sum_{T\in\mathcal{T}^{yp}} h_T^2\eta^{-1}\|r_T\|_T^2 \sum_{T\in\mathcal{T}^{yp}} \eta\|\nabla v\|_{\omega_T}^2} + \sqrt{\sum_{T\in\mathcal{T}^{yp}} \eta\|d_T\|_T^2 \sum_{T\in\mathcal{T}^{yp}} \eta^{-1}\|q\|_T^2}\\
&\qquad + \sqrt{\sum_{T\in\mathcal{T}^{yp}} \xi^{-1}\|f_T\|_T^2 \sum_{T\in\mathcal{T}^{yp}} \xi\|v\|_T^2} + \sqrt{\sum_{E\in\mathcal{E}^{yp}} h_E\eta^{-1}\|j_E\|_E^2 \sum_{E\in\mathcal{E}^{yp}} \eta\|\nabla v\|_{\omega_E}^2} \Bigg\}\\
&\quad\le C \left\{ \sum_{T\in\mathcal{T}^{yp}} \left\{ h_T^2\eta^{-1}\|r_T\|_T^2 + \eta\|d_T\|_T^2 + \xi^{-1}\|f_T\|_T^2 \right\} + \sum_{E\in\mathcal{E}^{yp}} h_E\eta^{-1}\|j_E\|_E^2 \right\}^{1/2}\\
&\qquad\qquad \times \underbrace{\left\{ \sum_{T\in\mathcal{T}^{yp}} \eta\|\nabla v\|_T^2 + \sum_{T\in\mathcal{T}^{yp}} \xi\|v\|_T^2 + \sum_{T\in\mathcal{T}^{yp}} \eta^{-1}\|q\|_T^2 \right\}^{1/2},}_{=|||(v,q)|||}
\end{aligned}$$

where we use the Cauchy–Schwarz inequality for both integrals and sums together with the error estimations for Clément interpolation provided in (8.4)–(8.5). The constant C only depends on the domain Ω and the smallest angle in $\mathcal{T}^{yp}$ ([Clé75, H4]) and especially is independent of ν, ξ and y_{old}. Thus we have

$$\begin{aligned}
&\frac{R((v,q))}{|||(v,q)|||}\\
&\le C \left\{ \sum_{T\in\mathcal{T}^{yp}} \left\{ h_T^2\eta^{-1}\|r_T\|_T^2 + \eta\|d_T\|_T^2 + \xi^{-1}\|f_T\|_T^2 \right\} + \sum_{E\in\mathcal{E}^{yp}} h_E\eta^{-1}\|j_E\|_E^2 \right\}^{1/2},
\end{aligned}$$

which implies

$$\begin{aligned}
\|R\|_* &= \sup_{(v,q)\in H_0^1(\Omega)^d\times L^2_{(0)}(\Omega)} \frac{R((v,q))}{|||(v,q)|||}\\
&\le C \left\{ \sum_{T\in\mathcal{T}^{yp}} \left\{ h_T^2\eta^{-1}\|r_T\|_T^2 + \eta\|d_T\|_T^2 + \xi^{-1}\|f_T\|_T^2 \right\} + \sum_{E\in\mathcal{E}^{yp}} h_E\eta^{-1}\|j_E\|_E^2 \right\}^{1/2}.
\end{aligned}$$

We summarize our finding in

Theorem 8.7. *Let (y,p) denote the solution to (4.11) and (y^h,p^h) denote the solution to (7.1). Let β^*, denote the constant defined in Theorem 4.4.*

Then there exists a constant $C>0$ depending on Ω and the smallest angle of $\mathcal{T}^{yp}$ such that

$$\beta^* \,|||\, (y-y^h, p-p^h)||| \leq C\eta_\Omega$$

holds, where η_Ω is given by

$$\eta_\Omega^2 = \sum_{T\in\mathcal{T}^{yp}} \left\{ h_T^2\eta^{-1}\|r_T\|_T^2 + \eta\|d_T\|_T^2 + \xi^{-1}\|f_T\|_T^2 \right\} + \sum_{E\in\mathcal{E}^{yp}} h_E\eta^{-1}\|j_E\|_E^2 \quad (8.43)$$

with

$$\begin{aligned} r_T &:= \xi y_{old} - Kc^h\nabla w^h - \xi y^h + \eta\Delta y^h - y_{old}\nabla y^h - \nabla p^h, \\ d_T &:= -div\; y^h, \\ f_T &:= K(c^h\nabla w^h - c\nabla w), \\ j_E &:= \begin{cases} -\left[\nu_E\cdot\left(\eta\nabla y^h - p^h I\right)\right]_E & \text{if } E\notin\partial\Omega, \\ 0 & \text{if } E\in\partial\Omega. \end{cases} \end{aligned}$$

Remark 8.8. Through appearance of the term $c\nabla w$ in f_T the estimator is not fully practical. The norm of f_T can be further estimated as

$$\begin{aligned} K^{-2}\sum_{T\in\mathcal{T}^{yp}} \|f_T\|_T^2 =& \|(c-c^h)\nabla w + c^h(\nabla w - \nabla w^h)\|_{L^2(\Omega)}^2 \\ \leq& \frac{3}{2}\left(\|e_c\nabla w\|_{L^2(\Omega)}^2 + \|c^h\nabla e_w\|_{L^2(\Omega)}^2 \right). \end{aligned}$$

From Lemma 5.8 we have $\|w\|_{H^2(\Omega)} \leq C$ independent of s, and by using Sobolev embedding we have

$$\|e_c\nabla w\|_{L^2(\Omega)}^2 \leq \|\nabla w\|_{L^4(\Omega)}^2\|e_c\|_{L^4(\Omega)}^2 \leq C\|w\|_{H^2(\Omega)}^2\|e_c\|_{H^1(\Omega)}^2 \leq C\|\nabla e_c\|_{L^2(\Omega)}^2,$$

where for the last inequality we use Poincaré–Friedrichs' inequality together with $(e_c,1)=0$. For the second addend we have

$$\|c^h\nabla e_w\|_{L^2(\Omega)}^2 \leq \|c^h\|_{L^\infty(\Omega)}^2\|\nabla e_w\|_{L^2(\Omega)}^2.$$

Since $c\in H^2(\Omega)\hookrightarrow L^\infty(\Omega)$ and $\|c\|_{H^2(\Omega)} \leq C$ from Lemma 5.8, it further holds

$$\|c^h\|_{L^\infty(\Omega)} \leq \|c^h - c\|_{L^\infty(\Omega)} + \|c\|_{L^\infty(\Omega)} \leq \|c^h - c\|_{L^\infty(\Omega)} + C.$$

Let I_hc denote the Lagrange interpolation of c. We set $m = \int_\Omega c^h - I_hc\,dx$, fulfilling $|m|\leq C$, and proceed

$$\begin{aligned} \|c^h - c\|_{L^\infty(\Omega)} &\leq \|c^h - I_hc - m\|_{L^\infty(\Omega)} + |m| + \|I_hc - c\|_{L^\infty(\Omega)} \\ &\leq \sigma(d,h)\|\nabla(c^h - I_hc)\|_{L^2(\Omega)} + C + Ch^{2-d/2}|c|_{H^2(\Omega)}, \end{aligned}$$

with

$$\sigma(d,h) = C \begin{cases} |\log h|^{1/2} & \text{if } d = 2, \\ h^{-1/2} & \text{if } d = 3. \end{cases}$$

Here for the first term we use discrete Sobolev inequalities from [HPUU09, Prop. 3.1], while for the third term we use approximation results for the Lagrange interpolation (see e.g. [BS08, Th. 4.4.20]). We proceed with

$$\begin{aligned} \|\nabla(c^h - I_h c)\|_{L^2(\Omega)} &\leq \|\nabla(c^h - c)\|_{L^2(\Omega)} + \|\nabla(c - I_h c)\|_{L^2(\Omega)} \\ &\leq \|\nabla e_c\|_{L^2(\Omega)} + Ch|c|_{H^2(\Omega)}, \end{aligned}$$

where we again use approximation results for the Lagrange interpolation from [BS08, Th. 4.4.20]. Putting everything together we have

$$\|c^h\|_{L^\infty(\Omega)} \leq C(1 + h^{2-d/2}) + \sigma(d,h) \left(\|\nabla e_c\|_{L^2(\Omega)} + Ch\right) = \xi(d,h)$$

and

$$K^{-2} \sum_{T \in \mathcal{T}^{yp}} \|f_T\|_T^2 \leq C \left(\|\nabla e_c\|^2_{L^2(\Omega)} + \xi(d,h)^2 \|\nabla e_w\|^2_{L^2(\Omega)} \right).$$

Thus we can use the results from Theorem 8.1 to estimate the error terms $\|f_T\|_T^2$. We note, that $\xi(d,h)$ is bounded independently of h.

We further note that the error indicators from Theorem 8.1 are defined on the mesh $\mathcal{T}^{cw}$ while $\|f_T\|_T$ is defined on $\mathcal{T}^{yp}$. The treatment of this situation is discussed in Section 8.3.

8.2.2 Efficiency of the estimator – a posteriori lower bound

We next show that our residual based error estimator is efficient up to higher order terms arising due to the presence of $\|f_T\|$.

For this we again use the bubble-technique which was also used to establish efficiency of the estimator for the Cahn–Hilliard part and we proceed similar as in showing the efficiency of the estimator for the Cahn–Hilliard part.

Lemma 8.9. *There exist a constant $C > 0$ independent of ξ, ν, y_{old} and h_T such that there*

$$\begin{aligned} \eta^{-1} h_T^2 \|r_T\|^2 \leq & C \max\left[\left(1 + \eta^{-1} h_T^2 \|y_{old}\|_{L^\infty(\Omega)}\right)^2, h_T^2 \eta^{-1} \xi, h_T^2 \right] |||\, (e_y, e_p) |||^2 \\ & + C\eta^{-1} osc_h^2(r_T, T) + C\eta^{-1} h_T^2 \|f_T\|^2 \end{aligned}$$

holds. Furthermore,

$$\eta \|d_T\|^2 \leq d \,|||\, (e_y, e_p) \,|||^2 .$$

Proof. Since div$y = 0$ and thus, div$y^h = \text{div}(y^h - y)$, the second statement follows from $\|\text{div}v\|^2 \leq d\|\nabla v\|^2$ for any $v \in H_0^1(\Omega)^d$.

To proof the first statement we start with

$$\eta^{-1}h_T^2\|r_T\|^2 \leq 2\eta^{-1}h_T^2\|\overline{r}_T\|^2 + 2\eta^{-1}\text{osc}_h^2(r_T, T) \tag{8.44}$$

where $\overline{r}_T$ is the mean value of r_T as defined in (8.3). Let $\psi_T := \overline{r}_T\lambda_T$, where λ_T is the canonical bubble function defined in Section 8.1.2. Then (8.16)–(8.22) imply

$$\|\overline{r}_T\|^2 \leq C(\overline{r}_T, \psi_T) \leq C(r_T, \psi_T) + Ch_T^{-1}\text{osc}_h(r_T, T)\|\psi_T\|. \tag{8.45}$$

Using $\xi y - \eta\Delta y + y_{old}\nabla y + \nabla p + Kc\nabla w - \xi y_{old} = 0$ we proceed with

$$\begin{aligned}(r_T, \psi_T) &= \left(\xi y_{old} - Kc^h\nabla w^h + \eta\Delta y^h - \xi y^h - \nabla p^h - y_{old}\nabla y^h, \psi_T\right)\\ &= (-f_T - \eta\Delta e_y + \xi e_y + \nabla e_p + y_{old}\nabla e_y, \psi_T)\\ &= \eta(\nabla e_y, \nabla\psi_T) + (\xi e_y + \nabla e_p + y_{old}\nabla e_y - f_T, \psi_T)\\ &\leq \eta\|\nabla e_y\|\|\nabla\psi_T\| + \|y_{old}\|_{L^\infty(\Omega)}\|\nabla e_y\|\|\psi_T\|\\ &\quad + (\xi\|e_y\| + \|\nabla e_p\| + \|f_T\|)\,\|\psi_T\|.\end{aligned}$$

Using this together with (8.16)–(8.22) in (8.45) we obtain

$$\begin{aligned}\|\overline{r}_T\|^2 \leq& C\left(\eta\|\nabla e_y\|\|\nabla\psi_T\| + \|y_{old}\|_{L^\infty(\Omega)}\|\nabla e_y\|\|\psi_T\|\right.\\ &\left.+ \left(\xi\|e_y\| + \|\nabla e_p\| + \|f_T\| + h_T^{-1}\text{osc}_h(r_T, T)\right)\|\psi_T\|\right)\\ \leq& C\left\{\left(h_T^{-1}\eta + \|y_{old}\|_{L^\infty(\Omega)}\right)\|\nabla e_y\| + \xi\|e_y\|\right.\\ &\left.+ \|\nabla e_p\| + \|f_T\| + h_T^{-1}\text{osc}_h(r_T, T)\right\}\|\overline{r}_T\|.\end{aligned}$$

Thus there holds

$$\begin{aligned}\|\overline{r}_T\| \leq C&\left\{\left(h_T^{-1}\eta + \|y_{old}\|_{L^\infty(\Omega)}\right)\|\nabla e_y\| + \xi\|e_y\| + \|\nabla e_p\|\right.\\ &\left.+\|f_T\| + h_T^{-1}\text{osc}_h(r_T, T)\right\},\end{aligned}$$

with some positive C which is independend of ξ, ν, y_{old} and h_T. Inserting this into (8.44) we arrive at

$$\begin{aligned}\eta^{-1}h_T^2\|r_T\|^2 \leq& 2\eta^{-1}h_T^2\|\overline{r}_T\|^2 + 2\eta^{-1}\text{osc}_h^2(r_T, T)\\ \leq& C\left\{\eta^{-1}\text{osc}_h^2(r_T, T) + \eta^{-1}\left(\eta + h_T\|y_{old}\|_{L^\infty(\Omega)}\right)^2\|\nabla e_y\|^2\right.\\ &\left.+ h_T^2\eta^{-1}\xi^2\|e_y\|^2 + h_T^2\eta^{-1}\|\nabla e_p\|^2 + h_T^2\eta^{-1}\|f_T\|^2\right\}\\ \leq& C\max\left(\left(1 + \eta^{-1}h_T\|y_{old}\|_{L^\infty(\Omega)}\right)^2, h_T^2\eta^{-1}\xi, h_T^2\right)|||\,(e_y, e_p)|||^2\\ &+ C\eta^{-1}\text{osc}_h^2(r_T, T) + C\eta^{-1}h_T^2\|f_T\|^2,\end{aligned}$$

where Young's inequality is used several times. □

Lemma 8.10. *There exists a constant $C > 0$ independend of ξ, ν, y_{old} and h such that there holds*

$$\begin{aligned} h_E\eta^{-1}\|j_E\|^2 \leq & C\Big\{\left(1+\xi\eta^{-1}h_E^2\right) \|\!|\!| (e_y,e_p)\|\!|\!|_{\omega_E}^2 \\ & + \max\left(\left(1+\eta^{-1}h_T\|y_{old}\|_{L^\infty(\Omega)}\right)^2, h_T^2\eta^{-1}\xi, h_T^2\right) \|\!|\!| (e_y,e_p)\|\!|\!|_{\omega_E}^2 \\ & + \sum_{T\in\omega_E} \eta^{-1}h_T^2\|f_T\|^2 + \eta^{-1} osc_h^2(r_T,\omega_E)\Big\}. \end{aligned}$$

Proof. Let E be an arbitrary edge of $\mathcal{E}^{yp}$ and define

$$\psi_E := \tilde{j}_E\lambda_E, \quad j_E = \left[\left(\eta\nabla y^h - p^h I\right)\nu_E\right]_E.$$

We use the properties of R defined in (8.40). Using (8.41) we have

$$R(\psi_E,0) = \sum_{T\in\omega_E} \{(r_T,\psi_E)_T + (f_T,\psi_E)_T\} + (j_E,\psi_E)_E.$$

From Theorem 4.4 we also obtain

$$R(\psi_E,0) \leq C \,\|\!|\!|\, (e_y,e_p) \,\|\!|\!|_{\omega_E} \|\!|\!|(\psi_E,0)\,\|\!|\!|_{\omega_E}.$$

Combining this with

$$\begin{aligned} \|\!|\!|(\psi_E,0)\|\!|\!|^2 \leq & \eta \sum_{T\in\omega_E} \|\nabla\psi_E\|_T^2 + \xi \sum_{T\in\omega_E} \|\psi_E\|^2 \\ \leq & C\left(\eta h_E^{-1} + \xi h_E\right)\|j_E\|^2, \end{aligned}$$

we obtain

$$\begin{aligned} \eta^{-1}h_E\|j_E\|^2 \leq & C\eta^{-1}h_E(j_E,\psi_E)_E \\ \leq & \eta^{-1}h_E C\Bigg\{ \,\|\!|\!|\, (e_y,e_p) \,\|\!|\!|_{\omega_E} \|\!|\!|(\psi_E,0)\|\!|\!|_{\omega_E}) \\ & + \sum_{T\in\omega_E} \|r_T\|\|\psi_E\| + \sum_{T\in\omega_E} \|f_T\|\|\psi_E\|\Bigg\} \\ \leq & \eta^{-1}h_E C\Bigg\{\sqrt{\eta h_E^{-1} + \xi h_E}\,\|\!|\!|\, (e_y,e_p) \,\|\!|\!|_{\omega_E} \|j_E\| \\ & + \sum_{T\in\omega_E} h_E^{1/2}\|r_T\|\|j_E\| + \sum_{T\in\omega_E} h_E^{1/2}\|f_T\|\|j_E\|\Bigg\} \\ = & C\Bigg\{\sqrt{1+\xi\eta^{-1}h_E^2}\,\|\!|\!|\, (e_y,e_p)\|\!|\!|_{\omega_E} \\ & + \eta^{-1/2}\sum_{T\in\omega_E} h_E\|r_T\| + \eta^{-1/2}\sum_{T\in\omega_E} h_E\|f_T\|\Bigg\}\eta^{-1/2}h_E^{1/2}\|j_E\|_E. \end{aligned}$$

Due the regularity conditions supposed for $\mathcal{T}^{yp}$ we have $h_E \leq Ch_T$ with some positive constant C independent of the mesh size. Thus,

$$\begin{aligned}\eta^{-1/2}h_E^{1/2}\|j_E\| \leq & C\left\{\sqrt{1+\xi\eta^{-1}h_E^2}\ |||\ (e_y,e_p)|||_{\omega_E}\right.\\ & \left. + \eta^{-1/2}\sum_{T\in\omega_E} h_T\|r_T\| + \eta^{-1/2}\sum_{T\in\omega_E} h_T\|f_T\|\right\},\end{aligned}$$

which implies

$$\begin{aligned}\eta^{-1}h_E\|j_E\|^2 \leq & C\left\{\ \left(1+\xi\eta^{-1}h_E^2\right)\ |||\ (e_y,e_p)|||^2_{\omega_E}\right.\\ & \left. + \eta^{-1}\sum_{T\in\omega_E} h_T^2\|r_T\|^2 + \eta^{-1}\sum_{T\in\omega_E} h_T^2\|f_T\|^2\right\}.\end{aligned}$$

Now we use the estimate from Lemma 8.9 and obtain

$$\begin{aligned}&\eta^{-1}h_E\|j_E\|^2\\ &\quad \leq C\left\{\alpha\ |||\ (e_y,e_p)\ |||^2_{\omega_E} + \sum_{T\in\omega_E}\eta^{-1}h_T^2\|f_T\|^2 + \eta^{-1}\mathrm{osc}_h^2(r_T,\omega_E)\right\},\end{aligned}$$

where

$$\alpha = \left(1+\xi\eta^{-1}h_E^2\right) + \max\left(\left(1+\eta^{-1}h_T\|y_{old}\|_{L^\infty(\Omega)}\right)^2, h_T^2\eta^{-1}\xi, h_T^2\right)$$

□

Combining the previous two lemmas we for $0 < h_T < 1$ have proven

Theorem 8.11. *There exists a constant $C > 0$ independend of h but depending on ξ, η, y_{old}, d such that there holds*

$$\eta_\Omega^2 \leq C\left(|||(e_y,e_p)\ |||^2 + osc_h^2(r_T,\Omega) + \sum_{T\in\mathcal{T}^{yp}}\|f_T\|^2\right).$$

Remark 8.12. From Remark 8.8 we deduce, that $\|f_T\|$ is bounded by the reliable and efficient estimator introduced in Theorem 8.1. Let h_y denote the gridsize of $\mathcal{T}^{yp}$ and h_c denote the gridsize of $\mathcal{T}^{cw}$. Then for fix h_c and for $h_y \to 0$ this upper bound on $\|f_T\|$ is independend of h_y and thus we cannot deduce $\|f_T\| \to 0$, i.e. the efficiency of the estimator.

However, in real situations one would always tend $h_y \to 0$ and $h_c \to 0$ simultaneously. Thus we expect that the overestimation introduced by using a not efficient estimator is moderate in real applications.

8.3 The adaptive mesh refinement cycle

Having the error estimators (8.15) and (8.43) at hand, we next describe the refinement cycles that we use during the numerical simulation. We use the classic refinement cycle

$$\text{SOLVE} \rightarrow \text{ESTIMATE} \rightarrow \text{MARK} \rightarrow \text{REFINE/COARSE}$$

once in every time step of our simulation. Thus, our overall algorithm for solving the discrete Cahn–Hilliard Navier–Stokes system (7.1)–(7.3) is given by:

1. INIT: Obtain initial meshes $\mathcal{T}^{cw}_{(0)}$ for the concentration and the potential, and $\mathcal{T}^{yp}_{(0)}$ for the flowfield and the pressure. Set $k = 0$.
2. SOLVE: Perform one step of simulation, i.e. at time t_{k+1} solve the system (7.1)–(7.3) on the meshes $\mathcal{T}^{cw}_{(k)}$ and $\mathcal{T}^{yp}_{(k)}$.
3. ESTIMATE: Calculate local error contributions on each triangle for $\mathcal{T}^{cw}_{(k)}$ and $\mathcal{T}^{yp}_{(k)}$ as described in Section 8.1 and Section 8.2.
4. MARK: Determine triangles in $\mathcal{T}^{cw}_{(k)}$ and $\mathcal{T}^{yp}_{(k)}$ for refinement and coarsening using the marking strategies defined below.
5. REFINE/COARSE: Obtain meshes $\mathcal{T}^{cw}_{(k+1)}$ and $\mathcal{T}^{yp}_{(k+1)}$ by refining and coarsening the meshes $\mathcal{T}^{cw}_{(k)}$ and $\mathcal{T}^{yp}_{(k)}$ according to the marking obtained in step 4.
6. Set $k := k + 1$, go to 2.

Note that the solution on the current time instance is calculated in step 2. The subsequent steps calculate a new mesh that is used for the next time step.

Now we have a look at the local error contributions for the Cahn–Hilliard and the Navier–Stokes system.

For each $T \in \mathcal{T}^{cw}$ we define

$$\eta_T = \left(\frac{\tau}{\text{Pe}}\right)^{-1}\left(\eta_T^{(1)}\right)^2 + \gamma^{-2}\left(\eta_T^{(2)}\right)^2 \quad \text{and}$$

$$\eta_{TE} = \sum_{E\in\mathcal{E}(T)}\left(\frac{\tau}{\text{Pe}}\left(\eta_E^{(1)}\right)^2 + \gamma^2\left(\eta_E^{(2)}\right)^2\right).$$

Here $E \in \mathcal{E}(T)$ means, that E is an edge of the triangle T.

For each $S \in \mathcal{T}^{yp}$ we define

$$\eta_S = h_S^2\eta^{-1}\|r_T\|^2, \tag{8.46}$$

$$\eta_{SE} = h_S\eta^{-1}\sum_{E\subset\mathcal{E}(S)}\|j_E\|^2, \tag{8.47}$$

$$\eta_{SD} = \eta\|d_T\|^2, \tag{8.48}$$

$$\eta_{SF} = \xi^{-1}\sum_{T\in\mathcal{C}}\left(\eta_T + \eta_{TE}\right). \tag{8.49}$$

Here $\mathcal{C} \subset \mathcal{T}^{cw}$ denotes the smallest set of triangles $T \in \mathcal{T}^{cw}$ such that $S \subset \bigcup_{T \in \mathcal{C}} T$

Using these error indicaters we next describe our marking strategy.

Marking strategy for $\mathcal{T}^{cw}$

Since we expect large errors mainly in the interfacial region, we use bulk marking, see e.g. [Dör96]. We define the set

$$\mathcal{A} = \{T \in \mathcal{T}^{cw} \mid a_{\min} \leq |T| \leq a_{\max}\}$$

of all admissible triangles for adaptation. The positive constants $a_{\min}$ and $a_{\max}$ denote the minimal and the maximal size of elements we allow in our meshes. The marking strategy performs the following steps:

(i) Fix constants θ^r and θ^c in $(0, 1)$.

(ii) Find a set $\mathcal{M}^T \subset \mathcal{T}^{cw}$ such that

$$\sum_{T \in \mathcal{M}^T} \eta_T \geq \theta^r \sum_{T \in \mathcal{T}^{cw}} \eta_T.$$

(iii) Find a set $\mathcal{M}^E \subset \mathcal{T}^{cw}$ such that

$$\sum_{T \in \mathcal{M}^E} \eta_{TE} \geq \theta^r \sum_{T \in \mathcal{T}^{cw}} \eta_{TE}.$$

(iv) Mark each $T \in (\mathcal{M}^E \cup \mathcal{M}^T) \cap \mathcal{A}$ for refinement.

(v) Find the set $\mathcal{C}^T \subset \mathcal{T}^{cw}$ such that

$$\eta_T \leq \frac{\theta^c}{N_T} \sum_{T \in \mathcal{T}^{cw}} \eta_T$$

holds for all $T \in \mathcal{C}^T$. Here and below N_T denotes the number of elements of $\mathcal{T}^{cw}$.

(vi) Find the set $\mathcal{C}^E \subset \mathcal{T}^{cw}$ such that

$$\eta_{TE} \leq \frac{\theta^c}{N_T} \sum_{T \in \mathcal{T}^{cw}} \eta_{TE}$$

holds for all $T \in \mathcal{C}^E$.

(vii) Mark all $T \in \left(\mathcal{C}^T \cup \mathcal{C}^E\right) \cap \mathcal{A}$ for coarsening.

Remark 8.13.

- Note that the marking for refinement of elements is splitted up into the two separate steps ((ii))–((iii)), and the marking for coarsening is splitted up into the separate steps ((v))–((vi)). This offers the possibility to properly consider the different scalings in η_T and η_{TE} introduced by $\frac{\tau}{\mathrm{Pe}}$ and γ.

- This strategy does not prevent a triangle from being both marked for refinement and for coarsening. In this case it is refined only.

- In our numerics, this strategy performs well whenever we choose τ small in comparison to the movement of the interface. This corresponds to the restriction on τ introduced in Theorem 6.3 to guarantee the convergence of Newton's method, and in Theorem 8.1 to guarantee the reliability of the estimator η_Ω. In our tests $\tau = \mathcal{O}(\gamma^2)$ turns out to be a suitable choice. Especially in spinodal decomposition the time step τ has to be chosen quite small to capture the system dynamics at the beginning of the evolution. For this an adaptive choice of τ would be desirable, see [BN09] for a heuristic approach to adapt the time step size τ.

Marking strategy for $\mathcal{T}^{yp}$

As is indicated by our numerical simulations errors for the velocity and pressure field are not clustered as it is observed on $\mathcal{T}^{cw}$. Although in our numerical tests there was not one best strategy, we again propose a bulk marking with separate treatment of all four error indicators.

Again we define the set of admissable triangles by

$$\mathcal{A} = \{T \in \mathcal{T}^{yp} \,|\, a_{\min} \leq |T| \leq a_{\max}\}.$$

with minimum and maximum allowed triangle sizes $a_{\min} > 0$ and $a_{\max} > 0$. The algorithm for marking triangles is given as follows:

(i) Fix constants θ^r and θ^c in $(0, 1)$.

(ii) Find a set $\mathcal{M}^S \subset \mathcal{T}^{yp}$ such that

$$\sum_{S \in \mathcal{M}^S} \eta_S \geq \theta^r \sum_{S \in \mathcal{T}^{yp}} \eta_S.$$

(iii) Find a set $\mathcal{M}^{SE} \subset \mathcal{T}^{yp}$ such that

$$\sum_{S \in \mathcal{M}^{SE}} \eta_{SE} \geq \theta^r \sum_{S \in \mathcal{T}^{yp}} \eta_S.$$

(iv) Find a set $\mathcal{M}^{SD} \subset \mathcal{T}^{yp}$ such that

$$\sum_{S \in \mathcal{M}^{SD}} \eta_{SD} \geq \theta^r \sum_{S \in \mathcal{T}^{yp}} \eta_S.$$

(v) Find a set $\mathcal{M}^{SF} \subset \mathcal{T}^{yp}$ such that

$$\sum_{S \in \mathcal{M}^{SF}} \eta_{SF} \geq \theta^r \sum_{S \in \mathcal{T}^{yp}} \eta_S.$$

(vi) Mark each $S \in (\mathcal{M}^S \cup \mathcal{M}^{SE} \cup \mathcal{M}^{SD} \cup \mathcal{M}^{SF}) \cap \mathcal{A}$ for refinement.

(vii) Find the set $\mathcal{C}^S \subset \mathcal{T}^{yp}$ such that

$$\eta_S \leq \frac{\theta^c}{N_T} \sum_{S \in \mathcal{T}^{yp}} \eta_S$$

holds for all $S \in \mathcal{C}^S$. Here and below N_T denotes the number of elements of $\mathcal{T}^{yp}$.

(viii) Find the set $\mathcal{C}^{SE} \subset \mathcal{T}^{yp}$ such that

$$\eta_{SE} \leq \frac{\theta^c}{N_T} \sum_{S \in \mathcal{T}^{yp}} \eta_{SE}$$

holds for all $S \in \mathcal{C}^{SE}$.

(ix) Find the set $\mathcal{C}^{SD} \subset \mathcal{T}^{yp}$ such that

$$\eta_{SD} \leq \frac{\theta^c}{N_T} \sum_{S \in \mathcal{T}^{yp}} \eta_{SD}$$

holds for all $S \in \mathcal{C}^{SD}$.

(x) Find the set $\mathcal{C}^{SF} \subset \mathcal{T}^{yp}$ such that

$$\eta_{SF} \leq \frac{\theta^c}{N_T} \sum_{S \in \mathcal{T}^{yp}} \eta_{SF}$$

holds for all $S \in \mathcal{C}^{SF}$.

(xi) Mark all $S \in (\mathcal{C}^S \cup \mathcal{C}^{SE} \cup \mathcal{C}^{SD} \cup \mathcal{C}^{SF}) \cap \mathcal{A}_h$ for coarsening.

The indicator η_{SF} is defined using the indicators η_T and η_{TE} for the Cahn–Hilliard part. In our numerical test we observe that these indicators are clustered at the interface and thus also the indicator η_{SF} is clustered at the interface. However, numerical tests indicate that the overall L^2 velocity error is only slightly reduced when refining at the interface, at least in the case of a dominant boundary velocity field. Thus to obtain small overall velocity errors with less triangles one should modify the proposed marking procedure. Since we are interested in the dynamics at the interface it is reasonable to use this indicator anyway to obtain a better resolved velocity field at the interface.

Again, if a triangle is both marked for refinement and coarsening it is refined only.

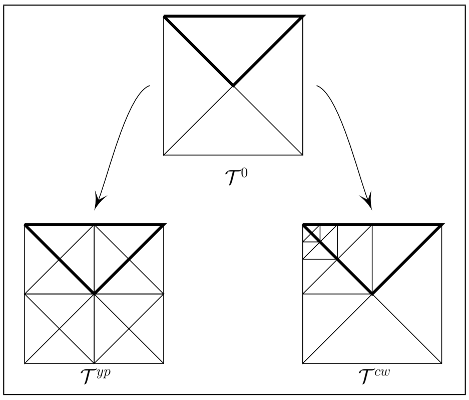

Figure 8.1: The refinement of the macro triangulation.

8.4 The two-mesh strategy

In the previous sections we described the adaptive concept for obtaining the Cahn–Hilliard and the Navier–Stokes meshes. We now describe the numerical treatment of the forcing term $c\nabla w$ in the Navier–Stokes equation, and of the transport term $y_{old}\nabla c$ in the Cahn–Hilliard equation.

Since $c_{old} \notin \mathcal{V}^{cw}$, the question of integrating products of functions defined on different meshes already arises when solving solely the Cahn–Hilliard equation. In this context we note that prolonging c_{old} and y_{old} from the mesh of the old time step to the mesh for the new time step needs extra care due to mass conservation. In our approach we do not prolonge solutions from the old time step and thus have to handle integrals defined on more than one mesh.

In the following we describe the evaluation of $(c^h\nabla w^h, v)$. The terms $(c^h y_{old}, v)$, (c_{old}, v) and (y_{old}, v) can be treated analogously. We note that for evaluating the error estimator up to three meshes are involved. Since $\lambda_s(c^h)$ is not contained in the Ansatz space $\mathcal{V}^{cw}$ further implementation work is necessary for its proper numerical treatment. Let us also refer to the discussion on the use of adaptive meshes in time dependent simulation in [Grä11, Sec. 6.2.1].

In order to initialize the overall adaptive procedure (1)–(6) we construct the meshes $\mathcal{T}^{cw}_{(0)}$ and $\mathcal{T}^{yp}_{(0)}$ starting from a common macro-triangulation $\mathcal{T}^0$ of Ω for both meshes. The macro mesh is refined to obtain $\mathcal{T}^{cw}_{(0)}$ and $\mathcal{T}^{yp}_{(0)}$, respectively, see Figure 8.1. The mesh $\mathcal{T}^{cw}_{(0)}$ especially may also be adapted to the initial concentration. The numerical solvers are then employed on the meshes resulting from this refinement procedure.

We evaluate $(c^h\nabla w^h, v)$ triangle-wise over all triangles of $\mathcal{T}^{yp}$. On a triangle $Y \in \mathcal{T}^{yp}$ determine the set of all triangles $\{C_k\}_{k=1}^{n_c} \subset \mathcal{T}^{cw}$, $n_c \in \mathbb{N}$, such that $\overset{\circ}{Y} \cap \overset{\circ}{C}_k \neq \emptyset$ for $k = 1, \dots n_c$, and perform the integration on each $\overset{\circ}{Y} \cap \overset{\circ}{C}_k$ exactly. Here $\overset{\circ}{\cdot}$ denotes the interior of the corresponding domain. The set $\{C_k\}_{k=1}^{n_c}$ can be determined easily by exploiting the fact that $\mathcal{T}^{cw}$ and $\mathcal{T}^{yp}$ stem from the same macro triangulation $\mathcal{T}^0$.

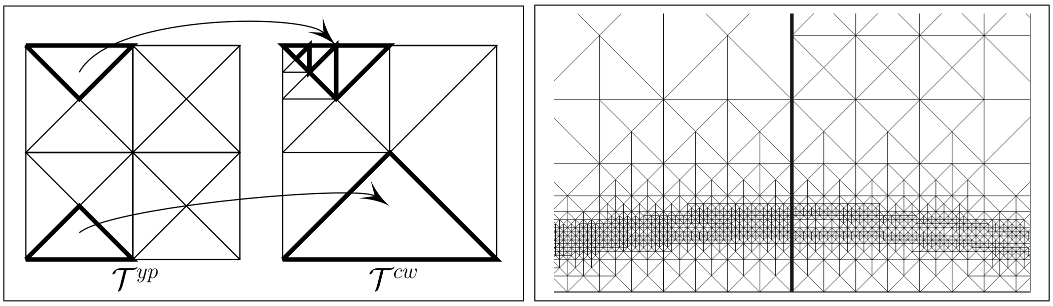

Figure 8.2: Left: Situations resulting from Assumption 8.14. Right: Typical relation of $\mathcal{T}^{cw}$ (left) and $\mathcal{T}^{yp}$ (right) after refinement.

Given an integration point (x, y) in triangle $Y \in \mathcal{T}^{yp}$, we know the corresponding mother triangle $T^0 \in \mathcal{T}^0$, i.e. the triangle from which Y is obtained by refinement. Thus (x, y) is also located in a triangle $C \in \mathcal{T}^{cw}$ which has $\mathcal{T}^0$ as mother triangle. The search of the triangle C in $\mathcal{T}^{cw}$ which contains (x, y) can therefore be restricted to all children of $\mathcal{T}^0$. This registration can be performed a priori before integration and is required only once per time step.

To perform this integration with reasonable computational effort we pose the following assumption

Assumption 8.14. For each pair of triangles $Y \in \mathcal{T}^{yp}$ and $C \in \mathcal{T}^{cw}$ there either holds $Y \subseteq C$ or $C \subseteq Y$ or $Y \cap C$ is a lower dimensional facette.

Using this assumption exactly two situations may occure:

$n_c = 1$: We either have $Y \subset C_1$ or $Y \equiv C_1$. In both cases integration can be performed by evaluating c^h and w^h at the integration points in Y.

$n_c > 1$: We have $Y = \bigcup_{k=1}^{n_c} C_k$. Exact integration now can be performed by integrating over each C_k and summing up for $k = 1, \ldots, n_c$.

This is shown in Figure 8.2, left plot. The case $n_c > 1$ is illustrated with the top triangle, while the case $n_c = 1$ is illustrated with the bottom triangle. For the bold triangle $Y \in \mathcal{T}^{yp}$ the corresponding set $\{C_k\}_{k=1}^{n_c}$ on the right is marked bold. In Figure 8.2, right plot, we show the resolution of $\mathcal{T}^{cw}$ (left) and $\mathcal{T}^{yp}$ (right) resulting from a typical simulation.

We note that Assumption 8.14 is satisfied if *bisection by newest vertex* ([Che08]) is used as refinement strategy and the initial meshes $\mathcal{T}^{cw}_{(0)}$ and $\mathcal{T}^{yp}_{(0)}$ are constructed from the same macro mesh $\mathcal{T}^0$.

9 Numerical examples

In this section we present the behaviour of our adaptive finite element solver for the simulation of the Cahn–Hilliard Navier–Stokes system. The implementation is done in C++. The refinement and coarsenening algorithms are based on *i*FEM [Che08]. As direct solvers we use SuperLU [DEG$^+$99] for non symmetric matrices and cholmod [CDHR08] for symmetric linear systems.

Due to our adaptive finite element method a direct solving using a LU decomposition turned out to be feasible for the systems arising in our Newton solver for the Cahn–Hilliard system. In [HHT11] a Schur complement based algorithm with incomplete LU factorization from SuperLU is used together with a BiCGSTAB iteration [van92, Mei08]. For more details on preconditioning techniques for the Cahn–Hilliard equation we refer to [BDQN12, BSB14]. Multigrid algorithms for the numerical solution of related problems are presented in [GK07, KW06].

We stop our semi-smooth Newton iteration as soon as a residuum of

$$tol_{SSN} = 10^{-6} + 10^{-12}\|F^{(1)}(c^0, w^0), F^{(2)}(c^0, w^0)\|$$

is reached, where c^0, w^0 denote the initial iterates of Newton's method. Stop typically is achieved after less than five steps.

For the solution of the saddle point problems (7.11) arising in the solution of the Navier-Stokes system we use a right sided preconditioned restarted gmres iteration presented in [SS86] with restart after 10 iterations. We use an upper triangular preconditioner ([BGL05, 10.1.2] [BP88]). This requires to find preconditioners for A and the Schur complement $S = -BA^{-1}B^t$, where A and B are denoted in (7.11). The A-block is preconditioned using a LU decomposition, while the Schur complement is preconditioned by the F_p preconditioner from [KLW02]. In our tests this solution strategy typically converges in less than 20 iterations if an absolute residual less than 10^{-8} is requiered.

We mention some work related to the solution of saddle point problems and especially of saddle point problems arising in the solution of Navier–Stokes equations. In [PRR05] a comparison of three different common numerical approaches is presented. Concerning the constructing of appropriate preconditioners we mention [DGSW10] and the book ([GR11]) on the simulation of two-phase flows. For an extensive overview of approaches for solving general saddle-point problems we refer the reader to [BGL05] and the many references therein.

In the following we demonstrate how our solution concept for the Cahn–Hilliard Navier–Stokes system performs in numerical practice. We start by showing the behaviour of our Newton solver with respect to the parameter s and the mesh size h in Section 9.1 and propose a coupling between the size of the smallest triangle and the parameter s. The resolution of the interfacial region obtained by our adaptive approach then is compared with results obtained by classical heuristic refinement strategies. We end the section with a comparison of solutions obtained with lumped inner product $(\lambda_s(c^h), v)$ as performed in [HHK13] and exactly evaluated inner product as proposed here.

In Section 9.2 a comparison of results obtained using adaptive meshes with results obtained by using homogeneous meshes is presented. We finish with some results concerning spinodal decomposition in Section 9.3.

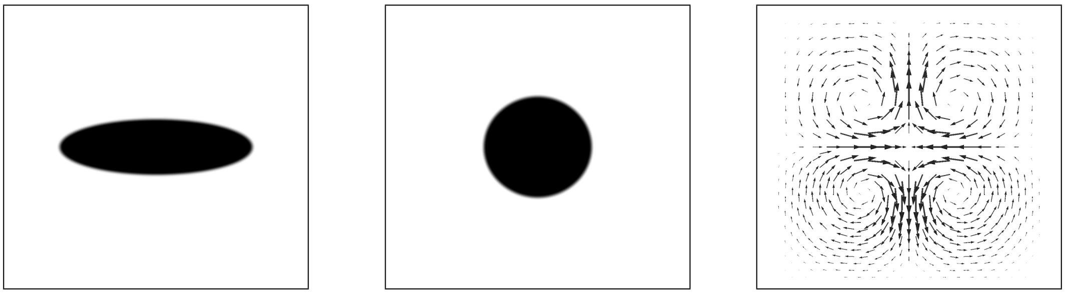

Figure 9.1: Phase field at $t = 0$ (left) and $t = 1000\tau$ (middle) together with the flowfield at $t = 1000\tau$ (right).

9.1 Test problem: Ellipse to circle

We now report on the behaviour of the Newton solver for solving (7.2)–(7.3) with respect to the parameters h, γ and s. We also compare meshes and concentrations obtained with and without lumping of the penalization. We further show how our adaptive concept aligns the meshes to the interface. As test example we here consider the following initial phase field

$$c_0(x, y) := -\tanh\left\{1000 \cdot \left[\left(\frac{x - 0.5}{0.35}\right)^2 + \left(\frac{y - 0.5}{0.1}\right)^2 - 1\right]\right\}.$$

This gives an ellipse centered at $(0.5, 0.5)$ with half axes of size 0.35 and 0.1 inside the unit square $\Omega = (0, 1)^2$. It holds $c^0(x, y) \approx 1$ in the interior of the ellipse and $c^0(x, y) \approx -1$ in its complement. For the flow we set $y^0(x, y) = 0$ and prescribe $y = 0$ on the boundary. Thus the flowfield is only driven by the interface. The parameters are set to $\tau = 0.01$, $\mathrm{Pe} = 1$, $K = 1$, $\mathrm{Re} = 100$ and $\gamma = (50\pi)^{-1}$. The parameters for the adaptive process are given by $\theta^r := 0.5$, $\theta^c := 0.1$, $a_{\min} := 2.5\times 10^{-6}$ and $a_{\max} := 0.01$ for $\mathcal{T}^{cw}$, and $\theta^r := 0.1$, $\theta^c := 0.05$, $a_{\min} := 2.5 \times 10^{-6}$ and $a_{\max} := 6.25 \times 10^{-4}$ for $\mathcal{T}^{yp}$.

In Figure 9.1 we show snapshots of the concentration together with the flowprofile at time $t = 1000\tau$. Here and in the following darker gray indicates higher values and lighter gray indicates smaller values. Especially when the phase field is displayed, black indicates values close to 1 and white indicates values close to -1.

Performance of the semi-smooth Newton solver with respect to the parameters h and s.

We show the performance of the Newton solver with respect the resolution of the mesh and the penalisation parameter. Since Newton's method for fixed s is formulated in function space we expect mesh independent behaviour of the iterative process, see e.g. [ABPR86, HU04]. To check the mesh independent behaviour of our solver, we simulate ten time steps on homogeneous meshes of gridsize h for various values of s and count the number of Newton steps

s	h	Newton steps	$\|c - c^*\|_{L^2(\Omega)}$
2.00×10^2	0.008839	2	2.193511e-2
	0.006250	2	2.173868e-2
	0.004419	2	2.175180e-2
	0.003125	2	2.178268e-2
	0.002210	2	2.180643e-2
1.60×10^3	0.008839	3	5.340040e-3
	0.006250	2	3.381581e-3
	0.004419	2	2.757271e-3
	0.003125	2	2.638523e-3
	0.002210	2	2.619292e-3
1.28×10^4	0.008839	4	4.787429e-3
	0.006250	4	2.287193e-3
	0.004419	4	1.055169e-3
	0.003125	3	5.633583e-4
	0.002210	3	3.657870e-4
1.02×10^5	0.008839	4	4.790669e-3
	0.006250	6	2.280048e-3
	0.004419	5	1.021374e-3
	0.003125	5	4.766472e-4
	0.002210	5	1.830728e-4
8.19×10^5	0.008839	4	4.792279e-3
	0.006250	7	2.281653e-3
	0.004419	6	1.022756e-3
	0.003125	10	4.767968e-4
	0.002210	8	1.801123e-4

Table 9.1: Number of Newton steps needed for various h and s.

needed to solve the system in the tenth time step. We perform these steps to exclude the influence of the initial value. We also compare the solutions against a reference solution c^* obtained on a very fine adaptive mesh with smallest triangle of size $a^*_{\text{min}} = 5 \times 10^{-7}$ and relaxation parameter $s^* = 3 \times 10^6$.

In Table 9.1 our results are depicted. In the first two columns the values of s and h are given. In the third column we show the number of Newton steps needed to obtain an absolute residual of size 10^{-6}, and in the fourth column we show the L^2-difference between the solution and the reference solution.

We see that the number of Newton steps indeed seems to be independent of the actual meshsize h and only does mildly increase with increasing s.

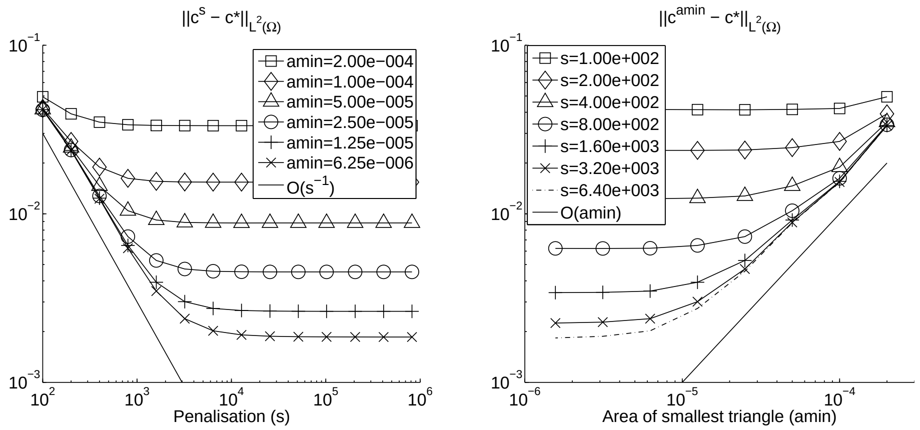

Figure 9.2: Error decay in c with respect to s and $a_{\min}$.

The coupling between s and $a_{\min}$

Column 4 of Table 9.1 indicates that increasing the penalisation parameter s reduces the error in the solution only up to a threshold $\overline{s}$, while for larger values of s the error is dominated by the error introduced by the spatial discretization. Thus further increasing the parameter s only increases numerical effort while not reducing the overall error. We thus next investigate the dependence of the error in c with respect to variation of s and $a_{\min}$, where $a_{\min}$ denotes the area of the smallest triangle allowed during the adaptive cycle. As will be shown later errors are clustered at the interface where the mesh is refined to the finest level. Thus we can expect that the total error using a homogeneously refined mesh of the same meshsize would only slightly increase the quality of the solution and we can use the area of the smallest triangle as a reference size.

We compare solutions obtained on adaptively refined meshes with $a_{\min} \in [2\times10^{-4}, 8\times10^{-7}]$ and penalty parameters $s \in [1\times10^2, 8\times10^5]$ with the solution obtained on a fine adaptive grid with $a^*_{\min} = 2 \times 10^{-7}$ and $s^* = 8 \times 10^7$. In Figure 9.2 we show the L^2 error in c in dependence of s (left) and $a_{\min}$ (right).

From the left plot in Figure 9.2 we deduce the relation $\|c^s - c^*\|_{L^2(\Omega)} \sim s^{-1}$, where c^* denotes the fine reference solution and c^s denotes the solution for the specific value of s. The right plot indicates $\|c^{a_{\min}} - c^*\|_{L^2(\Omega)} \sim a_{\min}$ where $c^{a_{\min}}$ denotes the solution for the specific value of $a_{\min}$. Thus, for equilibrating the error we propose to couple the penalisation parameter s and the size of the smallest triangle by $s = \mathcal{O}(a_{\min}^{-1})$.

A comparison with heuristic mesh adaptation

Frequently meshes for solving Cahn–Hilliard equations are constructed using heuristic strategies. We now compare the meshes obtained by the adaptation strategy presented in the present work with two heuristic strategies exploiting the knowledge of the location of the interface. The following comparison is taken from [HHK13].

The first approach uses the fact that $|\nabla c|$ is large in the transition region from $c = 1$ to $c = -1$ and constructs fine meshes with a local grid size which is related to the value of $|\nabla c|$ (small T where $|\nabla c|$ is large). The second heuristic approach constructs a fine mesh in regions where $1 - |c|$ is larger than a certain threshold, as proposed in [BBG11, KSW08]. For our comparison we use the parameters $\theta^r = 0.5$, $\theta^c = 0.1$, $a_{\min} = 10^{-6}$ and $a_{\max} = 0.01$. The parameter $thres > 0$ needed below is set equal to γ.

An adaptive concept based on $\|\nabla c\|_{L^2(T)}$: In this approach, on $T \in \mathcal{T}_{cw}$ we define the (local) indicator $\eta_G^T = \|\nabla c_T\|^2_{L^2(T)}$. We test two different marking strategies. The first approach is a marking according to the same rules and with the same parameters θ^r, θ^c, $a_{\min}$ and $\alpha_{\max}$ as used in our residual based approach. We also employ a marking which is based on balancing of indicators. For this we define an overall tolerance $tol > 0$ and mark a triangle T for refinement if $\eta_G^T > tol^2/NT$ holds, and mark it for coarsening, if $\eta_G^T < \theta^c tol^2/NT$ is satisfied. Here we use $tol = thres$ and the same θ^c as for the bulk-type strategy. NT denotes the number of triangles in $\mathcal{T}^{cw}$.

An adaptive concept based on $|c|$ on T: In this approach, a triangle $T \in \mathcal{T}^{cw}$ is marked for refinement whenever the indicator $\eta_V^T = \min_T(1 - |c|)$ satisfies $\eta_V^T > thres$. It is marked for coarsening if $\eta_V^T < 0$ holds. If $0 \le \eta_V^T \le thres$ is satisfied the triangle is left untouched.

In Figure 9.3 we compare the refinement obtained by the four adaptation strategies under consideration at the left arc of the ellipse at $t = 100\tau$ and show the distribution of the error indicators η_T and η_{TE} across the interface. The top left mesh is obtained with our strategy while the top middle plot shows the indicator η_T and the top right plot shows the indicator η_{TE}. The mesh on the bottom left results from the refinement based on the size of $1-|c|$, and the last two meshes on the bottom are obtained by the approach based on the size of $\|\nabla c\|_{L^2(T)}$ using either the bulk type marking (middle) or the tolerance marking (right). The left bold line indicates the isoline $c = -1$, the right bold line the discrete isoline $c = 1$.

Our approach performs as expected: refinement takes place in the neighourhood of the discrete isolines $c = 1$ and $c = -1$ since the function c develops a kink at these locations. This is reflected in the localization of the error indicators η_T and η_{TE} which are mainly located around the discrete isolines. Between these isolines the function c is smooth and the constructed grid is coarser as in the neighborhood of the isolines. The error indicators are small between the isolines corresponding to the smooth solution.

The adaptive approach based on the value of $1 - |c|$ refines the interface to the smallest triangle possible, regardless of how large the actual error contribution on the corresponding triangle is. This results in a uniformly refined mesh between the isolines $c = 1$ and $c = -1$.

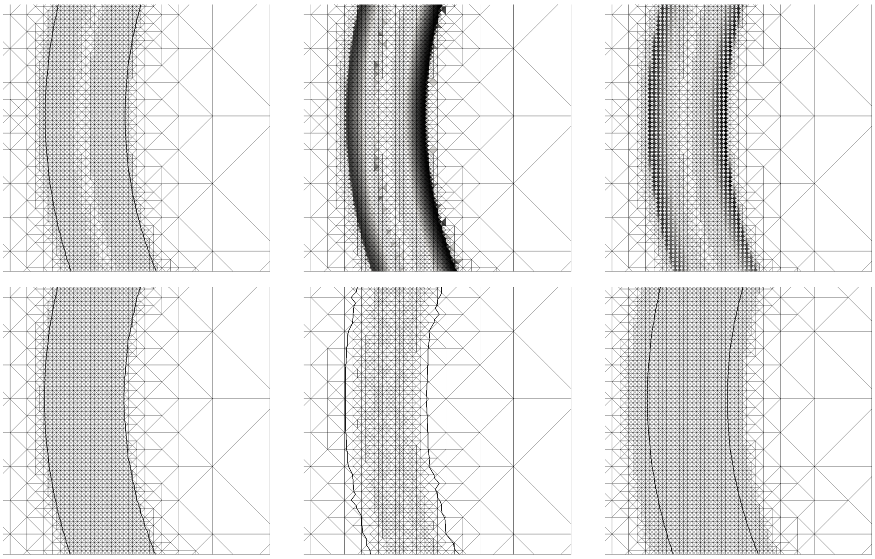

Figure 9.3: The triangulation of the interface at $t = 100\tau$ using residual based adaptation (top, left), value based adaptation (bottom, left), gradient based adaptation using the bulk marking (bottom, middle) and gradient based adaptation using the tolerance marking (bottom, right). The bold lines indicate the discrete isolines $c = -1$ and $c = 1$. The error indicators η_T and η_{TE} are shown at the top middle, resp. top right.

The triangulation obtained by adaptation based on ∇c with bulk marking delivers a fine mesh around $c = 0$ and tends to construct coarse meshes around the isolines $c = 1$ and $c = -1$, i.e. where the largest numerical error is expected. Finally, tolerance marking delivers a uniformly refined mesh in the whole interface and also in the region around the interface.

In summary, the triangulation obtained by our residual approach for sufficiently small $a_{\min}$ delivers meshes with the smallest mesh size at the boundary of the interface and of medium mesh size in the interface region. The meshes obtained by the approach based on the size of $1 - |c|$ look similar but are uniform in the interfacial region and therefore contain more triangles than our approach. The bulk-type marking strategy using ∇c seems not to be useful, while the tolerance strategy delivers meshes similar to our strategy but again with more triangles and a uniform resolution in the interfacial region. We further mention that the tolerance to choose for the ∇c based strategy is not an intrinsic size, thus we see no guideline for choosing it, while in the case of the value based adaptation the threshold gives a distance to the discrete isolines and in the residual based case it is an amount of error contribution. Finally, we emphasize that our approach delivers reliable error bounds.

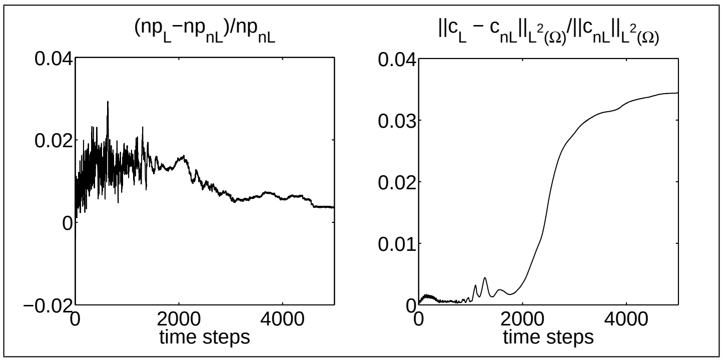

Figure 9.4: Relative number of nodes with lumping (left) and relative L^2-difference between the solutions (right).

Comparison of meshes and solutions obtained with and without lumping.

We also report on the influence of the error introduced through lumping of the inner product $(\lambda_s(c^h), v)$ on the adaptive concept proposed in [HHT11, HHK13], see Remark 8.2.

To begin with we by $\mathcal{T}^L$ we denote the meshes obtained by using the adaptive strategy in [HHK13] based on a simulation using lumping and by $\mathcal{T}^{nL}$ we denote the meshes obtained from the adaptive concept presented in Section 8.3 based on a simulation without lumping. The respective phase fields are denoted by c_L and c_{nL}. Due to the additional terms in the a-posteriori error estimator in [HHT11, HHK13] we expect $\mathcal{T}^L$ to contain more triangles than $\mathcal{T}^{nL}$. We also expect that these extra triangles are located around the interface as the extra indicator is concentrated there. In Figure 9.4 we show the relative difference between the degrees of freedom in $\mathcal{T}^L$ and $\mathcal{T}^{nL}$ as a function of the time step number (left plot). In the right plot we show the relative L^2-difference between c_L and c_{nL} plotted over time.

We see that $\mathcal{T}^L$ in this example contains about one percent more nodes than $\mathcal{T}^{nL}$. During the first pahse of the simulation we observe a large variation in the number of nodes obtained by the two approaches. If the length of the interface is large, as it is the case in spinodal decomposition (see Section 9.3), this relative difference reaches approximately five percent. The constant difference at the end of the simulation arises from reaching the final state which in this example is given by a circle. Since lumping induces a further error contribution we expect that the solutions without lumping are more accurate. The higher numerical effort caused by exact evaluation of $(\lambda_s(c^h), v)$ in total pays off, since the extra numerical work only has to be performed on a few number of triangles and thus is small. The reduction of the number of triangles reduces the effort in the numerical expensive part of solving the systems arising in the Newton iteration. We note that both, the estimator for the simulation with lumping and without lumping, are reliable.

9.2 Test problem: Circle in lid

We now compare numerical results obtained on homogeneous meshes with numerical results obtained on adapted meshes. Since the estimator for the Navier–Stokes system depends on the flow field y_{old} we use a setting with inhomogeneous Dirichlet boundary data for the flowfield, so that we can expect a relevant transport contribution in the system.

Here in $\Omega = (0,1)^2$ the initial concentration corresponds to a circle centered at $m_c = (0.5, 0.5)$ with radius $r = 0.25$, i.e. we set

$$c_0(x,y) := -\tanh\left\{1000 \cdot \left[\left(\frac{x-0.5}{0.25}\right)^2 + \left(\frac{y-0.5}{0.25}\right)^2 - 1\right]\right\}.$$

Furthermore, in (3.7) we set

$$g(x,y) := \begin{cases} (16x^2(1-x)^2, 0)^t & \text{if } y \equiv 1, \\ 0 & \text{else.} \end{cases}$$

The initial flowfield then is calculated as the solution of a stationary Stokes equation with this boundary data. The parameters are given as $\tau = 0.01$, $\mathrm{Pe} = 200$, $K = 1$, $\mathrm{Re} = 400$ and $\gamma = (50\pi)^{-1}$. The adaptation is controlled by $\theta^c = 0.1$ and $\theta^r = 0.5$ for both $\mathcal{T}^{cw}$ and $\mathcal{T}^{yp}$. The lower bounds for the triangles in $\mathcal{T}^{cw}$ and $\mathcal{T}^{yp}$ are given by $a^{cw}_{\min} = 5.7 \times 10^{-6}$ and $a^{yp}_{\min} = 2 \times 10^{-5}$. The respective upper bounds are set to $a^{cw}_{\max} = 0.01$ and $a^{yp}_{\max} = 6.25 \times 10^{-4}$, respectively.

In Figure 9.5 we show snapshots of the evolution of the bubble. We note that the interface does not touch the boundary of Ω during its evolution. We see that the bubble in general follows the velocity field driven by the boundary data, while the velocity field gets taylored to the bubble and near the interface is parallel to the interface.

Error decay on uniform and adaptively refined meshes

We next compare the errorlevel obtained on structured uniform meshes with the error level obtained on our adaptive meshes. Since no analytical solution is available, we compare solutions obtained on coarser grids to a solution on a fine homogeneous mesh. The outline of the test reads as follows. From the Cahn–Hilliard part we calculate a stationary bubble using our adaptive method, and compute a flowfield from the stationary Stokes system with boundary data as specified above, where we use a fine mesh. Then we perform five time steps with stepsize $\tau = 0.01$ to align the flowfield to the interface. With this data, we perform the next time step, where we compare the numerical results obtained on successively refined homogeneous meshes with those obtained in the adaptive meshes. We use the same initial grid for the Cahn–Hilliard and the Navier–Stokes system.

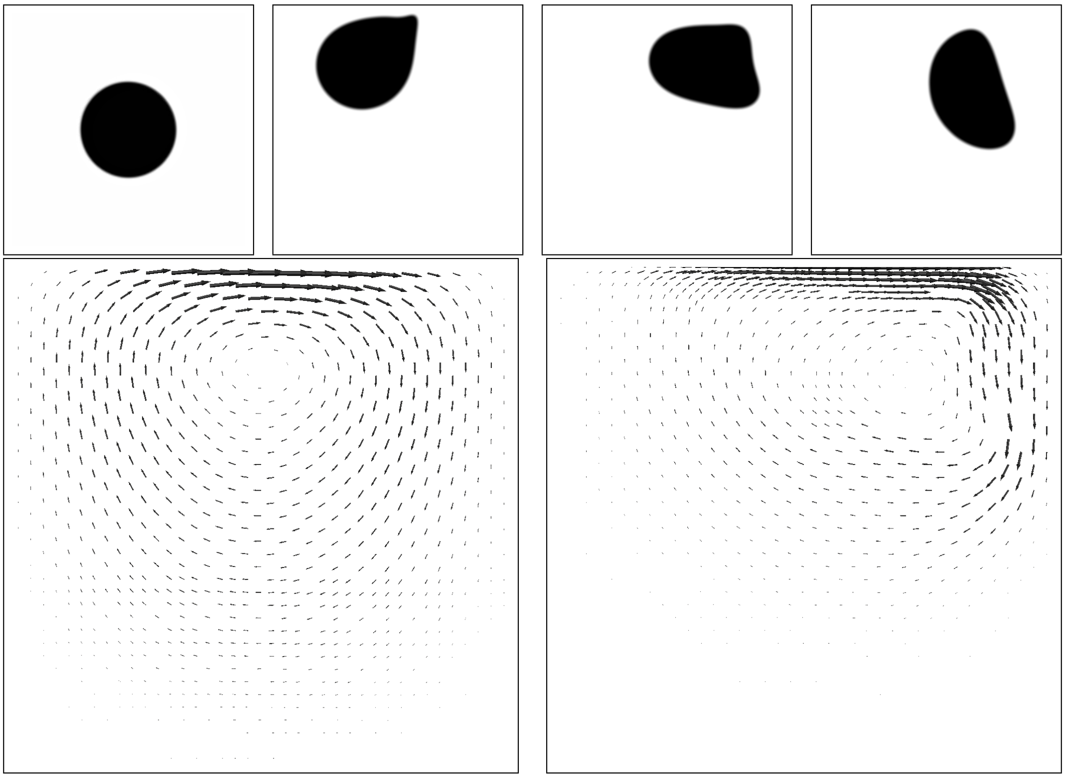

Figure 9.5: Phase field c at $t = 0$, $t = 2500\tau$, $t = 5000\tau$ and $t = 7500\tau$ (top, left to right), and the flowfield at $t = 0$ and $t = 5000\tau$ (bottom, left to right).

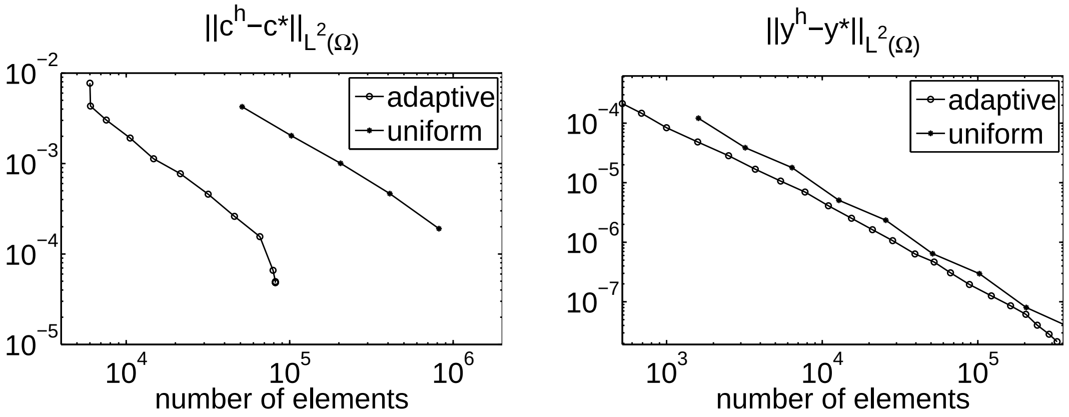

Figure 9.6: Error decay in c (left) and y (right) both for a homogeneously and an adaptively refined mesh.

In Figure 9.6 we show the decay of the error in the phase field (left) and in the flow field (right).

We observe that the adaptive method achieves the same accuracy as the method with uniform refinement with about 10% of the number of elements used for the uniform refinement. This clearly shows the benefit of the adaptive concept in the simulation of the Cahn–Hilliard part.

On right we show the decay in error plotted over number of triangles for the velocity mesh $\mathcal{T}^{yp}$. We stress that we omit the error indicator η_{SF} defined in (8.49), since it is concentrated at the interface, which in this example is of minor interest for the flowfield. The error in the velocity field using adapted meshes is only slightly smaller than the error obtained on homogeneous meshes, since the flow is quite laminar, so that the error is distributed homogeneously over the domain.

In Figure 9.7 we show snapshots of the evolution of the resulting meshes $\mathcal{T}^{cw}$ (top) and $\mathcal{T}^{yp}$ (bottom). We see that $\mathcal{T}^{cw}$ is refined to the coarsest level outside the interface and is refined to the finest level at the borders of the interface as already observed in the test comparing different adapt strategies for the construction of $\mathcal{T}^{cw}$ in Section 9.1. For $\mathcal{T}^{yp}$ due to the tangential boundary data we expect the adaptive concept to refine at the top part of the domain as can be seen in Figure 9.7 (bottom).

9.3 Test problem: Spinodal decomposition

Our last test problem is the simulation of spinodal decomposition. Spinodal decomposition describes the demixing of a two-phase fluid from a stable homogeneous mixture into its two phases by phase separation. A typical example how to make a stable mixture separating into its two phases is a change of temperature, see e.g. [FM08, Sig79]. John W. Cahn and John E. Hilliard investigate this process for metals in [CH58]. The process of phase separation

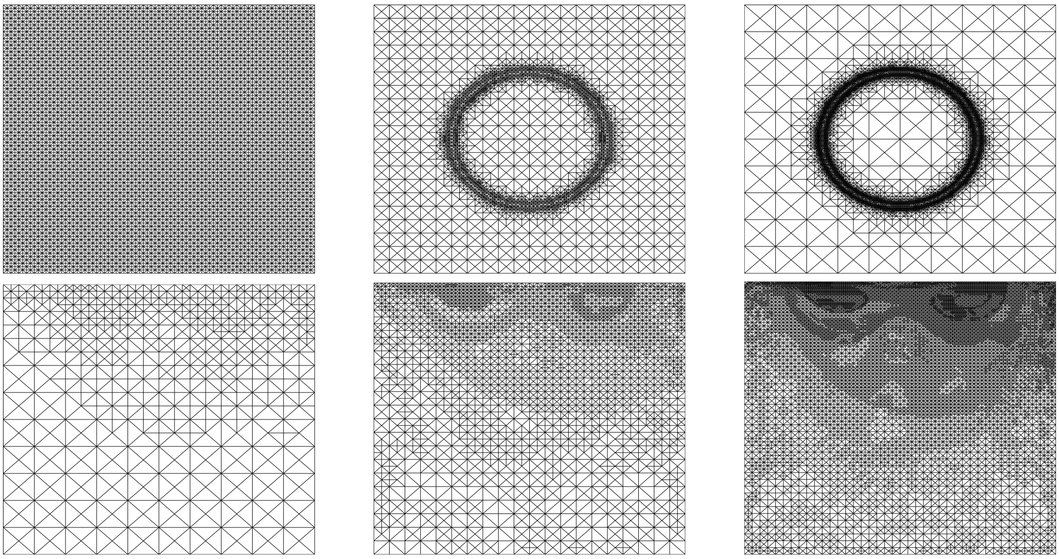

Figure 9.7: The evolution of $\mathcal{T}^{cw}$ (top) and $\mathcal{T}^{yp}$ (bottom) through the adaptive cycle. For $\mathcal{T}^{cw}$ we show the initial mesh, the mesh obtained after three cycles and the one obtained after six cycles (left to right). For $\mathcal{T}^{yp}$ we show the mesh obtained after four, eight and twelve cycles (left to right).

consists of two stages. In a first very rapid stage the mixture separates into its two phases yielding small scaled structures. After this initial decomposition the two phases further evolve driven by diffusion to minimize the Ginzburg–Landau free energy (2.1), which results in the reduction of the length of the interface. In material science this process is very important, see e.g. [EAK+01, BPC+07]. Alloys are often used to obtain materials with better properties than pure materials. Thus dealloying by phase separation may lead to reasonable threats.

In the original work [CH58] the separation is only driven by diffusion. In many cases the movement is also driven by advection of the particles. In this setting the motion of the interface is influenced by the resulting velocity field, see e.g. [Sig79, BOS11], so that extensions of the Cahn–Hilliard model with the Navier–Stokes equations are used to describe the physics in this situation.

In Figure 9.8 we show snapshots of such a spinodal decomposition, where we used $\tau = 1e-5$, $\mathrm{Pe} = \mathrm{Re} = 1$, and $\gamma = 1/(50\pi)$. For the adaptation of $\mathcal{T}^{cw}$ we use $a_{\min} = 5e-6$, $a_{\max} = 0.05$ and $\theta^r = 0.7$ and $\theta^c = 0.1$. For $\mathcal{T}^{yp}$ we use a homogeneous mesh with $h_y = 0.005$.

Since our adaptive strategy is mainly resolving the interface where the Ginzburg–Landau energy is concentrated we observe that the reduction in the number of triangles is proportional to the loss of energy. With the parameters used here we have a diffusion driven regime and thus we expect the energy E to scale like $E \sim t^{-1/3}$ (see [BOS11, Sig79]).

In Figure 9.9 we show the evolution of the Ginzburg–Landau energy and of the number of elements. We observe that both the Ginzburg–Landau energy and the number of elements decay like $t^{-1/3}$.

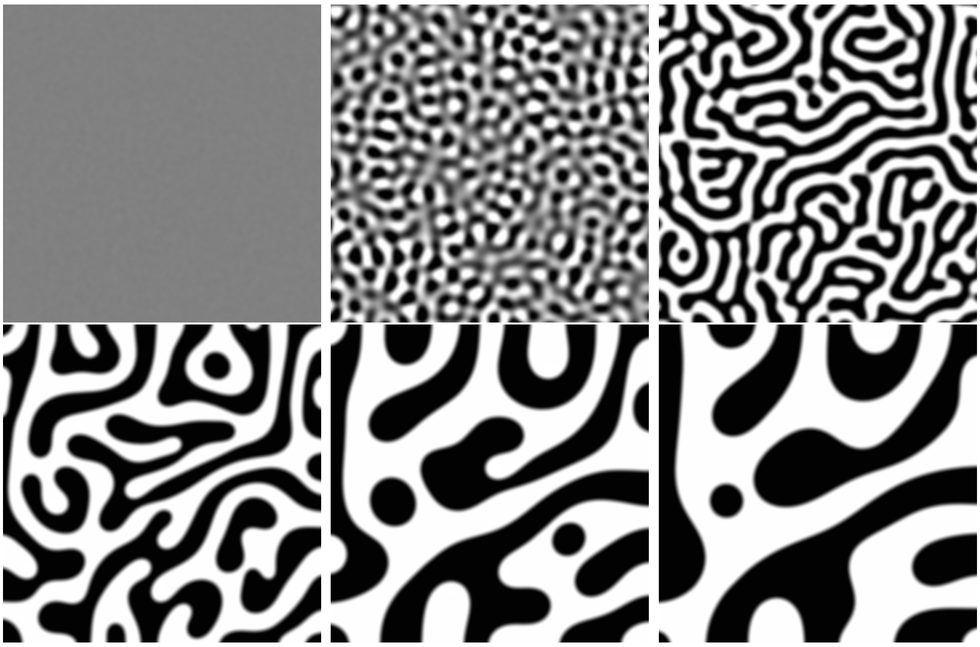

Figure 9.8: Spinodal decomposition at time $t = 0$, 100τ, 200τ, 1000τ, 4000τ, 8000τ (top, left) to (bottom, right).

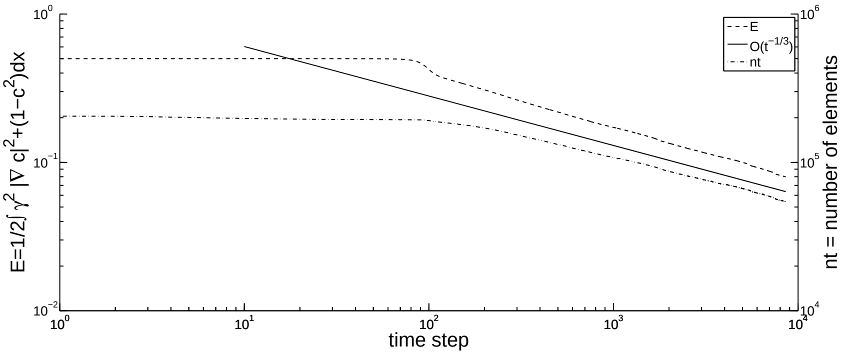

Figure 9.9: Evolution of Ginzburg–Landau energy and number of elements over time steps.

10 An application of the residual based adaptive concept to a fluid with different densities

The Navier–Stokes model considered in the previous sections has the drawback of assuming equal densities for the two fluids. In this section we apply our residual based adaptive finite element method to a model that works for different densities, namely we use the model presented in [AGG12]. There are several other models for the case of different densities available, see e.g. [Boy02, DSS07, LT98] and [AMW98] and the many references therein.

The model from [AGG12] has the advantage that our adaptive method for creating meshes for the simulation of the two-phase structure can be easily extended to it with only minor changes, and that it gives rise to energy estimates that we will exploit in Section 11.

10.1 The mathematical model for two-phase flow with different densities

The mathematical model we consider is derived in [AGG12, Ch. 3]. It reads:

$$\rho\partial_t y + ((\rho y + j)\cdot\nabla)\, y - \operatorname{div}(2\eta D y) + \nabla p = -\sigma\gamma \operatorname{div}(\nabla c\otimes\nabla c) + G, \tag{10.1}$$
$$\operatorname{div} y = 0, \tag{10.2}$$
$$\partial_t c - \operatorname{div}(m\nabla w) + y\cdot\nabla c = 0, \tag{10.3}$$
$$-\sigma\gamma\Delta c + \sigma\gamma^{-1}\Psi'(c) = w. \tag{10.4}$$

Here $Dy = (\nabla y + (\nabla y)^t)/2$ denotes the symmetrized gradient and y denotes the volume avaraged velocity. The pressure is denoted by p and the phase-field and chemical potential are denoted by c and w, respectively. An additional transport j arises from diffusion inside the interface and is given by $j = -\frac{\partial\rho}{\partial c}(c)m(c)\nabla w$, where

$$\rho = \rho(c) = \frac{\rho_2-\rho_1}{2}c + \frac{\rho_2+\rho_1}{2}$$

denotes the density of the fluid, where ρ_1, resp. ρ_2, denote the density of fluid A and B, respectively. By $\eta = \eta(c)$ we denote the viscosity of the fluid and by $m = m(c)$ its mobility. We assume that the mobility is bounded away from zero and from above, i.e. there exists $0 < \underline{m} \le \overline{m} < \infty$ such that $\underline{m} \le m(c) \le \overline{m}$ holds for all $c\in\mathbb{R}$. We note that Ostwald ripening effects might occure due to the non degenerate mobility (see [AGG12]). σ is a constant related to the surface energy density, see [AGG12, Sec. 4.3.4], and γ again is related to the thickness of the interface. Note that if we choose $\sigma = \gamma$ and $m(c) \equiv \frac{1}{\mathrm{Pe}}$ equations (10.3)–(10.4) are equal to (3.3)–(3.4), i.e. the Cahn–Hilliard system, that we investigated in the previous sections. The volume force G is given by $G = \rho g$, where g is the gravitational force. We further note that, for equal densities ($\rho_1 \equiv \rho_2$) and equal viscosities ($\eta_1 \equiv \eta_2$), the model (10.1)–(10.2) in absence of the gravitational force differs from (3.1)–(3.9) only in the definition of the pressure. Let p^{phys} denote the pressure in the physical system and $\tilde{p}$ denote the pressure in (3.1)–(3.9). Then we have the relations

$$\begin{aligned} p^{phys} =& p - \left(\frac{\sigma\gamma}{2}|\nabla c|^2 + \frac{\sigma}{2\gamma}(1-c^2) + \frac{s}{2}\lambda(c)^2 \right), \\ p^{phys} =& \tilde{p} - cw, \end{aligned}$$

where the latter holds if we choose $K = 1$ in (3.1).

We use the boundary conditions

$$\nabla w \cdot \nu_\Omega = \nabla c \cdot \nu_\Omega = 0 \qquad \text{on } \partial\Omega.$$

For the flow we either use no-slip conditions $y = 0$ on $\partial\Omega$, or free-slip conditions, i.e. $y \cdot \nu_\Omega = 0$, $\left(\nu_\Omega^\perp\right)_k \eta(c) D y \nu_\Omega = 0$ for $k = 1, \dots, d-1$, where ν_Ω denotes the outer normal to Ω and $\left(\nu_\Omega^\perp\right)_k$ denotes a basis for the tangential plane. Also a combination of these conditions might be used.

The free-slip condition means that no flow through the boundary is allowed and tangential to the boundary the natural boundary condition $(\eta(c)Dy - pI)\nu_\Omega = 0$ holds. Different types of boundary data with slip are described in [SS12]. For sake of simplicity we in the following use no-slip data.

Concerning the existence of solutions to (10.1)–(10.4) we refer to [ADG13a] for the case of non degenerated mobility, and to [ADG13b] for the case of degenerated mobility.

In [GK14] a thermodynamical consistent discretization of (10.1)–(10.4) is presented. In [AV12] simulations of this model for the [HTK$^+$09] benchmark were compared with simulations of the models from [Boy02] and [DSS07].

10.2 Time discretization and Moreau–Yosida relaxation

We next state a time discretization for (10.1)–(10.4) similiar to that derived in Section 4, where from here onwards we replace Ψ by the Moreau–Yosida regularization of the double-obstacle potential, see Section 5.

Let $\tau > 0$ again denote the time discretization parameter. We set $\xi = 1/\tau$ and use the semi-implicit Euler scheme for time discretization. Variables with the subscript $_{old}$ refer to the previous time instance. The weak formulation of the system to solve in the actual time step then reads:
Find $(y, p, c, w) \in H_0^1(\Omega)^d \times L^2_{(0)}(\Omega) \times H^1(\Omega) \times H^1(\Omega)$ fulfilling

$$\begin{aligned} \xi\,(\rho(y - y_{old}), v) + (((\rho y + j)\cdot\nabla)\, y, v)& \\ +(2\eta Dy : \nabla v) - (p, \operatorname{div} v)& \\ -\sigma\gamma(\nabla c \otimes \nabla c, \nabla v) - (G, v) &= 0 \quad \forall v \in H_0^1(\Omega)^d, \quad (10.5) \\ -(\operatorname{div} y, q) &= 0 \quad \forall q \in L^2_{(0)}(\Omega), \quad (10.6) \\ (c - c_{old}, v) + \tau(m_{old}\nabla w, \nabla v) - \tau(y_{old} c, \nabla v) &= 0 \quad \forall v \in H^1(\Omega), \quad (10.7) \\ \sigma\gamma(\nabla c, \nabla v) - (w, v) + (\lambda_s(c), v) - \sigma\gamma^{-1}(c_{old}, v) &= 0 \quad \forall v \in H^1(\Omega). \quad (10.8) \end{aligned}$$

Here $m_{old} := m(c_{old})$, and $j := -\frac{\partial \rho}{\partial c}(c)m(c)\nabla w$. We note that this discretization again decouples the Navier–Stokes system (10.5)–(10.6) from the Cahn–Hilliard system (10.7)–(10.8) in a way that we can first solve (10.7)–(10.8) using y_{old} to obtain c and w on the current time step, and then solve (10.5)–(10.6) to obtain y and p.

The choice of τ is adapted to a CFL condition. After each time step we calculate the local CFL values $CFL_T = \max_{x\in T} |y(x)|\tau/h_T$ on each triangle T and choose the next timestep τ such that $\max_T CFL_T \leq \theta$ holds, where $\theta > 0$ is a given threshold that we typically set to $\theta = 0.1$. In our numerical results we with $\theta = 0.1$ observe a stable behaviour of our time discretization scheme.

10.3 Spatial discretization

As in Section 7 we introduce spaces $\mathcal{V}^{cw}$, $\mathcal{V}^y$ and $\mathcal{V}^p$ using $P1$ elements for c^h and w^h ($\mathcal{V}^{cw}$), as well as for p^h ($\mathcal{V}^p$), and using $P2$ elements for the flowfield y^h ($\mathcal{V}^y$).

In each timestep we solve the following discrete problem to approximate solutions to (10.5)–(10.8):
Find $(y_h, p_h, c_h, w_h) \in \mathcal{V}^y \times \mathcal{V}^p \times \mathcal{V}^{cw} \times \mathcal{V}^{cw}$ such that the following holds:

$$\begin{aligned}
&\xi \int_\Omega \rho_h (y_h - y_{old}) v_h \, dx \\
&+\frac{1}{2}\int_\Omega \rho_h y_h \nabla(y_h \cdot v_h)\, dx + \frac{1}{2}\int_\Omega \rho_h \left((y_h\cdot\nabla)y_h\right) v_h \, dx \\
&-\frac{1}{2}\int_\Omega \rho_h \left((y_h\cdot\nabla)v_h\right) y_h \, dx \\
&+\frac{1}{2}\int_\Omega j_h \nabla(y_h \cdot v_h)\, dx + \frac{1}{2}\int_\Omega \left((j_h\cdot\nabla)y_h\right) v_h \, dx \\
&-\frac{1}{2}\int_\Omega \left((j_h\cdot\nabla)v_h\right) y_h \, dx \\
&+\int_\Omega 2\eta_h Dy_h : Dv_h \, dx - \int_\Omega p_h \mathrm{div}\ v_h \, dx \\
&-\sigma\gamma \int_\Omega (\nabla c_h \otimes \nabla c_h) : \nabla v_h \, dx - \int_\Omega G v_h \, dx = 0 \quad \forall v_h \in V^y, \qquad (10.9)
\end{aligned}$$

$$-\int_\Omega q_h \mathrm{div}\ y_h \, dx = 0 \quad \forall q_h \in V^p, \qquad (10.10)$$

$$\begin{aligned}
&\int_\Omega (c_h - c_{old}, v_h) + \tau \int_\Omega m(c_{old}) \nabla w_h \nabla v_h \, dx \\
&-\tau \int_\Omega c_h y_{old} \nabla v_h \, dx = 0 \quad \forall v_h, \in \mathcal{V}^{cw},
\end{aligned} \qquad (10.11)$$

$$\sigma\gamma \int_\Omega \nabla c_h \nabla v_h \, dx + \int_\Omega (\lambda_s(c_h) - \sigma\gamma^{-1} c_{old} - w) v_h \, dx = 0 \quad \forall v_h \in \mathcal{V}^{cw}. \qquad (10.12)$$

Here, $j_h := -\frac{\partial\rho}{\partial c}(c_h)m(c_h)\nabla w_h$, $\eta_h := \eta(c_h)$, and $\rho_h := \rho_h$. The form of (10.9) is motivated by the identity

$$2\int_\Omega ((u\cdot\nabla)v)w\,dx = \int_\Omega u\nabla(v\cdot w)\,dx + \int_\Omega ((u\cdot\nabla)v)w\,dx - \int_\Omega ((u\cdot\nabla)w)v\,dx$$

which holds for all $u, v, w \in H^1(\Omega)^d$.

10.4 The adaptive concept

In Section 8.1 we derive an a-posteriori error estimator for the discretized Cahn–Hilliard system (7.2)–(7.3) of Section 7. Concerning equations (10.11)–(10.12) there are only minor changes and thus a residual based error estimator for (10.11)–(10.12) can be derived along the lines of Section 8.1. Since the derivation is straight forward we here only state the results. We again assume $y_{old} \in L^\infty(\Omega)$.

We define the following element residuals:

$$\begin{aligned} r_h^{(1)} &:= c^h - c_{old} + \tau y_{old}\nabla c^h - \tau\nabla w^h\cdot\nabla m(c_{old}),\\ r_h^{(2)} &:= \lambda_s(c^h) - w^h - \sigma\gamma^{-1}c_{old}, \end{aligned}$$

and the error indicators:

$$\begin{aligned} \eta_T^{(1)} &:= h_T\|r_h^{(1)}\|_T, & \eta_E^{(1)} &:= h_E^{1/2}\|m(c_{old})\left[\nabla w^h\right]_E\cdot\nu_E\|_E,\\ \eta_T^{(2)} &:= h_T\|r_h^{(2)}\|_T, & \eta_E^{(2)} &:= h_E^{1/2}\|\left[\nabla c^h\right]\cdot\nu_E\|_E. \end{aligned}$$

Theorem 10.1. *There exists a constant $C > 0$ depending only on the domain Ω and the smallest angle of the mesh $\mathcal{T}_h^{cw}$ such that there holds*

$$\tau\underline{m}\|\nabla e_w\|^2 + \left(\frac{1}{2}\sigma\gamma - \frac{\tau C_p^2}{\underline{m}}\|y_{old}\|_\infty^2\right)\|\nabla e_c\|^2 + s^{-1}\|e_{\lambda_s}\|^2 \le C\eta_\Omega^2$$

where η_Ω is given by

$$\begin{aligned} \eta_\Omega^2 =& (\tau\underline{m})^{-1}\sum_{T\in\mathcal{T}^{cw}}\left(\eta_T^{(1)}\right)^2 + \tau\underline{m}^{-1}\sum_{E\in\mathcal{E}^{cw}}\left(\eta_E^{(1)}\right)^2\\ &+ (\sigma\gamma)^{-1}\sum_{T\in\mathcal{T}^{cw}}\left(\eta_T^{(2)}\right)^2 + \sigma\gamma\sum_{E\in\mathcal{E}^{cw}}\left(\eta_E^{(2)}\right)^2. \end{aligned}$$

Thus the estimator is able to bound the error from above and thus is reliable.

Remark 10.2.
For $\sigma \equiv \gamma$ and $m(c_{old}) \equiv \underline{m} \equiv \mathrm{Pe}^{-1}$ this estimator coincides with that given in Theorem 8.1.

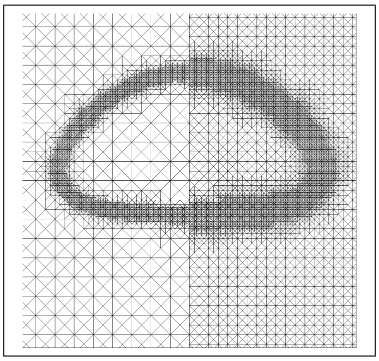

Figure 10.1: Relation between $\mathcal{T}^{cw}$ (left) and $\mathcal{T}^{yp}$(right).

Theorem 10.3. *There exists a constant $C > 0$ depending on s^{-1}, σ, γ, τ, $m(c_{old})$, Ω, $\|y_{old}\|_{\infty,\Omega}$ and the smallest angle of the mesh $\mathcal{T}^{cw}$ such that*

$$\begin{aligned} C\eta_\Omega^2 \leq & s^{-1}\|e_{\lambda_s}\|^2 + \sigma\gamma\|\nabla e_c\|^2 + \tau\overline{m}^2\underline{m}^{-1}\|\nabla e_w\|^2 \\ & + (\tau\underline{m})^{-1} osc_h^2(r_h^{(1)}, \Omega) + (\sigma\gamma)^{-1} osc_h^2(r_h^{(2)}, \Omega) \end{aligned}$$

holds.

Thus the estimator bounds the error from below, and thus also is efficient.

10.5 Meshes

For the construction of $\mathcal{T}^{cw}$ we use the strategy described in Section 8.3.

For the construction of $\mathcal{T}^{yp}$ we use an heuristic approach which we motivate in the following. Modell (10.1)–(10.4) contains the homogeneously distributed gravitational force as volume force. We do not consider further external forces. Thus it seems to be sufficient to match $\mathcal{T}^{yp}$ to $\mathcal{T}^{cw}$ on the interface to resolve the locallized force $\mathrm{div}(\nabla c \otimes \nabla c)$ well and to use a refined mesh outside the interface, see Figure 10.1. Since in this way $\mathcal{T}^{yp}$ is a locally refined version of $\mathcal{T}^{cw}$ we can represent c_h and w_h exactly on $\mathcal{T}^{yp}$, which simplifies numerical integration.

10.6 Numerics

We now present results we obtain by our simulations. Concerning the solution of the nonlinear system (10.9)–(10.12) we again use Newton's method to solve the Cahn–Hilliard equation as described in Section 7. For the Navier–Stokes part we use an Oseen fixpoint iteration. In the $k-th$ iteration of the fixpoint method we have to solve the saddle-point problem

$$\begin{pmatrix} A(y_h^k) & B^t \\ B & 0 \end{pmatrix} \begin{pmatrix} y_h^{k+1} \\ p_h^{k+1} \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix}.$$

This is the matrix representation of (10.9)–(10.10), where the transport terms are substituted by

$$\begin{aligned}&\frac{1}{2}\int_\Omega \rho_h y_h^k \nabla(y_h^{k+1}\cdot v_h)\,dx\\&+\frac{1}{2}\int_\Omega \rho_h\left((y_h^k\cdot\nabla)y_h^{k+1}\right)v_h\,dx-\frac{1}{2}\int_\Omega \rho_h\left((y_h^k\cdot\nabla)v_h\right)y_h^{k+1}\,dx.\end{aligned}$$

The matrix B denotes the discrete *div* operator, see (7.11).

We solve this system as in the equal densities case by a preconditioned gmres method ([SS86]) with restart, where we use the triangular preconditioner given in [BGL05, Sec. 10.1.2]. This requires to find preconditioner for $A(y_h^k)$ and the Schur complement $S=-BA(y_h^k)^{-1}B^t$. As preconditioner for $A(y_h^k)$ we use a LU factorization of the diagonal blocks obtained by UMFPACK ([Dav04]) and for S we use a modification of the F_p preconditioner presented in [KLW02], where the block triangular preconditioner was used together with a multigrid preconditioner for $A(y_h^k)$.

As numerical example we compare our numerical results with the results given in [AV12] for the rising bubble benchmarks from [HTK+09].

The shape of the interface and the scaling of the surface tension

In order to incorporate the surface tension coefficient correctly into our numerics we next calculate the shape of the interface up to first order as given in the sharp-interface analysis in [AGG12, Sec. 4.3].

Let $\phi_0(z)$ denote the distribution of c across the interface, ranging from $-\infty$ to $+\infty$ as described in [AGG12, Sec. 4.3]. Here $z := x/\gamma$ holds with x denoting the signed distance to the zero level line of c and γ corresponds to the interfacial thickness. We interpret the Moreau–Yosida regularization of the variational inequality as a regularization of the double-obstacle free-energy in the following form:

$$\Psi_s(c)=\frac{1}{2}\left((1-c^2)+s(\max(0,c-1)^2+\min(0,c+1)^2)\right). \tag{10.13}$$

This free energy takes its minima at $c=\pm\frac{s}{s-1}$. Thus this free energy takes its minima close to ±1 for s large enough. Following [AGG12, Sec. 4.3] ϕ_0 then is the unique solution of the following ordinary differential equation:

$$\begin{aligned}\partial_{zz}\phi_0-\Psi_s'(\phi)=&0,\\ \phi_0(0)=&0,\\ \phi_0(z)\to\pm\frac{s}{s-1},&\text{ for } z\to\pm\infty.\end{aligned}$$

The solution of this equation is given by

$$\begin{aligned} z_0 =& \arctan \sqrt{s-1}, \\ \phi_0(z) =& \begin{cases} \sqrt{\frac{s}{s-1}} \sin(z) & |z| \le z_0, \\ \frac{1}{s-1} \left(s - \exp\left(-\sqrt{s-1}(z-z_0) \right) \right) & z > z_0, \\ -\frac{1}{s-1} \left(s - \exp\left(\sqrt{s-1}(z+z_0) \right) \right) & z < -z_0. \end{cases} \end{aligned} \tag{10.14}$$

The relation between the physical surface tension σ^{phys} and the parameter σ encorporated in the model is due to [AGG12, Sec. 4.3.4] given by

$$\sigma^{phys} = \left(\int_{-\infty}^{\infty} \left(\partial_z \phi_0\right)^2 \, dz \right) \sigma. \tag{10.15}$$

Here we have

$$\int_{-\infty}^{\infty} \left(\partial_z \phi_0\right)^2 \, dz = (s-1)^{-3/2} + \frac{s}{s-1} \left(\arctan \sqrt{s-1} + \frac{\sqrt{s-1}}{s} \right).$$

We note that $\sigma^{phys} \to \frac{\pi}{2}\sigma$ for $s \to \infty$, which means that for the unrelaxed double-obstacle energy, we obtain $\sigma^{phys} = \frac{\pi}{2}\sigma$.

The first benchmark from [AV12, HTK+09]

In [HTK+09] a rigorous numerical setup is given for comparing rising bubble simulations. Numerical results are given for two configurations by three independent working groups. The model used in [HTK+09] is the sharp interface model for two-phase flows. Our model (10.1)–(10.4) at least formaly converges to this sharp interface model as $\gamma \to 0$, see [AGG12, Sec. 4]. As benchmark quantities the circularity, the rising velocity and the evolution of the center of mass are investigated.

The circularity is defined by

$$\Theta = \frac{\text{perimeter of area-equivalent circle}}{\text{perimeter of bubble}} \le 1.$$

The rising velocity is defined by

$$V_c = \frac{\int_{c>0} y \, dx}{\int_{c>0} 1 \, dx}.$$

The center of mass is given by

$$y_c = \frac{\int_{c>0} x_2 \, dx}{\int_{c>0} 1 \, dx}.$$

Here x_2 denotes the second component of $x = (x_1, x_2)$. Note that the process is symmetric with respect to the first component and thus we only investigate the second component.

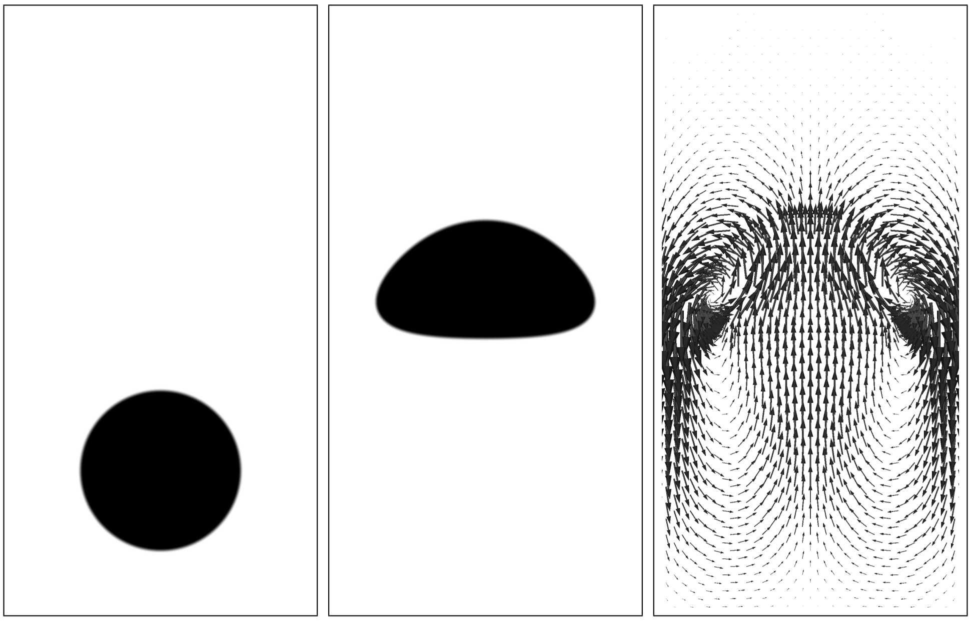

Figure 10.2: Initial bubble (left) and final bubble (center) as well as final velocity field (right) for first benchmark and $\gamma = 0.01$.

The parameters for the first benchmark are given as follows: The densities are $\rho_1 = 1000$ and $\rho_2 = 100$. The viscosities are $\eta_1 = 10$ and $\eta_2 = 1$ and the gravity is $g = -0.98$. The surface tension for the sharp interface modell is $\sigma^{phys} = 24.5$. We use $s = 10001$ and from (10.15) we obtain $\sigma \approx 15.6$. We complement the parameters by the diffuse-interface specific parameters γ which we vary, and by the mobility which we fix with $m(c) \equiv 0.001\gamma$. We note that in [AV12] a degenerate mobility is used. The domain of simulation is given by $\Omega = (0,1) \times (0,2)$. At the top and the bottom we use homogenous Dirichlet boundary data, thus $y = 0$ for the flowfield while on the left and right we apply free-slip conditions, i.e. $y \cdot \nu_\Omega = 0$, $\nu_\Omega^\perp \eta(c) D y \nu_\Omega = 0$. Corresponding to ϕ_0 the initial value for the simulation then is

$$c_0(x,y) = -\phi_0\left(\left(\sqrt{(x-0.5)^2 + (y-0.5)^2} - 0.25\right)/\gamma\right)$$

so that $c_0 \approx \frac{s}{s-1}$ corresponds to fluid 2 which forms the bubble, and $c_0 \approx -\frac{s}{s-1}$ corresponds to fluid 1 which is the surrounding fluid. c_0 describes a bubble centered at $M = (0.5, 0.5)$ with radius $r = 0.25$. In Figure 10.2 we show the initial bubble (left) and the bubble after three time instances (middle). On the right we show the velocityfield after three time instances.

In [AV12] three different diffuse interface models including [AGG12] are compared with the sharp interface model reported in [HTK$^+$09]. In Table 10.1 we compare our results to those obtained with the code MooNMD (ref) of the group of Lutz Tobiska at the University of Magdeburg (group 3 in [HTK$^+$09]) and those reported in [AV12] for $\gamma = 0.005$ (AV). We further show the temporal behaviour of the three benchmark parameters for our simulations

γ	$\Theta_{\min}$	$t\vert_{\Theta=\Theta_{\min}}$	$V_{c,\max}$	$t\vert_{V_c=V_{c,\max}}$	$y_c(t=3)$
0.0400	0.9110	1.9472	0.2322	0.9198	1.0694
0.0200	0.9035	1.9486	0.2370	1.0000	1.0759
0.0100	0.9019	1.9076	0.2402	0.9375	1.0782
0.0050	0.9015	1.9012	0.2412	0.9286	1.0788
ref	0.9013	1.9000	0.2417	0.9239	1.0817
AV	0.9045	1.9460	0.2401	0.9460	1.0785

Table 10.1: Benchmark values for the first benchmark.

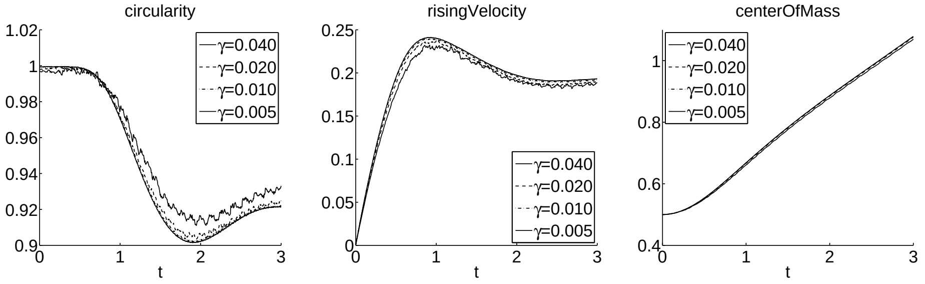

Figure 10.3: Evolution of circularity, rising velocity and center of mass over time for varying interfacial thicknesses for first benchmark.

in Figure 10.3.

One clearly observes the benefit of using the double-obstacle potential. For the same value of γ our numerical results are closer to the sharp interface reference solution than those reported in [AV12]. We note that using the polynomial free energy we are able to reproduce the results from [AV12].

The second benchmark from [HTK$^+$09, AV12].

In this benchmark we set $\rho_1 := 1000$, $\rho_2 := 1$, $\eta_1 := 10$ and $\eta_2 := 0.1$. Further we set $\sigma^{phys} := 1.96$, yielding $\sigma \approx 1.24$.

For this example the different sharp interface simulations presented in [HTK$^+$09] only agree up to approximately time $t = 2.0$. Especially it is not clear whether or not topological changes develop in the sharp interface simulations. For this reason we follow [AV12] and compare our results only up to this time instance. In Figure 10.4 we show the bubble after two (left) and after three (middle) time instances. We also show the velocity field after two time instances.

In Figure 10.5 we depict the temporal evolution of the benchmark values and in Table 10.2 we show our numerical results.

Again our results are in better agreement with the results obtained by the sharp interface simulation than the results obtained in [AV12] with the smooth free energy.

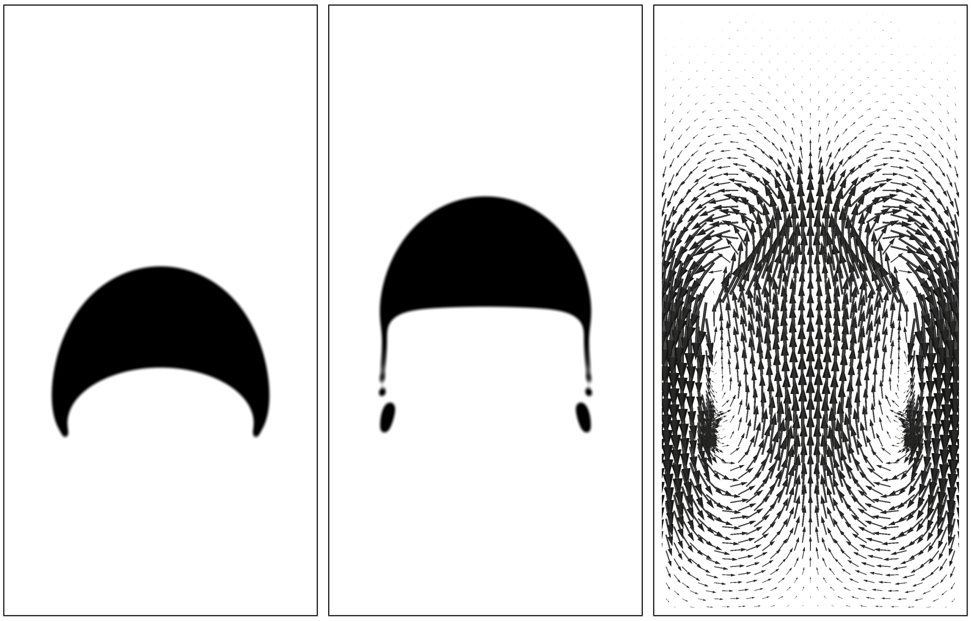

Figure 10.4: Bubble after two time instances (left) and final bubble (center) together with the final velocity field (right) for second benchmark ($\gamma = 0.01$).

γ	$\Theta_{\min}$	$t\vert_{\Theta=\Theta_{\min}}$	$V_{c,\max}$	$t\vert_{V_c=V_{c,\max}}$	$y_c(t=2)$
0.0400	0.6693	2.0000	0.2415	0.7901	0.9000
0.0200	0.6647	1.9998	0.2447	0.7013	0.9082
0.0100	0.6793	2.0000	0.2474	0.7221	0.9126
0.0050	0.6878	2.0000	0.2489	0.7394	0.9143
ref	0.6901	2.0000	0.2502	0.7300	0.9154
AV	0.6722	2.0000	0.2490	0.7540	0.9098

Table 10.2: Benchmark values for the second benchmark.

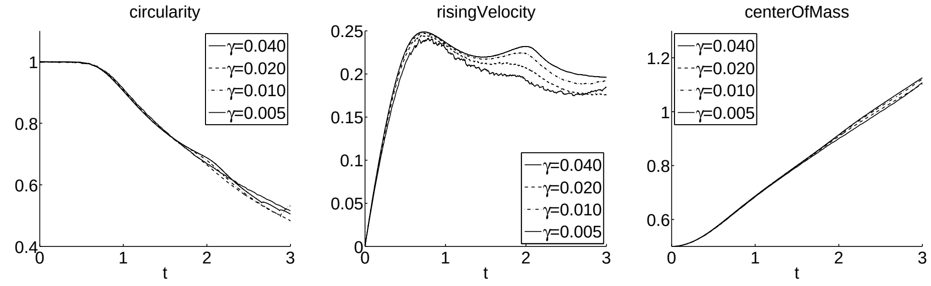

Figure 10.5: Evolution of circularity, rising velocity and center of mass over time for second benchmark.

11 A stable time discretization

A second benefit of the model [AGG12], beside being valid for different densities, is its thermodynamical consistency, thus the availability of energy estimates.

In the previous section we applied the concepts described in Sections 3–9 to this new model. Especially the sequential coupling proposed in Section 4 was used. However, the sequential coupling yields a discretization, that not preserves the thermodynamical consistency. In this section we present a new time discretization that gives rise to an energy estimate in the time discrete and fully discrete setting. This discretization is proposed by Professor Harald Garcke from the University of Regensburg and this section is published in [GHK14].

In the following we present the time discretization, that gives rise to a discrete in time energy estimate and present a discretization concept in space that is able to preserve the energy estimate in the fully discrete setting. Using the energy estimate we derive an a-posteriori adaptive concept estimating the error for the Navier–Stokes system and the Cahn–Hilliard system with one estimator.

A drawback of this new discretization is the appearance of a fully coupled Cahn–Hilliard Navier–Stokes system that has to be solved on each time instance. Additionally, a post processing of the set of marked triangles is required, that reduces the amount of triangles that are coarsened in the adaptive concept. This effectively increases the amount of triangles in the computational grids.

In the following we work with a general free energy of double well type, thus with exactly two minima, denoted by F. For F we subsequently use the splitting $F = F_+ + F_-$, where F_+ denotes the convex part of F and F_- denotes its concave part.

For convenience we here restate the model [AGG12, Sec. 3] in strong form

$$\begin{aligned}
\rho\partial_t y + ((\rho y + J)\cdot\nabla)\, y & & \qquad (11.1)\\
-\operatorname{div}(2\eta D y) + \nabla p &= w\nabla c + \rho g && \forall x\in\Omega,\ \forall t\in I, \qquad (11.2)\\
\operatorname{div}(y) &= 0 && \forall x\in\Omega,\ \forall t\in I, \qquad (11.3)\\
\partial_t c + y\cdot\nabla c - \operatorname{div}(m\nabla w) &= 0 && \forall x\in\Omega,\ \forall t\in I, \qquad (11.4)\\
-\sigma\gamma\Delta c + F'(c) - w &= 0 && \forall x\in\Omega,\ \forall t\in I, \qquad (11.5)\\
y(0,x) &= y_0(x) && \forall x\in\Omega, \qquad (11.6)\\
c(0,x) &= c_0(x) && \forall x\in\Omega, \qquad (11.7)\\
y(t,x) &= 0 && \forall x\in\partial\Omega,\ \forall t\in I, \qquad (11.8)\\
\nabla w(t,x)\cdot\nu_\Omega = \nabla c(t,x)\cdot\nu_\Omega &= 0 && \forall x\in\partial\Omega,\ \forall t\in I, \qquad (11.9)
\end{aligned}$$

where $J = -\frac{d\rho}{dc} m\nabla w$. The detailed description is given in Section 10.1.

Note that comparing to (10.1)–(10.4) the right hand side in the Navier–Stokes equation (11.2) changed and in the following we use $w\nabla c$ as interfacial force instead of $\sigma\gamma\mathrm{div}(\nabla c \otimes c)$ that we used in Section 10. Using $w\nabla c$ as interfacial force we obtain the physical pressure p, see the discussion in Section 10.1.

For our analysis we state a set of assumptions:

A1 There exists constants $\overline{\rho} \geq \underline{\rho} > 0$, $\overline{\eta} \geq \underline{\eta} > 0$, and $\overline{m} \geq \underline{m} > 0$ such that the following relations are satisfied:

- $\overline{\rho} \geq \rho(c) \geq \underline{\rho} > 0$,
- $\overline{\eta} \geq \eta(c) \geq \underline{\eta} > 0$,
- $\overline{m} \geq m(c) \geq \underline{m} > 0$.

Especially we assume that the mobility is non degenerated. In addition we assume, that ρ, μ, and m are continuous.

A2 $F : \mathbb{R} \to \mathbb{R}$ is continuously differentiable.

A3 F and the derivatives F'_+ and F'_- are polynomially bounded, i.e. there exists $C > 0$ such that $|F(x)| \leq C(1 + |x|^q)$, $|F'_+(x)| \leq C(1 + |x|^{q-1})$ and $|F'_-(x)| \leq C(1 + |x|^{q-1})$ holds for some $q \in [1, 4]$ if $n = 3$ and $q \in [1, \infty)$ if $n = 2$,

A4 F'_+ is Newton (sometimes called slantly) differentiable (see e.g. [HIK03]) regarded as nonlinear operator $F'_+ : H^1(\Omega) \to (H^1(\Omega))^*$ with Newton derivative G satisfying

$$(G(c)\delta c, \delta c) \geq 0$$

for each $c \in H^1(\Omega)$ and $\delta c \in H^1(\Omega)$.

To ensure Assumption **A1** we introduce a cut-off mechanism to ensure the bounds on ρ defined in Assumption **A1** independently of c. Note that $\eta(c)$ and $m(c)$ can be chosen arbitrarily fulfilling the stated bounds. We define the mass density as a smooth, monotone and strictly positive function $\rho(c)$ fulfilling

$$\rho(c) = \begin{cases} \frac{\tilde{\rho}_2 - \tilde{\rho}_1}{2} c + \frac{\tilde{\rho}_1 + \tilde{\rho}_2}{2} & \text{if } -1 - \frac{\tilde{\rho}_1}{\tilde{\rho}_2 - \tilde{\rho}_1} < c < 1 + \frac{\tilde{\rho}_1}{\tilde{\rho}_2 - \tilde{\rho}_1}, \\ \text{const} & \text{if } c > 1 + \frac{2\tilde{\rho}_1}{\tilde{\rho}_2 - \tilde{\rho}_1}, \\ \text{const} & \text{if } c < -1 - \frac{2\tilde{\rho}_1}{\tilde{\rho}_2 - \tilde{\rho}_1}. \end{cases}$$

For a discussion we refer to [Grü13, Remark 2.1].

Remark 11.1. Assumptions **A2**–**A4** are for example fulfilled by the polynomial free energy

$$F^{poly}(c) = \frac{\sigma}{4\gamma}\left(1 - c^2\right)^2,$$

and for the relaxed double-obstacle free energy used in the previous sections given by

$$F^{rel}(c) = \frac{\sigma}{2\gamma}\left(1 - c^2 + s\lambda^2(c)\right), \tag{11.10}$$

with

$$\lambda(c) := \max(0, c-1) + \min(0, c+1),$$

where $s \gg 0$ denotes the relaxation parameter. For convenience we here also restate the double-obstacle free energy

$$F^{obst}(c) = \begin{cases} \frac{\sigma}{2\gamma}(1-c^2) & \text{if } |c| \leq 1, \\ \infty & \text{else.} \end{cases}$$

As in the previous sections, in the numerical examples we use the free energy $F \equiv F^{rel}$. For this choice the splitting into convex and concave part reads

$$F_+(c) = s\frac{\sigma}{2\gamma}\lambda^2(c), \qquad F_-(c) = \frac{\sigma}{2\gamma}(1-c^2).$$

11.1 The time discrete setting

In the present section we formulate our time discretization scheme that is based on a weak formulation of (11.2)–(11.9) which we derive next. To begin with, note that for a sufficiently smooth solution (y, c, w) of (11.2)–(11.9) we can rewrite (11.2), using the linearity of ρ, as

$$\partial_t(\rho y) + \operatorname{div}(\rho y \otimes y) + \operatorname{div}(y \otimes J) - \operatorname{div}(2\eta D y) + \nabla p = w\nabla c + \rho g, \tag{11.11}$$

see [AGG12, p. 14].

We also note that the term $\rho y + J$ in (11.2) is not solenoidal (which might lead to difficulties both in the analytical and the numerical treatment) and that the trilinear form $(((\rho y + J)\cdot\nabla)u, v)$ is not anti-symmetric. To obtain a weak formulation yielding an anti-symmetric convection term we use a convex combination of (11.2) and (11.11) to define a weak formulation. We multiply equations (11.2) and (11.11) by the solenoidal test function $\frac{1}{2}w \in H(\operatorname{div}, \Omega)$, integrate over Ω, add the resulting equations and perform integration by parts. This gives

$$\frac{1}{2}\int_\Omega (\partial_t(\rho y) + \rho\partial_t y)\, v\, dx + a^t(\rho y + J, y, v) + \int_\Omega 2\eta D y : D v\, dx = \int_\Omega w\nabla c v + \rho g v\, dx,$$

with $a^t(\cdot,\cdot,\cdot)$ denoted in (4.5). Equations (11.4)–(11.5) are treated classically. This leads to

Definition 11.2. We call y, c, w a weak solution to (11.2)–(11.9) if $y(0) = y_0$, $c(0) = c_0$, $y(t) \in H(\mathrm{div}, \Omega)$ for *a.e.* $t \in I$ and

$$\begin{aligned} \frac{1}{2}\int_\Omega (\partial_t(\rho y) + \rho\partial_t y)\, v\, dx + \int_\Omega 2\eta Dy : Dv\, dx& \\ +a^t(\rho y + J, y, v) = \int_\Omega w\nabla c v + \rho g v\, dx \quad &\forall v \in H(\mathrm{div}, \Omega), \quad (11.12) \end{aligned}$$

$$\int_\Omega (\partial_t c + y \cdot \nabla c)\, \Phi\, dx + \int_\Omega m(c)\nabla w \cdot \nabla\Phi\, dx = 0 \quad \forall \Phi \in H^1(\Omega), \qquad (11.13)$$

$$\sigma\gamma \int_\Omega \nabla c \cdot \nabla\Psi\, dx + \int_\Omega F'(c)\Psi\, dx - \int_\Omega w\Psi\, dx = 0 \quad \forall \Psi \in H^1(\Omega), \qquad (11.14)$$

is satisfied for almost all $t \in I$. Here a^t is introduced in (4.5).

Theorem 11.3. *Let* y, c, w *be a sufficiently smooth solution to* (11.12)*–*(11.14)*. Then there holds*

$$\begin{aligned} &\frac{d}{dt}\left(\int_\Omega \frac{\rho|y|^2}{2} + \frac{\sigma\gamma}{2}|\nabla c|^2 + F(c)\, dx\right) \\ &\quad = -\int_\Omega 2\eta|Dy|^2 + m|\nabla w|^2\, dx + \int_\Omega \rho g y\, dx. \end{aligned}$$

Proof. By testing (11.12) with $v \equiv y$, (11.13) with $\Phi \equiv w$ and (11.14) with $\Psi \equiv \partial_t c$ and adding the resulting equations the claim follows. □

In [ADG13a, ADG13b] an alternative weak formulation of (11.2)–(11.9) is proposed, for which the authors show existence of weak solutions.

We now introduce a time discretization which mimics the energy inequality in Theorem 11.3 on the discrete level. Let

$$0 = t_0 < t_1 < \ldots < t_{k-1} < t_k < t_{k+1} < \ldots < t_M = T$$

denote an equidistant subdivision of the interval $\overline{I} = [0, T]$ with $\tau_{k+1} - \tau_k = \tau$. From here onwards the superscript k denotes the corresponding variables at time instance t_k.

Time integration scheme
Let $c_0 \in H^1(\Omega)$ and $y_0 \in H(\mathrm{div}, \Omega)$.

Initialization for $k = 0$*:*
Set $c^0 = c_0$ and $y^0 = y_0$.
Find $c^1 \in H^1(\Omega)$, $w^1 \in W^{1,q}(\Omega)$, $q > d$, $y^1 \in H(\mathrm{div}, \Omega)$, such that (10.5)–(10.8) holds, with $y \equiv y^1$, $y_{old} \equiv y^0$, $c \equiv c^1$, $c_{old} \equiv c^0$, and $w \equiv w^1$.

Two-step scheme for $k \geq 1$*:*
Given $c^{k-1} \in H^1(\Omega)$, $c^k \in H^1(\Omega)$, $w^k \in W^{1,q}(\Omega)$, $q > d$, $y^k \in H(\mathrm{div}, \Omega)$,

find $y^{k+1} \in H(\mathrm{div}, \Omega)$, $c^{k+1} \in H^1(\Omega)$, $w^{k+1} \in H^1(\Omega)$ satisfying

$$\begin{aligned}
\frac{1}{2\tau} \int_\Omega \left(\rho^k y^{k+1} - \rho^{k-1} y^k\right) v + \rho^{k-1}(y^{k+1} - y^k) v \, dx \\
+a(\rho^k y^k + J^k, y^{k+1}, v) + \int_\Omega 2\eta^k D y^{k+1} : Dv \, dx \\
- \int_\Omega w^{k+1} \nabla c^k v - \rho^k g v \, dx =& 0 \quad \forall v \in H(\mathrm{div}, \Omega), \quad (11.15) \\
\frac{1}{\tau} \int_\Omega (c^{k+1} - c^k) \Phi \, dx + \int_\Omega (y^{k+1} \cdot \nabla c^k) \Phi \, dx \\
+ \int_\Omega m(c^k) \nabla w^{k+1} \cdot \nabla \Phi \, dx =& 0 \quad \forall \Phi \in H^1(\Omega), \quad (11.16) \\
\sigma\gamma \int_\Omega \nabla c^{k+1} \cdot \nabla \Psi \, dx - \int_\Omega w^{k+1} \Psi \, dx \\
+ \int_\Omega ((F_+)'(c^{k+1}) + (F_-)'(c^k)) \Psi \, dx =& 0 \quad \forall \Psi \in H^1(\Omega), \quad (11.17)
\end{aligned}$$

where $J^k := -\frac{d\rho}{dc}(c^k) m^k \nabla w^k$.

We note that in (11.15)–(11.17) the only nonlinearity arises from F'_+ and thus only the equation (11.17) is nonlinear. Let us summarize properties of this scheme in the following remark.

Remark 11.4.

- The initialization is performed using the time disretization proposed in Section 10. Especially from regularity theory for the Laplace operator we have $\mu^1 \in H^2(\Omega) \hookrightarrow W^{1,q}, q > d$.
- In Theorem 11.12 we show existence and uniqueness of a solution to the time discrete model (11.15)–(11.17). Using the Assumption **A4** posed on F, it can be shown that Newton's method in function space can be used to compute a solution to (11.15)–(11.17) using the steps from Theorem 11.12.
- Through the use of ρ^{k-1}, (11.15)–(11.17) is a 2-step scheme. However, by replacing (11.15) with

$$\begin{aligned}
\frac{1}{2\tau} \int_\Omega \left(\rho^{k+1} y^{k+1} - \rho^k y^k\right) v + \rho^k (y^{k+1} - y^k) v \, dx \\
+a(\rho^k y^k + J^k, y^{k+1}, v) + \int_\Omega 2\eta D y^{k+1} : Dv \, dx \\
- \int_\Omega w^{k+1} \nabla c^k v + \rho^k g v \, dx = 0 \quad \forall v \in H(\mathrm{div}, \Omega),
\end{aligned}$$

one obtains an one-step scheme, which then also is nonlinear in the time discretization of (11.12). The resulting system is analyzed in future work.

In [GK14] Grün and Klingbeil propose a time-discrete solver for (11.2)–(11.9) which leads to strongly coupled systems for y, c and w at every time step and requires a fully nonlinear solver. For this scheme Grün in [Grü13] proves an energy inequality and the existence of so called generalized solutions.

11.2 The fully discrete setting and energy inequalities

For a numerical treatment we next discretize the weak formulation (11.15)–(11.17) in space. We aim at an adaptive discretization of the domain Ω, and thus to have a different spatial discretization in every time step.

Let $\mathcal{T}^k = \bigcup_{i=1}^{NT} T_i$ denote a conforming triangulation of $\overline{\Omega}$ with closed simplices $T_i, i = 1, \dots, NT$ and edges $E_i, i = 1, \dots, NE$, $\mathcal{E}^k = \bigcup_{i=1}^{NE} E_i$. Here k refers to the time instance t_k. On $\mathcal{T}^k$ we define the following finite element spaces:

$$\begin{aligned}\mathcal{V}^1(\mathcal{T}^k) =& \{v \in C(\mathcal{T}^k) \,|\, v|_T \in P^1(T) \,\forall T \in \mathcal{T}^k\} =: \operatorname{span}\{\Phi^i\}_{i=1}^{NP},\\ \mathcal{V}^2(\mathcal{T}^k) =& \{v \in C(\mathcal{T}^k) \,|\, v|_T \in P^2(T) \,\forall T \in \mathcal{T}^k,\, v|_{\partial\Omega} = 0\},\end{aligned}$$

where $P^l(S)$ denotes the space of polynomials up to order l defined on S.

We introduce the discrete analogon to the space $H(\operatorname{div}, \Omega)$:

$$\begin{aligned}H(\operatorname{div}, \mathcal{T}^k) = & \{v \in \mathcal{V}^2(\mathcal{T}^k)^d \,|\, (\operatorname{div} v, q) = 0 \,\forall q \in \mathcal{V}^1(\mathcal{T}^k) \cap L^2_{(0)}(\Omega),\, v|_{\partial\Omega} = 0\}\\ & := \operatorname{span}\{b^i\}_{i=1}^{NF},\end{aligned}$$

We introduce a weighted Leray projection ([CF88, Rem. 1.10]) $L_\rho^{k+1} : H(\operatorname{div}, \mathcal{T}^k) \to H(\operatorname{div}, \mathcal{T}^{k+1})$ by

$$(\rho^{k-1} L_\rho^{k+1} y^k, v) = (\rho^{k-1} y^k, v) \,\forall v \in H(\operatorname{div}, \mathcal{T}^{k+1}).$$

to prolonge velocity fields from former time steps. It is a weighted orthogonal projection onto $H(\operatorname{div}, \mathcal{T}^{k+1})$ and thus fulfills

$$(\rho^{k-1} L_\rho^{k+1} y^k, L_\rho^{k+1} y^k) \le (\rho^{k-1} y^k, y^k).$$

We further introduce a H^1-stable projection operator $\mathcal{P}^k : H^1(\Omega) \to \mathcal{V}^1(\mathcal{T}^k)$ satisfying

$$\|\mathcal{P}^k v\|_{L^p(\Omega)} \le \|v\|_{L^p(\Omega)} \text{ and } \|\nabla \mathcal{P}^k v\|_{L^r(\Omega)} \le \|\nabla v\|_{L^r(\Omega)}$$

for $v \in H^1(\Omega)$ with $r \in [1, 2]$ and $p \in [1, 6)$ if $n = 3$, and $p \in [1, \infty)$ if $n = 2$. Possible choices are the Clément operator ([Clé75]) or, by restricting the preimage to $C(\overline{\Omega}) \cap H^1(\Omega)$, the Lagrangian interpolation operator.

Using these spaces we state the discrete counterpart of (11.15)–(11.17): Let $k \ge 1$, given $c^{k-1} \in \mathcal{V}^1(\mathcal{T}^{k-1})$, $c^k \in \mathcal{V}^1(\mathcal{T}^k)$, $w^k \in \mathcal{V}^1(\mathcal{T}^k)$, $y^k \in H(\operatorname{div}, \mathcal{T}^k)$, find $y_h^{k+1} \in H(\mathrm{div}, \mathcal{T}^{k+1})$, $c_h^{k+1} \in \mathcal{V}^1(\mathcal{T}^{k+1})$, $w_h^{k+1} \in \mathcal{V}^1(\mathcal{T}^{k+1})$ such that for all $v \in H(\mathrm{div}, \mathcal{T}^{k+1})$, $\Phi \in \mathcal{V}^1(\mathcal{T}^{k+1})$, $\Psi \in \mathcal{V}^1(\mathcal{T}^{k+1})$ there holds:

$$\frac{1}{2\tau}(\rho^k y_h^{k+1} - \rho^{k-1} L_\rho^{k+1} y^k + \rho^{k-1}(y_h^{k+1} - L_\rho^{k+1} y^k), v)$$
$$+a^t(\rho^k y^k + J^k, y_h^{k+1}, v) + (2\eta^k D y_h^{k+1}, Dv) - (w_h^{k+1} \nabla c^k + \rho^k g, v) = 0, \quad (11.18)$$
$$\frac{1}{\tau}(c_h^{k+1} - \mathcal{P}^{k+1} c^k, \Phi) + (m(c^k)\nabla w_h^{k+1}, \nabla \Phi) + (y_h^{k+1} \nabla c^k, \Phi) = 0, \quad (11.19)$$
$$\sigma\gamma(\nabla c_h^{k+1}, \nabla \Psi) + (F'_+(c_h^{k+1}) + F'_-(\mathcal{P}^{k+1} c^k), \Psi) - (w_h^{k+1}, \Psi) = 0, \quad (11.20)$$

where $c^0 = Pc_0$ denotes the L^2 projection of c_0 in $\mathcal{V}^1(\mathcal{T}^0)$, and c_h^1, w_h^1, y_h^1 are obtained from (10.9)–(10.12) using $y^0 = Sy_0$, where S denotes the Stokes projection, see e.g. [GR86].

Existence of solution to the fully discrete system

We next show the existence of a unique solution to the fully discrete system (11.18)–(11.20).

Theorem 11.5. *There exist* $y_h^{k+1} \in H(div, \mathcal{T}^{k+1})$*,* $c_h^{k+1} \in \mathcal{V}^1(\mathcal{T}^{k+1})$*,* $w_h^{k+1} \in \mathcal{V}^1(\mathcal{T}^{k+1})$ *solving* (11.18)*–*(11.20)*.*

Proof. By testing (11.19) with $\Phi \equiv 1$, integration by parts in $(y_h^{k+1} \nabla c^k, 1)$ and using $y_h^{k+1} \in H(\mathrm{div}, \mathcal{T}^{k+1})$ we obtain

$$(c_h^{k+1}, 1) = (\mathcal{P}^{k+1} c^k, 1).$$

We define $\alpha = \frac{1}{|\Omega|}\int_\Omega \mathcal{P}^{k+1} c^k \, dx$ and set

$$V_{(0)} := \{v_h \in \mathcal{V}^1(\mathcal{T}^{k+1}) \mid (v_h, 1) = 0\}.$$

Then $z^{k+1} := c_h^{k+1} - \alpha$ fulfills $z^{k+1} \in V_{(0)}$. In the following we use z^{k+1} as unknown for the phase field, since the mean value of c is fixed. In addition we introduce $x^{k+1} := w_h^{k+1} - \frac{1}{|\Omega|}\int w_h^{k+1} \, dx$ and require (11.19)–(11.20) preliminarily only for test functions with zero mean value.

We define

$$X = H(\mathrm{div}, \mathcal{T}^{k+1}) \times V_{(0)} \times V_{(0)},$$

with the inner product

$$((y_1, x_1, z_1), (y_2, x_2, z_2))_X := (Dy_1, Dy_2) + (\nabla x_1, \nabla x_2) + (\nabla z_1, \nabla z_2),$$

and norm $\|\cdot\|_X^2 = (\cdot, \cdot)_X$. It follows from the inequalities of Korn and Poincaré that $(\cdot,\cdot)_X$ indeed forms an inner product on X. For $(y,x,z)\in X$ we define

$$\begin{aligned}(G(y,x,z),(\overline{y},\overline{x},\overline{z}))_X := &\left(\frac{1}{2}(\rho^k+\rho^{k-1})y-\rho^{k-1}L_\rho^{k+1}y^k,\overline{y}\right)\\ &+\tau a^t(\rho^k y^k+J^k,y,\overline{y})+\tau(2\eta^k Dy,D\overline{y})\\ &-\tau(x\nabla c^k,\overline{y})-\tau(\rho^k g,\overline{y})\\ &+(z-\mathcal{P}^{k+1}c^k,\overline{x})+\tau(m(c^k)\nabla x,\nabla\overline{x})+\tau(y\nabla c^k,\overline{x})\\ &+\sigma\gamma(\nabla z,\nabla\overline{z})-(x,\overline{z})\\ &+(F_+'(z+\alpha)+F_-'(\mathcal{P}^{k+1}c^k),\overline{z}).\end{aligned}$$

Now we show $(G(y,x,z),(y,x,z))_X>0$ for $\|(y,x,z)\|_X$ large enough and that G satisfies the supposition of [Tem77, Lem. II.1.4]. It then follows from [Tem77, Lem. II.1.4], that G admits a root $(y^*,x^*,z^*)\in X$. For convenience, we repeat [Tem77, Lem. II.1.4] in the appendix, see Lemma A8.

The function G is obviously continuous. We now estimate

$$\begin{aligned}(G(y,x,z),(y,x,z))_X \geq &\,\underline{\rho}(y,y)+2\tau\underline{\eta}(Dy,Dy)+\tau\underline{m}(\nabla x,\nabla x)\\ &+\sigma\gamma(\nabla z,\nabla z)+(F_+'(z+\alpha),z)\\ &-(\rho^{k-1}L_\rho^{k+1}y^k,y)-\tau(\rho^k g,y)\\ &-(\mathcal{P}^{k+1}c^k,x)+(F_-'(\mathcal{P}^{k+1}c^k),z).\end{aligned}\tag{11.21}$$

Using the convexity of F_+, which implies that F_+' is monotone, we obtain

$$(F_+'(z+\alpha),z)=(F_+'(z+\alpha)-F_+'(\alpha),z)+(F_+'(\alpha),z)\geq(F_+'(\alpha),z).$$

By using Hölder's and Poincaré's inequality and the stability of the projections L_ρ^{k+1} and $\mathcal{P}^{k+1}$ in (11.21) we obtain

$$(G(y,x,z),(y,x,z))_X>0$$

for $\|(y,x,z)\|_X\geq R$ if R is large enough. Now [Tem77, Lem. II.1.4] implies the existence of $(y^*,x^*,z^*)\in X$ such that $G(y^*,x^*,z^*)=0$. Defining

$$(y_h^{k+1},w_h^{k+1},c_h^{k+1})=(y^*,x^*+\beta,z^*+\alpha)$$

with β such that $(\beta,1)=(F_+'(c_h^{k+1})+F_-'(\mathcal{P}^{k+1}c^k),1)$ holds we obtain that $(y_h^{k+1},w_h^{k+1},c_h^{k+1})$ solves (11.18)–(11.20). □

Remark 11.6. Note that we do not need that the variables from old time instances are defined on the mesh used on the current time instance. We further do not need any smallness requirement on the mesh size h or on the time step length τ.

Theorem 11.7. *Let* $(c_h^{k+1}, w_h^{k+1}, y_h^{k+1})$ *be a solution to* (11.18)–(11.20). *Then for* $k \geq 1$*:*

$$\begin{aligned}\frac{1}{2}\int_\Omega \rho^k \left|y_h^{k+1}\right|^2 dx + \frac{\sigma\gamma}{2}\int_\Omega |\nabla c_h^{k+1}|^2 dx + \int_\Omega F(c_h^{k+1})\, dx \\ +\frac{1}{2}\int_\Omega \rho^{k-1}|y_h^{k+1} - L_\rho^{k+1}y^k|^2 dx + \frac{\sigma\gamma}{2}\int_\Omega |\nabla c_h^{k+1} - \nabla\mathcal{P}^{k+1}c^k|^2 dx \\ +\tau\int_\Omega 2\eta^k|Dy_h^{k+1}|^2 dx + \tau\int_\Omega m^k|\nabla w_h^{k+1}|^2 dx \qquad (11.22) \\ \leq \frac{1}{2}\int_\Omega \rho^{k-1}\left|y^k\right|^2 dx + \tau\int_\Omega \rho^k g y_h^{k+1} \\ +\frac{\sigma\gamma}{2}\int_\Omega |\nabla\mathcal{P}^{k+1}c^k|^2 dx + \int_\Omega F(\mathcal{P}^{k+1}c^k)\, dx.\end{aligned}$$

Proof. We have

$$\begin{aligned}&\frac{1}{2}\left(\rho^k y_h^{k+1} - \rho^{k-1}L_\rho^{k+1}y^k\right)\cdot y_h^{k+1} + \frac{1}{2}\rho^{k-1}\left(y_h^{k+1} - L_\rho^{k+1}y^k\right)\cdot y_h^{k+1} \\ &\quad = \frac{1}{2}\rho^k\left|y_h^{k+1}\right|^2 + \frac{1}{2}\rho^{k-1}\left|y_h^{k+1} - L_\rho^{k+1}y^k\right|^2 - \frac{1}{2}\rho^{k-1}\left|L_\rho^{k+1}y^k\right|^2, \qquad (11.23)\end{aligned}$$

$$\begin{aligned}&\nabla c_h^{k+1}\cdot\left(\nabla c_h^{k+1} - \nabla c^k\right) \\ &\quad = \frac{1}{2}|\nabla c_h^{k+1}|^2 - \frac{1}{2}|\nabla c^k|^2 + \frac{1}{2}|\nabla c_h^{k+1} - \nabla c^k|^2, \qquad (11.24)\end{aligned}$$

and since F_+ is convex and F_- is concave,

$$F_+(c_h^{k+1}) - F_+(c^k) \leq F_+'(c_h^{k+1})(c_h^{k+1} - c^k), \qquad (11.25)$$

$$F_-(c_h^{k+1}) - F_-(c^k) \leq F_-'(c^k)(c_h^{k+1} - c^k). \qquad (11.26)$$

The inequality is now obtained from testing (11.15) with y_h^{k+1}, (11.16) with w_h^{k+1}, (11.17) with $(c_h^{k+1} - \mathcal{P}^{k+1}c^k)/\tau$, and adding the resulting equations. This leads to

$$\begin{aligned}&\frac{1}{2\tau}(\rho^k y_h^{k+1} - \rho^{k-1}L_\rho^{k+1}y^k, y_h^{k+1}) + \frac{1}{2\tau}(\rho^{k-1}(y_h^{k+1} - L_\rho^{k+1}y^k), y_h^{k+1}) \\ &+a^t(\rho^k y^k + J^k, y_h^{k+1}, y_h^{k+1}) + (2\eta^k Dy_h^{k+1} : Dy_h^{k+1}) - (w_h^{k+1}\nabla c^k, y_h^{k+1}) \\ &+\frac{1}{\tau}(c_h^{k+1} - \mathcal{P}^{k+1}c^k, w_h^{k+1}) + (y_h^{k+1}\nabla c^k, w_h^{k+1}) + (m^k\nabla w_h^{k+1}, \nabla w_h^{k+1}) \\ &+\sigma\gamma\frac{1}{\tau}(\nabla c_h^{k+1}, \nabla(c_h^{k+1} - \mathcal{P}^{k+1}c^k)) - \frac{1}{\tau}(w_h^{k+1}, c_h^{k+1} - \mathcal{P}^{k+1}c^k) \\ &+\frac{1}{\tau}(F_+'(c_h^{k+1}), c_h^{k+1} - \mathcal{P}^{k+1}c^k) + \frac{1}{\tau}(F_-'(c^k), c_h^{k+1} - \mathcal{P}^{k+1}c^k) \\ &-(\rho^k g, y_h^{k+1}) = 0.\end{aligned}$$

The equalities (11.23) and (11.24) and the inequalities (11.25) and (11.26) now imply

$$\begin{aligned}\frac{1}{2\tau}\int_\Omega \left(\rho^k|y_h^{k+1}|^2+\rho^{k-1}|y_h^{k+1}-L_\rho^{k+1}y^k|^2-\rho^{k-1}|L_\rho^{k+1}y^k|^2\right)dx\\ +\int_\Omega 2\eta^k|Dy_h^{k+1}|^2\,dx+\int_\Omega m^k|\nabla w_h^{k+1}|^2\,dx\\ +\frac{\sigma\gamma}{2\tau}\int_\Omega |\nabla c_h^{k+1}|^2+|\nabla c_h^{k+1}-\nabla\mathcal{P}^{k+1}c^k|^2-|\nabla\mathcal{P}^{k+1}c^k|^2\,dx\\ +\frac{1}{\tau}\int_\Omega\left(F(c_h^{k+1})-F(\mathcal{P}^{k+1}c^k)\right)dx-\int_\Omega \rho^k g y_h^{k+1}\,dx\le 0,\end{aligned}$$

which is the claim, using $(\rho^{k-1}L_\rho^{k+1}y^k, L_\rho^{k+1}y^k)\le(\rho^{k-1}y^k,y^k)$, i.e. using the stability of L_ρ^{k+1}. □

Theorem 11.8. *System* (11.18)–(11.20) *admits a unique solution.*

Proof. Assume there exist two different solutions to (11.18)–(11.20) denoted by (y^1,c^1,w^1) and (y^2,c^2,w^2). We show that the difference $y=y^1-y^2, c=c^1-c^2, w=w^1-w^2$ is zero.

After inserting the two solutions into (11.18)–(11.20) and substracting the two sets of equations we perform the same steps as for the derivation of the discrete energy estimate, Theorem 11.7, and obtain

$$\begin{aligned}0=&\frac{1}{2}\int_\Omega(\rho^k+\rho^{k-1})y^2\,dx+2\tau\int_\Omega\eta^k|Dy|^2\,dx\\ &+\tau\|\sqrt{m^k}\nabla w\|^2+\sigma\gamma\|\nabla c\|^2+\left(F_+'(c^1)-F_+'(c^2),c^1-c^2\right).\end{aligned}$$

Since all these terms are non negative we obtain

$$\begin{aligned}\frac{1}{2}\int_\Omega(\rho^k+\rho^{k-1})y^2\,dx=&0, & \int_\Omega\eta^k|Dy|^2\,dx=&0,\\ \|\nabla w\|^2=&0, & \|\nabla c\|^2=&0.\end{aligned}$$

Since both $\eta(\cdot)$ and $\rho(\cdot)$ are strictly positive by Assumption **A1** we conclude $\|y\|_{H^1(\Omega)^d}=0$ and thus the uniqueness of the velocity field.

By testing (11.19) by $\Phi\equiv 1$ we obtain $(c^1,1)=(c^2,1)=(\mathcal{P}^{k+1}c^k,1)$ and thus $(c^1-c^2,1)=0$. Poincaré-Friedrichs inequality then yields $\|c\|_{H^1(\Omega)}=0$, and thus the uniqueness of the phase field.

Last we directly obtain that the chemical potential is unique up to a constant. By testing (11.20) with $\Psi\equiv 1$ and inserting the two solutions we obtain $(w^1-w^2,1)=(F_+'(c^1)-F_+'(c^2),1)=0$ and thus $\|w\|_{H^1(\Omega)}=0$, again by using Poincaré-Friedrichs inequality. □

Theorem 11.7 estimates the Ginzburg Landau energy of the current phase field c^{k+1} against the Ginzburg Landau energy of the projection of the old phase field $\mathcal{P}^{k+1}c^k$. Our aim is to obtain global in time inequalities estimating the energy of the new phase field against the energy of the old phase field at each time step. For this purpose let us state an assumption that later will be justified.

Assumption 11.9. Let $c^k \in \mathcal{V}^1(\mathcal{T}^k)$ denote the phase field at time instance t_k. Let $\mathcal{P}^{k+1}c^k \in \mathcal{V}^1(\mathcal{T}^{k+1})$ denote the projection of c^k in $\mathcal{V}^1(\mathcal{T}^{k+1})$. We assume that there holds

$$F(\mathcal{P}^{k+1}c^k) + \frac{1}{2}\sigma\gamma|\nabla\mathcal{P}^{k+1}c^k|^2 \leq F(c^k) + \frac{1}{2}\sigma\gamma|\nabla c^k|^2. \tag{11.27}$$

This assumption means, that the Ginzburg Landau energy is not increasing through projection. Thus no energy is numerically produced.

Assumption 11.9 is in general not fulfilled for arbitrary sequences $(\mathcal{T}^k)_k$ of triangulations. To ensure (11.27) we add a post processing step to the adaptive space meshing, see Section 11.3.

Theorem 11.10. *Assume that for every $k = 0, 1, \ldots$ Assumption 11.9 holds. Then for every $1 \leq k < l$ we have*

$$\begin{aligned}
\frac{1}{2}(\rho_h^{k-1}y_h^k, y_h^k) + &\int_\Omega F(c_h^k)\,dx + \frac{1}{2}\sigma\gamma(\nabla c_h^k, \nabla c_h^k) + \tau\sum_{m=k}^{l-1}(\rho^m g, y_h^{m+1}) \\
&\geq \frac{1}{2}(\rho^{l-1}y_h^l, y_h^l) + \int_\Omega F(c_h^l)\,dx + \frac{1}{2}\sigma\gamma(\nabla c_h^l, \nabla c_h^l) \\
&+ \sum_{m=k}^{l-1}(\rho^{m-1}(y_h^{m+1} - L_\rho^{m+1}y_h^m), (y_h^{m+1} - L_\rho^{m+1}y_h^m)) \\
&+ \tau\sum_{m=k}^{l-1}(2\eta^m Dy_h^{m+1}, Dy_h^{m+1}) \\
&+ \tau\sum_{m=k}^{l-1}(m(c_h^m)\nabla w_h^{m+1}, \nabla w_h^{m+1}) \\
&+ \frac{1}{2}\sigma\gamma\sum_{m=k}^{l-1}(\nabla c_h^{m+1} - \nabla\mathcal{P}^{m+1}c_h^m, \nabla c_h^{m+1} - \nabla\mathcal{P}^{m+1}c_h^m).
\end{aligned}$$

Proof. The stated result is obtained immediately from the energy estimate over one time step (11.7) together with the Assumption 11.9. □

Remark 11.11. We note that using $\Phi = 1$ in (11.19) and using integration by parts only delivers $(c_h^{k+1}, 1) = (\mathcal{P}^{k+1}c^k, 1)$ instead of $(c_h^{k+1}, 1) = (c^k, 1)$. If we use the quasi interpolation operator Q^{k+1} introduced by Carstensen in [Car99] for our generic projection $\mathcal{P}^{k+1}$, we would obtain $(c_h^{k+1}, 1) = (c^k, 1)$ since Q^{k+1} preserves the mean value, i.e. $(\varphi, 1) = (Q^{k+1}\varphi, 1)\ \forall\varphi \in L^1(\Omega)$.

On the other hand if we use Lagrange interpolation $\mathcal{I}^{k+1}$ we have

$$|(\mathcal{I}^{k+1}c^k,1)_T-(c^k,1)_T|\le Ch_T^3\|c^k\|_T,$$

and the deviation of $(\mathcal{I}^{k+1}c^k,1)$ from $(c^k,1)$ remaines small if we use bisection as refinement strategy, since then $\mathcal{I}^{k+1}c^k\in\mathcal{V}^1(\mathcal{T}^{k+1})$ and $c^k\in\mathcal{V}^1(\mathcal{T}^k)$ only differ on coarsened patches.

Existence of a solution to the time discrete system

Now we have shown that there exists a unique solution to (11.18)–(11.20). The energy inequality can be used to obtain uniform bounds on the solution and will be used to obtain a solution to the time discrete system (11.15)–(11.17) by a Galerkin method.

Theorem 11.12. *Let $y^k\in H(div,\Omega)$, $c^{k-1}\in H^1(\Omega)$, $c^k\in H^1(\Omega)$, and $w^k\in W^{1,q}(\Omega), q>d$ be given data. Then there exists a weak solution to* (11.15) –(11.17)*. Moreover, $c^{k+1}\in H^2(\Omega)$ and $w^{k+1}\in H^2(\Omega)$ holds.*

Proof. We proceed as follows. We construct a sequence of meshes $(\mathcal{T}_l^{k+1})_{l\to\infty}$ with gridsize $h_l\overset{l\to\infty}{\longrightarrow}0$. We show that the sequence $(y_l^{k+1},c_l^{k+1},w_l^{k+1})$ of unique and discrete solutions to (11.18)–(11.20) is bounded independently of l, and thus a weakly convergent subsequence exists which we show to converge to a weak solution of (11.15)–(11.17).

Let us start with defining the sequence of meshes. Let $\mathcal{T}_0^{k+1}=\mathcal{T}^{k+1}$ and $\mathcal{T}_{l+1}^{k+1}$, $l=0,1,\ldots$, be obtained from $\mathcal{T}_l^{k+1}$ by bisection of all triangles. The projection onto $\mathcal{T}_l^{k+1}$ we denote by $\mathcal{P}_l^{k+1}$.

From the discrete energy inequality (11.22) we obtain

$$\begin{aligned}\frac{1}{2}\int_\Omega\rho^k\left|y_l^{k+1}\right|^2dx+\frac{\sigma\gamma}{2}\int_\Omega|\nabla c_l^{k+1}|^2\,dx+\int_\Omega F(c_l^{k+1})\,dx\\+\frac{1}{2}\int_\Omega\rho^{k-1}|y_l^{k+1}-L_\rho^{k+1}y^k|^2\,dx+\frac{\sigma\gamma}{2}\int_\Omega|\nabla c_l^{k+1}-\nabla\mathcal{P}_l^{k+1}c^k|^2\,dx\\+\tau\int_\Omega 2\eta^k|Dy_l^{k+1}|^2\,dx+\tau\int_\Omega m^k|\nabla w_l^{k+1}|^2\,dx\\\le\frac{1}{2}\int_\Omega\rho^{k-1}\left|L_\rho^{k+1}y^k\right|^2dx+\tau\int_\Omega\rho^k g y_l^{k+1}\\+\frac{\sigma\gamma}{2}\int_\Omega|\nabla\mathcal{P}_l^{k+1}c^k|^2\,dx+\int_\Omega F(\mathcal{P}_l^{k+1}c^k)\,dx.\end{aligned}$$

We have the stability of the projection operators and thus

$$\begin{aligned}\int_\Omega|\nabla\mathcal{P}_l^{k+1}c^k|^2\,dx\le\|\nabla c^k\|^2_{L^2(\Omega)},\\\int_\Omega\rho^{k-1}|L_\rho^{k+1}y^k|^2\,dx\le\int_\Omega\rho^{k-1}|y^k|^2\,dx.\end{aligned}$$

Due to Assumption **A3** on F there exists a constant $C > 0$ such that

$$\begin{aligned}\int_\Omega F(\mathcal{P}_l^{k+1}c^k)\,dx \leq &C \int_\Omega |\mathcal{P}_l^{k+1}c^k|^q + 1\,dx\\ \leq &C\left(\|\mathcal{P}_l^{k+1}c^k\|_{L^q(\Omega)}^q + 1\right)\\ \leq &C\left(\|c^k\|_{L^q(\Omega)}^q + 1\right),\end{aligned}$$

where we again use the L^q-stability of the projection operator together with the Sobolev embedding $H^1(\Omega) \hookrightarrow L^q(\Omega)$ with q as in Assumption **A3**. By using Hölder's inequality and Young's inequality we further have

$$\begin{aligned}\tau\int_\Omega \rho^k g y_l^{k+1}\,dx \leq &\tau\left(\int_\Omega \rho^k|g|^2\,dx\right)^{1/2}\left(\int_\Omega \rho^k|y_l^{k+1}|^2\,dx\right)^{1/2}\\ \leq &\tau^2\int_\Omega \rho^k|g|^2\,dx + \frac{1}{4}\int_\Omega \rho^k|y_l^{k+1}|^2\,dx\end{aligned}$$

Since $\rho^{k-1} > 0$, $\rho^k > 0$, $\eta^k > 0$, and $m^k > 0$ by Assumption **A1** we obtain that $\|y_l^{k+1}\|_{H^1(\Omega)^d}$, $\|\nabla c_l^{k+1}\|$ and $\|\nabla w_l^{k+1}\|$ are uniformly bounded independend of l.

By inserting $\Phi \equiv 1$ in (11.19) we obtain $(\mathcal{P}_l^{k+1}c^k, 1) = (c_l^{k+1}, 1)$ and by Poincaré-Friedrichs inequality thus

$$\begin{aligned}\|c_l^{k+1}\|_{H^1(\Omega)} &\leq C\left(\|\nabla c_l^{k+1}\| + (\mathcal{P}_l^{k+1}c^k, 1)\right)\\ &\leq C\left(\|\nabla c_l^{k+1}\| + \|\mathcal{P}_l^{k+1}c^k\|\right)\\ &\leq C\left(\|\nabla c_l^{k+1}\| + \|c^k\|\right).\end{aligned}$$

Thus $\|c_l^{k+1}\|_{H^1(\Omega)}$ is uniformly bounded.

We obtain $(w_l^{k+1}, 1) = (F_+'(c_l^{k+1}) + F_-'(\mathcal{P}_l^{k+1}c^k), 1)$ by inserting $\Psi \equiv 1$ in (11.20). Due to Assumption **A3** on F_+' the first part can be bounded by $C(\|c_l^{k+1}\|_{L^q(\Omega)}^q + 1)$ which is bounded by Sobolev embedding. Also due to Assumption **A3** on F_- and due to the L^q stability of $\mathcal{P}_l^{k+1}$ the second part can be bounded by $C(\|c^k\|_{L^q(\Omega)}^q + 1)$. Thus, by the same arguments as $\|c_l^{k+1}\|_{H^1(\Omega)}$, also $\|w_l^{k+1}\|_{H^1(\Omega)}$ is uniformly bounded.

Consequently there exist $\overline{y} \in H_0^1(\Omega)^d, \overline{c} \in H^1(\Omega), \overline{w} \in H^1(\Omega)$ and a subsequence l_i such that $y_{l_i}^{k+1} \rightharpoonup \overline{y}$ in $H_0^1(\Omega)^d$, $c_{l_i}^{k+1} \rightharpoonup \overline{c}$ in $H^1(\Omega)$, $w_{l_i}^{k+1} \rightharpoonup \overline{w}$ in $H^1(\Omega)$ for $l_i \to \infty$.

We show that this triple of functions indeed is a weak solution to (11.15)–

(11.17). Inserting the sequence into (11.15)–(11.17) yields

$$
\begin{aligned}
\frac{1}{\tau}\int_\Omega \left(\frac{\rho^k+\rho^{k-1}}{2}y_{l_i}^{k+1} - \rho^{k-1}L_\rho^{k+1}y^k\right) v\,dx& \\
+a^t(\rho^k y^k + J^k, y_{l_i}^{k+1}, v) + \int_\Omega 2\eta^k Dy_{l_i}^{k+1} : Dv\,dx& \\
-\int_\Omega w_{l_i}^{k+1}\nabla c^k v + \rho^k g v\,dx &= 0\,\forall v \in H(\operatorname{div},\Omega),\\
\tau^{-1}\int_\Omega (c_{l_i}^{k+1} - \mathcal{P}^{k+1}c^k)\Phi\,dx + \int_\Omega (y_{l_i}^{k+1}\cdot\nabla c^k)\Phi\,dx& \\
+\int_\Omega m(c^k)\nabla w_{l_i}^{k+1}\cdot\nabla\Phi\,dx &= 0\,\forall\Phi\in H^1(\Omega),\\
\sigma\gamma\int_\Omega \nabla c_{l_i}^{k+1}\cdot\nabla\Psi\,dx - \int_\Omega w_{l_i}^{k+1}\Psi\,dx& \\
+\int_\Omega ((F_+)'(c_{l_i}^{k+1}) + (F_-)'(\mathcal{P}^{k+1}c^k))\Psi\,dx &= 0\,\forall\Psi\in H^1(\Omega).
\end{aligned}
$$

Now there holds

$$
\frac{1}{2\tau}\int_\Omega \left(\rho^k+\rho^{k-1}\right) y_{l_i}^{k+1} v\,dx \le \frac{1}{\tau}\overline{\rho}\|y_{l_i}^{k+1}\|\|v\|
$$

and thus $\frac{1}{2\tau}\int_\Omega \left(\rho^k+\rho^{k-1}\right) v\cdot\,dx \in (H_0^1(\Omega)^d)^*$ yielding

$$
\frac{1}{2\tau}\int_\Omega \left(\rho^k+\rho^{k-1}\right) y_{l_i}^{k+1} v\,dx \to \frac{1}{2\tau}\int_\Omega \left(\rho^k+\rho^{k-1}\right)\overline{y} v\,dx.
$$

Since $w^k \in W^{1,q}(\Omega), q > d$ there holds $J^k \in L^q(\Omega)^d$ and thus by Sobolev embedding we obtain

$$
\begin{aligned}
\left|\int_\Omega \left(\left(\left(\rho^k y^k + J^k\right)\cdot\nabla\right) y_{l_i}^{k+1}\right) v\,dx\right| \le& C\|\left(\rho^k y^k + J^k\right) v\|\|\nabla y_{l_i}^{k+1}\|,\\
\left|\int_\Omega \left(\left(\left(\rho^k y^k + J^k\right)\cdot\nabla\right) v\right) y_{l_i}^{k+1}\,dx\right|& \\
\le C\|\left(\left(\rho^k y^k + J^k\right)\nabla\right) v\|_{L^{\frac{2q}{q+2}}(\Omega)^d}&\|y_{l_i}^{k+1}\|_{L^{\frac{2q}{q-2}}(\Omega)^d},
\end{aligned}
$$

and thus $a^t(\rho^k y^k + J^k,\cdot,v) \in (H_0^1(\Omega)^d)^*$. This gives

$$
a^t(\rho^k y^k + J^k, y_{l_i}^{k+1}, v) \to a^t(\rho^k y^k + J^k, \overline{y}, v)
$$

The convergence of the remaining terms can be concluded in a similar manner.

Since $c_{l_i}^{k+1} \rightharpoonup \overline{c}$ in $H^1(\Omega)$ there exists a subsequence, again denoted by l_i such that $c_{l_i}^{k+1} \to \overline{c}$ in $L^q(\Omega)$, q as in Assumption **A3**. From Assumption **A3** and the dominated convergence theorem we thus obtain

$$
\int_\Omega F_+'(c_{l_i}^{k+1})\Psi\,dx \to \int_\Omega F_+'(\overline{c})\Psi\,dx.
$$

Next we show the weak solenoidality of $\overline{y}$. To begin with we note that every $q \in L^2_{(0)}(\Omega)$ can be approximated by a sequence $(q_l)_{l\in\mathbb{N}} \subset \mathcal{V}^1(\mathcal{T}_l^{k+1}) \cap L^2_{(0)}(\Omega)$, so that for every $\xi > 0$ an index N_ξ exists, such that $\|q - q_l\| \leq \xi$ for $l \geq N_\xi$. Now we have for arbitrary $q \in L^2_{(0)}(\Omega)$

$$|(\text{div } \overline{y}, q)| \leq |(\text{div } \overline{y}, q - q_l)| + |(\text{div } \overline{y} - \text{div } y_{l_i}, q_l)| + |(\text{div } y_{l_i}, q_l)|.$$

Let $\xi > 0$ be given. For the first addend we have $|(\text{div } \overline{y}, q - q_l)| \leq \|\text{div } \overline{y}\|\|q - q_l\| \leq C\xi$ for $l \geq N_\xi$.

Since the sequence q_l is defined on the same hierarchy of meshes as y_l we may restrict q_l to the subsequence l_i and obtain that both q_{l_i} and y_{l_i} are defined on the same meshes. We set $n := \min\{l_i \,|\, l_i \geq N_\xi\}$. Now we have $(\text{div } y_{l_i}, q_n) = 0$ for $l_i \geq n$, since then $q_n \in \mathcal{V}^1(\mathcal{T}_{l_i}^{k+1})$, i.e. the third addend vanishes. By choosing l_i so large that $|(\text{div } \overline{y} - \text{div } y_{l_i}, q_n)| \leq C\xi$ holds by weak convergence of y_{l_i}, the weak solenoidality of $\overline{y}$ is shown, since $\xi > 0$ is chosen arbitrarily.

Thus the triple $\overline{y}, \overline{c}, \overline{w}$ indeed is a weak solution.

It remains to obtain the stated higher regularity for w^{k+1} and c^{k+1}. This directly follows by regularity results for the Laplacian, see [EG04, Thm. 3.10]. Since $w^{k+1} - F'_+(c^{k+1}) - F'_-(\mathcal{P}^{k+1}c^k) \in L^2(\Omega)$ it follows that $c^{k+1} \in H^2(\Omega)$ and thus, since we have $\tau^{-1}(c^{k+1} - \mathcal{P}^{k+1}c^k) + y^{k+1}\nabla c^k \in L^2(\Omega)$, we obtain $w^{k+1} \in H^2(\Omega)$. □

The uniqueness of the solution follows by the same steps as the uniqueness of the discrete solutions, see Theorem 11.8. Like the fully discrete scheme, also the time-discrete scheme fulfills an energy inequality.

Theorem 11.13. *Let $c^{k+1}, w^{k+1}, y^{k+1}$ be a solution to* (11.15)–(11.17)*. Then the following energy inequality holds.*

$$\begin{aligned}
&\frac{1}{2}\int_\Omega \rho^k \left|y^{k+1}\right|^2 dx + \frac{\sigma\gamma}{2}\int_\Omega |\nabla c^{k+1}|^2\, dx + \int_\Omega F(c^{k+1})\, dx \\
&\quad + \frac{1}{2}\int_\Omega \rho^{k-1}|y^{k+1} - y^k|^2\, dx + \frac{\sigma\gamma}{2}\int_\Omega |\nabla c^{k+1} - \nabla c^k|^2\, dx \\
&\quad + \tau\int_\Omega 2\eta^k|Dy^{k+1}|^2\, dx + \tau\int_\Omega m^k|\nabla w^{k+1}|^2\, dx \\
&\leq \frac{1}{2}\int_\Omega \rho^{k-1}\left|y^k\right|^2 dx + \frac{\sigma\gamma}{2}\int_\Omega |\nabla c^k|^2\, dx + \int_\Omega F(c^k)\, dx + \int_\Omega \rho^k g y^{k+1}.
\end{aligned}$$

Proof. The inequality is obtained from testing (11.15) with y^{k+1}, (11.16) with w^{k+1}, (11.17) with $(c^{k+1} - c^k)/\tau$ and using the same arguments as in the proof for Theorem 11.7. □

Remark 11.14. Let F denote the relaxed double-obstacle free energy introduced in Remark 11.1, with relaxation parameter s. Let $(y_s, c_s, w_s)_{s\in\mathbb{R}}$ denote the sequence of solutions of (11.15)– (11.17) for a sequence $(s_l)_{l\in\mathbb{N}}$. From the linearity of (11.15) and Theorem 5.4 it follows, that there exists a subsequence, still denoted by $(y_s, c_s, w_s)_{s\in\mathbb{R}}$, such that

$$(y_s, c_s, w_s)_{s\in\mathbb{R}} \to (y^*, c^*, w^*) \quad \text{ in } H^1(\Omega),$$

where (y^*, c^*, w^*) denotes the solution of (11.15)–(11.17), where F^{obst}, denoted in Remark 11.1, is chosen as free energy. Especially $|c^*| \le 1$ holds. In the following argumentation we concentrate on the phase field only. From the regularity $c_s \in H^2(\Omega)$ together with a-priori estimates on the solution of the Poisson problem and the energy inequality of Theorem 11.13, we obtain the existence of a strongly convergent subsequence $c_{s'} \to c^*$ in $C^{0,\alpha}(\overline{\Omega})$, where we use the compact embedding $H^2(\Omega) \hookrightarrow C^{0,\alpha}(\overline{\Omega})$ for $2\alpha < 4 - d$.

Thus for s large enough we have $|c_s| \le 1 + \theta$ with θ arbitrarily small. Currently we are not able to quantify how large s has to be chosen in dependence of θ to guarantee this bound. Therefore we use the cut-off procedure described before Remark 11.1.

11.3 The A-Posteriori Error Estimation

For an efficient solution of (11.18)–(11.20) we next describe an a-posteriori error estimator based mesh refinement scheme that is reliable and efficient up to terms of higher order and errors introduced by the projection. We also describe how Assumption 11.9 on the evolution of the free energy, given in (11.22), under projection is fulfilled in the discrete setting.

In the present section we propose an all-in-one adaptation concept for the fully coupled Cahn–Hilliard Navier–Stokes system, where we exploit the energy inequality of Theorem 11.7.

The fully discrete system used in the numerical realization

Since in our numerical realization we do not include the solenoidality of the velocity field y into the discrete Ansatz space we now introduce a weak formulation for the time discrete version of (11.12)–(11.14) in primitive variables, which by [GR86] is equivalent to (11.15)–(11.17):
For $k \ge 1$, given $c^{k-1} \in H^1(\Omega)$, $c^k \in H^1(\Omega)$, $w^k \in W^{1,q}(\Omega), q > d$, $y^k \in H^1_0(\Omega)^d$ find $y^{k+1} \in H^1_0(\Omega)^d$, $p^{k+1} \in L^2_{(0)}(\Omega)$, $c^{k+1} \in H^1(\Omega)$, and $w^{k+1} \in H^1(\Omega)$ satisfy-

ing

$$\begin{aligned}
&\frac{1}{2\tau}(\rho^k y^{k+1} - \rho^{k-1}y^k + \rho^{k-1}(y^{k+1} - y^k), v)\\
&+a^t(\rho^k y^k + J^k, y^{k+1}, v) + (2\eta Dy^{k+1} : Dv)\\
&\quad -(p^{k+1}, \mathrm{div}(v)) - (w^{k+1}\nabla c^k + \rho^k g, v) = 0 \qquad \forall v \in H_0^1(\Omega)^d, \qquad (11.28)\\
&\qquad -(\mathrm{div}(y^{k+1}), q) = 0 \qquad \forall q \in L^2_{(0)}(\Omega), \qquad (11.29)\\
&\frac{1}{\tau}(c^{k+1} - c^k, \Phi) + (y^{k+1} \cdot \nabla c^k, \Phi)\\
&\qquad +(m(c^k)\nabla w^{k+1}, \nabla\Phi) = 0 \qquad \forall \Phi \in H^1(\Omega), \qquad (11.30)\\
&\sigma\gamma(\nabla c^{k+1}, \nabla\Psi) - (w^{k+1}, \Psi)\\
&\qquad +((F_+)'(c^{k+1}) + (F_-)'(c^k), \Psi) = 0 \qquad \forall \Psi \in H^1(\Omega). \qquad (11.31)
\end{aligned}$$

The corresponding fully discrete system now reads:
For $k \geq 1$, given $c^{k-1} \in H^1(\Omega)$, $c^k \in H^1(\Omega)$, $w^k \in W^{1,q}(\Omega), q > d$, $y^k \in H_0^1(\Omega)^d$ find $y_h^{k+1} \in \mathcal{V}^2(\mathcal{T}^{k+1})$, $p_h^{k+1} \in \mathcal{V}^1(\mathcal{T}^{k+1})$, $\int_\Omega p_h^{k+1}\,dx = 0$, $c_h^{k+1} \in \mathcal{V}^1(\mathcal{T}^{k+1})$, $w_h^{k+1} \in \mathcal{V}^1(\mathcal{T}^{k+1})$ such that for all $v \in \mathcal{V}^2(\mathcal{T}^{k+1})$, $q \in \mathcal{V}^1(\mathcal{T}^{k+1})$, $\Phi \in \mathcal{V}^1(\mathcal{T}^{k+1})$, $\Psi \in \mathcal{V}^1(\mathcal{T}^{k+1})$ there holds:

$$\begin{aligned}
&\frac{1}{2\tau}(\rho^k y_h^{k+1} - \rho^{k-1}L_\rho^{k+1}y^k + \rho^{k-1}(y_h^{k+1} - L_\rho^{k+1}y^k), v)\\
&\qquad +a^t(\rho^k y^k + J^k, y_h^{k+1}, v) + (2\eta^k Dy_h^{k+1}, \nabla v)\\
&\qquad -(w_h^{k+1}\nabla c^k + \rho^k g, v) - (p_h^{k+1}, \mathrm{div} v) = 0, \qquad (11.32)\\
&\qquad -(\mathrm{div} y_h^{k+1}, q) = 0, \qquad (11.33)\\
&\frac{1}{\tau}(c_h^{k+1} - \mathcal{P}^{k+1}c^k, \Phi) + (m(c^k)\nabla w_h^{k+1}, \nabla\Phi) - (y_h^{k+1}c^k, \nabla\Phi) = 0, \qquad (11.34)\\
&\sigma\epsilon(\nabla c_h^{k+1}, \nabla\Psi) + (F_+'(c_h^{k+1}) + F_-'(\mathcal{P}^{k+1}c^k), \Psi) - (w_h^{k+1}, \Psi) = 0. \qquad (11.35)
\end{aligned}$$

Thus we use the famous Taylor–Hood LBB-stable $P2 - P1$ finite element for the discretization of the velocity - pressure field and piecewise linear and continuous finite elements for the discretization of the phase field and the chemical potential. For other kinds of possible discretizations of the velocity-pressure field we refer to e.g. [Ver10].

Note that we perform integration by parts in (11.34) in the transport term. As soon as $\mathcal{P}^{k+1}$ is a mass conservened projection we by testing equation (11.34) with $\Phi = 1$ obtain the conservation of mass in the fully discrete scheme.

The link between equations (11.32)–(11.35) and (11.18)–(11.20) is provided by the next theorem.

Theorem 11.15. *Let $y_h^{k+1}, c_h^{k+1}, w_h^{k+1}$ denote the unique solution to* (11.18)–(11.20)*. Then there exists a unique pressure $p_h^{k+1} \in \mathcal{V}^1(\mathcal{T}^{k+1})$, $\int_\Omega p_h^{k+1}\,dx = 0$ such that $(y_h^{k+1}, p_h^{k+1}, c_h^{k+1}, w_h^{k+1})$ is a solution to* (11.32)–(11.35)*. The opposite direction is obvious.*

Proof. Since we use LBB-stable finite elements, from [GR86, Thm. II 1.1] we obtain the stated result. □

Derivation of the error estimator

We begin with noting that the special structure of our time discretization gives rise to an error estimator which both estimates the error in the approximation of the velocity, and in the approximation of the phase field and the chemical potential. We are not able to estimate the error in the approximation of the pressure field and the estimator will only be reliable and efficient up to higher order terms.

In the derivation of the estimator we follow Section 8.1 and thus present only the reliability result.

We define the following error terms:

$$
\begin{aligned}
e_y &:= y_h^{k+1} - y^{k+1}, & e_p &:= p_h^{k+1} - p^{k+1},\\
e_c &:= c_h^{k+1} - c^{k+1}, & e_w &:= w_h^{k+1} - w^{k+1},
\end{aligned}
$$

as well as the discrete element residuals

$$
\begin{aligned}
r_h^{(1)} :=& \frac{\rho^k + \rho^{k-1}}{2} y_h^{k+1} - \rho^{k-1} L_\rho^{k+1} y^k + \tau(b^k\nabla)y_h^{k+1} + \frac{1}{2}\tau\mathrm{div}(b^k)y_h^{k+1}\\
&- 2\tau\mathrm{div}\left(\eta^k D y_h^{k+1}\right) + \tau\nabla p_h^{k+1} - \tau w_h^{k+1}\nabla c^k - \rho^k g,\\
r_h^{(2)} :=& c_h^{k+1} - \mathcal{P}^{k+1}c^k + \tau y_h^{k+1}\nabla c^k - \tau\mathrm{div}(m^k\nabla w_h^{k+1}),\\
r_h^{(3)} :=& F'_+(c_h^{k+1}) + F'_-(\mathcal{P}^{k+1}c^k) - w_h^{k+1},
\end{aligned}
$$

where $b^k := \rho^k y^k + J^k$. Furthermore we define the error indicators

$$
\begin{aligned}
\eta_T^{(1)} &:= h_T\|r_h^{(1)}\|_T, & \eta_E^{(1)} &:= h_E^{1/2}\|2\eta^k\left[Dy_h^{k+1}\right]_E\cdot\nu_E\|_E,\\
\eta_T^{(2)} &:= h_T\|r_h^{(2)}\|_T, & \eta_E^{(2)} &:= h_E^{1/2}\|m^k\left[\nabla w_h^{k+1}\right]_E\cdot\nu_E\|_E,\\
\eta_T^{(3)} &:= h_T\|r_h^{(3)}\|_T, & \eta_E^{(3)} &:= h_E^{1/2}\|\left[\nabla c_h^{k+1}\right]_E\cdot\nu_E\|_E.
\end{aligned}
$$

By using the same steps as in Section 8.1 we derive the following theorem.

Theorem 11.16. *There exists a constant $C > 0$ only depending on the domain Ω and the regularity of the mesh $\mathcal{T}^{k+1}$ such that*

$$
\begin{aligned}
\underline{\rho}\|e_y\|^2 + \tau\underline{\eta}\|\nabla e_y\|^2 + \tau\underline{m}\|\nabla e_w\|^2 + \sigma\gamma\|\nabla e_c\|^2 + (F'_+(c_h^{k+1}) - F'_+(c^{k+1}), e_c)&\\
\le C\left(\eta_\Omega^2 + \eta_{h.o.t} + \eta_C\right),&
\end{aligned}
$$

holds with

$$
\begin{aligned}
\eta_\Omega^2 =& \frac{1}{\tau\underline{\eta}}\sum_{T\in\mathcal{T}^{k+1}}\left(\eta_T^{(1)}\right)^2 + \frac{\tau}{\underline{\eta}}\sum_{E\in\mathcal{E}^{k+1}}\left(\eta_E^{(1)}\right)^2\\
&\frac{1}{\tau\underline{m}}\sum_{T\in\mathcal{T}^{k+1}}\left(\eta_T^{(2)}\right)^2 + \frac{\tau}{\underline{m}}\sum_{E\in\mathcal{E}^{k+1}}\left(\eta_E^{(2)}\right)^2\\
&\frac{1}{\sigma\gamma}\sum_{T\in\mathcal{T}^{k+1}}\left(\eta_T^{(3)}\right)^2 + \sigma\gamma\sum_{E\in\mathcal{E}^{k+1}}\left(\eta_E^{(3)}\right)^2,\\
\eta_{h.o.t.} =& \tau(div(e_y), e_p),\\
\textit{and } \eta_C =& (\mathcal{P}^{k+1}c^k - c^k, e_w) - (F'_-(\mathcal{P}^{k+1}c^k) - F'_-(c^k), e_c).
\end{aligned}
$$

Remark 11.17.

- The term $\eta_{h.o.t.}$ is of higher order. By approximation results it can be estimated in terms of h_T to a higher order then the orders included in $\eta_T^{(i)}$, $\eta_E^{(i)}$, $i = 1, 2, 3$. Thus it is neglected in the numerics.

- The term η_C arises due to the transfer of c^k from the old grid $\mathcal{T}^k$ to the new grid $\mathcal{T}^{k+1}$ through the projection $\mathcal{P}^{k+1}$. In our numerics presented in Section 11.4 we use Lagrangian interpolation $\mathcal{I}^{k+1}$ as projection operator. We note that $\mathcal{I}^{k+1}c^k$ and c^k do only differ in regions of the domain where coarsening in the last time step took place, if bisection is used as refinement strategy. Since it seems unlikly that elements being coarsened in the last time step are refined again in the present time step, this term is neglected in the numerics. We note that this term might be further estimated to obtain powers of h_T by approximation results for the Lagrange interpolation, see e.g. [EG04] and Remark 11.11.

- Due to these two terms involved the estimator is not fully reliable.

- Neglecting these two terms the estimator can be shown to be efficient by the standard bubble technique as in Section 8.1.

- An adaptation of the time step size is not considered, since it would conflict with the time discretization over three time instances. In our numerics we have to choose time steps small enough to sufficently well resolve the interfacial force $w_h^{k+1}\nabla c^k$.

- The marking of triangles is performed following the marking strategy proposed in Section 8.3.

Ensuring the validity of the energy estimate

To ensure the validity of the energy estimate during the numerical computations we ensure that Assumption 11.9 holds trianglewise. For the following considerations we restrict to bisection as refinement strategy combined with the iFEM coarsening strategy proposed in [Che08]. This strategy only coarsens patches consisting of four triangles by replacing them by two triangles if the central node of the patch is an inner node of $\mathcal{T}^k$, and patches consisting of two triangles by replacing them by one triangle if the central node of the patch lies on the boundary of Ω. A patch fulfilling one of these two conditions we call a nodeStar. By using this strategy, we do not harm the Assumption 11.9 on triangles that are refined. We note that this assumption can only be violated on patches of triangles where coarsening appears.

After marking triangles for refinement and coarsening and before applying refinement and coarsening to $\mathcal{T}^k$ we make a postprocessing of all triangles that are marked for coarsening.

Let M^C denote the set of triangles marked for coarsening obtained by the marking strategy described in Section 8.3. To ensure the validity of the energy estimate (11.22) we perform the following post processing steps:

```
PP-1 For each triangle $T \in M^C$:
         if $T$ is not part of a nodeStar
         then set $M^C := M^C \setminus T$.

PP-2 For each nodeStar $S \in M^C$:
         if Assumption 11.9 is not fulfilled on $S$
         then set $M^C := M^C \setminus S$.
```

The resulting set M^C does only contain triangles yielding nodeStars on which the Assumption 11.9 is fulfilled.

11.4 Numerics

Now we use the adaptive concept developed in Section 11.3 to investigate the evolution of the energy inequality on the numerical level.

The nonlinear system (11.32)–(11.35) appearing in every time step of our approach is solved using the semi-smooth Newton method. Let us first describe how the linear systems arising in Newton's method are solved. At each time step in the Newton iteration we have to solve systems with linear operators G of the form

$$G = \left(\begin{array}{c|c} \mathcal{F} & \mathcal{I} \\ \hline \mathcal{T} & \mathcal{C} \end{array}\right) = \left(\begin{array}{cc|cc} A & B & I & 0 \\ B^t & 0 & 0 & 0 \\ \hline T & 0 & C_{11} & C_{12} \\ 0 & 0 & C_{21} & C_{22} \end{array}\right).$$

Here $\mathcal{F}$ and $\mathcal{C}$ are the discrete realizations of linearized Navier–Stokes and Cahn–Hilliard systems, respectively, while $\mathcal{I}$ represents their coupling through the interfacial force, and $\mathcal{T}$ the coupling through the transport at the interface. The order of the unknowns is (y, p, w, c).

Unique solvability of the systems arising from Newton's method can be shown by using the energy method of Section 11.2 taking Assumption **A4** into account.

The system is solved by a preconditioned gmres iteration with restart after 10 iterations. As preconditioner we use the block diagonal preconditioner

$$\mathcal{P} = \begin{pmatrix} \tilde{\mathcal{F}} & 0 \\ 0 & \mathcal{C} \end{pmatrix}$$

where $\mathcal{C}$ is inverted by LU decomposition, while $\tilde{\mathcal{F}}$ is an upper triangular block preconditioner ([BP88]) for Oseen type problems. It uses the F_p preconditioner [KLW02] for the Schur complement, i.e.

$$\tilde{\mathcal{F}} = \begin{pmatrix} \tilde{A} & B \\ 0 & \tilde{S} \end{pmatrix},$$

where $\tilde{S}$ is the F_p preconditioner for the Schur complement of $\mathcal{F}$ and $\tilde{A}$ is composed of the diagonal blocks of A and is inverted by LU decomposition.

The implementation is done in C++, where the adaptive concept is build upon *i*FEM ([Che08]). As linear solvers we again use **umfpack** ([Dav04]) and **cholmod** ([CDHR08]). The Newton iteration is implemented in its inexact variant, ensuring local superlinear convergence.

Examples

We investigate the evolution of the free energy and the validity of the energy inequality. Since we use Lagrange interpolation as projection operator, we violate the conservation of mass whenever coarsening is performed. This is numerically investigated.

Thereafter we give results for a qualitative benchmark for rising bubble dynamics. For this example we also show the influence of the required post processing step concerning the evolution of the meshes.

Concerning the free energy F we use the relaxed double-obstacle free energy (11.10) and set the relaxation parameter to $s = 10000$.

Investigation of the free energy We start by investigating the evolution of the free energy and the validity of the energy inequality in Theorem 11.7. Here we use the classic example of spinodal decomposition [CH58, FM08] as test case. The parameters are chosen as: $\rho_1 = \rho_2 = \eta_1 = \eta_2 = 1$, $g \equiv 0$, and $m(c) \equiv 10^{-3}\gamma$, $\gamma = 0.01$, $\sigma = 0.01$ and $\tau = 10^{-5}$.

In absence of outer forces the spinodal decomposition admits a characteristic speed of demixing, see e.g. [Sig79, OSS13]. Especially in the case of a diffusion driven setting the Ginzburg–Landau energy E decreases with the rate $E \sim t^{-1/3}$.

In Figure 11.1 we show the time evolution of the monotonically decreasing Ginzburg–Landau energy (left plot). We obtain the expected rate of $E \sim t^{-1/3}$ and also observe a time span where $E \sim t^{-1}$ holds, as predicted in [OSS13].

Next we investigate the validity of the energy inequality, see Figure 11.1 (right plot). We there show the time evolution of the term

$$\begin{aligned}
\zeta =& \frac{1}{2}\int_\Omega \rho^k \left|y_h^{k+1}\right|^2 dx + \frac{\sigma\gamma}{2}\int_\Omega |\nabla c_h^{k+1}|^2\,dx + \int_\Omega F(c_h^{k+1})\,dx \\
&+ \frac{1}{2}\int_\Omega \rho^{k-1}|y_h^{k-1} - L_\rho^{k+1} y^k|^2\,dx + \frac{\sigma\gamma}{2}\int_\Omega |\nabla c_h^{k+1} - \nabla\mathcal{I}^{k+1}c^k|^2\,dx \\
&+ \tau\int_\Omega 2\eta^k|Dy_h^{k+1}|^2\,dx + \tau\int_\Omega m^k|\nabla w_h^{k+1}|^2\,dx \\
&- \left(\frac{1}{2}\int_\Omega \rho^k \left|y^k\right|^2 dx + \frac{\sigma\gamma}{2}\int_\Omega |\nabla\mathcal{I}^{k+1}c^k|^2\,dx + \int_\Omega F(\mathcal{I}^{k+1}c^k)\,dx\right) \\
&- \int_\Omega \rho^k g y_h^{k+1}.
\end{aligned}$$

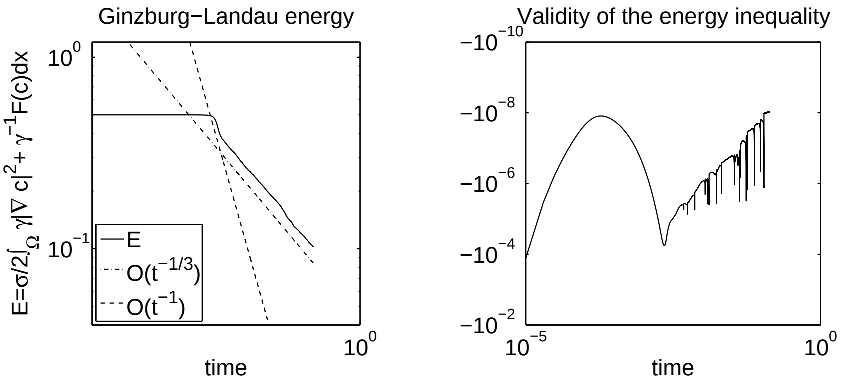

Figure 11.1: Time evolution of the Ginzburg–Landau energy (left), and validity of the energy inequality (right).

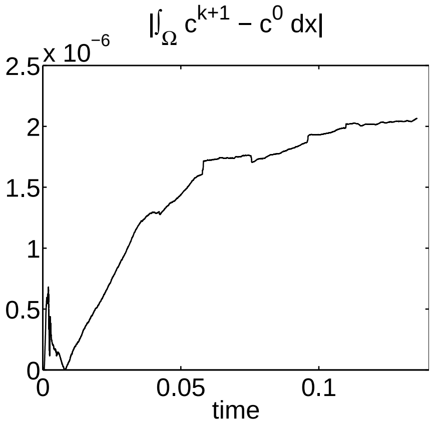

Figure 11.2: Time evolution of the deviation of the mean value of c.

The post processing (**PP-1**)–(**PP-2**) guarantees, that this term is always negative. The influence of (**PP-1**)–(**PP-2**) on the mesh quality is investigated later.

The violation in the conservation of mass Since we use Lagrange interpolation as projection operator between successive grids, we do not have full mass conservation, but have a violation in the mean value of c as discussed in Remark 11.11. In Figure 11.2 we depict the time evolution of the term $\left|\int_\Omega c^{k+1} - c^0\,dx\right|$, i.e. the difference between the mean value of c and the mean value of the initial phase field c^0. The numerical setup is the spinodal decomposition.

As can be observed, the violation increases with time, and the violation in mass conservation finally is of size 10^{-6}. We note that the order of the mean value is $|\Omega|$ and here we have $|\Omega| = 1$. Thus though we have deviation of mass, its size is small in comparison to the actual mean value.

	$(\Theta_c)_{\min}$	t_Θ	$(V_c)_{\max}$	t_V	$M_c(t=3)$
stab $\gamma = 0.02$	0.9080	1.9672	0.2388	0.9765	1.0786
div $\gamma = 0.02$	0.9035	1.9486	0.2370	1.0000	1.0759
ref	0.9013	1.9000	0.2417	0.9239	1.0817

Table 11.1: Results for the first benchmark from [HTK$^+$09].

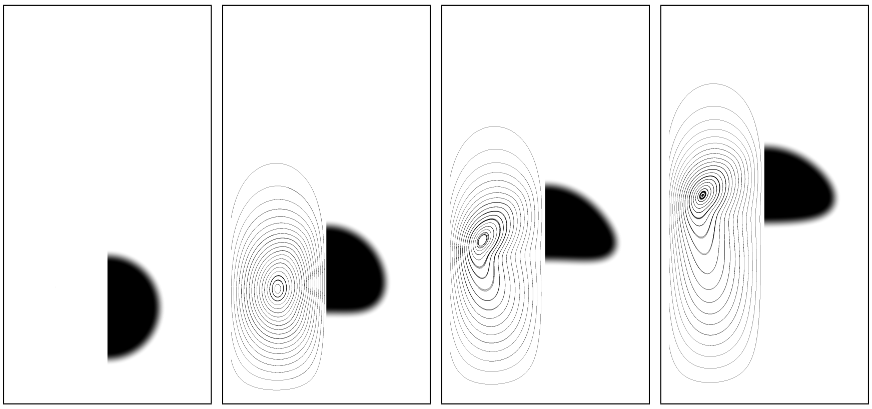

Figure 11.3: The evolution of the bubble at times $t \in \{0, 1, 2, 3\}$. The phase field is shown in the right part and streamlines of the velocity field in the left part of each plot.

Comparison with an existing benchmark We now again investigate the benchmark proposed in [HTK$^+$09], that we already simulated in Section 10.6.

For convenience we repeat the three benchmark values, for their definition see Section 10.6. The values we obtain from the simulation are the circularity, the rising velocity and the center of mass. As benchmark values the minimal circularity $(\Theta_c)_{\min}$ together with the time $t_\Theta := t(\Theta_c \equiv (\Theta_c)_{\min})$, the maximal rising velocity $(V_c)_{\max}$ together with the time $t_V := t(V_c \equiv (V_c)_{\max})$ and the center of mass $M_c(t = 3)$ at the final time $t = 3$ are chosen.

Our results with the new discretization are shown in Table 11.1, first row (stab $\gamma = 0.02$). For comparison we also restate the result obtained in Section 10.6 (div $\gamma = 0.02$) . The result corresponding to 'ref' again is a reference set of values taken from the sharp interface numerics in [HTK$^+$09].

We see that our results are in quite good aggrement with those obtained with sharp interface numerics. We also obtain that using the interfacial force $\operatorname{div}(\nabla c \otimes \nabla c)$ as used in Section 10.6 yields better values for the circularity while the maximal rising velocity and the total rise in the sence of the final location of the center of mass at final time are closer to sharp interface numerics in the new simulation.

In Figure 11.3 we show the evolution of the bubble for the benchmark setting.

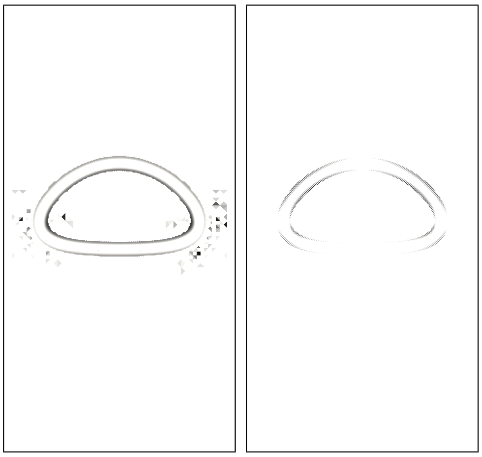

Figure 11.4: The distribution of the error indicators at time $t = 3$. η_T on the left, η_{TE} on the right. Black indicates higher errors.

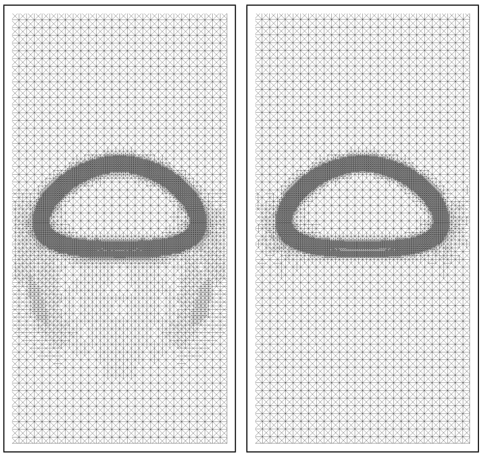

Figure 11.5: The mesh with (left) and without (right) postprocessing at final time $t = 3$.

Distribution of the error indicators Next we investigate the distribution of the error indicators. In Figure 11.4 we show the distribution of the error indicators η_T and η_{TE}. These are obtained from Theorem 11.16 and are given trianglewise as

$$\eta_T = \frac{1}{\underline{\eta}} \left(\eta_T^{(1)}\right)^2 + \frac{1}{\tau \underline{m}} \left(\eta_T^{(2)}\right)^2 + \frac{1}{\sigma\gamma} \left(\eta_T^{(3)}\right)^2,$$
$$\eta_{TE} = \sum_{E\in T} \left(\frac{\tau}{\underline{\eta}} \left(\eta_E^{(1)}\right)^2 + \frac{\tau}{\underline{m}} \left(\eta_E^{(2)}\right)^2 + \sigma\gamma \left(\eta_E^{(3)}\right)^2 \right).$$

We observe that a similar distribution is obtained as in Section 9.1 The errors are concentrated at the boundary of the interface. We further have additional error contributions from the Navier–Stokes part in a neighborhood of the bubble.

Influence of the post processing of the marked triangles Finally we investigate the spatial discretization obtained by our adaptive concept. Especially we show the influence of the post processing step (**PP-1**)–(**PP-2**) on reducing the number of triangles that are coarsened.

We simulate the rising bubble benchmark in the setting described above with and without the postprocessing steps. We note that without the postprocessing artificial energy is generated numerically through the coarsening process and the validity of the energy inequality can not be guaranteed, and in fact is not given.

In Figure 11.5 we show the final meshes at $t = 3$ with postprocessing (left) and without postprocessing (right). We see that there are regions in the bulk phase below the bubble where the postprocessing prevents the adaptive

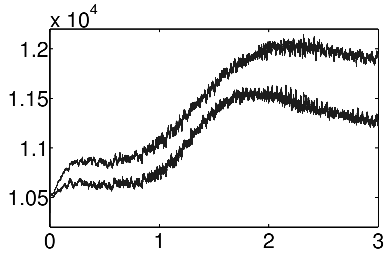

Figure 11.6: The evolution of the number of nodes for the described benchmark with (upper line) and without (lower line) postprocessing.

strategy from coarsening the triangles to the coarsest level. Thus we obtain a larger number of nodes if we use the post processing as is demonstrated in Figure 11.6 where we display the evolution of the number of mesh nodes with and without postprocessing.

We see that the number of nodes increases (by maximal 10% in this example) since not all triangles that are marked for coarsening are coarsened. On the other hand we note, that the energy inequality in the case without post processing is violated in 1692 of 60000 simulation steps and this violation takes place within the first 7000 time steps.

Let us note, that from the fluid mechanical point of view and if one considers the bubble as an obstacle in the channel flow, the region detected by the post processing is the wake, where the fluid is accelerated. Thus we expect a refined flow mesh there.

12 Summary and outlook

We proposed a diffuse interface approach for the simulation of two-phase flow based on a Moreau–Yosida relaxation of the double-obstacle free energy. In Sections 4–9 we proposed a discretization, which yields a sequential coupling of the Cahn–Hilliard and the Navier–Stokes system together with an adaptive concept which treats the resulting two systems separately.

In Section 10 we adapted this discretization concept to the model proposed in [AGG12] for two-phase flow, which is able to cope with fluids of different densities. We tested our approach by successfully passing a rising bubble benchmark.

In Section 11 we investigated a fully coupled time discretization over three time instances which delivers almost linear systems in every time step. The time discretization further gave rise to a discrete-in-time energy inequality.

We consider it as a drawback, that due to the discretization over three time instances we have to use fixed time step length in the simulation. In future work, this has to be overcome by using the more nonlinear time distretization presented in Remark 11.4.

Furthermore, the fully discrete setting in the current implementation is not mass conserving, since we use Lagrange interpolation to prolonge between successive spatial grids. This has to be changed to a mass conserving prolongation as discussed in Remark 11.11.

Since we use Moreau–Yosida relaxation to encorporate the bounds $|c| \leq 1$ the phase field violates this bound in the numerical simulation. Here, an estimate is required for the violation. In [?] promising results are presented that might be used to obtain results on this.

Part B
Closed-loop control of a Cahn–Hilliard Navier–Stokes system

In the second part of this thesis we develop model predictive control concepts for two-phase flows governed by the Cahn–Hilliard Navier–Stokes system. Model predictive control is a closed-loop control concept tailored to steer perturbed systems to a desired trajectory, or to stabilize the state of a perturbed system at a given reference trajectory, which for example describes an ideal evolution of the system obtained from an open-loop control run. For a discussion of model predictive control we refer to [GP11, NP97].

In [CHK99, HV02, Hin05a] instantaneous control is proposed as a variant of model predictive control and the approach taken in the present work extends the concept of instantaneous control to two-phase flows.

13 The general control concept

In this section we describe the general feedback control concept we apply to the Cahn–Hilliard Navier–Stokes system of Section 10. We describe the general concept of model predictiv control (MPC) following [GP11, NP97] and consider a variant of model predictive control called instantaneous control (IC), which in the context of control of incompressible flows is proposed in [Cho95, CHK99, HK00] and for the case of distributed control for Navier–Stokes equation is analyzed in [Hin05a]. In [HK13a] instantaneous control is applied to the control of two-phase flow of fluids with equal densities, and in [HK13b, Kah13] to the case of different densities. The following sections present the work contained in the latter publications.

We now state the concept of model predictive control in an abstract setting as for example given in [GP11]. We model the behavior of the real-time process using the following variables. By $x(t)$ we denote the state of the system at time t. A is a linear and time independent operator. The nonlinear system behavior is modelled by $b(x,t)$. With $u(t)$ we denote the (feedback) control at time t and B denotes the control operator. C is an observation operator mapping states $x(t)$ to observations $y(t)$. The general model we use is given as

$$\begin{aligned} \dot{x}(t) + Ax(t) &= b(x,t) + Bu(t), \\ y(t) &= Cx(t). \end{aligned} \tag{13.1}$$

As control goal we formulate

$$\|y(t) - y_d(t)\| \leq C \quad \forall t, \tag{13.2}$$

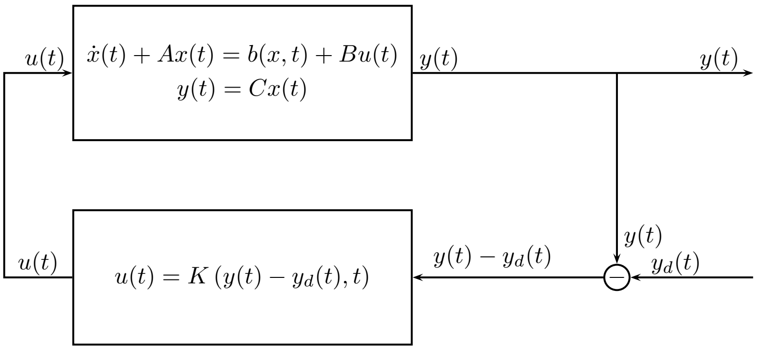
$u(t)$
$\dot{x}(t) + Ax(t) = b(x,t) + Bu(t)$
$y(t) = Cx(t)$
$y(t)$
$y(t)$
$y(t)$
$y_d(t)$
$y(t) - y_d(t)$
$u(t)$
$u(t) = K\,(y(t) - y_d(t), t)$

Figure 13.1: The general feedback control concept.

where $y_d(t)$ denotes a given desired trajectory and $C > 0$ is independent of t and moderately small.

The feedback control for the system is obtained by measuring the output y and setting $u = K(y - y_d)$, where K is the feedback control law. This control loop is depicted in Figure 13.1.

In the following we for simplicity assume $B = id$ and $C = id$, thus we assume fully distributed control and fully observable systems, i.e. $y(t) \equiv x(t)$. In Section 15 we address the more realistic case of Dirichlet boundary control, still with a fully observable system.

To describe our control approach in this abstract setting we introduce, for simplicity, an equidistant time grid

$$t_0 \le t_1 \le t_2 \le \ldots \le t_k \le t_{k+1} \le \ldots \tag{13.3}$$

with $t_{k+1} - t_k = \tau$ for $k = 0, 1, \ldots$, and state the general control loop:

For $k = 0, 1, \ldots$ do:

GCL1 Obtain the state $x^k := x(t_k)$.

GCL2 Calculate a feedback control $u_\star^{k+1} = K(x^k)$.

GCL3 Apply $u_\star^{k+1}$ to steer the system from time instance t_k to t_{k+1}.

In step **GCL1** the state of the system at time instance t_k is obtained by measurements. This data is used in step **GCL2** to obtain a control, that serves the control goal (13.2), and this control is applied to the real world process in step **GCL3** until time instance t_{k+1} is reached. Note that the control calculated in step **GCL2** gives a feedback to the current and actual state of the system measured in step **GCL1**.

In the following we are concerned with step **GCL2** of the general loop. In Section 13.1 we describe the model predictive control concept for step **GCL2** and in Section 13.2 we present a variant of model predictive control called instantaneous control.

In our examples the steps **GCL1**, which in a real world example corresponds to measurements of the state, and **GCL3**, which corresponds to the application of the control obtained in **GCL2**, are substituted by numerical simulation.

13.1 Model Predictive Control

In the present section we describe the model predictive control concept as e.g. proposed in [GP11]. The terminus ”model predictiv” arises from obtaining the control by predicting future behavior of the real-time process based on simulating a suitable model of the process and calculating a control suitable for controlling the modelled system in the future. In the following we describe how the control $u_\star^{k+1}$ in step **GCL2** is obtained for the general model (13.1).

Let the time grid (13.3) be given. Here x^k denotes the state at time instance t_k and $b^k = b(x^k, t_k)$ denotes the corresponding nonlinearity. By u^{k+1} we denote the control which steers the system from time instance t_k to t_{k+1}. Using a semi-implicit discretization in time we obtain the time discrete model

$$(I + \tau A)x^{k+1} = x^k + \tau b^k + u^{k+1}, \quad k = 0, 1, \ldots, \tag{13.4}$$

which enables us to predict the future behaviour of the system for given $u^{k+1}, u^{k+2}, \ldots$. Of course other discretizations in time are possible and for the case of control of the Navier–Stokes system are discussed e.g. in [Hin05a, Sec. 3].

To define the feedback control $u_\star^{k+1}$, used in step **GCL3** we state the following open-loop control problem over L time steps

$$\begin{aligned} &\min J(x^{k+1}, \ldots, x^{k+L}, u^{k+1}, \ldots, u^{k+L}) \\ &\qquad \text{s.t. } (I + \tau A)x^{j+1} = x^j + \tau b^j + u^{j+1} \text{ for } j = k, \ldots, k + L - 1, \end{aligned} \tag{$\mathcal{P}_k$}$$

where

$$J(x^{k+1}, \ldots, x^{k+L}, u^{k+1}, \ldots, u^{k+L}) := \sum_{i=1}^{L} \left(\frac{1}{2}\|x^{k+i} - x_d^{k+i}\|^2 + \frac{\alpha}{2}\|u^{k+i}\|^2 \right).$$

Here $x_d^k := x_d(t_k)$. We note that for $L = 1$ problem $(\mathcal{P}_k)$ for τ small enough admits a unique solution u^{k+1}. Anyway if $L > 1$ holds, this in general is not the case since then (13.4) admits a nonlinear constraint in problem $(\mathcal{P}_k)$. Here we assume that $(\mathcal{P}_k)$ admits at least one solution.

Let $(u^{k+1}, \ldots, u^{k+L})$ denote a solution to $(\mathcal{P}_k)$. The feedback control then is defined by $u_\star^{k+1} := u^{k+1}$ and we abbreviate the full process of obtaining $u_\star^{k+1}$ by

$$u_\star^{k+1} := K(u_0^{k+1}, x^k, t_k, L),$$

where K is the nonlinear feedback operator, u_0^{k+1} is the initial control for the optimization, x^k is the state at time t_k. The optimization is performed over L timesteps.

For real time application the full optimization of $(\mathcal{P}_k)$ might be too time consuming and thus an approximate solution might be used. This leads to the so called concept of instantaneous control [CHK99, Hin05a], which we describe next.

13.2 Instantaneous Control

Model predictive control is called instantaneous control, if $(\mathcal{P}_k)$ is solved inexactly by e.g. applying only one step of a steepest descent method to its solution using a suitable stepsize $\theta > 0$, and starting from an appropriate initial control u_0^{k+1}. In the case $L = 1$ the operator K performs the following steps:

IC1 Solve $(I + \tau A)z = x^k + \tau b(x^k, t_k) + u_0^{k+1}$.

IC2 Solve $(I + \tau A^*)\lambda = z - x_d^{k+1}$.,

IC3 Set $d = -(\alpha u_0^{k+1} + \lambda)$.

IC4 Determine $\theta > 0$.

IC5 Set $u_\star^{k+1} = u_0^{k+1} + \theta d = K(u_0^{k+1}, x^k, t_k, 1)$.

Here, we use the adjoint calculus to express the derivatives of the functional $J(x^{k+1}, u^{k+1})$ with respect to the control through the adjoint variable λ, see e.g. [HPUU09]. Instantaneous control with $L = 1$ is analytically investigated in [HV02] for control of Burger's equation, and in [CHK99] is used for the control of back facing step flows. In [Hin05a] it is proven that instantaneous control for incompressible flow in two spatial dimensions is able to reach a desired trajectory exponentially fast.

We further note that, due to the potential presence of local minima, the feedback controller for the model predictive control in general is not uniquely defined. However in the case of instantaneous control the controller is always unique, if the steps **IC1**–**IC5** can be performed, i.e. if the operator $I + \tau A$ is invertible. This does also apply for instantaneous control with $L > 1$.

14 The control problem

After having the instantaneous control policy at hand we apply it to the model for the simulation of two-phase flows with mass density contrast (10.1)–(10.4) [AGG12, Ch.3] already investigated in Section 10 and Section 11. For this it is convenient to restate the system here. For a given time interval I it reads: Find a flowfield y with a pressure field p as well as an order-parameter c

together with a chemical potential w such that the following equations hold:

$$\begin{aligned}
\rho\partial_t y + ((\rho y + j)\cdot\nabla)\, y - \operatorname{div}(2\eta D y) + \nabla p = \\
-\sigma\gamma\operatorname{div}(\nabla c\otimes\nabla c) + \rho G + Eu \quad & \text{in } I\times\Omega, && (14.1)\\
\operatorname{div} y = 0 \quad & \text{in } I\times\Omega, && (14.2)\\
y = 0 \quad & \text{on } I\times\partial\Omega, && (14.3)\\
y(0,x) = y_0(x) \quad & \text{in } \Omega, && (14.4)\\
\partial_t c - \operatorname{div}(m\nabla w) + y\cdot\nabla c = 0 \quad & \text{in } I\times\Omega, && (14.5)\\
-\sigma\gamma\Delta c + \lambda_s(c) - \sigma\gamma^{-1} c = w, \quad & \text{in } I\times\Omega, && (14.6)\\
\partial_{\nu_\Omega} c = \partial_{\nu_\Omega} w = 0 \quad & \text{on } I\times\partial\Omega, && (14.7)\\
c(0,x) = c_0(x) \quad & \text{in } \Omega. && (14.8)
\end{aligned}$$

The control is denoted by the volume force Eu acting on the Navier–Stokes equation. Here $E : U \to L^2(I,(H_0^1(\Omega)^d)^*)$ is a control operator mapping from the space of admissible controls U to the space of right hand sides. We note that by the operator E we can both realize distributed control ($U = L^2(I,L^2(\Omega)^d)$, $E = id_{L^2(I,L^2(\Omega)^d)\hookrightarrow L^2(I,(H_0^1(\Omega)^d)^*)}$) and finite dimensional (or parameterized) control ($U = L^2(I,\mathbb{R}^M)$, $Eu = \sum_{i=1}^M f_i(x)u_i(t)$ for given $f_i \in (H_0^1(\Omega)^d)^*$). We assume that U is a Hilbert space with inner product denoted by $(\cdot,\cdot)_U$, and norm induced by the inner product.

As in Section 10 $\rho = \rho(c)$ denotes the density of the fluid, $\eta = \eta(c)$ denotes the viscosity and $m = m(c)$ denotes the mobility. The gravitational force is denoted by G. By j we abbreviate the transport term arising from the two-phase structure, namely $j = -\rho'(c)m(c)\nabla w$.

The control goal consists in steering the order parameter c to a given desired phasefield c_d, i.e.

$$\|c(t) - c_d(t)\|_{L^2(\Omega)} \overset{t\to\infty}{\longrightarrow} 0.$$

14.1 The time discrete problem

Following the model predictive control approach, we next discretize (14.1)–(14.7) in time. We stress that the time discretization stated in the following is used only for calculating the control using the model predictive control approach. The simulation for steps **GCL1** and **GCL3** in the control loop is performed using the discretization proposed in Section 10, extended by the corresponding control.

Let t denote the current time instance and τ the fixed time step length. We set $\xi = \tau^{-1}$. By the superscript ^{k} we denote terms defined on the old time instance $t^k = t - \tau$. By abuse of notation in the following we set $U := U|_t$ and $E := E|_t$.

We use the following time discretization:
Given $y^k \in H_0^1(\Omega)^d$, $c^k \in H^1(\Omega)$, $w^k \in H^1(\Omega)$, and $u \in U$. At time instance

t^{k+1}, find $y \in H_0^1(\Omega)^d$, $p \in L^2_{(0)}(\Omega)$, $c \in H^1(\Omega)$, $w \in H^1(\Omega)$ solving

$$\begin{aligned}
\xi\rho^k y - \operatorname{div}(\eta^k \nabla y) + \nabla p & \\
+ \left(\left(\rho^k y^k + j^k\right) \cdot \nabla\right) y^k - \xi\rho^k y^k & \\
+\sigma\gamma \operatorname{div}(\nabla c^k \otimes \nabla c^k) - \rho^k G - Eu &= 0 \quad \text{in } \Omega, && (14.9)\\
\operatorname{div} y &= 0 \quad \text{in } \Omega, && (14.10)\\
y &= 0 \quad \text{on } \partial\Omega, && (14.11)\\
(c - c^k) - \tau \operatorname{div}\left(m^k \nabla w\right) + \tau y \cdot \nabla c^k &= 0 \quad \text{in } \Omega, && (14.12)\\
-\sigma\gamma\Delta c + \lambda_s(c^k) - \sigma\gamma^{-1} c^k - w &= 0 \quad \text{in } \Omega, && (14.13)\\
\partial_{\nu_\Omega} c = \partial_{\nu_\Omega} w &= 0 \quad \text{on } \partial\Omega. && (14.14)
\end{aligned}$$

Here we substituted Dy by ∇y to decouple the spatial dimensions in the Navier–Stokes system (14.9)–(14.11). We note that this discretization again decouples the Navier–Stokes system (14.9)–(14.11) from the Cahn–Hilliard system (14.12)–(14.14). But here we first solve the Navier–Stokes system (14.9)–(14.11) to obtain the velocity field y and afterward we solve the Cahn–Hilliard system (14.12)–(14.14) to obtain the phase field c.

Lemma 14.1. *Given $u \in U$ together with $c^k, w^k \in H^1(\Omega)$ and $y^k \in H_0^1(\Omega)^d$ there exists a unique pair $(y, p) \in H_0^1(\Omega)^d \times L^2_{(0)}(\Omega)$ solving* (14.9) – (14.11). *There further exists a unique pair $(c, w) \in H^1(\Omega) \times H^1(\Omega)$ solving* (14.12) – (14.14).

Proof. The existence of a unique solution to (14.9) – (14.11) is achieved using Lax–Milgram's theorem (Theorem A1).

The existence of a unique solution to (14.12) – (14.14) can be achieved as in Section 4 using a helper problem. □

14.1.1 Adjoint representation of the gradient direction

To obtain the instantaneous control we next consider the optimal control problem

$$\begin{aligned}
\min_u J(c,u) = \frac{1}{2}\|c - c_d\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|u\|^2_U & \\
s.t.\, (14.9) - (14.14). &
\end{aligned} \tag{P_{dist}}$$

Using adjoint calculus [HPUU09] one verifies that

$$\nabla J(u) = \alpha u + E^* p_3.$$

Here we assume that U is selfadjoint and that the Riesz-isomorphism can be neglected. The variable p_3 is related to u through the adjoint system

$$
\begin{aligned}
p_2 - \tau \mathrm{div}\left(m^k \nabla p_1\right) =& 0, && (14.15)\\
-\sigma\gamma\Delta p_2 - p_1 =& c(u) - c_d, && (14.16)\\
\xi\rho^k p_3 - \mathrm{div}\left(\eta^k \nabla p_3\right) + \nabla p_4 =& \tau p_1 \nabla c^k, && (14.17)\\
\mathrm{div} p_3 =& 0, && (14.18)\\
p_3 = \nu_\Omega \cdot \nabla p_1 = \nu_\Omega \cdot \nabla p_2 =& 0 \text{ on } \partial\Omega, && (14.19)
\end{aligned}
$$

with $(p_1, p_2, p_3, p_4) \in H^1(\Omega) \times H^1(\Omega) \times H_0^1(\Omega)^n \times L^2_{(0)}(\Omega)$ denoting the adjoint variables, and $c(u)$ denoting the solution to (14.9)–(14.14) for a given u.

We note that existence and uniqueness of the adjoint variables p_1, p_2, p_3, p_4 follows from Lemma 14.1.

14.1.2 Obtaining the steepest descent stepsize

To achieve sufficient decrease in the value of $J(u)$ through

$$u := u - \theta \nabla J(u)$$

the choice of the stepsize θ is crucial. In the present situation the functional J is quadratic, since (14.9) –(14.14) forms a linear system with a well defined, linear and continuous solution operator. Now, for given $u \in U$, let $g := \nabla J(u)$ and denote by $c(g)$ the phase field defined through the system

$$
\begin{aligned}
\xi\rho^k y - \mathrm{div}(\eta^k \nabla y) + \nabla p =& Eg && (14.20)\\
\mathrm{div}\ y =& 0, && (14.21)\\
c - \tau \mathrm{div}\left(m^k \nabla w\right) =& -\tau y \nabla c^k, && (14.22)\\
-\sigma\epsilon\Delta c - w =& 0. && (14.23)
\end{aligned}
$$

If $g \neq 0$, the optimal stepsize θ is well defined and satisfies

$$\theta = \mathrm{argmin}_{t\in\mathbb{R}} J(u - tg).$$

A short calculation shows

$$\theta = \frac{(c(u) - c_d, c(g))_{L^2(\Omega)} + \alpha(u, g)_{L^2(\Omega)}}{\|c(g)\|^2_{L^2(\Omega)} + \alpha\|g\|^2_{L^2(\Omega)}} = \frac{\|g\|^2_{L^2(\Omega)}}{\|c(g)\|^2_{L^2(\Omega)} + \alpha\|g\|^2_{L^2(\Omega)}}, \quad (14.24)$$

so that in the present situation the computation of the optimal steepest descent stepsize requires one additional linear system solve.

14.1.3 The initial condition for the gradient step

We initialize the gradient step with $u^0 = 0$. We note that in [Hin05a] for the case of distributed control an initial control u^0 is defined that, if used in the instantaneous control approach, is able to steer the state of the two-dimensional Navier–Stokes system to a desired state exponentially fast.

We further note that due to starting with $u^0 = 0$ the stepsize strategy described in Section 14.1.2 is essential for steering the phase field towards the desired state.

14.2 Spatial Discretization

The spatial discretization is performed by linear finite elements for both the concentration and the chemical potential yielding approximations c^h, w^h. For the flowfield and the pressure we use the LBB-stable Taylor-Hood $P^2 - P^1$ finite element pair, see e.g. [HT74, Ver10], yielding approximations y^h, p^h. For the spatial treatment of the Cahn-Hilliard part (14.12)–(14.14) we use the adaptive approach presented in Section 10. In the case of distributed control, the control $u = -\theta g$ is implicitly discretized by p_3, see [Hin05b].

14.3 Obtaining the instantaneous control

We finish this section with denoting the allover procedure used to calculate the instantaneous control.

For a given initial control guess $u^0 \in U$ the controller now performs the steps:

1. Given y^k, c^k, w^k, compute $c(u^0)$,

2. compute $p_3(c(u^0))$,

3. set $g = \nabla J(u^0) = \alpha u^0 - E^* p_3(u^0)$,

4. compute $c(g)$,

5. compute θ with $u = u^0, c(g)$ and g,

6. set $u := u^0 - \theta g$.

In total the numerical amount of work consists of three linear system solves.

15 Dirichlet boundary control

In this section we consider tangential Dirichlet boundary control for two-phase flows. Tangential Dirichlet Control is technological feasible (see [BLK01, MK02]) and mass conserving. Tangential Dirichlet boundary control for incompressible flows is for example investigated in [Bar11, BLK01].

Tangential boundary control yields a minor number of changes in the procedure presented before that we will present in the following. We start with stating the time discrete optimization problem we consider here:

$$\min J(c,u) = \frac{1}{2}\|c-c_d\|^2_{L^2(\Omega)} + \frac{\alpha}{2}\|u\|^2_{\mathbb{R}^M}$$

$$\begin{aligned} & s.t. \\ \xi\rho^k(y-y^k) + \left(\left(\rho^k y^k + j^k\right)\cdot\nabla\right) y^k & \\ -\mathrm{div}\left(\eta^k\nabla y\right) + \nabla p = & \\ -\sigma\gamma\mathrm{div}(\nabla c^k \otimes \nabla c^k) + \rho^k G & \quad \text{in } \Omega, \\ \mathrm{div}\; y = 0 & \quad \text{in } \Omega, \\ y = Eu & \quad \text{on } \partial\Omega, \\ (c-c^k) - \tau\mathrm{div}\left(m^k\nabla w\right) + \tau y\nabla c^k = 0 & \quad \text{in } \Omega, \\ -\sigma\gamma\Delta c + \lambda_s(c^k) - \sigma\gamma^{-1}c^k = w & \quad \text{in } \Omega, \\ \partial_{\nu_\Omega} c = \partial_{\nu_\Omega} w = 0 & \quad \text{on } \partial\Omega. \end{aligned} \tag{P_{bnd}}$$

The operator $E : \mathbb{R}^M \to H^{1/2}(\partial\Omega)^d$ is a control operator of the form $E(u)(x) = \sum_{m=1}^M f_m(x)u_m$ with given $f_m \in H^{1/2}(\partial\Omega)^d$. To obtain mass conservation we assume $f_m \cdot \nu_\Omega = 0$ for $m = 1,\ldots,M$, where ν_Ω denotes the outer normal on Ω.

The corresponding adjoint system again is (14.15)–(14.19), while the gradient changes to

$$\nabla J(u) = \alpha u + E^*\left(\eta^k\nabla p_3 \cdot \nu_\Omega\right).$$

Here the adjoint operator $E^* : (H^{1/2}(\partial\Omega))^* \to \mathbb{R}^M$ of E is given by

$$E^*(g) = \left(\int_{\partial\Omega} f_1 g\, dS, \ldots, \int_{\partial\Omega} f_M g\, dS\right)^t.$$

Concerning the optimal step size θ, the term (14.24) stays valid, where now $c(g)$ is the solution of

$$\xi\rho^k y - \mathrm{div}(\eta^k\nabla y) + \nabla p = 0 \text{ in } \partial\Omega, \tag{15.1}$$

$$\mathrm{div}\; y = 0 \text{ in } \partial\Omega, \tag{15.2}$$

$$c - \tau\mathrm{div}\left(m^k\nabla w\right) = -\tau y\nabla c^k \text{ in } \partial\Omega, \tag{15.3}$$

$$-\sigma\epsilon\Delta c - w = 0 \text{ in } \partial\Omega, \tag{15.4}$$

$$u = Eg \text{ on } \partial\Omega. \tag{15.5}$$

16 The resulting feedback controller

In this section we construct the state feedback controller which realizes the control concept described the previous sections.

We set $Y := H(\mathrm{div},\Omega)\times H^1(\Omega)\times H^1(\Omega)$ and introduce a linear and bounded operator

$$B_k : Y \to Y^*$$

by

$$B_k(y,c,w)^t := \begin{pmatrix} \xi\rho^k y - \mathrm{div}\left(\eta^k\nabla y\right) & 0 & 0 \\ y\nabla c^k & \xi c & -\mathrm{div}(m^k\nabla w) \\ 0 & -\sigma\epsilon\Delta c & -w \end{pmatrix},$$

where $\rho^k := \rho(c^k)$, $\eta^k := \eta(c^k)$, and $m^k := m(c^k)$. We stress that the operator does depend on the phase field from the old time instance. Now the time discretization (14.9)–(14.14) can be written in the form

$$B_k(y^{k+1},c^{k+1},w^{k+1})^t = \begin{pmatrix} \xi\rho^k y^k - t^k\nabla y^k - \sigma\epsilon\mathrm{div}(\nabla c^k\otimes\nabla c^k) + \rho^k g + Eu \\ \xi c^k \\ -f(c^k) \end{pmatrix}.$$

Here $f(c^k) = \lambda_s(c^k) - \frac{\sigma}{\gamma}c^k$ denotes the free energy evaluated at the old time instance, $\xi := 1/\tau$. Furthermore $t^k := \rho^k y^k + \rho'^k m^k\nabla w^k$.

The dual operator $B_k^* : Y^* \to Y^{**} = Y$ of B_k is given by

$$B_k^*(p_3,p_1,p_2)^t := \begin{pmatrix} \xi\rho^k p_3 - \mathrm{div}\left(\eta^k\nabla p_3\right) & p_1\nabla c^k & 0 \\ 0 & \xi p_1 & -\sigma\epsilon\Delta p_2 \\ 0 & -\mathrm{div}(m^k\nabla p_1) & -p_2 \end{pmatrix}.$$

The adjoint equation thus can be written as

$$B_k^*(p_3,p_1,p_2)^t = \begin{pmatrix} 0 \\ c^{k+1} - c_d^{k+1} \\ 0 \end{pmatrix}.$$

Here c_d^{k+1} denotes the desired state at time instance t_{k+1}.

Let us define the restriction and extension operators P_k and E_k, $k = 1,2,3$ by $P_1((y,c,w)) := y$, $P_2((y,c,w)) := c$, $P_3((y,c,w)) := w$, $E_1(y) := (y,0,0)$, $E_2(c) := (0,c,0)$, and $E_3(w) := (0,0,w)$.

Now we have everything at hand to construct the resulting controller. The control u after one gradient step with stepsize $\theta > 0$ is given by

$$\begin{aligned} u =& u^0 - \theta\nabla J(u^0), \\ =& u^0 - \theta\alpha u^0 + \theta E^* p_3, \\ =& (1-\theta\alpha)u^0 + \theta E^* P_1 B_k^*(E_2 c^{k+1} - E_2 c_d^{k+1}), \\ =& (1-\theta\alpha)u^0 + \theta E^* P_1 B_k^*\left(E_2 P_2 B_k\left(f^k + P_1 E u^0\right) - E_2 c_d^{k+1}\right), \\ =& (1-\theta\alpha)u^0 + \theta E^* P_1 B_k^* E_2 P_2 B_k(\underbrace{f^k - B_k^{-1}E_2 c_d^{k+1}}_{F^k} + P_1 E u^0), \end{aligned}$$

where

$$f^k := \begin{pmatrix} \xi\rho^k y^k - t^k\nabla y^k - \sigma\epsilon \mathrm{div}(\nabla c^k \otimes \nabla c^k) + \rho^k g \\ \xi c^k \\ -f(c^k) \end{pmatrix},$$

and

$$F^k := f^k + \begin{pmatrix} 0 \\ -\xi c_d^{k+1} \\ \sigma\epsilon\Delta c_d \end{pmatrix}.$$

Especially in the case $u^0 = 0$ we obtain the controller

$$u = \theta E^* P_1 B_k^* E_2 P_2 B_k F^k$$

We see that the controller directly scales with the stepsize θ of the gradient step. In the case of Dirichlet boundary control the controller changes accordingly and is given by

$$u = \theta E^* \eta^k \nabla(P_1 B_k^* E_2 P_2 B_k F^k) \cdot \nu_\Omega.$$

Remark 16.1. We note that the controller obtained in this way does depend on the whole state at the current time instance t_k obtained from measurements. In practice measurements of the full system state are not available. Thus surrogates have to be provided. This goes beyond the scope of this work.

17 Numerical investigation of the feedback controller

In this section we report on the behaviour of the instantaneous control concept described in the previous sections. We present two examples covering the main aspects of instantaneous control. In the first example we move a circle of fluid and morph it to a square. In the second example we revisit the benchmark investigated in Section 10 where the control gain consists in pushing down the bubble and preventing it from rising. We start with describing the numerical implementation of the controller in the discrete setting.

17.1 Numerical implementation

We here only describe the case of distributed controls and add comments for the case of boundary control where appropriate.

We note that the controller depends on two parameters, namely the time step size $\tau = \tau_u$, and the weight α for the cost of applying the control. We note that there is no need to align the parameter τ_u, which is the model predictive control step size used for the controller, and the time step size τ_s used for the simulation to substitute the measurements. We especially note that a change of τ_u over time would change the controller and will not be considered here, while τ_s can be adapted to numerical requirements, say for fulfilling a CFL-condition.

For the spatial discretization of the primal equation (14.9)–(14.14) we use the same Ansatz as for the simulation of the controlled system as described in Section 10. Especially we use the same meshes. We also use these meshes for the simulation of the adjoint equations (14.15) – (14.19).

In the case of distributed control we for the discretization follow the variational discretization proposed in [Hin05b] yielding that the distributed control u is discretized by the discretization of the adjoint velocity p_3.

The linear systems arising in solving the primal and dual equations are solved as described in Section 10.

17.2 Test problem: Circle to square

In this example we test the ability of the resulting controller to control the dissipative system without outer forces. This means, that we use equal densities and equal viscosities for the two fluids involved and thus neglect the gravitational force. We further use constant mobility and note that in this setting the model used here results in the model investigated in [HK13a].

The domain of computation is the unit square $\Omega = (0,1)^2$. In this domain a circle is located in the center, thus the initial value for the phase-field is given by

$$\begin{aligned} c^0(x_1,x_2) =& \phi_0(z),\\ z =& \frac{1}{\gamma}\left(\sqrt{(x_1-0.5)^2+(x_2-0.5)^2}-0.25\right), \end{aligned}$$

where ϕ_0 is defined in (10.14) and is the first order approximation of the phase-field across the interface. Thus c^0 describes the first order approximation of a bubble with center located at $m = (0.5, 0.5)$ and radius $r = 0.25$. The goal in this example is to move this circle to a new center at $M = (0.35, 0.35)$ and to morph it into a square centered at M of corresponding width, such that mass is conserved. The desired phasefield c_d thus is given by a square with center at M, such that $\int_\Omega c_d(x)\,dx = \int_\Omega c^0(x)\,dx$ holds. Without control the square would evolve to a circle, thus is instable and the controller has to stabilize the square. In Figure 17.1 we depict the initial and the desired phase field.

The further parameters are given as follows: The densities are set to $\rho_1 = \rho_2 = 1$. The viscosities are $\eta_1 = \eta_2 = 0.02$ and $g := (0,0)^t$. We use $\sigma = \gamma = 0.005$. Furthermore $m := \gamma/1000$, $\tau := 0.01$ and $\alpha = 0.001$.

Distributed Control

We first show the behavior of our controller in the case of distributed control. In Figure 17.2 snapshots of the evolution of the phase field are shown.

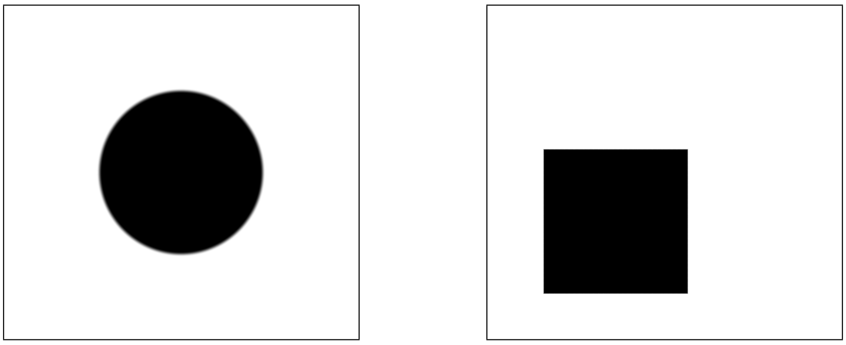

Figure 17.1: Initial (left) and desired (right) distribution for example circle to square.

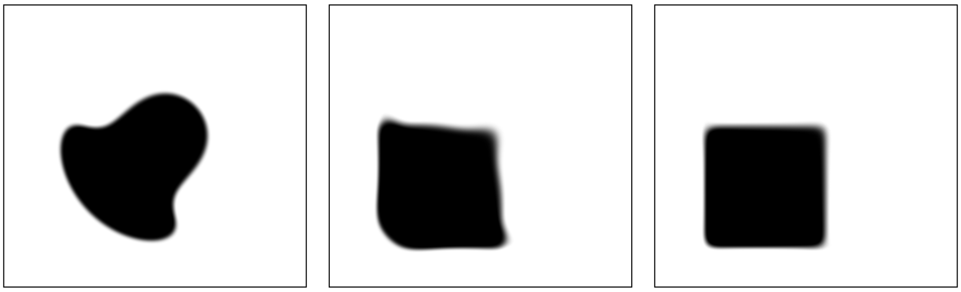

Figure 17.2: Evolution of the phase field for example circle to square ($t \in \{1, 2, 4\}$).

Behaviour of the controller with respect to τ_u We start with investigating the dependence of the controller on τ_u. Since the aim of the controller is to minimize the difference to the desired state our measure for the quality of the controller will be the difference $\|c - c_d\|_{L^2(\Omega)}$.

Since τ_u is the amount of time the controller looks into the future we expect that larger values of τ_u give better properties of the controller. This can be seen in the numerics where the steering properties for large τ_u are significantly better then for small τ_u, see Figure 17.3.

Behaviour of the controller with respect to α Next we investigate the influence of the parameter α on the controller. Since this is a weight in a penalty term we expect that the smaller this value is chosen the larger the control will be and thus the faster the controller will steer the system into the desired state. This is what we can observe from Figure 17.4. We see that for smaller α the controller steers the system faster into the desired state but we also observe, that reducing α beyond a specific value does nearly not affect further the controller properties. This is due to the fact, that for very small α the influence of τ_u on the controller is dominating the quality of the controller.

We note that the oscillations of $\|c - c_d\|$ reported in Figure 17.4 can be explained by the inertia of the fluid and the coupling of the fluid to the phase field c. We note that this effect is the more significant the smaller the viscosity

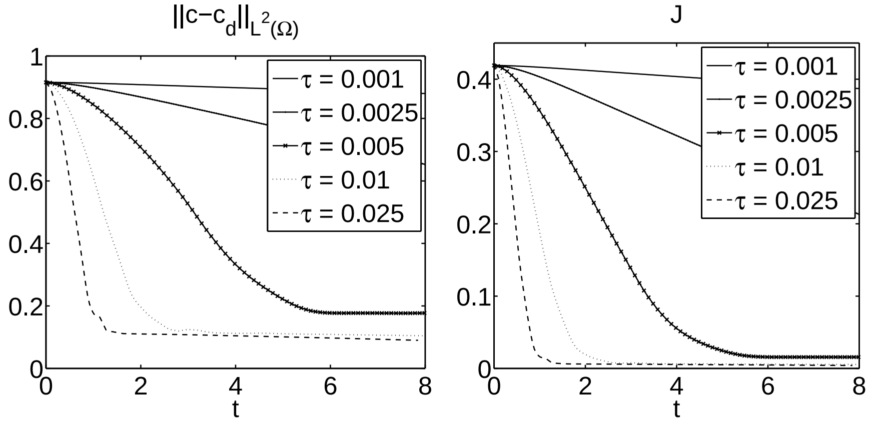

Figure 17.3: The evolution of $\|c - c_d\|_{L^2(\Omega)}$ and J for various control parameters τ_u and distributed control.

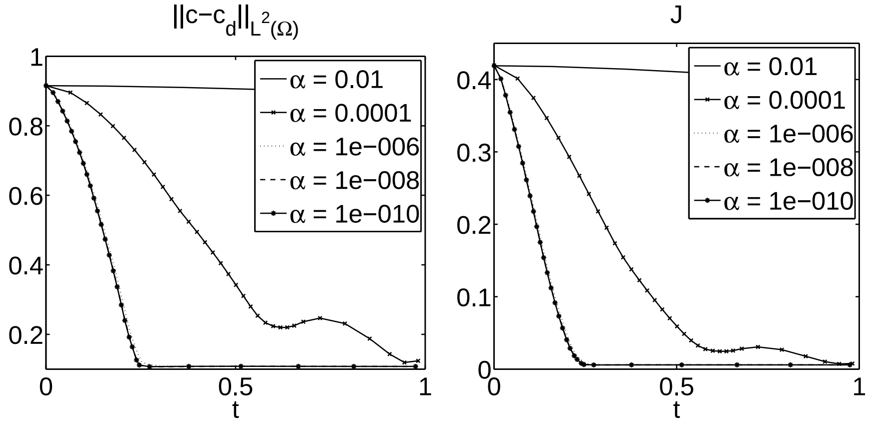

Figure 17.4: The evolution of $\|c - c_d\|_{L^2(\Omega)}$ and J for various control parameters α and distributed control.

of the fluid is and might be overcome with a larger prediction horizon for the control with more then one time step.

The stepsize θ for the gradient step According to (14.24) the stepsize for the gradient step is given by

$$\theta = \frac{\|g\|^2_{L^2(\Omega)}}{\|c(g)\|^2_{L^2(\Omega)} + \alpha\|g\|^2_{L^2(\Omega)}} \leq \frac{1}{\alpha}.$$

For calculating this stepsize we have to perform an additional simulation. This effort might be reduced by substituting θ with $\tilde{\theta} = \alpha^{-1}$ which is an upper bound on the optimal stepsize. In Figure 17.5 we depict the relative difference between θ and $\tilde{\theta} = \alpha^{-1}$ for two values of τ_u, namely $\tau_u = 0.01$ (left) and $\tau_u = 0.001$ (right). As can be seen the relative difference is the smaller the larger we choose α. We further obtain that the relative differences are larger

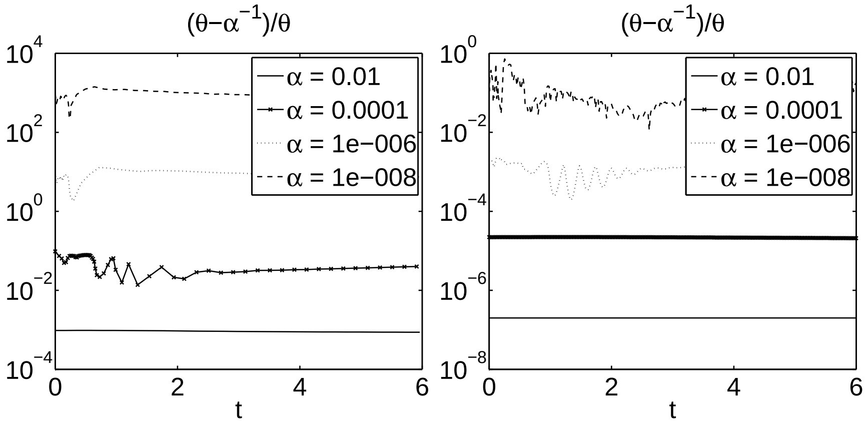

Figure 17.5: The evolution of the relative difference between θ and $\tilde{\theta}$ for $\tau_u = 0.01$ (left) and $\tau_u = 0.001$ (right) and distributed control.

for larger horizons τ_u. This can be explained as in the case of testing the behavior with respect to changes in τ_u. For larger τ_u the allover influence of the control increases and thus the ratio $\|c(g)\|_{L^2(\Omega)}/(\alpha\|g\|_{L^2(\Omega)})$. This implies that the deterioration between θ and α^{-1} increases. On the other hand, the smaller τ_u is the smaller is the impact of the control and thus the smaller $\|c(g)\|$ with the result, that θ tends to α^{-1}.

The distribution of the control From the fact that we start the optimization process with zero control, we obtain that the actual control applied is equal to $u = \theta p_3$ in the case of distributed control. The equations (14.17)–(14.18) for p_3 can be regarded as a time step of a time discrete Stokes problem starting from the initial condition $(p_3)_{old} = 0$. The volume force is equal to $p_1 \nabla c_{old}$ and thus is expected to be located on the interface of c_{old}, since ∇c_{old} is expected to be very small outside of the interface of c_{old}. In fact the volume force is located at the intersection of the interface of c_{old} and the interface of c_d is as shown in Figure 17.6 on the left side. We show an overlay of the meshes for c_{old} and c_d. The locally refined areas correspond to the corresponding interfaces. On the right hand side we show the magnitude of the volume force $p_1 \nabla c_{old}$ and the resulting adjoint velocity field p_3 depicted by arrows. Note that the adjoint velocity and thus the control is strongly located on the intersection of the interfaces corresponding to c_{old} and c_d.

Finite Dimensional Boundary Control

We next investigate the finite dimensional boundary control. Let o_m, $m = 1, \ldots, M$ denote $M = 160$ equidistantly distributed points on $\partial\Omega$ with $\|o_m - o_{m-1}\| = \beta$. For $x \in \partial\Omega$ we define functions f_m, $m = 1, \ldots, M$ on $\partial\Omega$ satisfying

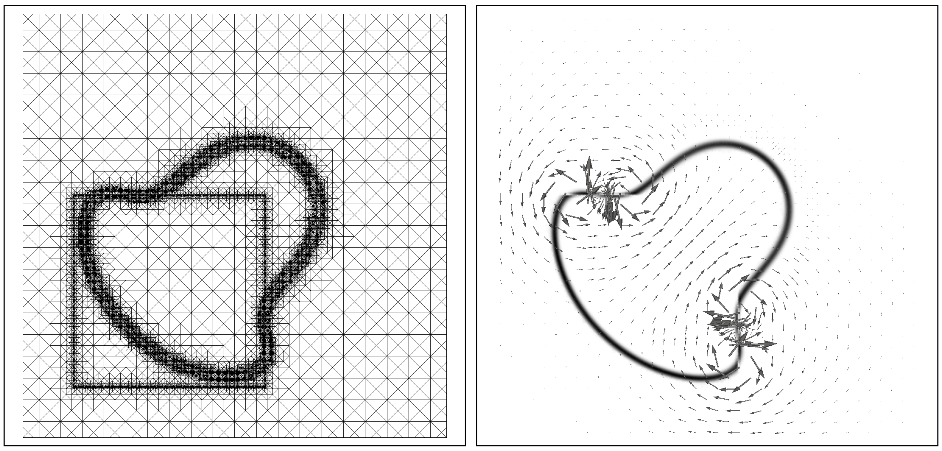

Figure 17.6: Overlay of the meshes for c_d and c_{old} (left) and the magnitude of the volume force $p_1 \nabla c_{old}$ (right) together with the adjoint velocityfield p_3 (arrows).

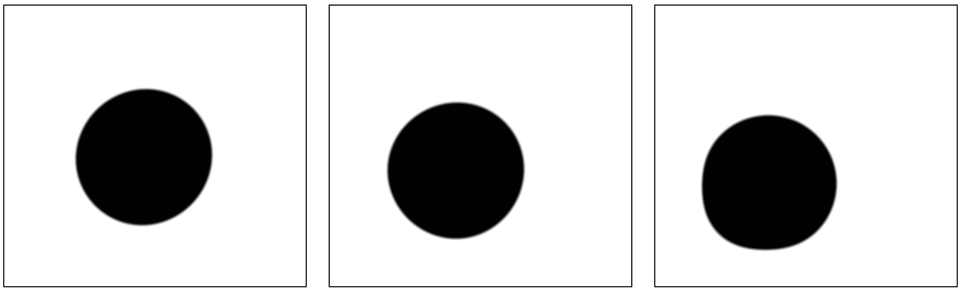

Figure 17.7: Evolution of the phase field for example circle to square with boundary control (t = 20,40,100).

$f_m \cdot \nu_\Omega = 0$, $m = 1, \ldots, M$, where ν_Ω denotes the outer normal at $\partial\Omega$, and

$$f_m(x) \cdot \nu_\Omega^\perp := \begin{cases} \cos^2(z\pi/2), z = \|x - o_m\|/\beta & z \leq 1, \\ 0 & z > 1, \end{cases} \tag{17.1}$$

in tangential direction $\nu_\Omega^\perp$.

The corresponding evolution of the phase field is depicted in Figure 17.7. We note that allowing the control only to act on the boundary we do not obtain the sharp corners of the square. Furthermore, we note that the corner on the left bottom is a littel bit more pronounced than the others since the boundary and thus the boundary control is closer to the interface there.

Behaviour of the controller with respect to τ_u In the case of boundary control we expect to obtain the same properties of the controller with respect to the parameter τ_u. Thus we expect the controller to steer $c \to c_d$ the faster, the larger τ_u is chosen. This can be seen in Figure 17.8.

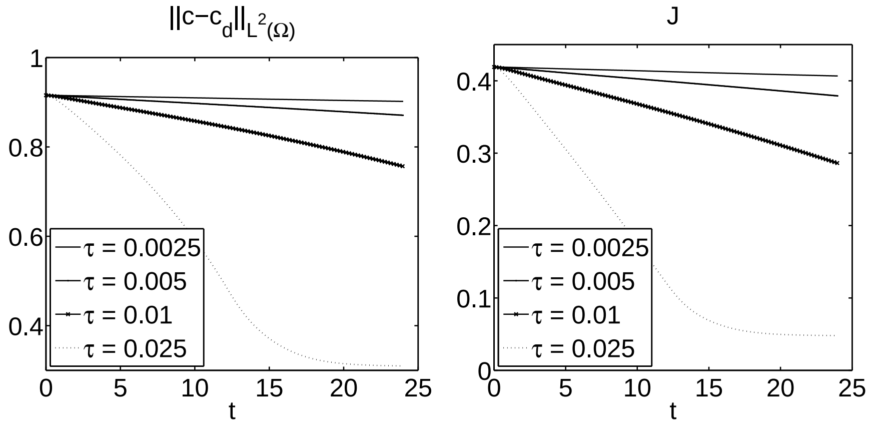

Figure 17.8: The evolution of $\|c - c_d\|_{L^2(\Omega)}$ and J for various control parameters τ_u and boundary control.

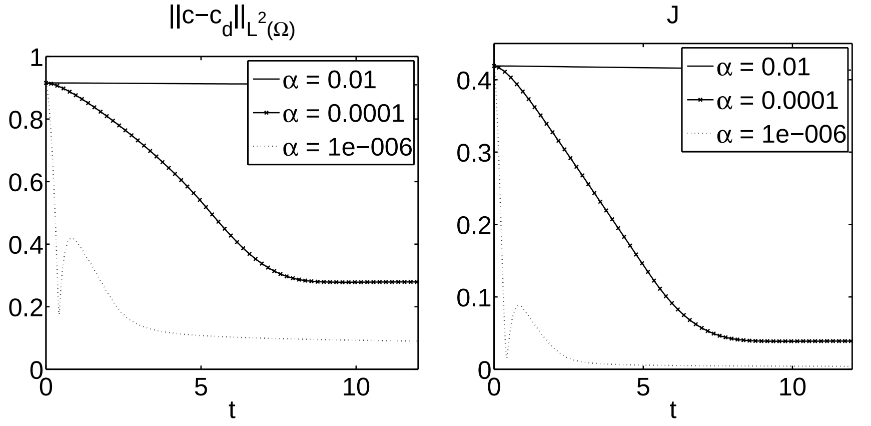

Figure 17.9: The evolution of $\|c - c_d\|_{L^2(\Omega)}$ and J for various control parameter α and boundary control.

Behaviour of the controller with respect to α As expected also in the case of boundary control the controller steers $\|c - c_d\|_{L^2(\Omega)}$ the faster to zero the smaller α is chosen, thus the less we penalise large controls. This is depicted in Figure 17.9.

Stepsize θ for the gradient step As in the case of distributed control we investigate the relative difference between the optimal stepsize θ and the approximation $\tilde{\theta} = \alpha^{-1}$. In Figure 17.10 we show the relative difference between θ and $\tilde{\theta}$ for $\tau_u = 0.01$ (left) and $\tau_u = 0.001$ (right). Again we see that for larger prediction horizon τ_u for given α the relative error increases. We further obtain that the relative difference is quite small if α is quite large and decreases for smaller τ_u. Thus for small τ_u and relatively large α substituting θ by the cheap approximation $\tilde{\theta} = \alpha^{-1}$ seems resonable.

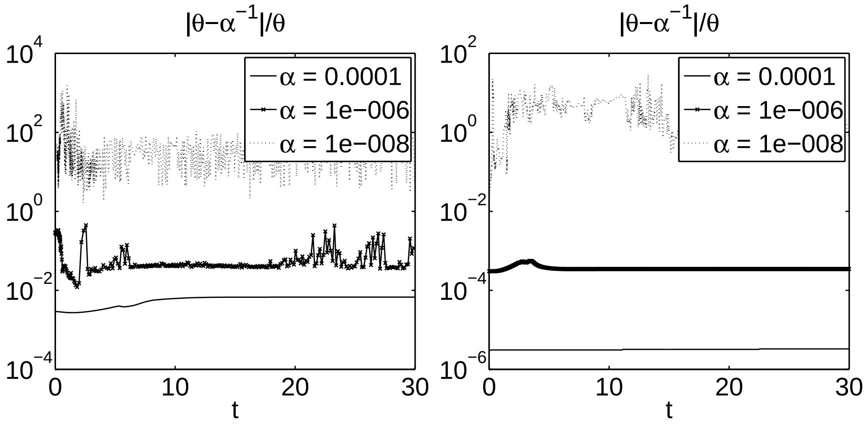

Figure 17.10: Relative difference between the optimal stepsize θ and its approximation $\tilde{\theta}$ for $\tau_u = 0.01$ (left) and $\tau_u = 0.001$.

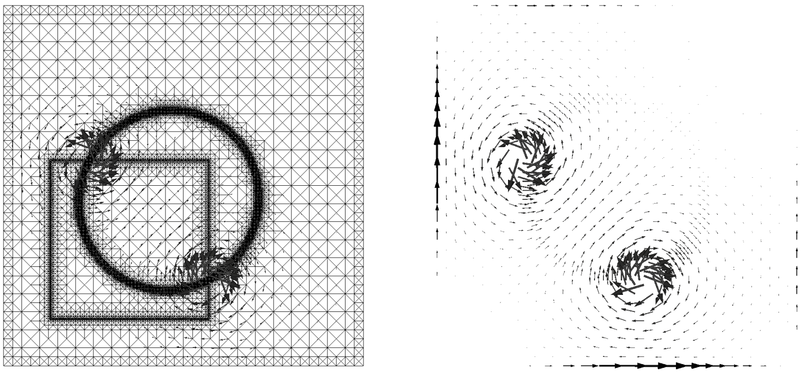

Figure 17.11: The adjoint velocity p_3 (left) and the boundary control Eu (right).

The distribution of the control In the case of discrete boundary control, due to starting the optimization with zero control, the actual control applied is $u = \theta E^* \left(\eta(c_{old})\nabla p_3 \nu_\Omega\right)$ and thus is a weighted integral over the tangential component of the normal derivative of p_3 on the boundary. Since the boundary does not coincide with the intersection of the interfaces of c_{old} and c_d we do not see the strong locality of p_3 in the discrete boundary control. In Figure 17.11 on the left we show the adjoint velocityfield p_3 together with an overlay of the meshes for c and c_d. Again, as expected, p_3 is strongest on the intersection of the interfaces of c_d and c. On the right we again show the adjoint velocity field p_3 together with the resulting control $u = \theta E^* \left(\eta(c_{old})\nabla p_3 \nu\right)$ depicted as Eu on the boundary.

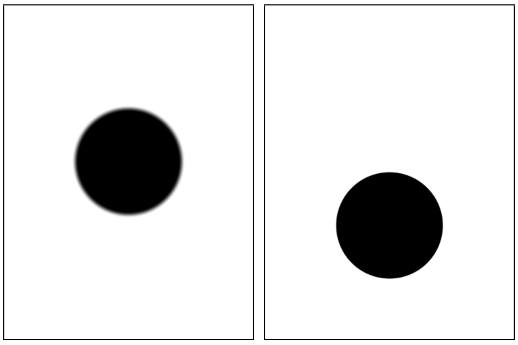

Figure 17.12: Initial (left) and desired (right) distribution for example rising bubble.

17.3 Test problem: The rising bubble

In this example we revisit the benchmark from Section 10. We investigate the cases of stabilising the bubble and thus preventing it from rising while conserving the shape of the bubble (case I), and also the case of even steering the bubble down against the rising forces (case II). The computational domain is given by $\Omega = (0, 1) \times (0, 1.5)$. The initial distribution of the phase field is given as

$$\begin{aligned} c^0(x_1, x_2) =& \phi_0(z), \\ z =& \frac{1}{\gamma}\left(\sqrt{(x_1 - 0.5)^2 + (x_2 - \xi)^2} - 0.25\right), \end{aligned}$$

which defines a bubble centered at $m = (0.5, \xi)$ with radius $r = 0.25$ and ϕ_0 again is the first order approximation to the interface as defined in (10.14). Here $\xi := 0.5$ (case I) or $\xi := 0.7$ (case II). The desired distribution in both cases is a bubble centered at $M = (0.5, 0.5)$. In Figure 17.12 we depict the initial phase field for case II and the desired phase field which also is the initial distribution for case I.

The further parameters are given as in the benchmark case, i.e. $\rho_1 = 1000$, $\rho_2 = 100$, $\eta_1 = 10$ and $\eta_2 = 1$. The gravitational force is $g = (0, -0.98)^t$. The surface tension is $\sigma^{phys} = 24.5$ resulting in $\sigma \approx 15.6$, and $\gamma := 0.01$. The mobility is set to $m := \gamma/1000$. If not mentioned differently we use $\tau_u = 0.001$ and $\alpha = 1e - 7$.

Here we only investigate the case of finite dimensional boundary control, thus $U = L^2(I, \mathbb{R}^M)$ and $Eu(t, x) = \sum_{i=1}^{M} f_i(x)u_i(t)$ for given f_i, $i = 1, \ldots, M$. Distributed control in the presense of gravity is addressed in Section 17.4.

Stabilization (case I)

Here we want to stabilize the rising bubble simulated in Section 10 with wall tangential Dirichlet boundary control, see e.g. [BLK01]. In a practical application this setup can be realized by moving boundary parts established through arrays of rotating disks, see [Kee98].

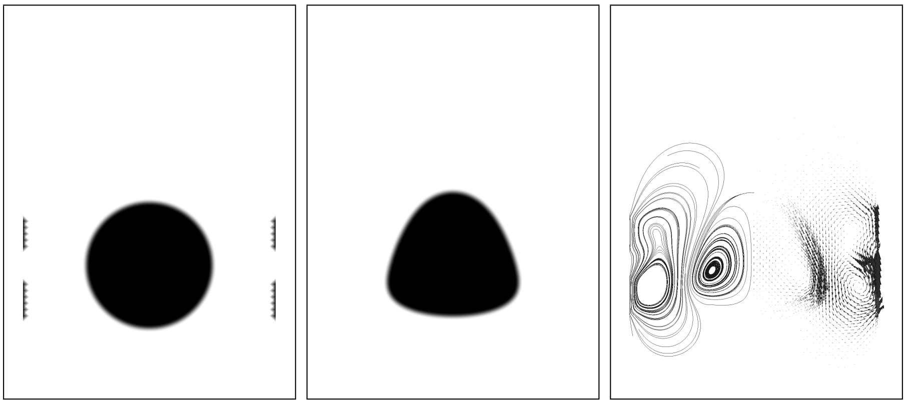

Figure 17.13: Initial phase field with location of the boundary controls (left), phase field at time $t = 8$ (middle) and the velocity field at time $t = 8$ (right). The velocity field is displayed by streamlines on left half plane and by a vector field on right half plane.

The functions f_i defining the control operator E are then piecewise constant functions, which are amplified by the control u, so that E does not map into $H^{1/2}(\partial\Omega)$. On the analytical level one has to use a very weak formulation of the Navier–Stokes system then, see e.g. [Ber04].

We use two disks on each wall, resulting in the control $u = (u_1, u_2, u_3, u_4)$ containing four time dependent control functions. The control operator E is given by

$$\begin{aligned} Eu = &\chi_{\{0\}\times[0.3,0.45]}(x)u_1(t) + \chi_{\{0\}\times[0.55,0.7]}(x)u_2(t) \\ &+ \chi_{\{1\}\times[0.3,0.45]}(x)u_3(t) + \chi_{\{1\}\times[0.55,0.7]}(x)u_4(t), \end{aligned}$$

where $\chi_{\{a\}\times[b,c]}(x)$ is the characteristic function of the set $\{a\} \times [b, c]$.

Since the influence of α and τ_u are already investigated in the previous example we not again investigate this.

In Figure 17.13 on the left we show the initial bubble together with the location of the boundary controls. In the middle plot we show the final shape of the bubble at time $t = 8$. We see that this control is able to prevent the bubble from rising, which is shown in Figure 17.14 where we display the temporal evolution of $\|c - c_d\|_{L^2(\Omega)}$. In Figure 17.13 on the right we further show the velocity field y at time $t = 8$. In the left half plane we display the velocity field by streamlines and in the right half plane we dipict it by a vector field. Note that the velocity corresponds to the actual control on the boundary.

We observe that the chosen control is able to prevent the bubble from rising and stabilizes the bubble at the position shown in Figure 17.13 (middle plot). Unfortunatly the shape of the controlled bubble is not retained and we thus should think of possibilities to improve the control mechanism to retain also the shape of the bubble.

We expect that the deformation of the controlled bubble arises from insufficient control action below the bubble. We therefore add two further controls

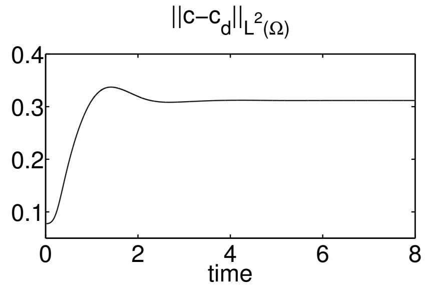

Figure 17.14: Evolution of $\|c - c_d\|_{L^2(\Omega)}$ for four control areas.

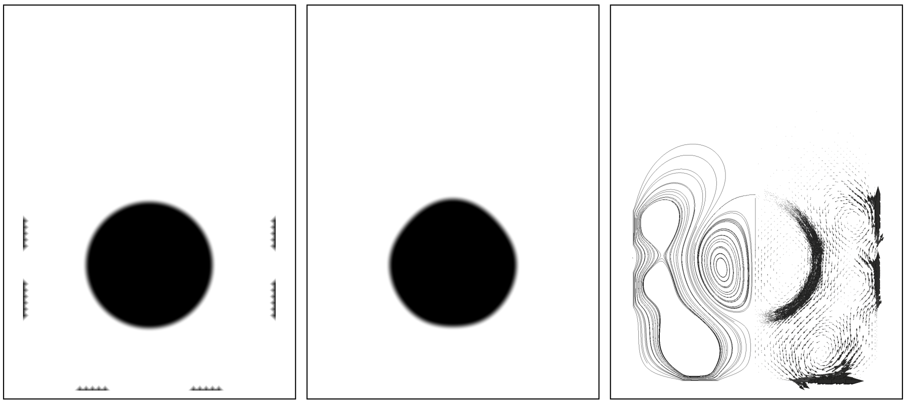

Figure 17.15: The initial phase field together with the six boundary controls (left plot). The stabilized phase field at $t = 8$ (middle) and the stabilizing velocity field (right), depicted as magnitude (left half plane) and as a vector field (right half plane).

on the bottom, yielding a six dimensional control $u^2 = (u_1^2, \dots, u_6^2)$ and the control operator

$$\begin{aligned} E^2 u_2 =& \chi_{\{0\}\times[0.3,0.45]}(x)u_1(t) + \chi_{\{0\}\times[0.55,0.7]}(x)u_2(t) \\ &+ \chi_{\{1\}\times[0.3,0.45]}(x)u_3(t) + \chi_{\{1\}\times[0.55,0.7]}(x)u_4(t) \\ &+ \chi_{[0.2,0.35]\times\{0\}}(x)u_5(t) + \chi_{[0.65,0.8]\times\{0\}}(x)u_6(t). \end{aligned}$$

This gives significant better results, as can be seen in Figure 17.15. In the left plot we depict the six control areas together with the initial phase field. In the middle plot we again show the final stablized bubble and in the right plot we see the stabilizing velocity, again depicted as streamlines and as vector field.

The two additional controls significantly improve the shape of the controlled

In Figure 17.16 we see that in the beginning the difference between the actual bubble and the desired one increases and then approaches the constant value 0.11 which is smaller than the respective value 0.31 in the case with four controls shown in Figure 17.14.

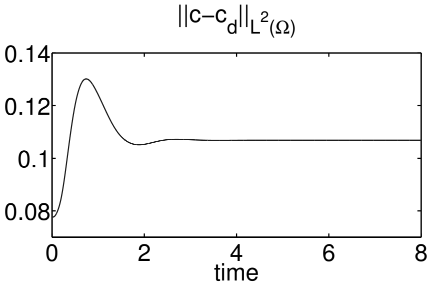

Figure 17.16: Evolution of $\|c - c_d\|_{L^2(\Omega)}$ for six control areas.

Steering down the bubble against gravitation (case II)

The control goal now consists in steering down the bubble down against the rising forces. Thus the initial phase field is a bubble centered at $M = (0.5, 0.7)$, see Figure 17.12.

Since the initial bubble is located above the target bubble we provide additional controls to increase the control impact on the bubble also in the upper regions of the configuration. We use seven controls at each wall and two controls on the bottom of the configuration, resulting in 16 control functions.

To illustrate what can be achieved with wall tangential boundary control for the present configuration, we for comparative studies also investigate a configuration with 20 overlapping control at each wall of the configuration, which results to a control problem with 80 time dependent control functions.

The 16 dimensional control $u = (u_1, \dots u_{16})$ is given by

$$\begin{aligned} Eu =& \chi_{\{0\}\times[0.25,0.39]}(x)u_1(t) + \chi_{\{0\}\times[0.41,0.49]}(x)u_2(t) \\ &+ \chi_{\{0\}\times[0.51,0.59]}(x)u_3(t) + \chi_{\{0\}\times[0.61,0.69]}(x)u_4(t) \\ &+ \chi_{\{0\}\times[0.71,0.79]}(x)u_5(t) + \chi_{\{0\}\times[0.81,0.89]}(x)u_6(t) \\ &+ \chi_{\{0\}\times[0.91,1.05]}(x)u_7(t) \\ &+ \chi_{\{1\}\times[0.25,0.39]}(x)u_8(t) + \chi_{\{1\}\times[0.41,0.49]}(x)u_9(t) \\ &+ \chi_{\{1\}\times[0.51,0.59]}(x)u_{10}(t) + \chi_{\{1\}\times[0.61,0.69]}(x)u_{11}(t) \\ &+ \chi_{\{1\}\times[0.71,0.79]}(x)u_{12}(t) + \chi_{\{1\}\times[0.81,0.89]}(x)u_{13}(t) \\ &+ \chi_{\{1\}\times[0.91,1.05]}(x)u_{14}(t) \\ &+ \chi_{[0.20,0.35]\times\{1\}}(x)u_{15}(t) + \chi_{[0.65,0.80]\times\{1\}}(x)u_{16}(t). \end{aligned}$$

In Figure 17.17 we show the evolution of the 16 dimensional control steering down the bubble against the gravitational forces. We depict the evolution at $t \in \{0.2, 0.4, 0.6, 0.8, 1.0, 4.0\}$ (from top left to bottom right) and show the controlled bubble on the right half plane together with the stream lines of the velocity field y on the left half plane.

We see that the bubble is steered down by a downwards pointing velocity field in the middle of the domain. The bubble is getting flat and thus the interface gets closer to the boundary where the velocity field points upwards

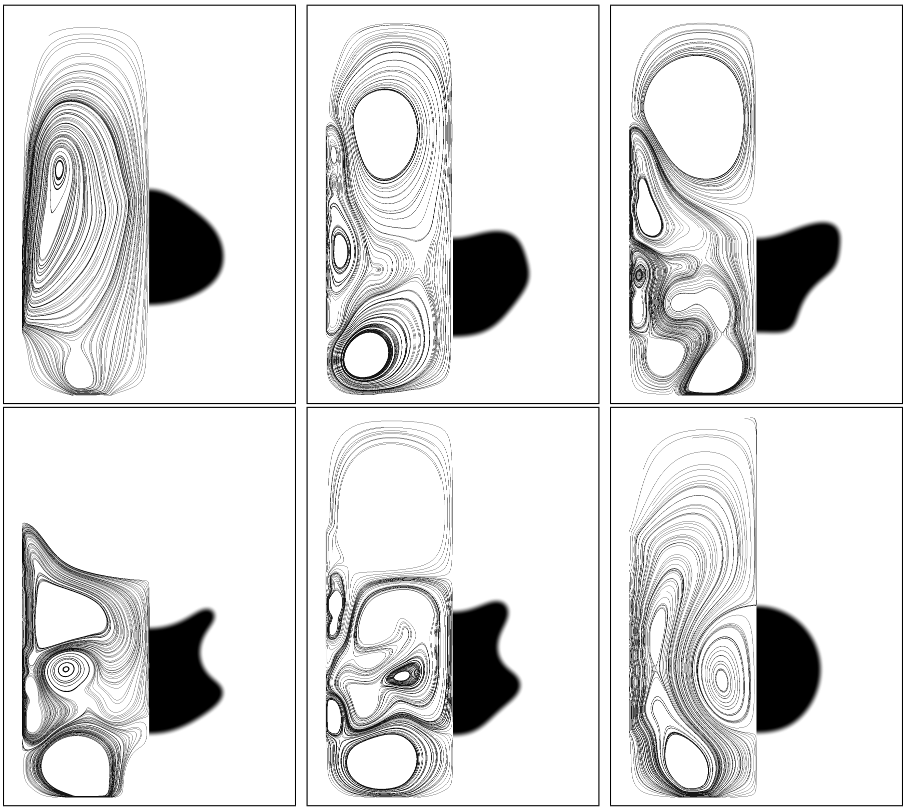

Figure 17.17: The controlled bubble at times $t = 0.2, 0.4, 0.6, 0.8, 1.0, 4.0$ (top left to bottom right) together with streamlines of y for 16 control areas.

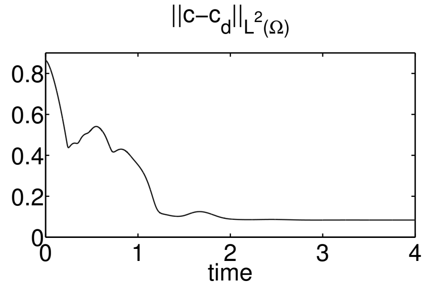

Figure 17.18: Evolution of $\|c - c_d\|_{L^2(\Omega)}$ for 16 control areas.

thus we obtain the kinks in the bubble displayed on the top. Smoothing these kinks yields very local vortices especially seen at time $t = 0.8$ and $t = 1.0$. The vortices are pronounced by inertia of the velocity field since the bubble at its way down has to be stopped at its final position resulting in a control pointing downwards close to the bubble. The shape at time $t = 1.0$ is close to the desired bubble and at $t = 1.2$ (not shown) the bubble is reached and thereafter stabilized by the constant velocity field shown for $t = 4.0$.

In Figure 17.18 we again show the evolution of the difference $\|c - c_d\|_{L^2(\Omega)}$. We see that the bubble is rapidly steered down in a first phase, then reshaping of the bubble increases $\|c - c_d\|_{L^2(\Omega)}$ again, and finally a decay to a stationary value is achieved.

As last test we compare the 16 dimensional control to a smooth 80 dimensional control modeling full Dirichlet boundary control. In Figure 17.19 we show the temporal evolution of the bubble and the velocity field for $t \in \{0.4, 0.8, 1.2, 1.6, 2.0, 4.0\}$ (left top to bottom right). In Figure 17.20 we depict the evolution of $\|c - c_d\|_{L^2(\Omega)}$. We see that the 80 dimensional control is steering the bubble down faster but then has more problems pushing the bubble to the final shape. This is reflected in Figure 17.20, where the evolution of $\|c - c_d\|_{L^2(\Omega)}$ in this case is shown.

We obtain $\|c - c_d\|_{L^2(\Omega)} = 0.84$ in the case of 16 control functions, and $\|c - c_d\|_{L^2(\Omega)} = 0.87$ in the case of 80 control functions.

17.4 Limitation and special aspects

We finish this section with the description of limitations of the presented approach that we found during the numerical investigation.

The interfaces of c_0 and c_d have to intersect

The control is determined by the adjoint velocity field which is driven by the volume force $p_1 \nabla c_{old}$. In Figure 17.11 we show that this force is concentrated at the intersection of the interfaces of c_{old} and p_1. Thus, if c_{old} and p_1 do not intersect, the adjoint velocity field p_3 is approximately zero and thus the control is very small.

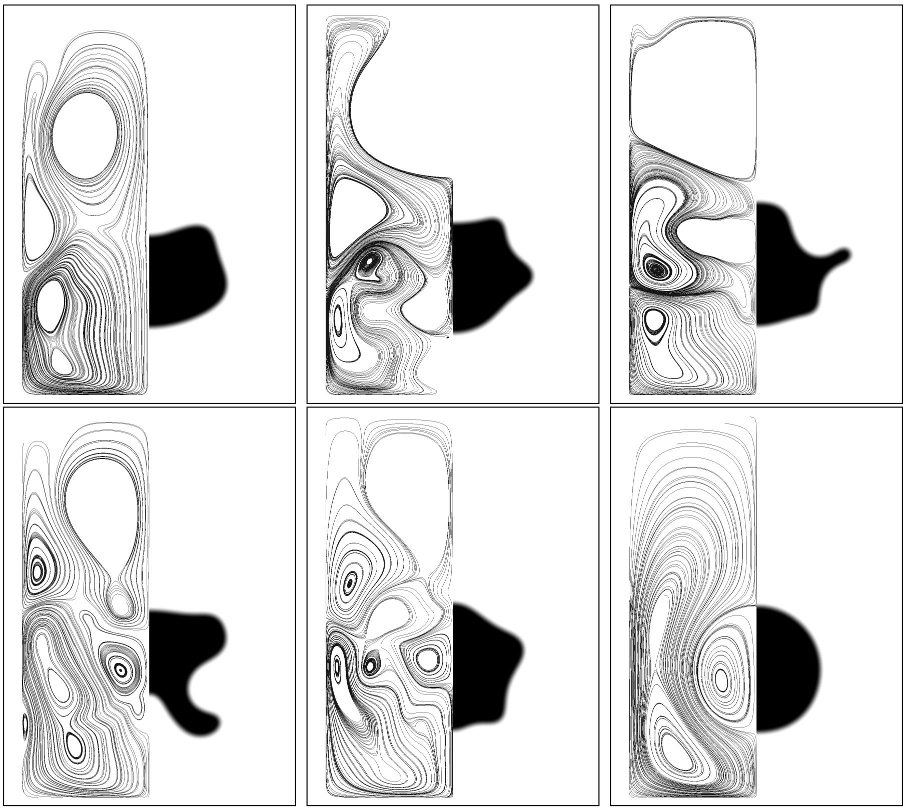

Figure 17.19: The controlled bubble at times $t = 0.4, 0.8, 1.2, 1.6, 2.0, 4.0$ (top left to bottom right) together with streamlines of y for 80 control areas.

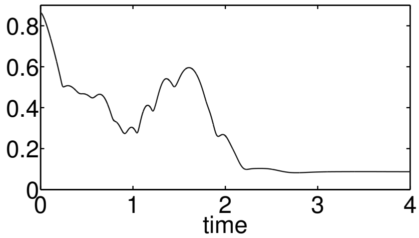

Figure 17.20: Evolution of $\|c - c_d\|_{L^2(\Omega)}$ for 80 control areas.

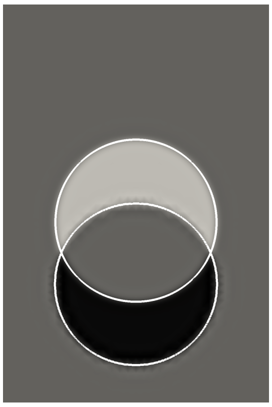

Figure 17.21: The adjoint variable p_1 for case II at the very beginning. The controlled bubble and the desired bubble are marked by their corresponding zero level line in white. The top bubble is the controlled bubble, the bottom bubble is the desired bubble.

In Figure 17.21 we show p_1 at the beginning of the rising bubble example (case II). We see that p_1 is a phase field taking approximately the three distinct values $-2, 0, 2$, following the structure of the volume force $c - c_d$ in the adjoint equation, which takes these values. In particular gray indicates $p_1 \approx 0$.

To overcome this problem one might use additional terms in the minimization functional. For example one could penalise the distance between the centers of mass of c and c_d. Anyway since it is reasonable to assume that c and c_d are close together this limitation is not a severe restriction.

The bubble can rip

Especially in the case of different densities it often appears that the bubble is riping during the optimization progress. This appears whenever parts of the bubble are moved to fast, while other parts are moving with a quite slow velocity. In the case of the bubble being steered down there also appears the problem that at the boundary the fluid moves upwards while in the middle it moves downwards. If parts of the bubble move into the outer region, for example since the bubble is moved to fast, the bubble rips.

Too fast movement of the bubble results from choosing inappropriate parameters for the controllers, thus τ_u and α have to be adapted. This effect can not be detected at runtime by the mean-square-difference functional used here so far. Other kinds of functionals might help, for example penalising the deviation in the circumference of the desired shape and the current shape.

18 First results for the general model predictive control

In Section 13 we described the general model predictive control concept by a sequence of optimal control problem over short time horizons. We later reduced this to an instantaneous variant with solving the optimal control problem approximately by one gradient step and using a prediction horizon whose lenght is chosen as one time step.

We now briefly describe how the general model predictive control concept is implemented and give first numerical results. For fully specifying the concept we have to give the time discrete system that is used for predicting the future behaviour of the fluid and its adjoint system. We further have to specify parameters and meshes. Here we only use distributed control and compare distributed control for the circle to square example with the results for the instantaneous control in Section 17.2. We also investigate the influence of the length of the prediction horizon on the quality of the feedback control.

For convenience of the reader we recall the model for two-phase flow with different densities which we use in the present section

$$\begin{aligned}
\rho\partial_t y + ((\rho y + j)\cdot\nabla)\, y - \operatorname{div}(2\eta Dy) + \nabla p = \qquad & \\
-\sigma\gamma \operatorname{div}(\nabla c\otimes\nabla c) + \rho g + u \quad & \text{in } I\times\Omega, && (18.1)\\
\operatorname{div} y = 0 \quad & \text{in } I\times\Omega, && (18.2)\\
y = 0 \quad & \text{on } I\times\partial\Omega, && (18.3)\\
y(0,x) = y_0(x) \quad & \text{in } \Omega, && (18.4)\\
\partial_t c - \operatorname{div}(m\nabla w) + y\cdot\nabla c = 0 \quad & \text{in } I\times\Omega, && (18.5)\\
-\sigma\gamma\Delta c + F'(c) = w \quad & \text{in } I\times\Omega, && (18.6)\\
\partial_\nu c = \partial_\nu w = 0 \quad & \text{on } I\times\partial\Omega, && (18.7)\\
c(0,x) = c_0(x) \quad & \text{in } \Omega, && (18.8)
\end{aligned}$$

where $j = -\rho'(c)m(c)\nabla w$. Here $F'(c)$ denotes the free energy. In the case of instantaneous control we used the relaxed double-obstacle free energy and in the time discrete setting evaluated it at the old time instance, see Section 14.1. We note that this linearisation is only used for the construction of the controller and is not recommended for simulations over more than one time instance since evaluating the free energy at the old time gives rise to concentration blow up at $\pm\infty$, so that after few time instances the concentration will attend values with absolute value much larger then 1. On the other hand, using the splitting proposed in previous sections would result in a nonlinear equation for the two-phase system which we would like to circumvent due to the higher numerical costs.

We again note, that the numerical concept used for predicting the future behaviour can be chosen quite independend of the concept used for simulation. In this section the simulation is based on the concept described in Section 11 but with a more nonlinear discretization in the Navier–Stokes system, see Remark 11.4. In the following we describe the time discretization used for predicting the future behavior of the concentration.

18.1 The time discrete system for predicting the future behaviour

We for predicting the future behavior of the system here use the polynomial free energy $F'(c) = \sigma\gamma^{-1}(c^3 - c)$, since it is smoother and simplier to handle. In this case time discretizations are available, which yield linear systems in every time step. In [AV12] a Taylor expansion is used for linearization without further proof. In [GT13] linear schemes for the Cahn–Hilliard equation are presented and derived. Here we use the linear scheme from Eyre [Eyr98] based on $F'(c^{k+1}, c^k) = \sigma\gamma^{-1}((c^k)^3 - 3c^k + 2c^{k+1})$. For denoting the time discrete formulation we introduce the superscript ^{k} indicating the k-th time instance.

The discretization of the free energy yields the following time discrete system for the Cahn–Hilliard equation in strong form:

$$\frac{1}{\tau}(c^{k+1} - c^k) - \operatorname{div}\left(m^k \nabla w^{k+1}\right) + y^{k+1} \cdot \nabla c^k = 0, \tag{18.9}$$

$$-\sigma\gamma\Delta c^{k+1} + \sigma\gamma^{-1}((c^k)^3 - 3c^k + 2c^{k+1}) - w^{k+1} = 0. \tag{18.10}$$

The velocity structure is discretized as before, thus

$$\begin{aligned} \frac{1}{\tau}(\rho_{old}(y^{k+1} - y^k)) + \operatorname{div}\left(\eta^k \nabla y^{k+1}\right) + \nabla p^{k+1} & \\ = -\sigma\gamma \operatorname{div}\left(\nabla c^k \otimes \nabla c^k\right) - \left(\left(\rho^k y^k + J^k\right) \cdot \nabla\right) y^k + \rho^k g + u^{k+1}, & \qquad (18.11) \\ -\operatorname{div} y^{k+1} = 0. & \qquad (18.12) \end{aligned}$$

The optimal control problem is given by

$$\begin{aligned} \min J(c^{j+1}, \ldots, &c^{j+L}, u^{j+1}, \ldots, u^{j+L}) \\ &\text{s.t. } (18.9) - (18.12) \text{ for } k = j, \ldots, j+L-1, \end{aligned} \tag{$\mathcal{P}_k$}$$

where

$$J(c^{j+1}, \ldots, c^{j+L}, u^{k+1}, \ldots, u^{k+L}) := \sum_{i=1}^{L} \left(\frac{1}{2}\|c^{j+i} - c_d^{j+i}\|^2 + \frac{\alpha}{2}\|u^{j+i}\|^2 \right).$$

We note that we assume a fixed step length τ for the time discretization and rescaled the functional with this value.

Using formal Lagrange calculus we obtain the following adjoint system

$(\xi = \tau^{-1})$:

$$
\begin{aligned}
\alpha u^{k+1} - p_3^{k+1} &= 0 \\
\xi \rho^k p_3^{k+1} - (J^k \nabla) p_3^{k+1} - \operatorname{div}\left(\eta^k D p_3^{k+1}\right) + \nabla p_4^{k+1} + p_1^{k+1} \nabla c^k & \\
-\xi \rho^{k+1} p_3^{k+2} + \rho^{k+1} \left(\nabla y^{k+2}\right)^t p_3^{k+2} &= 0 \\
\operatorname{div} p_3^{k+1} &= 0 \\
-\operatorname{div}(m^k \nabla p_1^{k+1}) - p_2^{k+1} + \operatorname{div}\left(\rho'^{k+1} m^{k+1} \left(\nabla y^{k+2}\right)^t p_3^{k+2}\right) &= 0 \\
\xi p_1^{k+1} - \sigma\gamma \Delta p_2^{k+1} + F_a'(c^{k+1}, c^k) p_2^{k+1} - \xi p_1^{k+2} - y^{k+2} \nabla p_1^{k+2} & \\
+F_b'(c^{k+2}, c^{k+1}) p_2^{k+2} & \\
+\xi (y^{k+2} - y^{k+1}) \rho'^{k+1} p_3^{k+2} & \\
+\left((\rho'^{k+1} y^{k+1} + \rho''^{k+1} m^{k+1} \nabla w^{k+1} + \rho'^{k+1} m'^{k+1} \nabla w^{k+1}) \nabla\right) y^{k+2} p_3^{k+2} & \\
-\eta'^{k+1} D y^{k+2} \nabla p_3^{k+2} & \\
+\sigma\gamma \operatorname{div}(\nabla p_3^{k+2} \nabla c^{k+1}) + \sigma\gamma \operatorname{div}(\left(\nabla p_3^{k+2}\right)^t \nabla c^{k+1}) - \rho'^{k+1} g p_3^{k+2} & \\
+\left(c^{k+1} - c_d^{k+1}\right) &= 0.
\end{aligned}
$$

Here $F_a'(c^{k+1}, c^k)$ denotes the derivative of $F'(c^{k+1}, c^k)$ with respect to the first argument and $F_b'(c^{k+1}, c^k)$ denotes the derivative of $F'(c^{k+1}, c^k)$ respect to the second argument.

18.2 The fully discrete system

The spatial discretization is again performed by finite elements with piecewise quadratic and globally continous functions for the velocity field and piecewise linear functions for pressure, phase field and chemical potential. We note that the control through variational discretization ([Hin05b]) is implicitly discretized through the discrete adjoint velocity field p_3.

Since the numerical realization is along the lines of the previous sections we here only comment on some special aspects of the numerical concept and implementation.

- We solve the optimal control problem by a steepest descent method with exact minimization in the descent direction. Since we thanks to the linear discretization of the Cahn–Hilliard free energy have a linear-quadratic problem finding the optimal stepsize parameter again is possible with one simulation of the predicting system.

- Concerning the temporal discretization we use a fixed time step length during the prediction. This length is chosen such that a CFL-condition in the first time step is fulfilled. Thus the time discretization step size might differ for each prediction step but is fixed during each prediction. Due to the varying step size lengths we do not use a fixed number of time steps, resulting in a variable length of the prediction horizon, but use a fixed length of the prediction horizon. A comparable concept with fixed time horizon and variable chosen time discretization for example is used in [WKG97] for optimal feeding of a bio-reactor.

- We do not use adaptation for the spatial meshes during prediction and perform the prediction step on the mesh obtained by adapting c^k, thus the mesh that the concentration c^{k+1} is defined on. This is necessary since if adaptation is used, after each prediction we obtain a new set of meshes and thus the gradient $\alpha u + p_3(u)$ might not be evaluable since u and p_3 are defined on different sets of meshes then.

- The gradient method for solving the optimal control problem is stopped as soon as $\|\nabla J(u_k)\| \leq 10^{-4}\|\nabla J(u_0)\| + 10^{-8}$ holds, but at most 10 minimization steps with exact minimization are applied. Typically the minimization stops after 10 steps with $\|\nabla J(u_k)\| \leq 10^{-2}\|\nabla J(u_0)\|$. We note that in fact this still is an instantaneous control concept, but due to the slow convergence of the gradient method and the high numerical effort to obtain the gradient this simplification is needed. For future work we plan to go for some conjugate gradient or quasi-Newton methods to be able to solve the full optimal control problem more efficiently

18.3 First numerical results

Let us present first numerical results for comparing the instantaneous control strategy presented in Section 14 and the model predictive control sketched above.

We use the circle to square example from Section 17.2. Here a bubble located at the center of the domain is moved to the left bottom and deformed to a square. In Section 17.2 we use instantaneous control over a time horizon with length $H = \tau_u = 0.01$ which is resolved by exactly one time step. Here we compare this strategy with model predictive control on larger horizons of length $H \in \{0.015, 0.025, 0.05\}$ that are resolved with more then one time step.

Due to the fact that the strength of the control not only depends on the length of the horizon but also on the length of the first time step during prediction, we do not expect that the MPC in the current used implementation yields a faster steering towards the desired distribution since the first time step is adapted to fulfill a CFL condition, which might be violated in the instantaneous control case. We further note that solving the optimal control problem to a higher accuracy might increase the quality of the MPC. As aforementioned the optimal control problem currently is not solved to a high accurancy.

In Figure 18.1 we depict the evolution of the term $\|c - c_d\|_{L^2(\Omega)}$ for the aforementioned time horizons and $\alpha = 0.001$ on the left hand side. On the right hand side we depict the evolution of this term for the instanteneous controller using $H = \tau_u = 0.01$ and the same value of α.

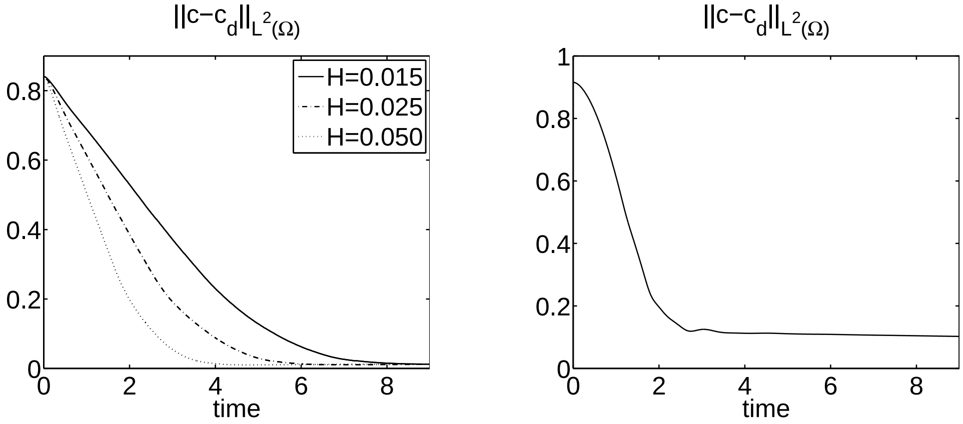

Figure 18.1: The evolution of $\|c - c_d\|$ for MPC with 10 gradient steps and various time horizons H (left) and the corresponding value for IC using $H = \tau_u = 0.01$ (right).

In the case of the model predictive control we see that the difference between c and c_d is reduced the faster the larger the prediction horizon is chosen and that for all horizons we end at the same value for $\|c - c_d\|$.

We further obtain that the oszillation between time 2 and 4 in the simulation with instantaneous control does not appear in the simulation with model predictive control. This can be explained by the optimization over a larger time horizon.

We see that the final value for $\|c - c_d\|$ in the case of instantaneous control is larger than in the case of the MPC. This might arises due to the fact, that the desired function c_d is used as finite element function in the instantaneous control case, but is used as analytic function in the model predictive control case. Thus we can not decide which of the two concepts under investigation delievers the smaller deviation.

The question of incorporation of adaptivity, both in time and space, has to be answered in future work. Also the possible influence of the time adaptation on the actually obtained control has to be considered.

19 Summary and outlook

We applied the instantaneous control concept to the control of two-phase flow. The focus was on fast simulation for the prediction step and we numerically showed that the concepts seems to work at least for some examples, while also shortcomings of the concept were discussed. Concerning future work one has to go for an analytical investigation to justify the practicability of the concept. I expect that this might force to change the focus from fast evaluation to other aspects like stable time discretizations or different functionals to minimize. Combining the stable time discretization from Section 11 with the results from

[HW14] might yield a first promising way of tackling the analytical questions.

We further showed first numerical results for the general model predictive control concept. At the current state the incorporation of data in instantaneous control and model predictive control have to be aligned to be able to compare these to concepts. For the model predictive control approach we have to go for a better optimization algorithm than the currently applied steepest descent method. Also the question of adaptation of space and time during the prediction process has to be addressed and especially the possible influence of the adaptation of time on the actual control has to be considered.

Bibliography

[Abe07] H. Abels. Diffuse Interface Models for Two-Phase flows of Viscous Incompressible Fluids. *Max-Planck Institut für Mathematik in den Naturwissenschaften, Leipzig, Lecture Note*, 36, 2007.

[ABH$^+$13] S. Aland, S. Boden, A. Hahn, F. Klingbeil, M. Weismann, and S. Weller. Quantitative comparison of Taylor Flow simulations based on sharp- and diffuse-interface models. *International Journal for Numerical Methods in Fluids*, 73(4):344–361, October 2013.

[ABPR86] E. L. Allgower, K. Böhmer, F. A. Potra, and W. C. Rheinboldt. A Mesh-Independence Principle for Operator Equations and Their Discretizations. *SIAM Journal on Numerical Analysis*, 23(1):160–169, 1986.

[ADG13a] H. Abels, D. Depner, and H. Garcke. Existence of weak solutions for a diffuse interface model for two-phase flows of incompressible fluids with different densities. *Journal of Mathematical Fluid Mechanics*, 15(3):453–480, September 2013.

[ADG13b] H. Abels, D. Depner, and H. Garcke. On an incompressible Navier–Stokes / Cahn–Hilliard system with degenerate mobility. *Annales de l'Institut Henri Poincaré (C) Non Linear Analysis*, 30(6):1175–1190, 2013.

[AF03] R. A. Adams and J. H. F. Fournier. *Sobolev Spaces, second edition*, volume 140 of *Pure and Applied Mathematics*. Elsevier, 2003.

[AGG12] H. Abels, H. Garcke, and G. Grün. Thermodynamically consistent, frame indifferent diffuse interface models for incompressible two-phase flows with different densities. *Mathematical Models and Methods in Applied Sciences*, 22(3):40, March 2012.

[AMW98] D. M. Anderson, G. B. McFadden, and A. A. Wheeler. Diffuse-interface methods in fluid mechanics. *Annual Review of Fluid Mechanics*, 30:139–165, 1998.

[AO00] M. Ainsworth and J. T. Oden. *A Posteriori Error Estimation in Finite Element Analysis.* Wiley, September 2000.

[AV12] S. Aland and A. Voigt. Benchmark computations of diffuse interface models for two-dimensional bubble dynamics. *International Journal for Numerical Methods in Fluids*, 69:747–761, 2012.

[Bar11] Viorel Barbu. *Stabilization of Navier–Stokes Flows.* Communications and Control Engineering. Springer, 2011.

[BBG99] J. W. Barrett, J. F. Blowey, and H. Garcke. Finite element approximation of the Cahn–Hilliard equation with degenerate mobility. *SIAM Journal on Numerical Analysis*, 37(1):286–318, 1999.

[BBG11] L. Blank, M. Butz, and H. Garcke. Solving the Cahn–Hilliard variational inequality with a semi-smooth Newton method. *ESAIM: Control, Optimisation and Calculus of Variations*, 17(4):931–954, Oktober 2011.

[BDQN12] P. Boyanova, M. Do-Quang, and M. Neytcheva. Efficient preconditioners for large scale binary Cahn–Hilliard models. *Computational Methods in Applied Mathematics*, 12(1):1–22, 2012.

[BE91] J. F. Blowey and C. M. Elliott. The Cahn–Hilliard gradient theory for phase separation with non-smooth free energy. Part I: Mathematical analysis. *European Journal of Applied Mathematics*, 2:233–280, 1991.

[BE92] J. F. Blowey and C. M. Elliott. The Cahn–Hilliard gradient theory for phase separation with non-smooth free energy. Part II: Numerical analysis. *European Journal of Applied Mathematics*, 3:147–179, 1992.

[Ber04] Martin Berggren. Approximation of Very Weak Solution to Boundary-Value Problems. *SIAM Journal on Numerical Analysis*, 42(2):860–877, 2004.

[BGL05] M. Benzi, G.H. Golub, and J. Liesen. Numerical solution of saddle point problems. *Acta Numerica*, 14:1–137, 2005.

[BLK01] A. Balogh, W.-J. Liu, and M. Krstic. Stability Enhancement by Boundary Control in 2D Channel Flow. *IEEE Transactions on Automatic Control*, 46(11):1696–1711, 2001.

[BMT01] T. R. Bewley, P. Moin, and R. Temam. DNS-based predictive control of turbulence: an optimal benchmark for feedback algorithms. *Journal of Fluid Mechanics*, 447:179–225, November 2001.

[BN09] L. Bañas and R. Nürnberg. A posteriori estimates for the Cahn–Hilliard equation. *Mathematical Modelling and Numerical Analysis*, 43(5):1003–1026, September 2009.

[BNN13] L. Bañas, A. Novick-Cohen, and R. Nürnberg. The degenerate and non-degenerate deep quench obstacle problem: A numerical comparison. *Networks and Heterogeneous Media*, 8(1):37–64, March 2013.

[BOS11] Y. Brennier, F. Otto, and C. Seis. Upper bounds on coarsening rates in demixing binary viscous liquids. *SIAM Journal on Mathematical Analysis*, 43(1):114–134, 2011.

[Boy99] F. Boyer. Mathematical study of multiphase flow under shear through order parameter formulation. *Asymptotic Analysis*, 20(2):175–212, 1999.

[Boy02] F. Boyer. A theoretical and numerical model for the study of incompressible mixture flows. *Computers & Fluids*, 31(1):41–68, January 2002.

[BP88] J. H. Bramble and J. E. Pasciak. A preconditioning technique for indefinite systems resulting from mixed approximations of elliptic problems. *Mathematics and Computation*, 50(181):1–17, 1988.

[BPC$^+$07] A. E. Bailey, W. C. K. Poon, R. J. Christianson, A. B. Schofield, U. Gasser, V. Prasad, S. Manley, P. N. Segre, L. Cipelletti, W. V. Meyer, M. P. Doherty, S. Sankaran, A. L. Jankovsky, W. L. Shiley, J. P. Bowen, J. C. Eggers, C. Kurta, T. Lorik, P. N. Pusey, and D. A. Weitz. Spinodal Decomposition in a Model Colloid-Polymer Mixture in Microgravity. *Physical Review Letters*, 99:4, 2007.

[BS08] S. C. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*, volume 15 of *Texts in Applied Mathematics.* Springer, 2008.

[BSB14] J. Bosch, M. Stoll, and P. Benner. Fast solution of Cahn-Hilliard Variational Inequalities using Implicit Time Discretization and Finite Elements. *Journal of Computational Physics*, 262:38–57, 2014.

[Car99] C. Carstensen. Quasi-interpolation and a-posteriori error analysis in finite element methods. *Mathematical Modelling and Numerical Analysis*, 33(6):1187–1202, 1999.

[CDHR08] Y. Chen, T. A. Davis, W. W. Hager, and S. Rajamanickam. Algorithm 887: Cholmod, supernodal sparse cholesky factorization and update/downdate. *ACM Transactions on Mathematical Software*, 35(3):1–14, 2008.

[CF88] P. Constantin and C. Foias. *Navier-Stokes-Equations.* The University of Chicago Press, 1988.

[CH58] J. W. Cahn and J. E. Hilliard. Free Energy of a Nonuniform System. I. Interfacial Free Energy. *The Journal of Chemical Physics*, 28(2):258–267, 1958.

[Che08] L. Chen. *iFEM: An Innovative Finite Element Method Package in Matlab, available at: ifem.wordpress.com*, 2008.

[CHK99] H. Choi, M. Hinze, and K. Kunisch. Instantaneous control of backward-facing step flows. *Applied numerical mathematics*, 31(2):133–158, October 1999.

[Cho95] H. Choi. Suboptimal Control of Turbulent Flow Using Control Theory. In *Proceedings of the International Symposium on Mathematical Modelling of Turbulent Flows, Tokyo, Japan*, 1995.

[Clé75] P. Clément. Approximation by finite element functions using local regularization. *RAIRO Analyse numérique*, 9(2):77–84, August 1975.

[CNQ00] X. Chen, Z. Nashed, and L. Qi. Smoothing methods and semismooth methods for nondifferentiable operator equations. *SIAM Journal on Numerical Analysis*, 38(4):1200–1216, 2000.

[Dav04] T. A. Davis. Algorithm 832: Umfpack v4.3 - an unsymmetric-pattern multifrontal method. *ACM Transactions on Mathematical Software*, 30(2):196–199, 2004.

[DEG$^+$99] J. W. Demmel, S. C. Eisenstat, J. R. Gilbert, X. S. Li, and J. W. H. Liu. A supernodal approach to sparse partial pivoting. *SIAM Journal on Matrix Analysis and Applications*, 20(3):720–755, 1999.

[DGSW10] H. S. Dollar, N. I. M. Gould, M. Stoll, and A. J. Wathen. Preconditioning saddle-point systems with applications in optimization. *SIAM Journal on Scientific Computing*, 32(1):249–270, 2010.

[Dör96] W. Dörfler. A convergent adaptive algorithm for Poisson's equation. *SIAM Journal on Numerical Analysis*, 33(3):1106–1124, 1996.

[DSS07] H. Ding, P. D. M. Spelt, and C. Shu. Diffuse interface model for incompressible two-phase flows with large density ratios. *Journal of Computational Physics*, 226(2):2078–2095, October 2007.

[EAK$^+$01] J. Erlebacher, M. J. Aziz, A. Karma, N. Dimitrov, and K. Sieradzki. Evolution of nanoporosity in dealloying. *Nature*, 410:450 – 453, March 2001.

[EG96] C. M. Elliott and H. Garcke. On the Cahn–Hilliard Equation with Degenerate Mobility. *SIAM Journal on Mathematical Analysis*, 27(2):404–423, 1996.

[EG04] A. Ern and J.-L. Guermond. *Theory and practice of finite elements*, volume 159 of *Applied mathematical sciences.* Springer Verlag, New York, 2004.

[Ell89] C.M. Elliott. *The Cahn–Hilliard model for the kinetics of phase separation. 'Mathematical Models for Phase Change Problems'*, volume 88 of *International Series of Numerical Mathematics*, pages 35–73. Birkhäuser Verlag, Basel, 1989.

[Eyr98] D. J. Eyre. Unconditionally gradient stable time marching the Cahn–Hilliard equation. In *Computational and Mathematical Models of Microstructural Evolution*, volume 529 of *MRS Proceedings*, 1998.

[FM08] E. P. Favvas and A. C. Mitropoulos. What is spinodal decomposition? *Journal of Engineering Science and Technology Review*, 1:25–27, 2008.

[Gar07] H. Garcke. Optimization problems and Cahn-Hilliard systems, in: Miniworkshop on control of free boundaries. In *Oberwolfach Reports*, volume 4(1), pages 447–486. European Mathematical Society Publishing House, 2007.

[GHK14] H. Garcke, M. Hinze, and C. Kahle. A stable and linear time discretization for a thermodynamically consistent model for two-phase incompressible flow. *arXiv: 1402.6524*, 2014.

[GK07] C. Gräser and R. Kornhuber. *On preconditioned Uzawa-type iterations for a saddle point problem with inequality constraints*, volume 55 of *Lecture Notes in computational science and engineering*, pages 91–102. Springer, 2007.

[GK14] G. Grün and F. Klingbeil. Two-phase flow with mass density contrast: Stable schemes for a thermodynamic consistent and frame indifferent diffuse interface model. *Journal of Computational Physics*, 257(A):708–725, January 2014.

[GP11] L. Grüne and J. Pannek. *Nonlinear Model Predictive Control.* Communications and Control Engineering. Springer, 2011.

[GR79] V. Girault and P. A. Raviart. *Finte Element Approximation of the Navier-Stokes equations.* Springer, Lecture Notes in mathematics 749, 1979.

[GR86] V. Girault and P. A. Raviart. *Finite Element Methods for Navier–Stokes Equations*, volume 5 of *Springer series in computational mathematics.* Springer, 1986.

[GR11] S. Gross and A. Reusken. *Numerical methods for two-phase incompressible flows*, volume 40 of *Springer Series in Computational Mathematics*. Springer, 2011.

[Grä11] C. Gräser. *Convex Minimization and Phase Field Models.* PhD thesis, Freie Universität Berlin, 2011.

[Grü13] G. Grün. On convergent schemes for diffuse interface models for two-phase flow of incompressible fluids with general mass densities. *SIAM Journal on Numerical Analysis*, 51(6):3036–3061, 2013.

[GT13] F. Guillén-González and G. Tierra. On linear schemes for a Cahn–Hilliard diffuse interface model. *Journal of Computational Physics*, 234:140–171, 2013.

[HH77] P. C. Hohenberg and B. I. Halperin. Theory of dynamic critical phenomena. *Reviews of Modern Physics*, 49(3):435–479, 1977.

[HHK13] M. Hintermüller, M. Hinze, and C. Kahle. An adaptive finite element Moreau–Yosida-based solver for a coupled Cahn–Hilliard/Navier–Stokes system. *Journal of Computational Physics*, 235:810–827, February 2013.

[HHT11] M. Hintermüller, M. Hinze, and M. H. Tber. An adaptive finite element Moreau–Yosida-based solver for a non-smooth Cahn–Hilliard problem. *Optimization Methods and Software*, 25(4-5):777–811, 2011.

[HIK03] M. Hintermüller, K. Ito, and K. Kunisch. The primal-dual active set strategy as a semi-smooth Newton method. *SIAM Journal on Optimization*, 13(3):865–888, 2003.

[Hin05a] M. Hinze. Instantaneous closed loop control of the Navier–Stokes system. *SIAM Journal on Control and Optimization*, 44(2):564–583, 2005.

[Hin05b] M. Hinze. A variational discretization concept in control constrained optimization: the linear quadratic case. *Computational Optimization and Applications*, 30(1):45–61, 2005.

[HK00] M. Hinze and K. Kunisch. *Three control methods for time - dependent fluid flow*, volume 60 of *Flow, Turbulence and Combustion*, pages 273–298. Springer, 2000.

[HK13a] M. Hinze and C. Kahle. A nonlinear Model Predictive Concept for the Control of Two-Phase Flows governed by the Cahn–Hilliard Navier–Stokes System. In *System Modeling and Optimization*, volume 391 in IFIP Advances in Information and Communication Technology, 2013.

[HK13b] M. Hinze and C. Kahle. Model Predictive Control of Variable Density Multiphase Flows Governed by Diffuse Interface Models. In *Proceedings of the first IFAC Workshop on Control of Systems Modeled by Partial Differential Equations*, volume 1, pages 127–132, 2013.

[HM07] M. Hinze and U. Matthes. Optimal and Model predictive control of the Boussinesq approximation. In *Control of Coupled partial differential equations*, volume 155 of *International Series of Numerical Mathematics*, pages 149–174. Springer, 2007.

[HPUU09] M. Hinze, R. Pinnau, M. Ulbrich, and S. Ulbrich. *Optimization with PDE constraints*, volume 23 of *Mathematical Modelling: Theory and Applications.* Springer, 2009.

[HT74] P. Hood and G. Taylor. *Navier–Stokes equations using mixed interpolation.* Finite Element Methods in Flow Problems. UAH Press, 1974.

[HTK$^+$09] S. Hysing, S. Turek, D. Kuzmin, N. Parolini, E. Burman, S. Ganesan, and L. Tobiska. Quantitative benchmark computations of two-dimensional bubble dynamics. *International Journal for Numerical Methods in Fluids*, 60(11):1259–1288, 2009.

[HU04] M. Hintermüller and M. Ulbrich. A mesh-independence result for semi-smooth Newton methods. *Math. Program.*, 101:151–184, 2004.

[HV02] M. Hinze and S. Volkwein. Instantaneous control for the Burgers equation: Convergence analysis and numerical implementation. *Nonlinear Analysis*, 50(1):1–26, July 2002.

[HW14] M. Hintermüller and D. Wegner. Optimal control of a semidiscrete cahn–hilliard navier–stokes system. *SIAM Journal on Control and Optimization*, 52(1):747–772, 2014.

[Jus11] U. Juschkat. A posteriori Fehlerschätzer für die Oseen Gleichung. Master's thesis, Ruhr Universität Bochum, July 2011.

[Kah13] C. Kahle. Instantaneous control of two-phase flow with different densities. *Oberwolfach Reports, Chapter: Interfaces and Free Boundaries: Analysis, Control and Simulation*, 10(1):898–901, 2013.

[Kee98] L. R. Keefe. Method and apparatus for reducing the drag of flows over surfaces. *US Patent US5 803 409*, 1998.

[KKL04] J. Kim, K. Kang, and J. Lowengrub. Conservative multigrid methods for Cahn-Hilliard fluids. *J. Comp. Phys.*, 193:511–543, 2004.

[KLW02] D. Kay, D. Loghin, and A. Wathen. A preconditioner for the steady state Navier–Stokes equations. *SIAM Journal on Scientific Computing*, 24(1):237–256, 2002.

[KSW08] D. Kay, V. Styles, and R. Welford. Finite element approximation of a Cahn–Hilliard–Navier–Stokes system. *Interfaces and Free Boundaries*, 10(1):15–43, 2008.

[KW06] D. Kay and R. Welford. A multigrid finite element solver for the Cahn–Hilliard equation. *Journal of Computational Physics*, 212:288–304, 2006.

[LT98] J. Lowengrub and L. Truskinovsky. Quasi-incompressible Cahn–Hilliard fluids and topological transitions. *Proceedings of the royal society A*, 454(1978):2617–2654, 1998.

[Mei08] A. Meister. *Numerik linearer Gleichungssysteme. Eine Einführung in moderne Verfahren.* Springer Vieweg, 2008.

[MK02] M. Milano and P. Koumoutsakos. A clustering genetic algorithm for cylinder drag optimization. *Journal of Computational Physics*, 175(1):79–107, January 2002.

[NMT11] B. R. Noack, M. Morzyński, and G. Tadmor, editors. *Reduced Order Modelling for Flow Control*, volume 528 of *International Centre for Mechanical Sciences, Courses and Lectures.* Springer, 2011.

[NP97] V. Nevistic and J. A. Primbs. Finite Receding Horizon Control: A General Framework for Stability and Performance Analysis. Technical Report 6, Automatic control laboratory, ETH Zürich, 1997.

[OP88] Y. Oono and S. Puri. Study of phase-separation dynamics by use of cell dynamical systems. I. Modeling. *Physical Review A*, 38(1):434–463, 1988.

[OSS13] F. Otto, C. Seis, and D. Slepčev. Crossover of the coarsening rates in demixing of binary viscous liquids. *Communications in Mathematical Sciences*, 11(2):441–464, 2013.

[Pro02] B. Protas. On the 'vorticity' formulation of the adjoint equations and its solution using the vortex method. *Journal of Turbulence*, 3:N48, 2002.

[PRR05] J. Peters, V. Reichelt, and A. Reusken. Fast iterative solvers for the discrete Stokes equations. *SIAM Journal on Scientific Computing*, 27(2):646–666, 2005.

[PW60] L. E. Payne and H. F. Weinberger. An optimal Poincaré inequality for convex domains. *Archive for Rational Mechanics and Analysis*, 5(1):286–292, Januar 1960.

[Ray92] L. Rayleigh. On the theory of surface forces - II. compressible fluids. *Philosophical Magazine Series 5*, 33(201):209–220, 1892.

[Sig79] E. D. Siggia. Late stages of spinodal decomposition in binary mixtures. *Physical Review A*, 29(2):595–605, August 1979.

[SS86] Y. Saad and M. H. Schultz. GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM Journal on Scientific and Statistical Computing*, 7(3):856–869, July 1986.

[SS12] I. W. Seo and C. G. Song. Numerical simulation of laminar flow past a circular cylinder with slip conditions. *nternational Journal for Numerical Methods in Fluids*, 68(12):1538–1560, April 2012.

[SY10] J. Shen and X. Yang. A Phase-Field Model ant Its Numerical Approximation for Two-Phase Incompressible Flows with Different Densities and Viscosities. *SIAM Journal on Scientific Computing*, 32(3):1159–1179, 2010.

[Tay96] M. E. Taylor. *Partial Differential Equations I: Basic Theory*, volume 115 of *Applied Mathematical Sciences*. Springer, 1996.

[Tem77] R. Temam. *Navier–Stokes equations - Theory and numerical analysis*. North-Holland Publishing Company, Amsterdam, New York, Oxford, 1977.

[van92] H. A. van der Vorst. Bi-cgstab: A fast and smoothly converging variant of bi-cg for the solution of nonsymmetric linear systems. *SIAM Journal on Scientific Computing*, 13(2):631–644, March 1992.

[vanon] J. van der Waals. The thermodynamic theory of capillarity under the hypothesis of a continuous density variation. *Journal of Statistical Physics*, 20(2):197–200, February 1979 (year of english translation).

[Ver84] R. Verfürth. Error estimates for a mixed finite element approximation of the Stokes equation. *RAIRO Analyse numérique*, 18(2):175–182, 1984.

[Ver89] R. Verfürth. A posteriori error estimates for the Stokes equation. *Numerische Mathematik*, 55(3):309–325, 1989.

[Ver10] R. Verfürth. A posteriori error analysis of space-time finite element discretizations of the time-dependent Stokes equations. *Calcolo*, 47:149–167, 2010.

[WKG97] W. Waldraff, R. King, and E. D. Gilles. Optimal feeding strategies by adaptive mesh selection for fed-batch bioprocesses. *Bioprocess Engineering*, 17(4):221–227, September 1997.

[ZK79] J. Zowe and S. Kurcyusz. Regularity and stability for the mathematical programming problem in Banach spaces. *Applied Mathematics and Optimization*, 5(1):49–62, 1979.

Appendix

Theorem A1 (Lax-Milgram,[EG04, Lem. 2.2])**.** *Let V be a Hilbert space, let $a : V \times V \to \mathbb{R}$ be a continuous bilinear form and $f : V \to \mathbb{R}$ a continuous linear form. Assume that a is coercive, i.e.*

$$\exists \alpha > 0 \,:\, a(u,u) \geq \alpha \|u\|_V^2 \, \forall u \in V.$$

Then the problem

$$\begin{cases} \text{Seek}\, u \in V \, \text{s.t.} \\ a(u,v) = f(v), \, \forall v \in V \end{cases}$$

admitts a unique solution an there holds

$$\|u\|_V \leq \frac{1}{\alpha} \|f\|_{V'} \, \forall f \in V'.$$

Theorem A2 ([ZK79])**.** *Let X and Y be real Banach spaces. Let C be a convex and closed subset of X and K a closed cone in Y with vertex at 0. Let*

$$\begin{array}{rl} F : X \to \mathbb{R} & \textit{Fréchet-differentiable,} \\ g : X \to Y & \textit{continuously Fréchet-differentiable} \end{array}$$

and consider the problem

$$\textit{(P) minimize } F(x) \textit{ subject to } x \in C \textit{ and } g(x) \in K.$$

We assume that there exists a unique optimal solution for (P) denoted by $\hat{x} \in C$ with corresponding $\hat{y} = g(\hat{x})$. We define the canonical hulls of $C \setminus \{\hat{x}\}$ and $K \setminus \{\hat{y}\}$ as

$$\begin{aligned} C(\hat{x}) =& \{x \in X \,|\, \exists \beta \geq 0, \, \exists c \in C, x = \beta(c - \hat{x})\}, \\ K(\hat{y}) =& \{y \in Y \,|\, \exists \lambda \geq 0, \, \exists k \in K, y = k - \lambda \hat{y}\}. \end{aligned}$$

The polar cone of a subset $A \subset X$ is denoted by

$$A^+ = \{x^* \in X^* \,|\, \langle x^*, a \rangle_{X^*,X} \geq 0 \, \forall a \in A\}.$$

Assume that the following constraint qualification holds:

$$g'(\hat{x}) C(\hat{x}) - K(\hat{y}) = Y. \tag{a1}$$

Then there exists a Lagrange multiplier $\mu^* \in Y^*$ such that there holds

$$\begin{aligned} \mu^* \in& K^+, \\ \langle \mu^*, \hat{y} \rangle_{Y^*,Y} =& 0, \\ F'(\hat{x}) - \mu^* \circ g'(\hat{x}) \in& C(\hat{x})^+. \end{aligned}$$

Definition A3 ([HIK03, CNQ00]). Let X and Z be Banach spaces, $D \subset X$ an open subset. A mapping $F : D \subset X \to Z$ is called Newton differentiable or slantly differentiable in $U \subset D$ if there exists a family of mappings $G : U \to \mathcal{L}(X, Z)$ such that

$$\lim_{d \to 0} \frac{1}{\|d\|_X} \|F(x+d) - F(x) - G(x+d)d\|_Z = 0 \quad \forall x \in U.$$

The operator G is called a Newton derivative of F on U.

Theorem A4 ([HIK03]). *Let X and Z be Banach spaces, $D \subset X$ an open subset. Let $F : D \to Z$ and $x^\star \in D$ fulfill $F(x^\star) = 0$. Let F be Newton differentiable with Newton derivative G in a neighbourhood $U(x^\star)$ of $x^\star$. Let G be non singular on $U(x^\star)$ and $\|G(x)^{-1}\|_{\mathcal{L}(Z,X)} \leq C \, \forall x \in U(x^\star)$. Let the sequence $\{x^k\}_{k\in\mathbb{N}}$ be generated by Newton's method, thus $x^0 \in D$ is given and $x^{k+1} = x^k - G(x^k)^{-1}F(x^k)$.*

Then the sequence $\{x^k\}_{k\in\mathbb{N}}$ converges superlinearly to x^* provided that $\|x^0 - x^*\|_X$ is sufficiently small.

Theorem A5 (Discrete Sobolev inequalities [HPUU09, Prop. 3.1]). *Let $\mathcal{T}$ denote a quasi-uniform, regular triangulation of $\Omega \subset \mathbb{R}^n$ $(n = 1, 2, 3)$. Then for every piecewise linear, continuous finite element function $v_h \in H_0^1(\Omega)$ there holds*

$$\|v_h\|_{L^\infty(\Omega)} \leq C\sigma(d, h)\|\nabla v\|_{L^2(\Omega)},$$

where

$$\sigma(d,h) = \begin{cases} 1 & \text{if } d = 1, \\ |\log h|^{1/2} & \text{if } d = 2, \\ h^{-1/2} & \text{if } d = 3. \end{cases}$$

A general saddle point problem

Let X, M denote two Hilbert spaces with corresponding dual spaces X^*, M^* with dualities $\langle \cdot, \cdot \rangle_{X^*,X}$ and $\langle \cdot, \cdot \rangle_{M^*,M}$ and norms $\|\cdot\|_X$ and $\|\cdot\|_M$. We introduce two continuous bilinear forms

$$a : X \times X \to \mathbb{R}, \quad b : X \times M \to \mathbb{R}$$

and consider the following variational problem:
For $l \in X^*$ and $\chi \in M^*$ find a pair $(u, \lambda) \in X \times M$ such that there holds

$$\begin{cases} a(u,v) + b(v,\lambda) = \langle l, v\rangle_{X^*,X} & \forall v \in X, \\ \qquad\qquad b(u,\mu) = \langle \chi, \mu\rangle_{M^*,M} & \forall \mu \in M. \end{cases} \tag{SP}$$

Theorem A6 ([GR79, Th. I.4.1]). *Assume that there holds:*

1. *The bilinear form a is coerzive, i.e. there exists a constant $\alpha > 0$ such that*
$$a(u,u) \geq \alpha \|u\|_X^2 \quad \forall u \in X,$$
2. *The bilinear form b satisifies the inf-sup condition, i.e. there exists a constant $\beta > 0$ such that*
$$\inf_{\mu \in M} \sup_{v \in X} \frac{b(v,\mu)}{\|v\|_X \|\mu\|_M} \geq \beta$$
holds.

Then there exists a unique $u \in V = \{v \in X, s.t.\, b(v,\mu) = \langle \chi, \mu \rangle_{M^*,M}\ \forall \mu \in M\}$ and a unique $\lambda \in M$ such that the pair (u, λ) is the unique solution of problem SP.

Theorem A7 (Pressure reconstruction, [GR86, Lem. I 2.1])**.**
Let $f \in H^{-1}(\Omega)^d$ satisfy
$$\langle f, v \rangle_{(H^{-1}(\Omega))^d, \left(H_0^1(\Omega)\right)^d} = 0 \qquad\qquad \forall v \in V.$$
Here $V = \{v \in H_0^1(\Omega)^d \,|\, (div(v), q) = 0\, \forall q \in L^2_{(0)}(\Omega)\}$.

Then there exists $p \in L^2(\Omega)$ such that
$$f = \nabla p$$
holds. If Ω is connected, p is unique up to an additive constant.

Lemma A8 ([Tem77, Lem. II.1.4])**.** *Let X be a finite dimensional Hilbert space with scalar product $(\cdot,\cdot)$ and norm $\|\cdot\|$ and let P be a continuous mapping from X into itself such that*
$$(P(x), x) > 0\, \forall \|x\| = k > 0.$$
Then there exists $\xi \in X$, $\|\xi\| \leq k$, such that
$$P(\xi) = 0.$$

Summary

This work consists of two parts. The first part deals with the simulation of two-phase flow using *diffuse interface* models. In the second part the presented concept is used to control two-phase flow.

The numerical concept is presented for a well investigated model for two-phase flow with equal densities and viscosities (model 'H' ([HH77])) is used. It consists of a coupled system for describing the two-phase structure (Cahn–Hilliard equation [CH58]) and the Navier–Stokes equation for the fluid structure. A time discretization decouples these two systems on each time instance and allows to treat them separately.

The two-phase structure leads to a variational inequality that is treated by Moreau–Yosida relaxation. It is shown that the unique solution of the system can be found by Newton's method in function space and that for vanishing relaxation the solution of the variational inequality is obtained. A reliable and efficient error estimator is proposed that is used to control the discretization error during the simulation of the two-phase structure.

The time discretization of the fluid structure yields a linear equation. Existence of a solution is proven and a reliable and efficient error estimator is provided. This yields a discretization of the fluid structure using different meshes for the spatial discretization and for the two-phase structure.

The numerical concepts are extensively tested numerically. Especially the behaviour if the solver with respect to parameters is investigated as well as the spatial discretization resulting from the adaptive concept.

The presented concepts are therafter used to simulate a model which allows for fluids with different densities and viscosities.

This part ends with presenting a new time discretization which allows for time discrete energy estimates that are also conserved in the fully discret setting.

The control of two-phase fluids is investigated in the second part of this work. The concept of model predictive control is presented. We use a variant of this, called *instantaneous control*. The concept is based on solving optimal control problems over short time horizons. We solve these problems approximately by only one gradient step. The concept is extensively investigated numerically. Here both distributed control and Dirchlet boundary control with finitely many controls is investigated.

Furthermore, we present first numerical results for the general model predictive control.

Zusammenfassung

Die vorliegende Arbeit besteht aus zwei Teilen. Der erste Teil befasst sich mit der Simulation von Zwei-Phasen Strömungen mittels *diffuse-interface* Modellen. Im zweiten Teil wird das bereit gestellte Konzept verwendet um Zwei-Phasen Strömungen zu steuern.

Das numerische Konzept wird zunächst an Hand eines gut untersuchten Modells für Zwei-Phasen Strömungen (Modell 'H' ([HH77]) mit gleichen Dichten und gleichen Viskositäten dargestellt. Es besteht aus einem gekoppelten System für die Beschreibung der Zwei-Phasen Struktur (Cahn–Hilliard Gleichungen [CH58]) und den Navier–Stokes Gleichungen für die Fluidstruktur. Eine Zeitdiskretisierung ermöglicht es, die Fluidstruktur in jedem Zeitschritt von der Zwei-Phasen Struktur zu trennen.

Die Zwei-Phasen Struktur führt auf eine Variationsungleichung, die mit Moreau–Yosida Relaxierung behandelt wird. Es wird gezeigt, dass die eindeutige Lösung des Systems mittels Newton Verfahrens im Funktionenraum gefunden werden kann, und dass für verschwindene Relaxierung die Lösung der Variationsungleichung gefunden wird. Es wird ein zulässiger und effizienter Fehlerschätzer hergeleitet, mit dem der Diskretisierungsfehler bei der Simulation der Zwei-Phasen Struktur kontrolliert werden kann.

Die Zeitdiskretisierung der Fluidstruktur führt auf eine lineare Gleichung. Existenz von Lösungen wird gezeigt und ein zulässiger und effizienter Fehlerschätzer wird dargestellt. Dies führt zu einer Diskretisierung auf anderen räumlichen Gittern als die Diskretisierung der Zwei-Phasen Struktur.

Die numerischen Verfahren werden ausgiebig numerisch untersucht. Insbesondere wird auf das Verhalten der Löser bezüglich der Parameter eingegangen und auf die durch das adaptive Konzept entstehenden Gitter.

Anschließend werden die dargestellten Konzepte verwendet um ein Modell zu simulieren welches auch zulässt, dass die Flüssigkeiten verschiedene Dichten und Viskositäten aufweisen.

Dieser Teil endet mit der Vorstellung einer neuen Zeitdiskretisierung die es erlaubt zeitdiskrete Energieabschätzungen zu erhalten, die auch im volldiskreten Modell erhalten bleiben.

Die Steuerung zweiphasiger Fluide wird im zweiten Teil der Arbeit untersucht. Hier wird das Konzept der modellprädiktiven Kontrolle vorgestellt. Wir verwenden eine Variante hiervon, die *instantane Kontrolle*. Das Konzept basiert auf dem Lösen von Kontrollproblemen über kurze Zeithorizonte. Wir lösen diese Probleme nur approximativ mittels eines Gradientschrittes. Das Konzept wird ausgiebig numerisch untersucht. Hierbei wird sowohl räumlich verteilte Kontrolle als auch diskrete am Rand des Gebiets angebrachte Kontrolle untersucht.

Des weiteren werden erste numerische Ergebnisse für die allgemeine modellprädiktiven Kontrolle vorgestellt.