Weierstraß-Institut
für Angewandte Analysis und Stochastik
Leibniz-Institut im Forschungsverbund Berlin e. V.

Preprint ISSN 2198-5855

A goal-oriented dual-weighted adaptive finite element approach for the optimal control of a nonsmooth Cahn–Hilliard–Navier–Stokes system

Michael Hintermüller, [1,3] Michael Hinze, [2] Christian Kahle,[2] Tobias Keil [3]

submitted: September 30, 2016

[1] Weierstrass Institute
Mohrenstr. 39
10117 Berlin
Germany

[2] Department of Mathematics
Universität Hamburg
Bundesstr. 55
20146 Hamburg
Germany
E-Mail: michael.hinze@uni-hamburg.de
christian.kahle@uni-hamburg.de

[3] Department of Mathematics
Humboldt-Universität zu Berlin
Unter den Linden 6
10099 Berlin
Germany
E-Mail: hint@math.hu-berlin.de
tkeil@math.hu-berlin.de

No. 2311

Berlin 2016

2010 *Mathematics Subject Classification.* 49K20, 49M25, 65K15, 76T10, 90C33.

Key words and phrases. Cahn-Hilliard, C-stationarity, mathematical programming with equilibrium constraints, Navier-Stokes, non-matched densities, non-smooth potentials, optimal control, adaptive finite element method, goal-oriented error estimation.

Acknowledgments. This research was supported by the German Research Foundation DFG through the SPP 1506 and the SPP 1962 and by the Research Center MATHEON through project C-SE5 funded by the Einstein Center for Mathematics Berlin.

Edited by
Weierstraß-Institut für Angewandte Analysis und Stochastik (WIAS)
Leibniz-Institut im Forschungsverbund Berlin e. V.
Mohrenstraße 39
10117 Berlin
Germany

Fax: +49 30 20372-303
E-Mail: preprint@wias-berlin.de
World Wide Web: http://www.wias-berlin.de/

Abstract

This paper is concerned with the development and implementation of an adaptive solution algorithm for the optimal control of a time-discrete Cahn–Hilliard–Navier–Stokes system with variable densities. The free energy density associated to the Cahn-Hilliard system incorporates the double-obstacle potential which yields an optimal control problem for a family of coupled systems in each time instant of a variational inequality of fourth order and the Navier–Stokes equation. A dual-weighted residual approach for goal-oriented adaptive finite elements is presented which is based on the concept of C-stationarity. The overall error representation depends on primal residuals weighted by approximate dual quantities and vice versa as well as various complementarity mismatch errors. Details on the numerical realization of the adaptive concept and a report on numerical tests are given.

1 Introduction

In this paper we develop an efficient numerical solver for the optimal control of two-phase flows which includes an intelligent mesh refinement technique. More precisely, we consider a diffuse interface model of phase separation which involves a nonsmooth version of the well-known Cahn-Hilliard (CH) system, which is due to Cahn and Hilliard's seminal work [14]. Phase field models are appreciated for their ability to overcome both, analytical difficulties of topological changes, such as, e.g., droplet break-ups or the coalescence of interfaces, as well as numerical challenges in capturing the interface dynamics. In the presence of hydrodynamic effects, the CH system has to be enhanced by an equation which captures the behavior of the fluid. In [34], Hohenberg and Halperin introduced a basic model for immiscible, viscous two-phase flows. Their so-called 'model H' combines the Cahn-Hilliard system with the Navier-Stokes equation. It is, however, restricted to the case where the two fluids possess nearly identical densities, i.e., matched densities. Recently, Abels, Garcke and Grün [2] obtained the following diffuse interface model for two-phase flows with non-matched densities:

$$\partial_t \varphi + v\nabla\varphi - \mathrm{div}(m(\varphi)\nabla\mu) = 0, \tag{1.1a}$$

$$-\sigma\epsilon\Delta\varphi + \frac{\sigma}{\epsilon}(\partial\Psi_0(\varphi) - \tilde{\kappa}\varphi) - \mu = 0, \tag{1.1b}$$

$$\begin{aligned}\partial_t(\rho(\varphi)v) + \mathrm{div}(v\otimes\rho(\varphi)v) - \mathrm{div}(2\eta(\varphi)D_{sy}(v)) + \nabla p \\ +\mathrm{div}(v\otimes J) - \mu\nabla\varphi = 0,\end{aligned} \tag{1.1c}$$

$$\mathrm{div}v = 0, \tag{1.1d}$$

$$v_{|\partial\Omega} = 0, \tag{1.1e}$$

$$\partial_n\varphi_{|\partial\Omega} = \partial_n\mu_{|\partial\Omega} = 0, \tag{1.1f}$$

$$(v,\varphi)_{|t=0} = (v_a,\varphi_a). \tag{1.1g}$$

The system is considered in the space-time cylinder $\Omega \times (0, \infty)$, where $\partial\Omega$ denotes the boundary of Ω. It is thermodynamically consistent in the sense that it allows for the derivation of local entropy or free energy inequalities.

In the above model, v represents the velocity of the fluid and p describes the fluid pressure. The symmetric gradient of v is defined by $D_{sy}(v) := \frac{1}{2}(\nabla v + \nabla v^\top)$. The density ρ of the mixture of the fluids depends on the order parameter φ, which reflects the mass concentration of the fluid phases. More precisely,

$$\rho(\varphi) = \frac{\rho_1 + \rho_2}{2} + \frac{\rho_2 - \rho_1}{2}\varphi, \tag{1.2}$$

where φ ranges in the interval $[-1, 1]$, and $0 < \rho_1 \leq \rho_2$ are the given densities of the two fluids under consideration. The relative flux $J := -\frac{\rho_2 - \rho_1}{2} m(\varphi)\nabla\mu$, which corresponds to the diffusion of the two phases, involves the gradient of the chemical potential μ. The viscosity and mobility coefficients of the system, η and m, depend on the actual concentration of the two fluids at each point in time and space. The initial states are given by v_a and φ_a, and $\sigma, \epsilon, \tilde{\kappa} > 0$ are positive constants. Furthermore, Ψ_0 represents the convex part of the homogeneous free energy density Ψ contained in the Ginzburg-Landau energy model which is associated with the Cahn-Hilliard part of (1.1). Usually, the homogeneous free energy density serves the purpose of restricting the order parameter φ to the physically meaningful range $[-1, 1]$ and to capture the spinodal decomposition of the phases. For this reason, it is typically non-convex and maintains two local minima near or at -1 and 1.

Different choices have been investigated in the literature, depending on the underlying applications. In [42], Oono and Puri found that in the case of deep quenches of, e.g., binary alloys, the double-obstacle potential proves to be the best choice for modeling the separation process. A similar observation appears to be true in the case of polymeric membrane formation under rapid wall hardening. The double-obstacle potential $\Psi(\varphi) = I_{[-1,1]}(\varphi) - \frac{\tilde{\kappa}}{2}\varphi^2$, with $I_{[-1,1]}$ denoting the indicator function of the interval $[-1, 1]$ in $\mathbb{R}$, combines the advantages of the existence of pure phases and the exclusiveness of the interval $[-1, 1]$ at the cost of losing differentiability (when compared for instance to the double-well potential). The presence of a non-smooth homogeneous free energy density gives rise to a variational inequality in (1.1b) which complicates the analytical and numerical treatment of the overall model.

The Cahn-Hilliard-Navier-Stokes system is used to model a variety of situations. These range from the aforementioned solidification process of liquid metal alloys, cf. [17], the simulation of bubble dynamics, as in Taylor flows [4], or the pinch-offs of liquid-liquid jets [37], to the formation of polymeric membranes [50] and protein crystallization, see e.g. [38] and references within. Furthermore, the model can be easily adapted to include the effects of surfactants such as colloid particles at fluid-fluid interfaces in gels and emulsions used in food, pharmaceutical, cosmetic, or petroleum industries [5, 44]. In many of these situations an optimal control context is desirable in order to influence the system in such a way that a prescribed system behavior is guaranteed.

Therefore we investigate the optimal control of the coupled Cahn-Hilliard-Navier-Stokes (CHNS) system. We point out that, due to the presence of the variational inequality constraint, the mapping between the control and the state is in general not differentiable. As a consequence, classical constraint qualifications for optimal control problems (see, e.g., [51]) fail, preventing the application of the Karush-Kuhn-Tucker (KKT) theory in Banach space for a primal-dual first-order characterization of an optimal solution. In fact, it is

known [28, 31] that the resulting problem falls into the realm of mathematical programs with equilibrium constraints (MPECs) in function space. A problem class, which even in finite dimensions, is well-known for its constraint degeneracy [40, 43]. As a result, stationarity conditions are no longer unique (in contrast to KKT conditions); compare [28, 29] in function space and, e.g., [46] in finite dimensions. Rather they depend on the underlying problem structure and/or on the chosen analytical approach.

Our work is based on the analytical results obtained in [27], where the problem has been discretized in time and a Yosida regularization technique yielding a sequence of approximating problems with a subsequent passage to the limit with the Yosida parameter has been utilized in order to derive stationarity conditions of C-stationarity type. In this paper, we develop and implement a solution algorithm based on the constructive nature of the former approach which solves each approximating problem by a Newton method applied to a suitable finite element discretization in space.

As the solution of a sequence of large-scale nonlinear optimization problems might cause an immense numerical expense, it is desirable to reduce the computational effort by choosing a beneficial adaptation process for the underlying space mesh. The general idea of adaptive finite element methods is to refine the discretization locally only in regions with large errors while keeping elements coarse wherever possible. This is especially useful in the context of variational inequalities where the analytical solution usually has a smooth structure on large parts of the domain, whereas it is often nonsmooth only in the small region where the active and the inactive sets meet. In the presence of an optimal control problem this approach can be modified. In our optimal control context we modify the method in order to guarantee an accurate evaluation of the objective functional. While this method has been successfully applied in PDE constrained optimization [8, 10, 22, 24, 25, 45, 49], to the best of our knowledge the literature concerning MPECs is rather scarce. However, recent work on adaptivity for elliptic MPECs indicates a good numerical behavior of these methods also for MPECs, cf. [13, 26].

Let us finally comment on further contributions to control and optimal control of Cahn–Hilliard Navier–Stokes systems. Model predictive control concepts for variable density Cahn–Hilliard Navier–Stokes systems are developed in [32, 33]. Optimal control of a related system is investigated in [9]. Phase field based shape and topology optimization concepts for flows are proposed and numerically implemented in [18]. This approach is extended to the minimization of surface functionals in [19].

The remainder of the paper is organized as follows. We start by formally introducing the optimal control problem under consideration and some additional concepts in section 2. This is followed by an explanation of the chosen discretization in space and the involved finite elements, respectively, in section 3. Section 4 provides a rigorous derivation of the goal-oriented error estimator. Finally, we present numerical results along with the details of the numerical implementation of the algorithm in section 5.

We end this introduction by defining some notation. Let $\Omega \subset \mathbb{R}^N, N = 2, 3$, be a bounded domain with smooth boundary $\partial\Omega \in C^2$. The smooth boundary ensures a higher regularity of the state in our subsequent analysis, cf. [27]. In our numerical tests, however, we observe that the subsequently developed algorithm achieves excellent results even for nonsmooth domains such as, e.g., the unit square.

We define the Sobolev spaces $H^1_{0,\sigma}(\Omega;\mathbb{R}^N) = \{f \in H^1_0(\Omega;\mathbb{R}^N) : \mathrm{div} f = 0, \text{ a.e. on } \Omega\}$ and $\overline{W}^{k,p}(\Omega) = \left\{f \in W^{k,p}(\Omega) : \int_\Omega f dx = 0\right\}$ for $k \in \mathbb{N}$ and $1 \leq p \leq \infty$, where 'a.e.' stands for 'almost everywhere'. Here, $W^{k,p}(\Omega)$ and $W^{k,p}_0(\Omega)$ denote the

usual Sobolev spaces, see [3]. For $p = 2$, we also write $H^k(\Omega)$ and $H_0^k(\Omega)$, respectively. By $(\cdot,\cdot)$ we denote the L^2-inner product, $\|\cdot\|$ is the induced norm, and $\langle\cdot,\cdot\rangle := \langle\cdot,\cdot\rangle_{\overline{H}^{-1},\overline{H}^1}$ represents the duality pairing between $\overline{H}^1(\Omega)$ and $\overline{H}^{-1}(\Omega)$. For a Banach space W, we denote by W^* its topological dual. In our notation for norms, we do not distinguish between scalar- or vector-valued functions. The inner product of vectors is denoted by '$\cdot$', the vector product is represented by '$\otimes$' and the tensor product for matrices is written as ':'.

2 The semi-discrete CHNS-system and the optimal control problem

In this paper, we study an optimal control problem for a semi-discrete variant of the Cahn-Hilliard-Navier-Stokes system. Concerning the mobility and viscosity coefficients, as well as the initial data for the velocity and the phase field parameter we invoke the following assumption.

Assumption 2.1. *1 The coefficient functions in (1.1a), (1.1c) satisfy $m, \eta \in C^2(\mathbb{R})$ and their derivatives up to second order are bounded, i.e. there exist constants $0 < b_1 \leq b_2$ such that for every $x \in \mathbb{R}$, it holds that $b_1 \leq \min\{m(x), \eta(x)\}$ and*

$$\max\{m(x), \eta(x), |m'(x)|, |\eta'(x)|, |m''(x)|, |\eta''(x)|\} \leq b_2.$$

2 The initial state satisfies $(v_a, \varphi_a) \in H^2_{0,\sigma}(\Omega;\mathbb{R}^N) \times \left(\overline{H}^2(\Omega) \cap \mathbb{K}\right)$ where

$$\mathbb{K} := \left\{ v \in \overline{H}^1(\Omega) : \psi_1 \leq v \leq \psi_2 \text{ a.e. in } \Omega \right\},\ \overline{\varphi_a} := \frac{1}{|\Omega|}\int_\Omega \varphi_a dx,$$

with $-1 - \overline{\varphi_a} =: \psi_1 < 0 < \psi_2 := 1 - \overline{\varphi_a}$.

3 The density ρ depends on the order parameter φ via

$$\rho(\varphi) = \max\left\{ \frac{\rho_1 + \rho_2}{2} + \frac{\rho_2 - \rho_1}{2}(\varphi + \overline{\varphi_a}), 0 \right\} \geq 0.$$

In the subsequent definition, $\tau > 0$ denotes the constant time step-size and $M \in \mathbb{N}$ the total number of equidistantly spaced time instances in the semi-discrete setting. For the sake of a simple notation, we further set $\sigma := \frac{1}{\epsilon}$ and $\kappa := \frac{\tilde{\kappa}}{\epsilon^2}$. Since the first equation of the Cahn-Hilliard system (1.1a) guarantees that the mean value of the order parameter remains constant, we consider a shifted Cahn-Hilliard-Navier-Stokes system where the mean value is set to zero. For each time step $i = 0, .., M-2$ we consider an arbitrary control $u_i \in U$ in a Banach space U which acts on the right-hand side of the Navier-Stokes equation via a bounded linear operator $B : U \to L^2(\Omega;\mathbb{R}^N)$. Note that U can be chosen as $L^2(\Omega;\mathbb{R}^N)$ and B as the identity operator. In our numerical tests, however, we choose a finite dimensional control, since the number of control parameters is usually limited in praxis, see Section 6.

Defintion 2.2 (Semi-discrete CHNS-system)**.** *Let $\Psi_0 : \overline{H}^1(\Omega) \to \mathbb{R}$ be the convex part of the double-obstacle potential with subdifferential $\partial\Psi_0$, i.e. the indicator function of $\mathbb{K}$. Fixing $(\varphi_{-1}, v_0) = (\varphi_a, v_a)$ we say that a triple*

$$(\varphi, \mu, v) = ((\varphi^i)_{i=0}^{M-1}, (\mu^i)_{i=0}^{M-1}, (v^i)_{i=1}^{M-1})$$

in $\overline{H}^1(\Omega)^M \times \overline{H}^1(\Omega)^M \times H^1_{0,\sigma}(\Omega;\mathbb{R}^N)^{M-1}$ *solves the semi-discrete CHNS system with respect to a given control* $u \in U^{M-1}$*, i.e.* $(\varphi,\mu,v) \in S_\Psi(u)$*, if there exists* $a^{i+1} \in \partial\Psi_0(\varphi^{i+1})$ *such that for all* $\phi \in \overline{H}^1(\Omega)$ *and* $\psi \in H^1_{0,\sigma}(\Omega;\mathbb{R}^N)$ *it holds that*

$$\left\langle \frac{\varphi^{i+1}-\varphi^i}{\tau}, \phi \right\rangle + \left\langle v^{i+1}\nabla\varphi^i, \phi \right\rangle + \left\langle m(\varphi^i)\nabla\mu^{i+1}, \nabla\phi \right\rangle = 0, \tag{2.1}$$

$$\left\langle \nabla\varphi^{i+1}, \nabla\phi \right\rangle + \left\langle a^{i+1}, \phi \right\rangle - \left\langle \mu^{i+1}, \phi \right\rangle - \left\langle \kappa\varphi^i, \phi \right\rangle = 0, \tag{2.2}$$

$$\begin{aligned} &\left\langle \frac{\rho(\varphi^i)v^{i+1}-\rho(\varphi^{i+1})v^i}{\tau}, \psi \right\rangle_{H^{-1},H^1_0} - \left\langle v^{i+1}\otimes\rho(\varphi^{i+1})v^i, \nabla\psi \right\rangle_{H^{-1},H^1_0} \\ &+ \left\langle v^{i+1}\otimes\frac{\rho_2-\rho_1}{2}m(\varphi^{i+1})\nabla\mu^i, \nabla\psi \right\rangle_{H^{-1},H^1_0} + (2\eta(\varphi^i)D_{sy}(v^{i+1}), D_{sy}(\psi)) \\ &- \left\langle \mu^{i+1}\nabla\varphi^i, \psi \right\rangle_{H^{-1},H^1_0} = \left\langle Bu^{i+1}, \psi \right\rangle_{H^{-1},H^1_0}. \end{aligned} \tag{2.3}$$

The first two equations are supposed to hold for every $0 \le i+1 \le M-1$ *and the last equation holds for every* $1 \le i+1 \le M-1$.

Remark 2.3. *For a more detailed explanation of the above assumptions, we refer to [27], where the problem was originally formulated.*

Introducing a Fréchet differentiable objective function $J : \mathcal{X} \to \mathbb{R}$ with

$$\mathcal{X} := \overline{H}^1(\Omega)^M \times \overline{H}^1(\Omega)^M \times H^1_{0,\sigma}(\Omega;\mathbb{R}^N)^{M-1} \times U^{M-1},$$

the optimal control problem reads

$$\begin{aligned} &\text{min } J(\varphi,\mu,v,u) \text{ over } (\varphi,\mu,v,u) \in \mathcal{X} \\ &\text{subject to (s.t.) } (\varphi,\mu,v) \in S_\Psi(u). \end{aligned} \tag{P_Ψ}$$

Although our subsequent analysis can be applied to this general setting, for the purpose of numerical realization we will later consider the case where J equals the following tracking-type function

$$J(\varphi,\mu,v,u) := \frac{1}{2}\|\varphi_M - \varphi_d\|^2 + \frac{\nu}{2}\|u\|^2. \tag{2.4}$$

Here, $\varphi_d \in H^1(\Omega)^M$ is a given desired state of the system.

In [27], it has been shown that the semi-discrete Cahn-Hilliard-Navier-Stokes system possesses a solution $(\varphi,\mu,v) \in S_\Psi(u)$ for every $Bu \in L^2(\Omega;\mathbb{R}^N)^{M-1}$. Furthermore, it was verified that the optimal control problem (P_Ψ) admits a solution. It can be characterized by the following stationarity conditions of $\mathcal{E}$-almost C-stationary type. Here, the notion of '$\mathcal{E}$-almost' is due to an application of Egorov's theorem; see [28].

Defintion 2.4. *A point* $(\varphi,\mu,v,a,u,p,r,q,\pi,\lambda^+,\lambda^-) \in \mathcal{Y}$ *with*

$$\begin{aligned} \mathcal{Y} :=& \overline{H}^1(\Omega)^M \times \overline{H}^1(\Omega)^M \times H^1_{0,\sigma}(\Omega;\mathbb{R}^N)^{M-1} \times \overline{L}^2(\Omega)^M \times U^{M-1} \times \overline{H}^1(\Omega)^M \\ &\times \overline{H}^1(\Omega)^M \times H^1_{0,\sigma}(\Omega;\mathbb{R}^N)^{M-1} \times \overline{H}^1(\Omega)^M \times \left(\overline{H}^1(\Omega)^*\right)^M \times \left(\overline{H}^1(\Omega)^*\right)^M \end{aligned}$$

is called $\mathcal{E}$*-almost C-stationary for* (P_Ψ) *if it satisfies the following system:*

Feasibility:

$$\left\langle \frac{\varphi^{i+1}-\varphi^i}{\tau}, \phi \right\rangle + \left\langle v^{i+1}\nabla\varphi^i, \phi\right\rangle + \left\langle m(\varphi^i)\nabla\mu^{i+1}, \nabla\phi\right\rangle = 0, \tag{2.5}$$

$$\left\langle \nabla\varphi^{i+1}, \nabla\phi\right\rangle + \left\langle a^{i+1}, \phi\right\rangle - \left\langle \mu^{i+1}, \phi\right\rangle - \left\langle \kappa\varphi^i, \phi\right\rangle = 0, \tag{2.6}$$

$$\begin{aligned}\left\langle \frac{\rho(\varphi^i)v^{i+1}-\rho(\varphi^{i+1})v^i}{\tau}, \psi \right\rangle_{H^{-1},H_0^1} - \left\langle v^{i+1}\otimes\rho(\varphi^{i+1})v^i, \nabla\psi\right\rangle_{H^{-1},H_0^1} \\ + \left\langle v^{i+1}\otimes\frac{\rho_2-\rho_1}{2}m(\varphi^{i+1})\nabla\mu^i, \nabla\psi\right\rangle_{H^{-1},H_0^1} + (2\eta(\varphi^i)D_{sy}(v^{i+1}), D_{sy}(\psi)) \\ - \left\langle \mu^{i+1}\nabla\varphi^i, \psi\right\rangle_{H^{-1},H_0^1} = \left\langle Bu^{i+1}, \psi\right\rangle_{H^{-1},H_0^1},\end{aligned} \tag{2.7}$$

$$\psi_1 \le \varphi^{i+1} \le \psi_2 \text{ a.e. on } \Omega, \tag{2.8}$$

$$(a^{i+1})^+ := \max(a^{i+1}, 0) \ge 0,\ (a^{i+1})^- := \max(-a^{i+1}, 0) \ge 0 \text{ a.e. on } \Omega, \tag{2.9}$$

$$\left\langle (a^{i+1})^+, \varphi^{i+1}-\psi_2\right\rangle = 0, \left\langle (a^{i+1})^-, \varphi^{i+1}-\psi_1\right\rangle = 0. \tag{2.10}$$

Adjoint system:

$$\begin{aligned}-\frac{1}{\tau}(p^{i+1}-p^i) + m'(\varphi^i)\nabla\mu^{i+1}\cdot\nabla p^{i+1} - \operatorname{div}(p^{i+1}v^{i+1}) - \Delta r^i \\ + \lambda^{i+1} - \kappa r^{i+2} - \frac{1}{\tau}\rho(\varphi^i)'v^{i+1}\cdot(q^{i+2}-q^{i+1}) \\ - (\rho(\varphi^i)'v^{i+1} - \frac{\rho_2-\rho_1}{2}m'(\varphi^i)\nabla\mu^{i+1})(Dq^{i+2})^\top v^{i+2} \\ + 2\eta(\varphi^i)'D_{sy}(v^{i+1}) : Dq^{i+1} + \operatorname{div}(\mu^{i+1}q^{i+1}) = \frac{\partial J}{\partial\varphi^i}(z),\end{aligned} \tag{2.11}$$

$$\begin{aligned}- r^i - \operatorname{div}(m(\varphi^{i+1})\nabla p^i) - \operatorname{div}(\frac{\rho_2-\rho_1}{2}m(\varphi^{i+1})(Dq^{i+1})^\top v^{i+1}) \\ - q^i\cdot\nabla\varphi^{i+1} = \frac{\partial J}{\partial\mu^i}(z),\end{aligned} \tag{2.12}$$

$$\begin{aligned}- \frac{1}{\tau}\rho(\varphi^{i+1})(q^{i+1}-q^i) - \rho(\varphi^{i+1})(Dq^{i+1})^\top v^{i+1} \\ - (Dq^i)(\rho(\varphi^{i-2})v^{i+1} - \frac{\rho_2-\rho_1}{2}m(\varphi^{i-2})\nabla\mu^{i+1}) \\ - \operatorname{div}(2\eta(\varphi^{i+1})D_{sy}(q^i)) + p^i\nabla\varphi^{i+1} = \frac{\partial J}{\partial v^i}(z),\end{aligned} \tag{2.13}$$

$$B^\star q^{i+1} = \frac{\partial J}{\partial u^i}(z), \tag{2.14}$$

$$r^i = \pi^i. \tag{2.15}$$

Complementarity conditions:

$$\left\langle (\lambda^i)^+, \varphi^i-\psi_2\right\rangle = 0, \left\langle (\lambda^i)^-, \varphi^i-\psi_1\right\rangle = 0, \tag{2.16}$$

$$\left\langle a^i, \pi^i\right\rangle = 0. \tag{2.17}$$

Moreover, for every $c > 0$ there exist a measurable subset M_c^i of $M^i := \{x \in \Omega : \psi_1 < \varphi^i(x) < \psi_2\}$ with $|M^i \setminus M_c^i| < c$ and

$$\langle \lambda^i, v\rangle = 0 \qquad \forall v \in \overline{H}^1(\Omega),\ v|_{\Omega\setminus M_c^i} = 0.$$

Remark 2.5. *Note that for the previous definition $(\lambda^i)^+$ and $(\lambda^i)^-$ have to be defined slightly differently than in [27]. More precisely, we define $(\lambda^i)^+ \in \overline{H}^1(\Omega)^*$ as the weak limit of $\left(\Psi_{0(m)}{}''\right)^+ (\varphi_{(m)}^{i+1})^* r_{(m)}^i$ in $\overline{H}^1(\Omega)^*$ and $(\lambda^i)^- \in \overline{H}^1(\Omega)^*$ as the weak limit of $\left(\Psi_{0(m)}{}''\right)^- (\varphi_{(m)}^{i+1})^* r_{(m)}^i$ in $\overline{H}^1(\Omega)^*$.*

Here, for a function f we define $(f)^+$ and $(f)^-$ as follows

$$(f)^+(x) := \begin{cases} 0 & \text{if } x \le 0 \\ f(x) & \text{if } x > 0 \end{cases}, (f)^-(x) := \begin{cases} f(x) & \text{if } x < 0 \\ 0 & \text{if } x \ge 0 \end{cases}.$$

By this definition it holds that $\lambda^i = (\lambda^i)^+ + (\lambda^i)^-$. Furthermore, we introduced the artificial variable π which allows us to fix a quantity in the complementarity mismatch in our subsequent error analysis, cf. Lemma 4.3 and Theorem 4.4.

3 Discretization of the problem

In order to treat the problem (P_Ψ) numerically, an additional discretization step is necessary. Here we follow the so called first optimize, then discretize approach in that we provide a discretization of the optimality conditions (2.5)–(2.17). For this let $(\mathcal{T}^i)_{i=0}^{M-1} = (\bigcup_{k=1}^{nt} T_k^i)_{i=0}^{M-1}$ denote a sequence of regular triangulations of Ω, cf. [12, Def. 4.4.13], such that $\mathcal{T}^i = \overline{\Omega}$, for $i = 0, \dots, M-1$, and such that the L^2-projection is stable in H^1, cf. [11]. On $\mathcal{T}^i$ we consider finite dimensional finite element subspaces

$$\begin{aligned} V_1^i &:= \{v \in C(\mathcal{T}^i) \,|\, v|_{T_k^i} \in P^1(T_k^i),\ k = 1, \dots, nt\} \\ &= \text{span}\{\phi_1^i, \dots, \phi_{N_1^i}^i\} \subset H^1(\Omega), \\ V_2^i &:= \{v \in C(\mathcal{T}^i)^N \,|\, v|_{\partial\Omega} = 0,\ v|_{T_k^i} \in P^2(T_k^i)^N,\ k = 1, \dots, nt\} \\ &= \text{span}\{\psi_1^i, \dots, \psi_{N_1^i}^i\} \subset H_0^1(\Omega)^N. \end{aligned}$$

We denote the fully discrete counterpart to a solution $(\varphi^i, \mu^i, v^i) \in H^1(\Omega) \times H^1(\Omega) \times H_0^1(\Omega, \mathbb{R}^N)$ of (2.1)–(2.3) by $(\varphi_h^i, \mu_h^i, v_h^i) \in V_1^i \times V_1^i \times V_2^i$. We note that we do not incorporate the solenoidality condition on the velocity into the discrete ansatz space, but introduce an additional pressure variable $\xi_h^i \in V_1^i$ and an adjoint pressure variable $\chi_h^i \in V_1^i$. For such a setting, the pair (V_2^i, V_1^i) is LBB-stable and thus admissible for this numerical realization of (2.3), cf., e.g., [21, 48].

The discrete variant of (2.1)–(2.3) we define as follows:

For $i = 0, \dots, M-1$ find, $(\varphi_h^{i+1}, \mu_h^{i+1}, a_h^{i+1}, v_h^{i+1}, \xi_h^{i+1}) \in (V_1^{i+1} \times V_1^{i+1} \times$

$V_1^{i+1} \times V_2^{i+1} \times V_1^{i+1}$) such that

$$\left\langle \frac{\varphi_h^{i+1} - \Pi^{i+1}\varphi_h^i}{\tau}, \phi \right\rangle + \left\langle v_h^{i+1}\nabla\varphi_h^i, \phi \right\rangle + \left\langle m(\varphi_h^i)\nabla\mu_h^{i+1}, \nabla\phi \right\rangle = 0 \quad \forall \phi \in V_1^{i+1}, \tag{3.1}$$

$$\left\langle \nabla\varphi_h^{i+1}, \nabla\phi \right\rangle + \left(a_h^{i+1}, \phi\right) - \left\langle \mu_h^{i+1}, \phi \right\rangle - \left\langle \kappa\Pi^{i+1}\varphi_h^i, \phi \right\rangle = 0 \quad \forall \phi \in V_1^{i+1}, \tag{3.2}$$

$$\frac{1}{\tau}\left\langle \rho(\varphi_h^i)v_h^{i+1} - \rho(\varphi_h^{i-1})v_h^i, \psi \right\rangle - \left\langle (t_h^i\nabla)\psi, v_h^{i+1} \right\rangle + \left\langle 2\eta(\varphi_h^i)D_{sy}(v_h^{i+1}), D_{sy}(\psi) \right\rangle$$
$$- \left\langle \mu_h^{i+1}\nabla\varphi_h^i, \psi \right\rangle_{H^{-1},H_0^1} - \left\langle \mathsf{div}\psi, \xi_h^i \right\rangle = (Bu^{i+1}, \psi) \quad \forall \psi \in V_2^{i+1}, \tag{3.3}$$

$$- (\mathsf{div} v_h^{i+1}, \phi) = 0, \tag{3.4}$$

with $t_h^i := \rho(\varphi_h^{i-1})v_h^i - \frac{\rho_2-\rho_1}{2}m(\varphi_h^{i-1})\nabla\mu_h^i$, together with the complementarity conditions for the Cahn–Hilliard problem:

$$\psi_1 \le \varphi_h^{i+1} \le \psi_2 \text{ a.e. on } \Omega, \tag{3.5}$$

$$(a_h^{i+1})^+ := \max(a_h^{i+1}, 0) \ge 0, (a_h^{i+1})^- := \max(-a_h^{i+1}, 0) \ge 0, \tag{3.6}$$

$$((a_h^{i+1})^+, \varphi_h^{i+1} - \psi_2) = 0, \quad ((a_h^{i+1})^-, \varphi_h^{i+1} - \psi_1) = 0. \tag{3.7}$$

Here and in the following $\Pi^{i+1} : L^2(\Omega) \to V_1^{i+1}$ denotes the orthogonal L^2 projection which is required for stability reasons, compare [20, 27].

Further, $\max(a_h^{i+1}, 0)$ and $\max(-a_h^{i+1}, 0)$ are understood pointwise in the nodes of $\mathcal{T}^{i+1}$, and $z = (\varphi, \mu, v, u)$.

The fully discrete counterpart to (2.5)–(2.17) is then defined as follows.

Defintion 3.1. *Let $\varphi_h^{-1} = \Pi_{H^1}(\varphi_a)$, $v_h^0 = \Pi_L(v_a)$ be given, where Π_{H^1} denotes the H^1 projection onto $\mathcal{T}^0$, while Π_L denotes the Leray projection ([15]) onto $\mathcal{T}^0$.*

We say that

$$(\varphi_h^i, \mu_h^i) \in (V_1^i)_{i=0}^{M-1} \times (V_1^i)_{i=0}^{M-1}, (v_h^i, \xi_h^i) \in (V_2^i)_{i=1}^{M-1} \times (V_1^i)_{i=1}^{M-1},$$
$$(a_h^i) \in \left(V_1^i\right)_{i=0}^{M-1}, (u^i) \in (U^i)_{i=1}^{M-1},$$
$$(p_h^i, r_h^i) \in (V_1^i)_{i=0}^{M-1} \times (V_1^i)_{i=0}^{M-1}, (q_h^i, \chi_h^i) \in (V_2^i)_{i=1}^{M-1} \times (V_1^i)_{i=1}^{M-1},$$
$$(\pi_h^i, \lambda_h^i) \in (V_1^i)_{i=0}^{M-1} \times \left(V_1^i\right)_{i=0}^{M-1}$$

is discrete C-stationary for (P_Ψ), if it satisfies (3.1)-(3.7) *together with the following equations:*

Adjoint system:

$$-\frac{1}{\tau}(\langle(p_h^{i+1},\Pi^{i+1}\phi\rangle-\langle p_h^i,\phi\rangle)+(m'(\varphi_h^i)\nabla\mu_h^{i+1}\cdot\nabla p_h^{i+1},\phi) \tag{3.8}$$

$$+\langle p_h^{i+1}v_h^{i+1},\nabla\phi\rangle+\langle\nabla r_h^i,\nabla\phi\rangle+(\lambda_h^{i-1},\phi)-\langle\kappa r_h^{i+1},\Pi^{i+1}\phi\rangle \tag{3.9}$$

$$-\left\langle\frac{1}{\tau}\rho'(\varphi_h^i)v_h^{i+1}(q_h^{i+2}-q_h^{i+1}),\phi\right\rangle \tag{3.10}$$

$$-\left\langle\left(\rho(\varphi_h^i)'v_h^{i+1}-\frac{\rho_2-\rho_1}{2}m'(\varphi_h^i)\nabla\mu_h^{i+1}\right)(Dq_h^{i+2})^\top v_h^{i+2},\phi\right\rangle \tag{3.11}$$

$$+\langle 2\eta(\varphi_h^i)'D_{sy}(v_h^{i+1}):Dq_h^{i+1},\phi\rangle-\langle\mu_h^{i+1}q_h^{i+1},\nabla\phi\rangle=\left\langle\frac{\partial J}{\partial\varphi_h^i}(z),\phi\right\rangle \tag{3.12}$$

$$-\langle r_h^i,\phi\rangle+\langle m(\varphi_h^{i-1})\nabla p_h^i,\nabla\phi\rangle \tag{3.13}$$

$$+\left\langle\frac{\rho_2-\rho_1}{2}m(\varphi_h^{i-1})(Dq_h^{i+1})^\top v_h^{i+1},\nabla\phi\right\rangle-(q_h^i\nabla\varphi_h^{i-1},\phi)=\left\langle\frac{\delta J}{\delta\mu_h^i}(z),\phi\right\rangle \tag{3.14}$$

$$-\frac{1}{\tau}\langle\rho(\varphi_h^{i-1})(q_h^{i+1}-q_h^i),\psi\rangle-\left\langle\rho(\varphi_h^{i-1})(Dq_h^{i+1})^\top v_h^{i+1},\psi\right\rangle \tag{3.15}$$

$$-\left\langle(Dq_h^i)(\rho(\varphi_h^{i-2})v_h^{i-1}-\frac{\rho_2-\rho_1}{2}m(\varphi_h^{i-2})\nabla\mu_h^{i-1}),\psi\right\rangle \tag{3.16}$$

$$+\langle 2\eta(\varphi_h^{i-1})D_{sy}(q_h^i),\nabla\psi\rangle+\langle p_h^i\nabla\varphi_h^{i-1},\psi\rangle-\langle\chi_h^i,\mathsf{div}\,\psi\rangle=\left\langle\frac{\partial J}{\partial v_h^i}(z),\psi\right\rangle \tag{3.17}$$

$$-\langle\mathsf{div}\,q_h^i,\psi\rangle=0, \tag{3.18}$$

$$\langle B^\star q_h^{i-1},\tilde{u}\rangle_{U^\star,U}=\left\langle\frac{\partial J}{\partial u^i}(z),\tilde{u}\right\rangle, \tag{3.19}$$

$$r_h^i=\pi_h^i. \tag{3.20}$$

Complementarity conditions:

$$((\lambda_h^i)^+,\varphi_h^i-\psi_2)=0,\quad((\lambda_h^i)^-,\varphi_h^i-\psi_1)=0, \tag{3.21}$$

$$(a_h^i,\pi_h^i)=0, \tag{3.22}$$

We note that the prolongation operator Π^{i+1} is applied to the test function in the adjoint equation.

We further define $(\lambda_h^i)^+$ and $(\lambda_h^i)^-$ as in Remark 2.5 and nodewise in the nodes x_j of $\mathcal{T}^i$ as

$$(\lambda_h^i)^+(x_j):=\begin{cases}\lambda_h^i & \text{if } \varphi_h^i(x_j)>0,\\ 0 & \text{else},\end{cases}\qquad(\lambda_h^i)^-(x_j):=\begin{cases}\lambda_h^i & \text{if } \varphi_h^i(x_j)<0,\\ 0 & \text{else}.\end{cases}$$

Remark 3.2. *We point out that '$\mathcal{E}$-almost' C-stationarity is an infinite dimensional concept which corresponds to the notion of C-stationarity in finite dimensions.*

4 Goal-oriented error estimator

This section is devoted to the derivation of an error estimator which is used to design an appropriate refinement technique for the space mesh. The derivation is based on the dual-weighted residual method which was already successfully transferred to the optimal control of an elliptic variational inequality in [26]. For this purpose, we start by defining the MPCC-Lagrangian of (P_Ψ) as follows:

Defintion 4.1. *The MPCC-Lagrangian $L : \mathcal{Y} \to \mathbb{R}$ corresponding to (P_Ψ) is given by*

$$\begin{aligned}
L(\varphi,\mu,v,a,u,p,r,q,\pi,\lambda^+,\lambda^-) &:= J(\varphi,\mu,v,u)\\
&+\sum_{i=-1}^{M-2}\left[\left\langle\frac{\varphi^{i+1}-\varphi^i}{\tau},p^{i+1}\right\rangle+\left\langle v^{i+1}\nabla\varphi^i,p^{i+1}\right\rangle+\left\langle m(\varphi^i)\nabla\mu^{i+1},\nabla p^{i+1}\right\rangle\right]\\
&+\sum_{i=-1}^{M-2}\left[\left\langle-\Delta\varphi^{i+1},r^{i+1}\right\rangle+\left\langle a^{i+1},r^{i+1}\right\rangle-\left\langle\mu^{i+1},r^{i+1}\right\rangle-\left\langle\kappa\varphi^i,r^{i+1}\right\rangle\right]\\
&+\sum_{i=0}^{M-2}\Bigg[\left\langle\frac{\rho(\varphi^i)v^{i+1}-\rho(\varphi^{i+1})v^i}{\tau},q^{i+1}\right\rangle-\left\langle v^{i+1}\otimes\rho(\varphi^{i+1})v^i,\nabla q^{i+1}\right\rangle_{H^{-1},H_0^1}\\
&+\left\langle v^{i+1}\otimes\frac{\rho_2-\rho_1}{2}m(\varphi^{i+1})\nabla\mu^i,\nabla q^{i+1}\right\rangle_{H^{-1},H_0^1}-\left\langle\mu^{i+1}\nabla\varphi^i,q^{i+1}\right\rangle_{H^{-1},H_0^1}\\
&+(2\eta(\varphi^i)D_{sy}(v^{i+1}),D_{sy}(q^{i+1}))-\left\langle Bu^{i+1},q^{i+1}\right\rangle_{H^{-1},H_0^1}\Bigg]\\
&-\sum_{i=0}^{M-1}\left\langle a^i,\pi^i\right\rangle-\sum_{i=0}^{M-1}\left\langle(\lambda^i)^+,\varphi^i-\psi_2\right\rangle-\sum_{i=0}^{M-1}\left\langle(\lambda^i)^-,\varphi^i-\psi_1\right\rangle.
\end{aligned}$$

For the sake of readability, we collect the primal variables in $y := (\varphi,\mu,a,v)$ which describes the state of the optimal control problem and the adjoint variables in $\Phi := (p,r,q)$. Furthermore, we define

$$\mathcal{Y}_h := (V_1^M)^3\times V_2^{M-1}\times U^{M-1}\times(V_1^M)^2\times V_2^{M-1}\times(V_1^M)^3. \tag{4.1}$$

Clearly, the MPCC-Lagrangian possesses the following saddle point property.

Remark 4.2. *If (y,u) is an $\mathcal{E}$-almost C-stationary point of (P_Ψ) with corresponding adjoints $(\Phi,\pi,\lambda^+,\lambda^-)$ then*

$$L(y,u,\Phi,\pi,\lambda^+,\lambda^-) = J(\varphi,\mu,v,u). \tag{4.2}$$

For an arbitrarily fixed triple $(\pi,\lambda^+,\lambda^-)$ the MPCC-Lagrangian $L(\,\cdot\,,\pi,\lambda^+,\lambda^-)$ is infinitely Gâteaux differentiable with respect to (y,u,Φ) and the corresponding second derivative is constant. This property gives rise to the subsequent lemma.

Lemma 4.3. *Let $(y_h,u_h,\Phi_h,\pi_h,\lambda_h^+,\lambda_h^-)\in\mathcal{Y}_h$ satisfy the discretized stationarity system derived in the previous section. Then for every point (y,u,Φ) it holds that*

$$\begin{aligned}
J(\varphi_h,\mu_h,v_h,u_h) =& L(y,u,\Phi,\pi_h,\lambda_h^+,\lambda_h^-)\\
&+\frac{1}{2}\nabla_x L(y,u,\Phi,\pi_h,\lambda_h^+,\lambda_h^-)((y_h,u_h,\Phi_h)-(y,u,\Phi))\\
&+\frac{1}{2}\nabla_x L(y_h,u_h,\Phi_h,\pi_h,\lambda_h^+,\lambda_h^-)((y_h,u_h,\Phi_h)-(y,u,\Phi)).
\end{aligned} \tag{4.3}$$

Proof. Since $(y_h, u_h, \Phi_h, \pi_h, \lambda_h^+, \lambda_h^-)$ is a discrete $\mathcal{E}$-almost C-stationary point, in particular due to equation (3.1)-(3.3), (3.18),(3.21) and (3.22), it holds that

$$J(\varphi_h, \mu_h, v_h, u_h) = L(y_h, u_h, \Phi_h, \pi_h, \lambda_h^+, \lambda_h^-). \tag{4.4}$$

Here, we additionally employed the fact that Π^{i+1} is the orthogonal projection onto V_1^{i+1}, i.e., $\left\langle \Pi^{i+1}\varphi_h^i, p_h^{i+1}\right\rangle = \left\langle \varphi_h^i, p_h^{i+1}\right\rangle$ for all $\varphi_h^i \in L^2(\Omega)$ and $p_h^{i+1} \in V_1^{i+1}$. Let X be a Banach space and $f : X \to \mathbb{R}$ be a twice Gâteaux differentiable function with constant second derivative. Applying Taylor's expansion at $x \in X$, then for an arbitrary $z \in X$ we derive $f(z) = f(x) + \nabla f(x)(z-x) + \frac{1}{2}\nabla^2 f(x)(z-x)^2$. Furthermore, the Taylor expansion of ∇f at x yields $\nabla f(z) = \nabla f(x) + \nabla^2 f(x)(z-x)$.

In summary, it holds that

$$\begin{aligned} f(z) &= f(x) + \nabla f(x)(z-x) + \frac{1}{2}(\nabla f(z) - \nabla f(x))(z-x) \\ &= f(x) + \frac{1}{2}\nabla f(x)(z-x) + \frac{1}{2}\nabla f(z)(z-x). \end{aligned}$$

Applying the last equation to $L(\,\cdot\,, \pi_h, \lambda_h^+, \lambda_h^-)$ with $x := (y, u, \Phi)$ and $z := (y_h, u_h, \Phi_h)$ proves the assertion. □

Using the previous lemma we present a first characterization of the difference of the objective values at stationary points of the semi-discrete and the fully discretized problem. Subsequently, the index δ denotes the difference of the discrete and the continuous variables, e.g. $(y_\delta, u_\delta, \Phi_\delta) := (y_h, u_h, \Phi_h) - (y, u, \Phi)$.

Theorem 4.4. *Let $(y_h, u_h, \Phi_h, \pi_h, \lambda_h^+, \lambda_h^-)$ be given as in Lemma 4.3 and $(y, u, \Phi, \pi, \lambda^+, \lambda^-)$ be a stationary point of the optimal control problem (P_Ψ). Then*

$$\begin{aligned} J(\varphi_h, \mu_h, v_h, u_h) - J(\varphi, \mu, v, u) = \frac{1}{2}&\left(\sum_{i=0}^{M-1}\left\langle a_h^i, \pi^i\right\rangle - \sum_{i=0}^{M-1}\left\langle a^i, \pi_h^i\right\rangle\right) \\ +\frac{1}{2}&\left(\sum_{i=0}^{M-1}\left\langle (\lambda^i)^+, \varphi_h^i - \psi_2\right\rangle - \sum_{i=0}^{M-1}\left\langle (\lambda_h^i)^+, \varphi^i - \psi_2\right\rangle\right) \\ +\frac{1}{2}&\left(\sum_{i=0}^{M-1}\left\langle (\lambda^i)^-, \varphi_h^i - \psi_1\right\rangle - \sum_{i=0}^{M-1}\left\langle (\lambda_h^i)^-, \varphi^i - \psi_1\right\rangle\right) \\ +\frac{1}{2}&\nabla_x L(y_h, u_h, \Phi_h, \pi_h, \lambda_h^+, \lambda_h^-)((y_h, u_h, \Phi_h) - (y, u, \Phi)) \end{aligned} \tag{4.5}$$

holds.

Proof. Since (y, u, Φ) is a stationary point, the gradient of the MPCC-Lagrangian with respect to a direction $(y_\delta, u_\delta, \Phi_\delta)$ reduces to

$$\begin{aligned} &\nabla_x L[y, u, \Phi, \pi_h, \lambda_h^+, \lambda_h^-](y_\delta, u_\delta, \Phi_\delta) \\ &\quad= \sum_{i=-1}^{M-2}\left\langle a_\delta^{i+1}, r^{i+1}\right\rangle - \sum_{i=-1}^{M-2}\left\langle (\lambda^{i+1})^+ + (\lambda^{i+1})^-, \varphi_\delta^{i+1}\right\rangle \\ &\qquad- \left(\sum_{i=0}^{M-1}\left\langle a_\delta^i, \pi_h^i\right\rangle + \sum_{i=0}^{M-1}\left\langle (\lambda_h^i)^+, \varphi_\delta^i\right\rangle + \sum_{i=0}^{M-1}\left\langle (\lambda_h^i)^-, \varphi_\delta^i\right\rangle\right) \\ &\qquad= \sum_{i=0}^{M-1}\left\langle a_\delta^i, \pi^i - \pi_h^i\right\rangle + \sum_{i=0}^{M-1}\left\langle (\lambda^i)^+ - (\lambda_h^i)^+, \varphi_\delta^i\right\rangle + \sum_{i=0}^{M-1}\left\langle (\lambda^i)^- - (\lambda_h^i)^-, \varphi_\delta^i\right\rangle. \end{aligned}$$

On the other hand, the feasibility of (y, u) implies that

$$\begin{aligned} L(y,u,\Phi,\pi_h,\lambda_h^+,\lambda_h^-) =& J(\varphi,\mu,v,u) - \sum_{i=0}^{M-1} \left\langle a^i, \pi_h^i \right\rangle \\ & - \sum_{i=0}^{M-1} \left\langle (\lambda_h^i)^+, \varphi^i - \psi_2 \right\rangle - \sum_{i=0}^{M-1} \left\langle (\lambda_h^i)^-, \varphi^i - \psi_1 \right\rangle. \end{aligned}$$

Inserting these equations into (4.3) leads to

$$\begin{aligned} J(\varphi_h,\mu_h,v_h,u_h) =& J(\varphi,\mu,v,u) - \sum_{i=0}^{M-1} \left\langle a^i, \pi_h^i \right\rangle - \sum_{i=0}^{M-1} \left\langle (\lambda_h^i)^+, \varphi^i - \psi_2 \right\rangle \\ & - \sum_{i=0}^{M-1} \left\langle (\lambda_h^i)^-, \varphi^i - \psi_1 \right\rangle + \frac{1}{2} \sum_{i=0}^{M-1} \left\langle a_h^i - a^i, \pi^i - \pi_h^i \right\rangle \\ & + \frac{1}{2} \sum_{i=0}^{M-1} \left\langle (\lambda^i)^+ - (\lambda_h^i)^+, \varphi_h^i - \varphi^i \right\rangle \\ & + \frac{1}{2} \sum_{i=0}^{M-1} \left\langle (\lambda^i)^- - (\lambda_h^i)^-, \varphi_h^i - \varphi^i \right\rangle \\ & + \frac{1}{2} \nabla_x L(y_h,u_h,\Phi_h,\pi_h,\lambda_h^+,\lambda_h^-)((y_h,u_h,\Phi_h) - (y,u,\Phi)). \end{aligned}$$

An appropriate rearrangement of the terms involving the complementarity conditions (2.16) and (2.17) yields the assertion. □

Remark 4.5. *Note that, since* $(y_h, u_h, \Phi_h, \pi_h, \lambda_h^+, \lambda_h^-)$ *satisfies the discrete stationarity system and taking into account the orthogonality of the projection* Π^{i+1}*, the direction* $(y_h, u_h, \Phi_h) - (y, u, \Phi)$ *in the last term of (4.5) can be replaced by any difference* $(y^\alpha, u^\alpha, \Phi^\alpha) - (y, u, \Phi)$ *involving arbitrary discrete variables* $(y^\alpha, u^\alpha, \Phi^\alpha) \in \mathcal{V}_1^2 \times \mathcal{V}_2 \times \mathcal{V}_1 \times \mathcal{V}_2 \times \mathcal{V}_1^2 \times \mathcal{V}_2$.

The last term on the right-hand side of equation (4.5) assembles the weighted dual and primal residuals, whereas the other terms display the mismatch in the complementarity between the discretized solution and the original one. For each time step $i \in \{0, .., M-1\}$, the mismatch is represented by the following four parts

$$\begin{aligned} \eta_{CM1,i} &:= \frac{1}{2} \left\langle a_h^i, \pi^i - \pi_h^i \right\rangle, \ \eta_{CM2,i} := \frac{1}{2} \left\langle (\lambda_h^i), \varphi^i - \varphi_h^i \right\rangle, \\ \eta_{CM3,i} &:= \frac{1}{2} \left\langle a^i, \pi_h^i - \pi^i \right\rangle, \ \eta_{CM4,i} := \frac{1}{2} \left\langle (\lambda^i)^+, \varphi_h^i - \psi_2 \right\rangle + \left\langle (\lambda^i)^-, \varphi_h^i - \psi_1 \right\rangle. \end{aligned}$$

Note that $\eta_{CM3,i}$ can be alternatively defined by

$$\begin{aligned} \eta_{CM3,i} = \left\langle a^i, \pi_h^i - \pi^i \right\rangle &= \left\langle \Delta\varphi^i + \mu^i + \kappa\varphi^{i+1}, \pi_h^i - \pi^i \right\rangle \\ &\approx \left\langle \Delta\varphi_h^i + \mu_h^i + \kappa\varphi_h^{i-1}, \pi_h^i - \pi^i \right\rangle. \end{aligned}$$

Next, we characterize the so-called dual-weighted primal residual $\eta_{CHNS,i} := \eta_{CH1,i} + \eta_{CH2,i} + \eta_{NS,i}$ by defining each of the three parts coming from the respective primal

equations (for $i = -1, .., M-2$)

$$\begin{aligned}
\eta_{CH1,i+1} &:= \left\langle \frac{\varphi_h^{i+1} - \varphi_h^i}{\tau}, p_\delta^{i+1} \right\rangle + \left\langle v_h^{i+1} \nabla \varphi_h^i, p_\delta^{i+1} \right\rangle + \left\langle m(\varphi_h^i) \nabla \mu_h^{i+1}, \nabla p_\delta^{i+1} \right\rangle, \\
\eta_{CH2,i+1} &:= \left\langle -\Delta \varphi_h^{i+1}, r_\delta^{i+1} \right\rangle + \left\langle a_h^{i+1}, r_\delta^{i+1} \right\rangle - \left\langle \mu_h^{i+1}, r_\delta^{i+1} \right\rangle - \left\langle \kappa \varphi_h^i, r_\delta^{i+1} \right\rangle, \\
\eta_{NS,i+1} &:= \left\langle \frac{\rho(\varphi_h^i) v_h^{i+1} - \rho(\varphi_h^{i-1}) v_h^i}{\tau}, q_\delta^{i+1} \right\rangle_{H^{-1}, H_0^1} \\
&\quad - \left\langle v_h^{i+1} \otimes \rho(\varphi_h^{i-1}) v_h^i, \nabla q_\delta^{i+1} \right\rangle_{H^{-1}, H_0^1} \\
&\quad - \left\langle v_h^{i+1} \otimes \frac{\rho_2 - \rho_1}{2} m(\varphi_h^{i-1}) \nabla \mu_h^i, \nabla q_\delta^{i+1} \right\rangle_{H^{-1}, H_0^1} \\
&\quad + (2\eta(\varphi_h^i) D_{sy}(v_h^{i+1}), D_{sy}(q_\delta^{i+1})) \\
&\quad - \left\langle \mu_h^{i+1} \nabla \varphi_h^i, q_\delta^{i+1} \right\rangle_{H^{-1}, H_0^1} - \left\langle B u_h^{i+1}, q_\delta^{i+1} \right\rangle_{H^{-1}, H_0^1}.
\end{aligned}$$

Since the Navier-Stokes equation is only defined for $i \in \{1, .., M-1\}$ based on the chosen discretization, we set $\eta_{NS,0} := 0$ for the sake of a brief notation.

In order to analyse the primal-weighted dual residual, we point out that the discrete stationary point satisfies

$$\left\langle \frac{\partial J}{\partial u^i}(\varphi_h, \mu_h, v_h, u_h) - B^\star q_h^i, u_\delta^i \right\rangle = 0. \tag{4.6}$$

Due to our specific choice of J in (2.4), the partial derivatives are given by

$$\frac{\partial J}{\partial \varphi^i}(\varphi_h, \mu_h, v_h, u_h) = \varphi_h^i - \varphi_d^i, \ \frac{\partial J}{\partial \mu^i}(\varphi_h, \mu_h, v_h, u_h) = 0, \ \frac{\partial J}{\partial v^i}(\varphi_h, \mu_h, v_h, u_h) = 0.$$

Incorporating the previous considerations, we define the primal-weighted dual residual in three parts for $i \in \{0, .., M-1\}$ (with $\eta_{ADv,0} := 0$) by

$$\begin{aligned}
\eta_{AD\varphi,i} := \Big[& \varphi_h^i - \varphi_d^i - \frac{1}{\tau}(p_h^{i+1} - p_h^i) + m'(\varphi_h^i) \nabla \mu_h^{i+1} \cdot \nabla p_h^{i+1} - \operatorname{div}(p_h^{i+1} v_h^{i+1}) \\
& - \Delta r_h^i + \lambda_h^i - \kappa r_h^{i+1} - \frac{1}{\tau} \rho'(\varphi_h^i) v_h^{i+1} \cdot (q_h^{i+2} - q_h^{i+1}) \\
& - (\rho'(\varphi_h^i) v_h^{i+1} - \frac{\rho_2 - \rho_1}{2} m'(\varphi_h^i) \nabla \mu_h^{i+1}) (D q_h^{i+2})^\top v_h^{i+2} \\
& + 2\eta'(\varphi_h^i) D_{sy}(v_h^{i+1}) : D q_h^{i+1} + \operatorname{div}(\mu_h^{i+1} q_h^{i+1}) \Big] (\varphi_\delta^i), \\
\eta_{AD\mu,i} := \Big[& - r_h^i - \operatorname{div}(m(\varphi_h^{i-1}) \nabla p_h^i) - \operatorname{div}(\frac{\rho_2 - \rho_1}{2} m(\varphi_h^{i-1}) (D q_h^{i+1})^\top v_h^{i+1}) \\
& - q_h^i \cdot \nabla \varphi_h^{i-1} \Big] (\mu_\delta^i), \\
\eta_{ADv,i} := \Big[& - \frac{1}{\tau} \rho(\varphi_h^{i-1}) (q_h^{i+1} - q_h^i) - \rho(\varphi_h^{i-1}) (D q_h^{i+1})^\top v_h^{i+1} \\
& - (D q_h^i) (\rho(\varphi_h^{i-2}) v_h^{i-1} - \frac{\rho_2 - \rho_1}{2} m(\varphi_h^{i-2}) \nabla \mu_h^{i-1}) \\
& - \operatorname{div}(2\eta(\varphi_h^{i-1}) D_{sy}(q_h^i)) + p_h^i \nabla \varphi_h^{i-1} \Big] (v_\delta^i).
\end{aligned}$$

By these definitions and Theorem 4.4, the discretization error with respect to the objective function is then given by

$$\begin{aligned} J(\varphi_h, \mu_h, v_h, u_h) - J(\varphi, \mu, v, u) \\ = \sum_{i=0}^{M-1} (\eta_{CM1,i} + \eta_{CM2,i} + \eta_{CM3,i} + \eta_{CM4,i} + \eta_{CH1,i} \\ + \eta_{CH2,i} + \eta_{NS,i} + \eta_{AD\varphi,i} + \eta_{AD\mu,i} + \eta_{ADv,i}). \end{aligned} \tag{4.7}$$

We point out that the integral structure of these error terms allows a patchwise evaluation on the underlying mesh.

Apart from the weights φ_δ^i, μ_δ^i and v_δ^i and p_δ^i, q_δ^i, r_δ^i, respectively, the primal-dual-weighted error estimators only contain discrete quantities. In order to obtain a fully a-posteriori error estimator the weights are approximated involving a local higher-order approximation based on the respective discrete variables.

5 The numerical realization

Next we describe how we employ the error representation (4.7) in order to find numerical approximations to solutions of the optimal control problem (P_Ψ). Our overall algorithm is based on solving an approximation of equations (3.1)–(3.22) for a given mesh sequence $(\mathcal{T}^1, \dots, \mathcal{T}^M)$, and then utilizing (4.7) to generate new grids that are better suited for representing the continuous optimal solution. The implementation is done in C++ using the finite element toolbox FEniCS [39] together with the PETSc linear algebra backend [7] and the linear solver MUMPS [6]. For the adaptation of the spatial meshes the toolbox ALBERTA [47] is used. Finite dimensional approximations of the minimization problem (P_Ψ) are solved by the steepest descent method from the GNU scientific library [1].

The relaxed equations

We introduce a smooth approximation of $\Psi_0(\varphi)$ using Moreau–Yosida relaxation, similar to [23], by

$$\Psi_0^s(\varphi) := \frac{s}{3} \left(|\max(0, \varphi - 1)|^3 + |\min(0, \varphi + 1)|^3 \right), \tag{5.1}$$

where $s \gg 0$ is a relaxation parameter. We point out that it is possible to use a Moreau-Yosida regularization combined with a semi-smooth Newton method here instead, as, e.g., in [28]. However, since we observe no singularities in our numerical tests and achieve a good approximation of feasibility already for moderate relaxation parameters, we choose the above approach for the ease if implementation.

Using Ψ_0^s we introduce the relaxed state equations (5.2)–(5.5)

$$\left\langle \frac{\varphi_h^{i+1} - \Pi^{i+1}\varphi_h^i}{\tau}, \phi \right\rangle + \left\langle v_h^{i+1}\nabla\varphi_h^i, \phi \right\rangle + \left\langle m(\varphi_h^i)\nabla\mu_h^{i+1}, \nabla\phi \right\rangle = 0,\ \forall \phi \in V_1^{i+1}, \tag{5.2}$$

$$\left\langle \nabla\varphi_h^{i+1}, \nabla\phi \right\rangle + ((\Psi_0^s)'(\varphi_h^{i+1}), \phi)_h - \left\langle \kappa\Pi^{i+1}\varphi_h^i, \phi \right\rangle - \left\langle \mu_h^{i+1}, \phi \right\rangle = 0,\ \forall \phi \in V_1^{i+1}, \tag{5.3}$$

$$\begin{aligned} &\frac{1}{\tau}\left\langle \rho(\varphi_h^i)v_h^{i+1} - \rho(\varphi_h^{i-1})v_h^i, \psi \right\rangle - \left\langle (t_h^i\nabla)\psi, v_h^{i+1} \right\rangle + \left\langle 2\eta(\varphi_h^i)D_{sy}(v_h^{i+1}), D_{sy}(\psi) \right\rangle \\ &- \left\langle \mu_h^{i+1}\nabla\varphi_h^i, \psi \right\rangle_{H^{-1},H_0^1} - \left\langle \mathsf{div}\psi, \xi_h^i \right\rangle = (Bu^{i+1}, \psi) \quad \forall \psi \in V_2^{i+1}, \end{aligned} \tag{5.4}$$

$$-(\mathsf{div}v_h^{i+1}, \phi) = 0 \quad \forall \phi \in V_1^{i+1}, \tag{5.5}$$

and the corresponding relaxed optimization problem (P_h^s).

$$\begin{aligned} &\min J(\varphi_h^i, \mu_h^i, v_h^i, u_h^i) \text{ over } (\varphi_h^i, \mu_h^i, v_h^i, u_h^i) \in (V_1^M)^2 \times V_2^{M-1} \times U^{M-1} \\ &s.t. \quad (5.2) - (5.5). \end{aligned} \tag{P_h^s}$$

Here, by $(\cdot,\cdot)_h$ we denote the lumped inner product

$$(f,g)_h := \int_\Omega I^i(fg)\,dx,$$

where I^i denotes the Lagrangian interpolation on V_1^i.

Remark 5.1. *The existence of feasible points for (P_h^s) and their boundedness with respect to u can be proven, e.g., by transferring the existence proof of Theorem 3.8 in [27] to the discretized problem. Since the equations* (5.2)–(5.5) *admit a unique solution for every control u, we can introduce the reduced functional $\hat{J}(u_h^i) := J(\varphi_h^i(u_h^i), \mu_h^i(u_h^i), v_h^i(u_h^i), u_h^i)$ and derive the existence of solutions and first order optimality conditions by standard arguments such as, e.g., in [51]. Consequently, we can apply a gradient descent method with respect to $\hat{J}$ in order so solve problem (P_h^s) numerically.*

Defintion 5.2. *Let $\varphi_h^{-1} = P(\varphi_a)$, $v_h^0 = L(v_a)$ be given, where P denotes the H^1 projection onto $\mathcal{T}^0$, while L denotes the Leray projection ([15]) onto $\mathcal{T}^0$.*

We say that

$$\begin{aligned} (\varphi_h^i, \mu_h^i) \in (V_1^i)_{i=0}^{M-1} \times (V_1^i)_{i=0}^{M-1}, (v_h^i, \xi_h^i) \in (V_2^i)_{i=1}^{M-1} \times (V_1^i)_{i=1}^{M-1}, \\ (u^i) \in (U^i)_{i=1}^{M-1}, \\ (p_h^i, r_h^i) \in (V_1^i)_{i=0}^{M-1} \times (V_1^i)_{i=0}^{M-1}, (q_h^i, \chi_h^i) \in (V_2^i)_{i=1}^{M-1} \times (V_1^i)_{i=1}^{M-1}, \end{aligned}$$

is a stationary point for (P_h^s)*, if it satisfies* (5.2)-(5.5) *together with the following equations:*

$$
\begin{aligned}
-\frac{1}{\tau}(\langle(p_h^{i+1},\Pi^{i+1}\phi\rangle - \langle p_h^i,\phi\rangle) + (m'(\varphi_h^i)\nabla\mu_h^{i+1}\cdot\nabla p_h^{i+1},\phi) + \langle p_h^{i+1}v_h^{i+1},\nabla\phi\rangle \\
+((\Psi_0^s)''(\varphi_h^i)r_h^i,\phi)_h - \langle \kappa r_h^{i+1},\Pi^{i+1}\phi\rangle - \left\langle \frac{1}{\tau}\rho'(\varphi_h^i)v_h^{i+1}(q_h^{i+2}-q_h^{i+1}),\phi\right\rangle \\
+\langle\nabla r_h^i,\nabla\phi\rangle - \left\langle\left(\rho(\varphi_h^i)'v_h^{i+1} - \frac{\rho_2-\rho_1}{2}m'(\varphi_h^i)\nabla\mu_h^{i+1}\right)(Dq_h^{i+2})^\top v_h^{i+2},\phi\right\rangle \\
+\langle 2\eta(\varphi_h^i)'D_{sy}(v_h^{i+1}):Dq_h^{i+1},\phi\rangle - \langle\mu_h^{i+1}q_h^{i+1},\nabla\phi\rangle = \left\langle\frac{\partial J}{\partial\varphi_h^i}(z),\phi\right\rangle \\
-\langle r_h^i,\phi\rangle + \langle m(\varphi_h^{i-1})\nabla p_h^i,\nabla\phi\rangle \\
+\left\langle\frac{\rho_2-\rho_1}{2}m(\varphi_h^{i-1})(Dq_h^{i+1})^\top v_h^{i+1},\nabla\phi\right\rangle - (q_h^i\nabla\varphi_h^{i-1},\phi) = \left\langle\frac{\delta J}{\delta\mu_h^i}(z),\phi\right\rangle \\
-\frac{1}{\tau}\langle\rho(\varphi_h^{i-1})(q_h^{i+1}-q_h^i),\psi\rangle - \left\langle\rho(\varphi_h^{i-1})(Dq_h^{i+1})^\top v_h^{i+1},\psi\right\rangle \\
-\left\langle(Dq_h^i)(\rho(\varphi_h^{i-2})v_h^{i-1} - \frac{\rho_2-\rho_1}{2}m(\varphi_h^{i-2})\nabla\mu_h^{i-1}),\psi\right\rangle \\
+\langle 2\eta(\varphi_h^{i-1})D_{sy}(q_h^i),\nabla\psi\rangle + \langle p_h^i\nabla\varphi_h^{i-1},\psi\rangle - \langle\chi_h^i,\mathsf{div}\ \psi\rangle = \left\langle\frac{\partial J}{\partial v_h^i}(z),\psi\right\rangle \\
-\langle\mathsf{div}\ q_h^i,\psi\rangle = 0, \\
\langle B^\star q_h^{i-1},\tilde u\rangle_{U^\star,U} = \left\langle\frac{\partial J}{\partial u^i}(z),\tilde u\right\rangle,
\end{aligned}
$$

Comparing the optimality system for (P_Ψ) and (P_h^s) and taking into account the convergence results from [27] it is reasonable to use the approximation (3.1)–(3.22)

$$
\begin{aligned}
(a_h^i,\phi) &\approx ((\Psi_0^s(\varphi_h^i))',\phi)_h,\\
(\lambda_h^i,\phi) &\approx ((\Psi_0^s(\varphi_h^i))''r_h^i,\phi)_h,
\end{aligned}
$$

also compare Remark 2.5.

Moreover by using s sufficiently large in (5.1) we guarantee, that the complementarity conditions (3.5)–(3.7) and (3.21)–(3.22) are sufficiently well fulfilled, when using these approximations for a and λ. For this we use the subsequent updating rule for the parameter s. In fact, in our numerical tests we observe that the complementarity conditions (3.21) and (3.22) are better fulfilled than (3.7) to at least 3 orders of magnitude. For this reason we base our update procedure in the following on (3.7) only, and next derive an estimate for the dependence of (3.7) with respect to s.

Exploiting the structure of Ψ_0^s we observe

$$
\begin{aligned}
&|(((\Psi_0^s(\varphi_h^i))',\varphi\pm 1)|\\
&\qquad\le \|(\Psi_0^s(\varphi_h^i))'\|_{L^1(\Omega)}\|\max(0,\varphi-1)+\min(0,\varphi+1)\|_{L^\infty(\Omega)},
\end{aligned}
$$

where we note, that in fact $\varphi\in L^\infty(\Omega)$ by elliptic regularity theory and Sobolev embeddings. Using $\Phi\equiv\pm 1$ in (5.3) we further observe that there exists $C>0$ independent of s such that

$$
\|(\Psi_0^s(\varphi_h^i))'\|_{L^1(\Omega)}\le C,
$$

and from [30, 36] we obtain

$$\| \max(0, \varphi - 1) + \min(0, \varphi + 1)\|_{L^\infty(\Omega)} \leq C s^{-1/2}$$

for this specific choice of Ψ_0^s. Together it holds

$$|(((\Psi_0^s(\varphi_h^i))', \varphi \pm 1)| \leq C s^{-1/2} \tag{5.6}$$

Let θ denote the maximum complementarity mismatch for (3.7) over all time instances. We find a new value s_{new} by estimating the unknown constant in (5.6) as $C = \theta s^{1/2}$, and, given a tolerance tol_c, we set $s_{new} := \left(\frac{C}{0.9 tol_c}\right)^2$, where the factor 0.9 is a guard to really get below the desired tolerance.

The local error indicator

Using these definitions we can evaluate the error indicators $\eta_\star$, as described in Section 4. For this it is important that the complementarity conditions (3.5)– (3.7) and (3.21)–(3.22) are sufficiently well approximated. We split the individual indicators canonically using the underlying meshes, e.g.

$$\eta_{CM1,i} = \sum_{T \in \mathcal{T}^i} \eta_{CM1,i}^T = \sum_{T \in \mathcal{T}^i} (\tilde{a}_h^{i+1}, \pi^i - \pi_h^i)_{|T}$$

and we define for each cell an indicator η_T^i as

$$\begin{aligned} \eta_T^i =& |\eta_{CM1,i}^T| + |\eta_{CM2,i}^T| + |\eta_{CM3,i}^T| + |\eta_{CM4,i}^T| \\ &+ |\eta_{CH1,i}^T| + |\eta_{CH2,i}^T| + |\eta_{NS,i}^T| \\ &+ |\eta_{AD\varphi,i}^T| + |\eta_{AD\mu,i}^T| + |\eta_{ADv,i}^T|. \end{aligned} \tag{5.7}$$

Note that the individual indicators might be negative, while we require a positive measure for the error, and thus sum up the absolute values of the individual terms.

We further set

$$(-\Delta \varphi_h^{i+1}, r_\delta^i) := \sum_{T \in \mathcal{T}^{i+1}} \left[(-\Delta \varphi_h^{i+1}, r_\delta^i)_{|T} + \sum_{E \subset T} \frac{1}{2} \left(\left[\nabla \varphi_h^{i+1} \right]_E, r_\delta^i \right)_{|E} \right],$$

compare, e.g., [26]. For an edge E contained in T the term $[f]_E$ denotes the jump of f across the edge E. More precisely, for a pair of cells T^+, T^- with $T^+ \cap T^- = E$ we define the jump as $[f(x_E)]_E := \left(\lim_{x \to x_E, x \in T^+} f(x) - \lim_{x \to x_E, x \in T^-} f(x)\right) \cdot \nu_{T^+,E}$, where $\nu_{T^+,E}$ is the unit normal on E pointing into T^+. Note that the definition of $[f]_E$ is independent of the permutation of T^+ and T^-.

The approximation of the continuous solution

Let us next discuss the approximation of the continuous solution $(y, u, \varphi, \mu, a, \lambda)$ using a higher-order finite element approximation of the discrete solution $(y_h, u_h, \varphi_h, \mu_h, a_h, \lambda_h)$.

For linear functions (i.e. φ, μ, p, r), we can use the procedure as described, e.g., in [26]. For a triangle T we use the nodes of the surrounding three triangles to define six points, with corresponding values of the finite element function under investigation. Then

we evaluate the unique quadratic polynomial that interpolates these six points, and use its restriction to T as quadratic finite element approximation to the continuous solution. If T lies on the boundary of Ω there are less than three surrounding triangles. Here we create virtual triangles by extending T as a parallelogram outside Ω, and we also extend the piecewise linear finite element function linearly on this virtual triangle to obtain again six points for the interpolation.

For quadratic finite elements we proceed analogously and evaluate a fourth-order polynomial on the given patch of cells, while for boundary cells we extend the given quadratic function as quadratic polynomial outside Ω.

For the multipliers a^i and λ^i we first calculate the representation $\tilde{a}_h^i$ and $\tilde{\lambda}_h^i$ in V_1^i and then use the extrapolation for linear elements as proposed above.

In any case we note that the resulting higher-order approximation is a trianglewise polynomial that is discontinuous across edges.

The final algorithm

Before stating the overall algorithm let us make the following simplification. Given φ_{-1} and v_0 on a sufficiently well resolved grid, we solve (3.1)–(3.2) to obtain φ_h^0 and μ_h^0 on the same grid and treat these functions as given data. Thus throughout the optimization process, the state that we optimize contains only the time instances $t_1, \dots, t_M$.

Including the Moreau–Yosida relaxation we use the following Algorithm 1.

Algorithm 1.

Data: Initial data: $\varphi_{-1}, \varphi_0, v_0, N_{\max}$

1 **repeat**
2 **for** $l = 1, \dots$ **do**
3 solve (P_h^s) using steepest descent method;
4 **if** complementarity conditions (3.7), (3.21), (3.22) are satisfied up to a tolerance tol_c **then**
5 break;
6 **else**
7 increase Moreau–Yosida parameter using (5.6);
8 **end**
9 **end**
10 calculate the error indicators and find the set $\mathcal{M}_r$ of cells to refine and the set $\mathcal{M}_c$ of cells to coarsen;
11 Adapt $(\mathcal{T}^i)_{i=1}^M$ using $\mathcal{M}_r$ and $\mathcal{M}_c$;
12 **until** $\sum_{i=1}^M |\mathcal{T}^i| < N_{\max}$;

Let us explain the steps of Algorithm 1 in detail.The outer loop describes the refinement of the grids $(\mathcal{T}^i)_{i=1}^M$ using (4.7). When the for-loop breaks, then we have found an approximate optimal control on the current sequence of grids that solves the system (3.1)–(3.22) sufficiently well, as it is required for our error indicators to be valid. Then, in line 10 we evaluate the error indicators η_T^i for all grids $\mathcal{T}^i$ and for all cells T and choose $\mathcal{M}_r$ as the

set with smallest cardinality such that

$$\sum_{T\in\mathcal{M}_r}\eta_T \geq \theta^r \sum_{i=1}^{M}\sum_{T\in\mathcal{T}^i}\eta_T$$

with a parameter $0<\theta^r<1$ using a greedy marking algorithm. We mark all cells in $\mathcal{M}_r$ for refinement. As in [23] we further choose $\theta^c\in(0,1)$ and define

$$M_c := \left\{ T\in(\mathcal{T}^i)_{i=1}^M \,|\, \eta_T \geq \frac{\theta^c}{N}\sum_{i=1}^{M}\sum_{T\in\mathcal{T}^i}\eta_T \right\},$$

where $N := \sum_{i=1}^{M}|\mathcal{T}^i|$. Thus, we use the well-known Dörfler marking procedure [16], where we refine a given proportion of the estimated error. We stress that we do not perform Dörfler marking on each time instance separately, but, as the representation (4.7) suggests, we perform a marking over all cells in the space-time cylinder. We mark cells for coarsening, if they contain an error that is smaller than θ^c times the mean error. We repeat this outer adaptation unless a given total amount of cells $N_{\max}$ is reached, summed over all cells, see line 12. We point out that we have to use a locally refined initial grid in order to get a meaningful initial resolution of the interface. This prevents us from using a very coarse grid initially. As a consequence, we also need to introduce a coarsening strategy.

The inner loop, i.e. lines 2–9, solves (P_h^s) using the steepest descent method from the GNU scientific library [1]. Thereafter we check whether the complementarity conditions are sufficiently well approximated by the current Moreau–Yosida relaxed system. For this we evaluate the terms (3.7), (3.21), (3.22) for all time instances. If the absolute value of all these terms is smaller then a given tolerance tol_c, we accept the solution and proceed with the adaptation step. If any of these terms has an absolute value larger than tol_c we increase the Moreau–Yosida parameter and solve the optimality problem again.

6 Numerical example: Splitting a bubble under gravity

Now we study a numerical example.Our aim is to prevent a bubble from rising and split it into two bubbles that are deformed to rounded squares.

The parameters of the fluid are $\rho_1 = 1000$, $\rho_2 = 100$, $\eta_1 = 10$, $\eta_2 = 1$, $g \equiv (0,-0.981)^\top$, and $\sigma = 24.5\cdot\frac{2}{\pi}$, where 24.5 is the physical surface tension and $\frac{2}{\pi}$ is a required scaling when using a phase field approximation with double-obstacle free energy density. These parameters arise from a benchmark for rising bubble dynamics in [35]. We further set $\epsilon = 0.02$ and $m(\varphi) = \epsilon/500$.

The initial phase field is given by a circle, located at $o = (0.5,0.5)^\top$ with radius $r = 0.2329$,

$$\varphi_0(x) = -1\cdot\begin{cases}\sin((\|x-o\|-r)/\epsilon) & \text{if } |\|x-o\|-r|/\epsilon \leq \pi/2,\\ \text{sign}(\|x-o\|-r) & \text{else.}\end{cases} \tag{6.1}$$

For the desired phase field we define

$$\varphi_d[Z,r](x) = \begin{cases}\sin((\|((x-Z)\|^6-r)/\epsilon) & \text{if } |\|(x-Z)\|-r|/\epsilon \leq \pi/2,\\ \text{sign}(\|(x-Z)\|^6-r) & \text{else,}\end{cases} \tag{6.2}$$

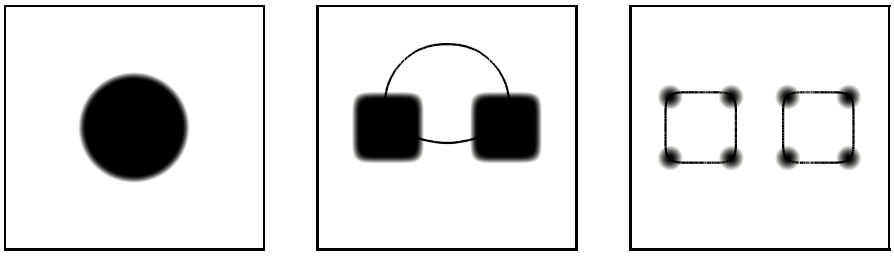

Figure 1: The initial shape φ_0, the desired shape φ_d together with the zero level line of the phase field at final time if no control is applied, the ansatz for the control, $|\sum_{i=1}^{16} f^i|$ together with the zero level line of φ_d (left to right).

which describes a square with smooth corners around Z with radius r and we use $\varphi_d(x) := \varphi_d[(0.25, 0.50)^\top, 0.15](x) \cdot \varphi_d[(0.75, 0.5)^\top, 0.15](x)$. The radius of φ_0 is chosen such that $\int_\Omega \varphi_d \, dx = \int_\Omega \varphi_0 \, dx$ since we have to respect mass conservation. We depict φ_0 and φ_d in Figure 1. Note that using the phase field approach we are not able to approach sharp corners and thus define φ_d with smooth corners.

To define Bu with the control operator $B : U \to L^2(\Omega, \mathbb{R}^N)$ we introduce the vector field

$$(f[o, \xi, c](x))^i = \begin{cases} \cos\left((\pi/2)\|\xi^{-1}(x-o)\|\right)^2 & \text{if } c \equiv i \text{ and } \|\xi^{-1}(x-o)\| \leq 1, \\ 0 & \text{else.} \end{cases}$$

This describes an approximation to the Gaussian with local support. The center is given by o and the diagonal matrix ξ describes the width of the Gaussian in coordinate directions. We identify a scalar value for ξ with ξI, where I denotes the identity matrix. The parameter c is the number of the component in which the vector field f is not zero.

We use 2×4 Ansatz functions for the control at the corners of each square. Thus we use the following 16 Ansatz functions $f[m_l^{ij}, \xi, c]$ with $m_k^{ij} = (0.5 + (-1)^k 0.25 + (-1)^i 0.13, 0.5 + (-1)^j 0.13)^\top$ with $1 \leq k, i, j \leq 2$, $\xi = 0.1$, $c \in \{0, 1\}$. Thus, $U = \mathbb{R}^{16}$ and $Bu := \sum_{c=0}^{1} \sum_{i,j,k=0}^{2} u_{ijkc} f[m_k^{ij}, \xi, c]$. In Figure 1 we show plots of φ_0,φ_d, and Bu together with the phase field at final time if no control is applied.

The optimization horizon is $T = 1.0$ and we use $\tau = 0.00125$ and we set $\alpha = 1e-11$. For Algorithm 1 we further set $tol_c = 1e-3$, $N_{max} = 8e6$, which means $1e4$ cells per time instance, and for the marking procedure we use $\theta^r = 0.7$ and $\theta^c = 0.01$.

On a single (time) sequence of grids we stop the optimization procedure as soon as

$$\|\nabla J(u)\|_U \leq 2e-7 + 0.1\|\nabla J(u^0)\|_U \tag{6.3}$$

holds, where u^0 denotes the initial control for the optimization procedure on the current sequence of meshes. The stopping criteria is motivated by the fact that we observe in our numerical tests that the optimal control of the previous adaptation step already constitutes a good guess for the optimal control on the refined mesh. Algorithm 1 is initialized with zero control, while subsequent optimization steps, i.e. line 3, are initialized with the control from the previous optimization run.

The optimal solution

In Figure 2 we depict the temporal evolution of the phase field $\varphi^\star$ corresponding to the optimal control $u^\star$, while in Figure 3 we depict the strength of the control over the time

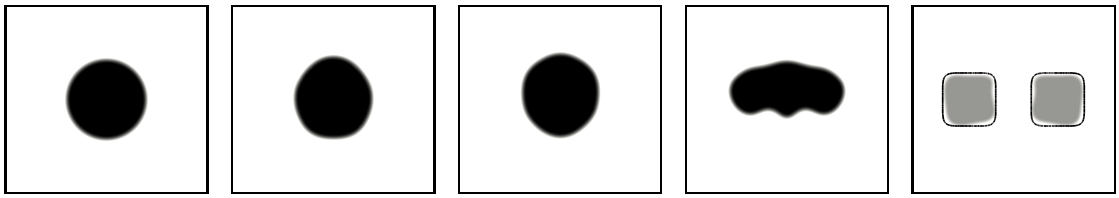

Figure 2: The evolution of the phase field φ with respect to the optimal control u for few Ansatz functions, $t = 0.00, 0.25, 0.50, 0.75, 1.00$ (left to right). For $t = 1.00$ we show φ in gray and in black the zero level line of the desired shape φ_d.

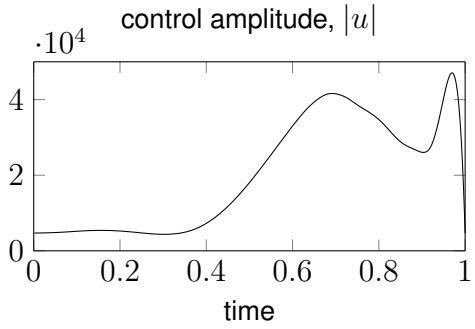

Figure 3: The amplitude of the control over time, $|u(t)|$. We observe that the control amplitude is increasing over time until $t = 0.7$ and thereafter is reduced again with a second maximum directly before a final strong reduction of the control at the final time.

horizon, i.e. $|u^\star(t)|$. To obtain this optimal control we solve the optimization problem (P_h^s) 10 times, i.e. line 3 of Algorithm 1 is executed 10 times. After the first two solves the Moreau–Yosida parameter was increased, i.e. line 7 in Algorithm 1 is executed, and after the next 8 solves the algorithm proceeded with evaluating the indicator (4.7) and the proposed Dörfler marking procedure, i.e. line 10 of Algorithm 1 was executed.

The number of cells

In Figure 4 (left) we present the evolution of the total number of cells over the adaptation steps. The right picture presents the distribution of the cells for the final sequence of grids. Here we show the number of cells of $\mathcal{T}^i$ for every time instance of the time horizon I . We observe that the number of cells increases as the "length" of the interface increases with time as the rising bubble is split up. We also observe that the cells are mainly refined inside and at the border of the diffuse interface $|\varphi| < 1$. Such a behavior is expected, as the phase field φ_d has a longer diffuse interface than φ_0. In this sense we discover a behavior that is similar to residual based error estimation, see e.g. [23, 20]. Since our dual weighted residual error estimator also contains terms from the Navier–Stokes and the adjoint equation, we further obtain refinement inside of the bulk domain if required. In Figure 5 we depict the subdomain $\Omega_u = (0, 1) \times (0.5, 1.0) \subset \Omega$ at $t = 0.7$. On the left we show $|v|$ in grayscale together with the isolines $\varphi \equiv \pm 1$ in black. On the right we show the corresponding mesh and note that the mesh is symmetric w.r.t. the central line.

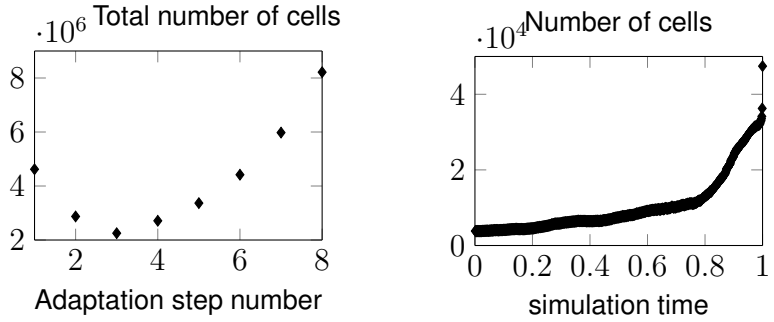

Figure 4: The evolution of the total number of cells, i.e. $\sum_{i=1}^{M} NC(\mathcal{T}^i)$, where $NC(\mathcal{T}^i)$ denotes the number of cells of the triangulation $\mathcal{T}^i$ over the adaptation steps (left). We note that we can not start with an arbitrarily coarse mesh, as the interface as least has to be resolved roughly at the initialization of the optimization procedure. On the right we depict the distribution of the number of cells over the time horizon. We observe that the mesh is refined most close to the final time instance, where our optimization aim is located.

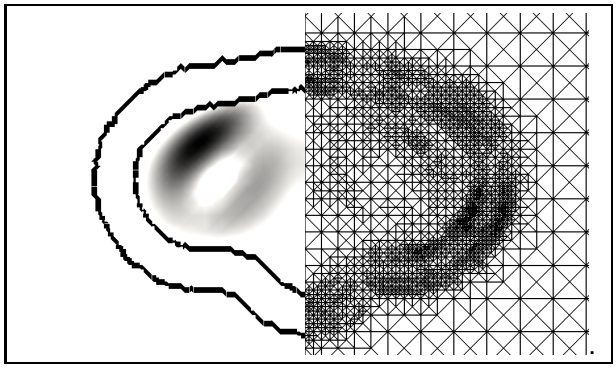

Figure 5: The subdomain $(0.0, 1.0) \times (0.25, 0.85) \subset \Omega$ at $t = 0.7$. On the left we show $|v|$ in grayscale together with the isolines $\varphi \equiv \pm 1$. On the right we show the corresponding triangulation. Note that the problem is symmetric w.r.t. $x \equiv 0.5$. We observe, that the mesh is refined inside the diffuse interface as expected, but also is refined in the regions with large velocity.

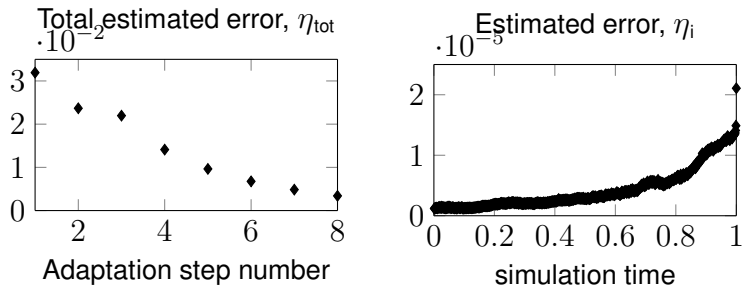

Figure 6: The evolution of the error estimator η_{tot} over the adaptation steps (left), the distribution of the estimated error over the time horizon (right)

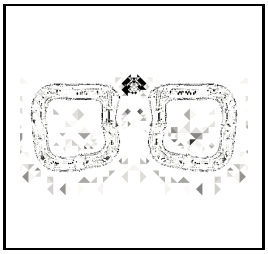

Figure 7: The distribution η_T^M at the time instance t_M for the optimal control $u^\star$. Black indicates large errors.

The error indicator

Let us next comment on the evolution of the error estimator η. In addition to (5.7), we introduce the notation $\eta^i := \sum_{T\in\mathcal{T}^i} \eta_T^i$ for the total estimated error at time instance i and $\eta_{tot} := \sum_{i=1}^M \eta^i$ for the total estimated error of a solution to the optimization problem.

In Figure 6 we depict the evolution of η_{tot} over the adaptation steps, i.e. for the optimal solution on the kth sequence of grids we show the estimator for the overall error on the left. We further show the distribution of η_T^M at the time instance t_M for the optimal solution $u^\star$ on the right. One observes a significant decay of the estimated error throughout the adaptation steps. Further, the largest error appears at later times of the simulation horizon. Especially the large error at final time can be explained by the term $\|\varphi_M - \varphi_d\|$ arising from the optimization aim, but we in general observe that larger time instances have a higher impact on the overall estimated error. Finally, in Figure 7, we depict the distribution of η_T^M for the optimal control. We observe that large errors mainly appear at the transition from diffuse interface to bulk, i.e. where $|\varphi| \approx 1$ holds. The large error component in the middle of the domain appears as an artefact of splitting the bubble into the two squares.

For a comparison of the error decay on a homogeneously refined grid and an adaptively refined grid, where residual based estimation is used for the pure phase field equation, we refer to [23].

The Moreau–Yosida relaxation

Finally we comment on the update procedure for the Moreau–Yosida parameter. In Figure 8 we show the evolution of the maximum complementarity mismatch over the optimization steps. Each column of the plot contains the maximum mismatch of the five complemen-

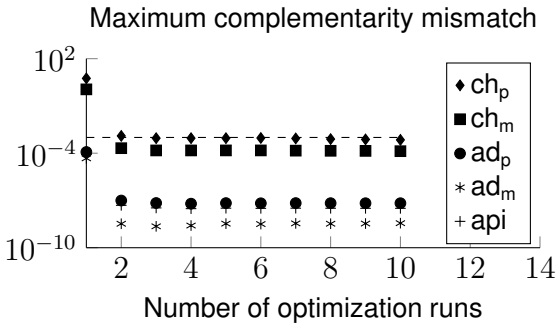

Figure 8: The maximum complementarity mismatch for each optimization run. Here ch$_p$ denotes $\max_{i=1,\dots,M} \max_{T\in\mathcal{T}^i}((a^i)^+, \varphi_h^i - 1)_T$, ch$_m$ denotes $\max_{i=1,\dots,M} \max_{T\in\mathcal{T}^i}((a^i)^-, \varphi_h^i + 1)_T$, ad$_p$ denotes $\max_{i=1,\dots,M} \max_{T\in\mathcal{T}^i}((\lambda^i)^+, \varphi_h^i - 1)_T$, ad$_m$ denotes $\max_{i=1,\dots,M} \max_{T\in\mathcal{T}^i}((\lambda^i)^-, \varphi_h^i + 1)_T$, and api denotes $\max_{i=1,\dots,M} \max_{T\in\mathcal{T}^i}(a^i, \pi^i)_T$. We observe that after only one optimization with subsequent increment of the Moreau–Yosida parameter s we reach the desired tolerance $tol_c = 1e-3$ indicated by the dashed line.

tarity relations (2.8)–(2.10), (2.16), and (2.17), where the maximum is taken over all time instances and all cells. The dashed line indicates the desired maximum mismatch tol_c, and we observe that already with the third value, i.e. after increasing s twice, the desired bound is reached. The corresponding values of s are $s = 8e6$ as initial value, $s = 3e14$ and finally $s = 6e14$ for subsequent steps. Thus, we observe that the results are insensitive w.r.t. s, as the parameter is not longer updated with decreasing h. For a rigorous analysis of the error introduced by using the Moreau–Yosida approximation in the case of control of the obstacle-problem we refer to [41]. We note that using a small initial value for s also results in well conditioned linear systems in Newton's method.

References

[1] *GSL - GNU Scientific Library*. `http://www.gnu.org/software/gsl/`.

[2] H. ABELS, H. GARCKE, AND G. GRÜN, *Thermodynamically consistent, frame indifferent diffuse interface models for incompressible two-phase flows with different densities*, Math. Models Methods Appl. Sci., 22 (2012), pp. 1150013, 40.

[3] R. A. ADAMS AND J. J. F. FOURNIER, *Sobolev spaces*, vol. 140 of Pure and Applied Mathematics (Amsterdam), Elsevier/Academic Press, Amsterdam, second ed., 2003.

[4] S. ALAND, S. BODEN, A. HAHN, F. KLINGBEIL, M. WEISMANN, AND S. WELLER, *Quantitative comparison of Taylor flow simulations based on sharp-interface and diffuse-interface models*, International Journal for Numerical Methods in Fluids, 73 (2013), pp. 344–361.

[5] S. ALAND, J. LOWENGRUB, AND A. VOIGT, *Particles at fluid-fluid interfaces: A new Navier-Stokes-Cahn-Hilliard surface-phase-field-crystal model*, Physical Review E, 86 (2012), p. 046321.

[6] P. AMESTOY, I. DUFF, J. KOSTER, AND J.-Y. L'EXCELLENT, *A fully asynchronous multifrontal solver using distributed dynamic scheduling*, SIAM Journal of Matrix Analysis and Applications, 23 (2001), pp. 15–41.

[7] S. BALAY, S. ABHYANKAR, M. ADAMS, J. BROWN, P. BRUNE, K. BUSCHELMAN, L. DALCIN, V. EIJKHOUT, W. GROPP, D. KAUSHIK, M. KNEPLEY, L. MCINNES, K. RUPP, B. SMITH, S. ZAMPINI, AND H. ZHANG, *PETSc Web page*. `http://www.mcs.anl.gov/petsc`, 2014.

[8] W. BANGERTH AND R. RANNACHER, *Adaptive finite element methods for differential equations*, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2003.

[9] L. BAŇAS, M. KLEIN, AND A. PROHL, *Control of interface evolution in multiphase fluid flows*, SIAM Journal on Control and Optimization, 52 (2014), pp. 2284–2318.

[10] R. BECKER, H. KAPP, AND R. RANNACHER, *Adaptive finite element methods for optimal control of partial differential equations: Basic concept*, SIAM Journal on Control and Optimization, 39 (2000), pp. 113–132.

[11] J. BRAMBLE, J. PASCIAK, AND O. STEINBACH, *On the Stability of the L^2 projection in $H^1(\Omega)$*, Mathematics of Computations, 71 (2001), pp. 147–156.

[12] S. C. BRENNER AND L. R. SCOTT, *The Mathematical Theory of Finite Element Methods*, vol. 15 of Texts in Applied Mathematics, Springer, 2008.

[13] C. BRETT, C. M. ELLIOTT, M. HINTERMÜLLER, AND C. LÖBHARD, *Mesh adaptivity in optimal control of elliptic variational inequalities with point-tracking of the state*, Interfaces Free Bound., 17 (2015), pp. 21–53.

[14] J. W. CAHN AND J. E. HILLIARD, *Free energy of a nonuniform system. i. interfacial free energy*, The Journal of chemical physics, 28 (1958), pp. 258–267.

[15] P. CONSTANTIN AND C. FOIAS, *Navier-Stokes-Equations*, The University of Chicago Press, 1988.

[16] W. DÖRFLER, *A convergent adaptive algorithm for Poisson's equation*, SIAM Journal on Numerical Analysis, 33 (1996), pp. 1106–1124.

[17] S. ECKERT, P. A. NIKRITYUK, B. WILLERS, D. RÄBIGER, N. SHEVCHENKO, H. NEUMANN-HEYME, V. TRAVNIKOV, S. ODENBACH, A. VOIGT, AND K. ECKERT, *Electromagnetic melt flow control during solidification of metallic alloys*, The European Physical Journal Special Topics, 220 (2013), pp. 123–137.

[18] H. GARCKE, C. HECHT, M. HINZE, AND C. KAHLE, *Numerical approximation of phase field based shape and topology optimization for fluids*, SIAM Journal on Scientific Computing, 37 (2015), pp. 1846–1871.

[19] H. GARCKE, C. HECHT, M. HINZE, C. KAHLE, AND K. F. LAM, *Shape optimization for surface functionals in Navier–Stokes flow using a phase field approach*, preprint in arXiv: 1504.06402, (2015).

[20] H. GARCKE, M. HINZE, AND C. KAHLE, *A stable and linear time discretization for a thermodynamically consistent model for two-phase incompressible flow*, Applied Numerical Mathematics, 99 (2016), pp. 151–171.

[21] V. GIRAULT AND P.-A. RAVIART, *Finite element methods for Navier-Stokes equations*, vol. 5 of Springer Series in Computational Mathematics, Springer-Verlag, Berlin, 1986. Theory and algorithms.

[22] A. GÜNTHER AND M. HINZE, *A posteriori error control of a state constrained elliptic control problem*, J. Numer. Math., 16 (2008), pp. 307–322.

[23] M. HINTERMÜLLER, M. HINZE, AND M. H. TBER, *An adaptive finite-element Moreau-Yosida-based solver for a non-smooth Cahn-Hilliard problem*, Optim. Methods Softw., 26 (2011), pp. 777–811.

[24] M. HINTERMÜLLER AND R. H. W. HOPPE, *Goal-oriented adaptivity in control constrained optimal control of partial differential equations*, SIAM J. Control Optim., 47 (2008), pp. 1721–1743.

[25] M. HINTERMULLER AND R. H. W. HOPPE, *Goal-oriented adaptivity in pointwise state constrained optimal control of partial differential equations*, SIAM J. Control Optim., 48 (2010), pp. 5468–5487.

[26] M. HINTERMÜLLER, R. H. W. HOPPE, AND C. LÖBHARD, *Dual-weighted goal-oriented adaptive finite elements for optimal control of elliptic variational inequalities*, ESAIM Control Optim. Calc. Var., 20 (2014), pp. 524–546.

[27] M. HINTERMÜLLER, T. KEIL, AND D. WEGNER, *Optimal control of a semidiscrete cahn-hilliard-navier-stokes system with non-matched fluid densities*, arXiv preprint arXiv:1506.03591, (2015).

[28] M. HINTERMÜLLER AND I. KOPACKA, *Mathematical programs with complementarity constraints in function space: C- and strong stationarity and a path-following algorithm*, SIAM J. Optim., 20 (2009), pp. 868–902.

[29] M. HINTERMÜLLER, B. S. MORDUKHOVICH, AND T. M. SUROWIEC, *Several approaches for the derivation of stationarity conditions for elliptic MPECs with upper-level control constraints*, Math. Program., 146 (2014), pp. 555–582.

[30] M. HINTERMÜLLER, A. SCHIELA, AND W. WOLLNER, *The length of the primal-dual path in Moreau-Yosida-based path-following methods for state constrained optimal control*, SIAM J. Optim., 24 (2014), pp. 108–126.

[31] M. HINTERMÜLLER AND D. WEGNER, *Optimal control of a semidiscrete Cahn-Hilliard-Navier-Stokes system*, SIAM J. Control Optim., 52 (2014), pp. 747–772.

[32] 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, vol. 391 in IFIP Advances in Information and Communication Technology, 2013.

[33] ——, *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, vol. 1, 2013, pp. 127–132.

[34] P. C. HOHENBERG AND B. I. HALPERIN, *Theory of dynamic critical phenomena*, Reviews of Modern Physics, 49 (1977), p. 435.

[35] 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 (2009), pp. 1259–1288.

[36] C. KAHLE, *A L^∞ bound for the Cahn–Hilliard equation with relaxed non-smooth free energy density*, arXiv:1511.02618, (2015).

[37] J. KIM, K. KANG, AND J. LOWENGRUB, *Conservative multigrid methods for Cahn-Hilliard fluids*, J. Comput. Phys., 193 (2004), pp. 511–543.

[38] J. KIM AND J. LOWENGRUB, *Interfaces and multicomponent fluids*, Encyclopedia of Mathematical Physics, (2004), pp. 135–144.

[39] A. LOGG, K.-A. MARDAL, AND G. WELLS, eds., *Automated Solution of Differential Equations by the Finite Element Method - The FEniCS Book*, vol. 84 of Lecture Notes in Computational Science and Engineering, Springer, 2012.

[40] Z.-Q. LUO, J.-S. PANG, AND D. RALPH, *Mathematical programs with equilibrium constraints*, Cambridge University Press, Cambridge, 1996.

[41] C. MEYER, A. RADEMACHER, AND W. WOLLNER, *Adaptive optimal control of the obstacle problem*, SIAM Journal on Scientific Computing, 37 (2015), pp. A918–A945.

[42] Y. OONO AND S. PURI, *Study of phase-separation dynamics by use of cell dynamical systems. i. modeling*, Physical Review A, 38 (1988), p. 434.

[43] J. OUTRATA, M. KOČVARA, AND J. ZOWE, *Nonsmooth approach to optimization problems with equilibrium constraints*, vol. 28 of Nonconvex Optimization and its Applications, Kluwer Academic Publishers, Dordrecht, 1998. Theory, applications and numerical results.

[44] S. PRAETORIUS AND A. VOIGT, *A phase field crystal model for colloidal suspensions with hydrodynamic interactions*, arXiv preprint arXiv:1310.5495, (2013).

[45] A. RÖSCH AND D. WACHSMUTH, *A-posteriori error estimates for optimal control problems with state and control constraints*, Numer. Math., 120 (2012), pp. 733–762.

[46] H. SCHEEL AND S. SCHOLTES, *Mathematical programs with complementarity constraints: stationarity, optimality, and sensitivity*, Math. Oper. Res., 25 (2000), pp. 1–22.

[47] A. SCHMIDT AND K. G. SIEBERT, *Design of adaptive finite element software: The finite element toolbox ALBERTA*, vol. 42 of Lecture Notes in Computational Science and Engineering, Springer, 2005.

[48] R. VERFÜRTH, *A posteriori error analysis of space-time finite element discretizations of the time-dependent Stokes equations*, Calcolo, 47 (2010), pp. 149–167.

[49] B. VEXLER AND W. WOLLNER, *Adaptive finite elements for elliptic optimization problems with control constraints*, SIAM J. Control Optim., 47 (2008), pp. 509–534.

[50] B. ZHOU ET AL., *Simulations of polymeric membrane formation in 2D and 3D*, PhD thesis, Massachusetts Institute of Technology, 2006.

[51] J. ZOWE AND S. KURCYUSZ, *Regularity and stability for the mathematical programming problem in Banach spaces*, Appl. Math. Optim., 5 (1979), pp. 49–62.