Numerical Analysis and Scientific Computing
Preprint Seria

High order approximations in space and time of a sixth order Cahn-Hilliard equation

O. Boyarkin R.H.W. Hoppe C. Linsenmann

Preprint #41

Department of Mathematics
University of Houston

October 2015

Russ. J. Numer. Anal. Math. Modelling, Vol. XX, No. X, pp. XX–XX (2013)
DOI 10.1515/ RJNAMM.2013.0..

High order approximations in space and time of a sixth order Cahn-Hilliard equation

Oleg Boyarkin* and Ronald H. W. Hoppe† and Christopher Linsenmann‡

25th October 2015

Abstract — We consider an initial-boundary value problem for a sixth order Cahn-Hilliard equation describing the formation of microemulsions. Based on a Ciarlet-Raviart type mixed formulation as a system consisting of a second order and a fourth order equation, the spatial discretization is done by a C^0 Interior Penalty Discontinuous Galerkin (C^0IPDG) approximation with respect to a geometrically conforming simplicial triangulation of the computational domain. The DG trial spaces are constructed by C^0 conforming Lagrangian finite elements of polynomial degree $k \geqslant 2$. This leads to an initial value problem for an index 1 Differential Algebraic Equation (DAE) which is further discretized in time by an s-stage Diagonally Implicit Runge-Kutta (DIRK) method of order $p \geqslant 2$. The resulting parameter dependent nonlinear algebraic system is solved numerically by a predictor-corrector continuation strategy with constant continuation as a predictor and Newton's method as a corrector featuring an adaptive choice of the continuation parameter. Numerical results illustrate the performance of the suggested approach.

Keywords: sixth order Cahn-Hilliard equation, interior penalty discontinuous Galerkin methods, diagonally implicit Runge-Kutta methods, automatic step-size selection in time

This paper is dedicated to the seventieth anniversary of Yuri A. Kuznetsov

1. Introduction

Microemulsions are thermodynamically stable colloidal dispersions of oil and water. They typically occur as oil-in-water, water-in-oil, or water/oil droplets with a

*Institute of Mathematics, University of Augsburg, D-86159 Augsburg, Germany (oleg.boyarkin@math.uni-augsburg.de). The author has been supported by the German National Science Foundation within the Priority Program SPP 1506.

†Department of Mathematics, University of Houston, Houston, TX 77204-3008, USA (rohop@math.uh.edu), and Institute of Mathematics, University of Augsburg, D-86159 Augsburg, Germany (hoppe@math.uni-augsburg.de). The author acknowledges support by the NSF grants DMS-1115658, DMS-1216857, DMS-1520886, and by the German National Science Foundation within the Priority Program SPP 1506.

‡Institute of Mathematics, University of Augsburg, D-86159 Augsburg, Germany (christopher.linsenmann@math.uni-augsburg.de)

diameter up to 200 nm and are thus considerably smaller than ordinary emulsions. Due to their efficient drug solubilization capacity and bioavailability, microemulsions have significant applications in pharmacology as drug carriers for the delivery of hydrophilic as well as lipophilic drugs (cf. [12,19,21,22,25,26]). Based on a Ginzburg-Landau free energy for ternary oil-water-microemulsions due to Gompper et al. [13,14,15,16], the dynamics of the microemulsification process can be described by an initial-boundary value problem for a sixth order Cahn-Hilliard equation which has been analytically investigated by Pawlow et al. [23,24,27] (cf. also [20]).

By introducing the chemical potential as an additional unknown, we consider a Ciarlet-Raviart type mixed formulation as a system consisting of a second and a fourth order equation. For its spatial discretization we use a C^0 Interior Penalty Discontinuous Galerkin (C^0IPDG) approximation with respect to a geometrically conforming simplicial triangulation of the computational domain where the DG trial spaces are constructed by C^0 conforming Lagrangian finite elements of polynomial degree $k \geqslant 2$. We note that IPDG methods for the standard fourth order Cahn-Hilliard equation have been studied in [28] based on IPDG approximations of fourth order problems including the biharmonic equation considered in [4,7] (cf. also [2,9,10,11]). The semidiscretization in space by the C^0IPDG method leads to an initial value problem for an index 1 Differential Algebraic Equation (DAE) which is discretized in time by an s-stage Diagonally Implicit Runge-Kutta method of order $p \geqslant 2$ with respect to a partitioning of the time interval (cf., e.g., [1,5,17]). The resulting parameter dependent nonlinear algebraic system is numerically solved by a predictor-corrector continuation strategy with the time step size as the continuation parameter. As a predictor we use constant continuation whereas the corrector is chosen as Newton's method [18]. The predictor-corrector continuation strategy features an adaptive choice of the continuation parameter based on an affine covariant convergence theory of the simplified Newton method [6].

The paper is organized as follows: In section 2, we present the initial-boundary value problem for the sixth order Cahn-Hilliard equation based on a Ginzburg-Landau free energy and introduce a Ciarlet-Raviart type mixed formulation as a system consisting of a second and a fourth order equation. Then, in section 3 we consider the semidiscretization in space by the C^0IPDG method leading to an initial value problem for an index 1 DAE. The discretization in time by an s-stage DIRK method is addressed in section 4 whereas section 5 is devoted to the numerical solution of the resulting parameter dependent nonlinear algebraic system by a predictor-corrector continuation strategy with an adaptive selection of the continuation parameter. In section 6, we present numerical results which show the formation of water-in-oil and oil-in-water droplets in a ternary water-oil-microemulsion system and illustrate the performance of the adaptive predictor-corrector continuation strategy. Finally, the last section 7 contains some concluding remarks.

2. The sixth order Cahn-Hilliard equation

Given a quadrilateral domain $\Omega = (a_1, b_1) \times (a_2, b_2), a_i < b_i, 1 \leqslant i \leqslant 2$, with boundary $\Gamma = \partial\Omega$, and exterior unit normal vector $\mathbf{n}$, denoting by $T > 0$ the final time, and setting $Q := \Omega \times (0,T)$, $\Sigma = \Gamma \times (0,T)$, we consider the following sixth order Cahn-Hilliard equation

$$\sigma \frac{\partial c}{\partial t} - M\Delta\Big(\varkappa\Delta^2 c - a(c)\Delta c - \frac{1}{2}a'(c)|\nabla c|^2 + f_0(c)\Big) = 0 \quad \text{in } Q \tag{2.1a}$$

with the boundary conditions

$$\mathbf{n}\cdot\nabla c = \mathbf{n}\cdot\nabla\mu(c) = \mathbf{n}\cdot\nabla\Delta c = 0 \quad \text{on } \Sigma \tag{2.1b}$$

and the initial concentration

$$c(\cdot,0) = c_0 \quad \text{in } \Omega. \tag{2.1c}$$

Here, σ is a surface energy density, M stands for the mobility which in the sequel will be assumed to be a positive constant, $\varkappa$ is a positive constant as well, and the coefficient function $a(c)$ is assumed to be of the form

$$a(c) = a_0 + a_2 c^2, \quad a_0 \in \mathbb{R},\ a_2 > 0.$$

The function $f_0(c) = F_0'(c)$ is the derivative of the multiwell free energy

$$F_0(c) = \int_\Omega \frac{\beta}{2}\,(c+1)^2(c^2+h_0)(c-1)^2\,dx, \quad h_0 \in \mathbb{R},$$

where β is another surface energy density. Moreover, $\mu(c)$ denotes the chemical potential which is the variation

$$\mu(c) = \frac{\delta F(c)}{\delta c}$$

of the total free energy

$$F(c) = F_0(c) + \int_\Omega \Big(\frac{1}{2}a(c)|\nabla c|^2 + \frac{1}{2}\varkappa|\Delta c|^2\Big)\,dx, \tag{2.2}$$

and c_0 is a given initial condition.

Remark 2.1 *As mentioned in section* 1*, the initial-boundary value problem* (2.1a)-(2.1c) *describes the dynamics of ternary oil-water-microemulsion systems where the solution c is an order parameter representing the local difference between the oil and water concentrations. We note that the Ginzburg-Landau free energy* (2.2) *for such systems has been suggested in* [14,15] *and* [13,16]. *Analytical results with regard to the existence and uniqueness of global solutions have been provided in* [23,24,27].

We introduce the chemical potential $\mu(c)$ as an additional unknown $w := \mu(c)$ and can equivalently formulate the sixth order Cahn-Hilliard equation (2.1a) as a system of a linear second order and a nonlinear fourth order equation in (c,w) according to

$$\sigma \frac{\partial c}{\partial t} - M\Delta w = 0 \quad \text{in } Q, \tag{2.3a}$$

$$\varkappa\Delta^2 c - a(c)\Delta c - a_2 c|\nabla c|^2 + f_0(c) - w = 0 \quad \text{in } Q \tag{2.3b}$$

with the boundary conditions

$$\mathbf{n}\cdot\nabla c = \mathbf{n}\cdot\nabla w = \mathbf{n}\cdot\nabla\Delta c = 0 \quad \text{on } \Sigma \tag{2.3c}$$

and the initial condition

$$c(\cdot,0) = c_0 \quad \text{in } \Omega. \tag{2.3d}$$

The initial-boundary value problem for the system (2.3a),(2.3b) has been analytically studied in [27] and it has been shown that the system admits a unique weak solution.

3. Semidiscretization in space by the C^0 interior penalty discontinuous Galerkin method

For semidiscretization in space of the coupled system (2.3a)-(2.3d) we will use the C^0IPDG method with respect to a simplicial triangulation of the computational domain. Due to the convexity of the computational domain, we can use the Ciarlet-Raviart mixed formulation of (2.3b) by introducing $z = \Delta c$ as an additional unknown so that (2.3b) can be written as the following system of two second order equations

$$z = \Delta c, \tag{3.1a}$$

$$\varkappa\Delta z + g(c) = w, \tag{3.1b}$$

where $g(c)$ denotes the nonlinear function

$$g(c) := -a(c)\Delta c - a_2 c|\nabla c|^2 + f_0(c). \tag{3.1c}$$

Multiplying (3.1a) by a test function $\varphi \in H^1(\Omega)$ and (3.1b) by a test function $\psi \in H^2(\Omega)$ and integrating over Ω, integration by parts and observing (2.3c) yields the weak formulation

$$(z,\varphi)_{0,\Omega} = -(\nabla c,\nabla\varphi)_{0,\Omega}, \tag{3.2a}$$

$$(\varkappa z,\Delta\psi)_{0,\Omega} - (\varkappa z,\mathbf{n}\cdot\nabla\psi)_{0,\Gamma} + (g(c),\psi)_{0,\Omega} = (w,\psi)_{0,\Omega}. \tag{3.2b}$$

We assume $\mathscr{T}_h(\Omega)$ to be a shape-regular simplicial triangulation of Ω. For $D \subseteq \overline{\Omega}$, we denote by $\mathscr{E}_h(D)$ the sets of nodal points of $\mathscr{T}_h$ in D. For $T \in \mathscr{T}_h$ and $E \in \mathscr{E}_h$ we

further refer to h_T and h_E as the diameter of T and the length of E.
Denoting by $P_k(T),\ k \in \mathbb{N}$, the linear space of polynomials of degree $\leqslant k$ on T, for $k \geqslant 2$ we refer to

$$V_h := \{v_h \in H^1(\Omega) \mid v_h|_T \in P_k(T),\ T \in \mathscr{T}_h\} \tag{3.3}$$

as the finite element space of Lagrangian finite elements of type k (cf., e.g., [3]). We refer to $\mathscr{N}_h(\Omega)$ as the set of nodal points such that any $v_h \in V_h$ is uniquely determined by its degrees of freedom $v_h(a),\ a \in \mathscr{N}_h(\Omega)$. We note that $V_h \not\subset H^2(\Omega)$ and hence, V_h is a nonconforming finite element space for the approximation of the fourth order equation (2.3b). In particular, for a function z_h on $\overline{\Omega}$ that is elementwise polynomial, we define averages and jumps according to

$$\{z_h\}_E := \begin{cases} \frac{1}{2}\Big(z_h|_{E\cap T_+} + z_h|_{E\cap T_-}\Big)\ , E \in \mathscr{E}_h(\Omega), \\ z_h|_E\ , E \in \mathscr{E}_h(\Gamma), \end{cases} \tag{3.4a}$$

$$[z_h]_E := \begin{cases} z_h|_{E\cap T_+} - z_h|_{E\cap T_-}\ , E \in \mathscr{E}_h(\Omega), \\ z_h|_E\ , E \in \mathscr{E}_h(\Gamma). \end{cases} \tag{3.4b}$$

The general C^0DG approximation of (3.2a),(3.2b) reads:
Given $w_h \in V_h$, find $(c_h, z_h) \in V_h \times V_h$ such that for all $(\varphi_h, \psi_h) \in V_h \times V_h$ it holds

$$\sum_{T\in\mathscr{T}_h(\Omega)} \Big((z_h, \varphi_h)_{0,T} + (\nabla c_h, \nabla\varphi_h)_{0,T}\Big) - \sum_{E\in\mathscr{E}_h(\bar{\Omega})} (\mathbf{n}_E \cdot \hat{\mathbf{c}}_E, \varphi_h)_{0,\partial T}\Big) = 0, \tag{3.5a}$$

$$\sum_{T\in\mathscr{T}_h(\Omega)} \Big((\varkappa z_h, \Delta\psi_h)_{0,T} + (g(c_h), \psi_h)_{0,T}\Big) - \sum_{E\in\mathscr{E}_h(\bar{\Omega})} (\hat{\mathbf{z}}_E, \nabla\psi_h)_{0,E} - (w_h, \psi_h)_{0,T}\Big) = 0, \tag{3.5b}$$

where $\hat{\mathbf{c}}_E$ and $\hat{\mathbf{z}}_E$ are suitably chosen numerical flux functions that determine the type of C^0DG approximation. In particular, for the C^0IPDG approximation we choose

$$\hat{\mathbf{c}}_E := \begin{cases} \{\nabla c_h\}_E\ , & E \in \mathscr{E}_h(\Omega) \\ 0\ , E \in \mathscr{E}_h(\Gamma) \end{cases}, \tag{3.5c}$$

$$\hat{\mathbf{z}}_E := \Big(\{\Delta c_h\}_E - \frac{\alpha}{h_E}[\frac{\partial c_h}{\partial n}]_E\Big)\mathbf{n}_E, \quad E \in \mathscr{E}_h(\bar{\Omega}), \tag{3.5d}$$

where $\alpha > 0$ is a penalization parameter. The choice (3.5c),(3.5d) has the advantage that for $\varphi_h = \varkappa\Delta\psi_h$ in (3.5a) we may eliminate the dual variable z_h from the system and thus arrive at the following primal variational formulation of the C^0IPDG approximation: Find $c_h \in V_h$ such that for all $\psi_h \in V_h$ it holds

$$a_h^{DG}(c_h, \psi_h) + \sum_{T\in\mathscr{T}_h(\Omega)} (g(c_h), \psi_h)_{0,T} = (w_h, \psi_h)_{0,\Omega}, \tag{3.6}$$

where $a_h^{DG}(\cdot,\cdot) : V_h \times V_h \to \mathbb{R}$ stands for the C^{0}IPDG bilinear form

$$a_h^{DG}(c_h,\psi_h) := \sum_{T\in\mathscr{T}_h(\Omega)} (\varkappa \Delta c_h, \Delta\psi_h)_{0,T} - \sum_{E\in\mathscr{E}_h(\overline{\Omega})} \Big((\varkappa \mathbf{n}_E \cdot [\nabla c_h]_E, \{\Delta\psi_h\}_E)_{0,E} + (\varkappa \{\Delta c_h\}_E, \mathbf{n}_E \cdot [\nabla\psi_h]_E)_{0,E} \Big) + \sum_{E\in\mathscr{E}_h(\overline{\Omega})} \frac{\alpha}{h_E} (\mathbf{n}_E \cdot [\nabla c_h]_E, \mathbf{n}_E \cdot [\nabla\psi_h]_E)_{0,E}. \tag{3.7}$$

The C^{0}IPDG approximation of (2.3b) has the advantage that we may approximate the variable w in (2.3a) by a function in V_h as well. Consequently, the C^{0}IPDG approximation of (2.3a),(2.3b) reads: Find $(c_h, w_h) \in H^1([0,T],V_h) \times L^2([0,T],V_h)$ such that for all $\varphi_h \in V_h$ it holds

$$\sigma\,(\frac{\partial c_h}{\partial t},\varphi_h)_{0,\Omega} - M\,(\nabla w_h, \nabla\varphi_h)_{0,\Omega} = 0, \tag{3.8a}$$

$$a_h^{DG}(c_h,\varphi_h) + \sum_{T\in\mathscr{T}_h(\Omega)} (g(c_h),\varphi_h)_{0,T} - (w_h,\psi_h)_{0,\Omega} = 0, \tag{3.8b}$$

$$(c_h(\cdot,0),\varphi_h)_{0,\Omega} - (c_0,\varphi_h)_{0,\Omega} = 0. \tag{3.8c}$$

We note that $V_h = \mathrm{span}\{\varphi_h^{(1)},\cdots,\varphi_h^{(N_h)}\}$, where $\varphi_h^{(i)}, 1 \leqslant i \leqslant N_h$, are the basis functions associated with the nodal points in $\mathscr{N}_h(\overline{\Omega})$. Hence, the unknowns $c_h, w_h \in V_h$ admit the representations

$$c_h = \sum_{j=1}^{N_h} c_{h,j}\,\varphi_h^{(j)}, \quad w_h = \sum_{j=1}^{N_h} w_{h,j}\,\varphi_h^{(j)}.$$

Introducing the vector-valued functions $\mathbf{c}_h(t) = (c_{h,1}(t),\cdots,c_{h,N_h}(t))^T$, $\mathbf{w}_h(t) = (w_{h,1}(t),\cdots,w_{h,N_h}(t))^T$, the matrices $\mathbf{M}_h = (m_{ij})_{i,j=1}^{N_h}, \mathbf{A}_h = (a_{ij}^{DG})_{i,j=1}^{N_h}, \mathbf{S}_h = (s_{ij})_{i,j=1}^{N_h}$, and the vectors $\mathbf{G}_h(\mathbf{c}_h) = (G_{h,1}(\mathbf{c}_h),\cdots,G_{h,N_h}(\mathbf{c}_h))^T, \mathbf{c}_h^0 = (c_{h,1}^0,\cdots,c_{h,N_h}^0)^T$ according to

$$m_{ij} := \sigma\,(\varphi_h^{(i)},\varphi_h^{(j)})_{0,\Omega},\ a_{ij}^{DG} := a_h^{DG}(\varphi_h^{(i)},\varphi_h^{(j)}),\ s_{ij} := M(\nabla\varphi_h^{(i)},\nabla\varphi_h^{(j)})_{0,\Omega},$$
$$G_{h,i}(\mathbf{c}_h) := \sum_{T\in\mathscr{T}_h(\Omega)} (g(c_h),\varphi_h^{(i)})_{0,T},\ c_{h,i}^0 := (c_0,\varphi_h^{(i)})^T\ , 1 \leqslant i \leqslant N_h,$$

the C^{0}IPDG approximation of (2.3a),(2.3b) can be written as an initial-value problem for the differential-algebraic system of index 1

$$\mathbf{M}_h\,\frac{d\mathbf{c}_h}{dt} + \mathbf{S}_h\mathbf{w}_h = \mathbf{0}, \tag{3.9a}$$

$$\mathbf{A}_h\mathbf{c}_h + \mathbf{G}_h(\mathbf{c}_h) - \mathbf{M}_h\mathbf{w}_h = \mathbf{0}, \tag{3.9b}$$

$$\mathbf{c}_h(0) = \mathbf{c}_h^0. \tag{3.9c}$$

Denoting by $\mathscr{M}_h$ the block diagonal matrix $\mathscr{M}_h := \mathrm{diag}(\mathbf{M}_h, \mathbf{0})$ and setting $\mathbf{z}_h := (\mathbf{c}_h, \mathbf{w}_h)^T$, the DAE (3.9a),(3.9b) can be written as

$$\mathscr{M}_h \frac{d\mathbf{z}_h}{dt} = -\begin{pmatrix} \mathbf{S}_h \mathbf{w}_h \\ \mathbf{A}_h \mathbf{c}_h + \mathbf{G}_h(\mathbf{c}_h) - \mathbf{M}_h \mathbf{w}_h \end{pmatrix} =: \mathbf{f}_h(\mathbf{z}_h). \tag{3.10}$$

4. Discretization in time by diagonally implicit Runge-Kutta methods

For the discretization in time of the index 1 DAE (3.9a)-(3.9c) implicit numerical integrators are mandatory (cf., e.g., [5,17]). Here, we will use (s,p) diagonally implicit Runge-Kutta (DIRK) methods of stage s and order p where the order is adjusted to the order of the C^0IPDG approximation in space. Given a partitioning of the time interval $[0,T]$ into subintervals $[t_{m-1}, t_m]$ of length $\tau_m := t_m - t_{m-1}, 1 \leqslant m \leqslant M$, we denote by $\mathbf{c}_h^m$ and $\mathbf{w}_h^m$ approximations of $\mathbf{c}_h$ and $\mathbf{w}_h$ at time level t_m and set $\mathbf{z}_h^m := (\mathbf{c}_h^m, \mathbf{w}_h^m)^T$. An s-stage DIRK method reads

$$\mathscr{M}_h \mathbf{z}_h^m = \mathscr{M}_h \mathbf{z}_h^{m-1} + \tau_m \sum_{i=1}^{s} b_i \mathbf{f}_h(\mathbf{z}_h^{m,i}),, \tag{4.1a}$$

$$\mathscr{M}_h \mathbf{z}_h^{(m,i)} = \mathscr{M}_h \mathbf{z}_h^{m-1} + \tau_m \sum_{j=1}^{i} a_{ij} \mathbf{f}_h(\mathbf{z}_h^{(m,j)}),\ 1 \leqslant i \leqslant s, \tag{4.1b}$$

where the coefficients $a_{ij}, 1 \leqslant j \leqslant i \leqslant s$, with $a_{ii} \neq 0, 1 \leqslant i \leqslant s$, and $b_i, c_i, 1 \leqslant i \leqslant s$, are given and the intermediate states $\mathbf{z}_h^{(m,i)}$ are approximations of $\mathbf{z}_h$ at $t_{m-1} + c_i \tau_m, 1 \leqslant i \leqslant s$. The method can be characterized by the Butcher scheme

Table 1. Butcher scheme of an s-stage DIRK method

c_1	a_{11}			
c_2	a_{21}	a_{22}		
$\vdots$	$\vdots$		$\ddots$	
c_s	a_{s1}	a_{s2}	$\cdots$	a_{ss}
	b_1	b_2	$\cdots$	b_s

Remark 4.1 *If $a_{ii} = a \neq 0, 1 \leqslant i \leqslant s$, then the method is called a singly diagonally implicit Runge-Kutta (SDIRK) method.*

In more explicit form (4.1b) reads

$$\mathbf{M}_h\mathbf{c}_h^{(m,i)} = \mathbf{M}_h\mathbf{c}_h^{m-1} - \tau_m \sum_{j=1}^{i} a_{ij}\mathbf{S}_h\mathbf{w}_h^{(m,j)},\ 1 \leqslant i \leqslant s \tag{4.2a}$$

$$\mathbf{0} = \tau_m \sum_{j=1}^{i} a_{ij}\big[\mathbf{A}_h\,\mathbf{c}_h^{(m,j)} + \mathbf{G}_h(\mathbf{c}_h^{(m,j)}) - \mathbf{M}_h\,\mathbf{w}_h^{(m,j)}\big],\ 1 \leqslant i \leqslant s. \tag{4.2b}$$

To avoid cancelations, we introduce the increments $\mathbf{u}_h^{(m,i)} := \mathbf{c}_h^{(m,i)} - \mathbf{c}_h^{m-1}$, $1 \leqslant i \leqslant s$. Then (4.2a), (4.2b) and the $\mathbf{c}_h$-component of (4.1a) can be written in the form

$$\mathbf{u}_h^{(m,i)} = -\tau_m \sum_{j=1}^{i} a_{ij}\mathbf{M}_h^{-1}\mathbf{S}_h\mathbf{w}_h^{(m,j)},\ \ 1 \leqslant i \leqslant s, \tag{4.3a}$$

$$\mathbf{0} = \sum_{j=1}^{i} a_{ij}\left(\mathbf{A}_h\mathbf{u}_h^{(m,j)} + \mathbf{G}_h(\mathbf{c}_h^{m-1} + \mathbf{u}_h^{(m,j)}) - \mathbf{M}_h\mathbf{w}_h^{(m,j)} + \mathbf{A}_h\mathbf{c}_h^{m-1}\right),\ 1 \leqslant i \leqslant s, \tag{4.3b}$$

$$\mathbf{c}_h^m = \mathbf{c}_h^{m-1} - \tau_m \sum_{i=1}^{s} b_i\mathbf{M}_h^{-1}\mathbf{S}_h\mathbf{w}_h^{(m,i)}. \tag{4.3c}$$

Setting $\mathbf{U}_h^m = \big(\mathbf{u}_h^{(m,1)} \cdots \mathbf{u}_h^{(m,s)}\big)^T \in \mathbb{R}^{N_h\times s}$ and $\mathbf{W}_h^m = \big(\mathbf{w}_h^{(m,1)} \cdots \mathbf{w}_h^{(m,s)}\big)^T \in \mathbb{R}^{N_h\times s}$, in matrix-vector notation (4.3a) and (4.3c) can be written as

$$\mathbf{U}_h^m = -\tau_m\,\mathbf{M}_h^{-1}\mathbf{S}_h\mathbf{W}_h^m\mathscr{A}^T, \quad \mathbf{c}_h^m = \mathbf{c}_h^{m-1} - \tau_m\,\mathbf{M}_h^{-1}\mathbf{S}_h\mathbf{W}_h^m\mathbf{b},$$

where $\mathbf{b} := (b_1,\cdots,b_s)^T$ and $\mathscr{A} \in \mathbb{R}^{s\times s}$ is the lower triangular matrix with entries $a_{ij}, 1 \leqslant j \leqslant i \leqslant s$. Further, introducing $\mathbf{d} = (d_1,\cdots,d_s)^T := \mathscr{A}^{-T}\mathbf{b} \in \mathbb{R}^s$ and using $\mathbf{U}_h^m\mathscr{A}^{-T} = -\tau_m\,\mathbf{M}_h^{-1}\mathbf{S}_h\mathbf{W}_h^m$, the $\mathbf{c}_h$-update (4.3c) can be simplified to

$$\mathbf{c}_h^m = \mathbf{c}_h^{m-1} + \sum_{i=1}^{s} d_i\,\mathbf{u}_h^{(m,i)}. \tag{4.4}$$

To summarize, at each time instant t_m the unknowns

$$\mathbf{x}_h^m := (\mathbf{u}_h^{m,1}, \mathbf{w}_h^{(m,1)}, \ldots, \mathbf{u}_h^{(m,s)}, \mathbf{w}_h^{(m,s)}, \mathbf{c}_h^m) \in \mathbb{R}^{2N_h s + N_h}$$

satisfy the following nonlinear system of equations

$$\mathbf{F}_h(\mathbf{x}_h^m, \tau_m) = \mathbf{0}, \tag{4.5}$$

where

$$\mathbf{F}_h(\mathbf{x}_h^m,\tau_m) := \begin{pmatrix} \mathbf{M}_h\,\mathbf{u}_h^{(m,1)} + \tau_m\, a_{11}\,\mathbf{S}_h\,\mathbf{w}_h^{(m,1)} \\ a_{11}\left(\mathbf{A}_h\,\mathbf{u}_h^{(m,1)} + \mathbf{G}_h(\mathbf{c}_h^{m-1}+\mathbf{u}_h^{(m,1)}) - \mathbf{M}_h\,\mathbf{w}_h^{(m,1)} + \mathbf{A}_h\,\mathbf{c}_h^{m-1}\right) \\ \vdots \\ \mathbf{M}_h\,\mathbf{u}_h^{(m,s)} + \tau_m \sum_{j=1}^{s} a_{sj}\,\mathbf{S}_h\,\mathbf{w}_h^{(m,j)} \\ \displaystyle\sum_{j=1}^{s} a_{sj}\left(\mathbf{A}_h\,\mathbf{u}_h^{(m,j)} + \mathbf{G}_h(\mathbf{c}_h^{m-1}+\mathbf{u}_h^{(m,j)}) - \mathbf{M}_h\,\mathbf{w}_h^{(m,j)} + \mathbf{A}_h\,\mathbf{c}_h^{m-1}\right) \\ \sum_{i=1}^{s} d_i\,\mathbf{u}_h^{(m,i)} - \mathbf{c}_h^m + \mathbf{c}_h^{m-1} \end{pmatrix}.$$

Dropping the second argument τ_m for notational convenience for the rest of this section, we partition $\mathbf{F}_h(\mathbf{x}_h^m)$ according to $(\mathbf{F}_{h,1}(\mathbf{x}_h^m),\ldots,\mathbf{F}_{h,s}(\mathbf{x}_h^m),\mathbf{F}_{h,s+1}(\mathbf{x}_h^m))$, where $\mathbf{F}_{h,i}(\mathbf{x}_h^m)\in\mathbb{R}^{2N_h}, 1\leqslant i\leqslant s$, and $\mathbf{F}_{h,s+1}(\mathbf{x}_h^m)\in\mathbb{R}^{N_h}$. The Jacobian $\mathbf{DF}_h(\mathbf{x}_h^m)$ has a lower triangular block structure

$$\begin{pmatrix} \mathbf{D}_1\mathbf{F}_{h,1}(\mathbf{x}_h^m) & & & \\ \mathbf{D}_1\mathbf{F}_{h,2}(\mathbf{x}_h^m) & \mathbf{D}_2\mathbf{F}_{h,2}(\mathbf{x}_h^m) & & \\ \vdots & \vdots & \ddots & \\ \mathbf{D}_1\mathbf{F}_{h,s}(\mathbf{x}_h^m) & \mathbf{D}_2\mathbf{F}_{h,s}(\mathbf{x}_h^m) & \cdots & \mathbf{D}_s\mathbf{F}_{h,s}(\mathbf{x}_h^m) \\ \mathbf{D}_1\mathbf{F}_{h,s+1}(\mathbf{x}_h^m) & \mathbf{D}_2\mathbf{F}_{h,s+1}(\mathbf{x}_h^m) & \cdots & \mathbf{D}_s\mathbf{F}_{h,s+1}(\mathbf{x}_h^m) \quad \mathbf{D}_{s+1}\mathbf{F}_{h,s+1}(\mathbf{x}_h^m) \end{pmatrix}$$

with blocks $\mathbf{D}_j\mathbf{F}_{h,i}(\mathbf{x}_h^m)$ given by

$$\mathbf{D}_i\mathbf{F}_{h,i}(\mathbf{x}_h^m) = \begin{pmatrix} \mathbf{M}_h & \tau_m\,a_{ii}\,\mathbf{S}_h \\ a_{ii}\left(\mathbf{A}_h + \mathbf{DG}_h(\mathbf{c}_h^{m-1}+\mathbf{u}_h^{(m,i)})\right) & -a_{ii}\,\mathbf{M}_h \end{pmatrix} \in \mathbb{R}^{2N_h\times 2N_h}$$

for $1\leqslant i\leqslant s$,

$$\mathbf{D}_j\mathbf{F}_{h,i}(\mathbf{x}_h^m) = \begin{pmatrix} \mathbf{0} & \tau_m\,a_{ij}\,\mathbf{S}_h \\ a_{ij}\left(\mathbf{A}_h + \mathbf{DG}_h(\mathbf{c}_h^{m-1}+\mathbf{u}_h^{(m,j)})\right) & -a_{ij}\,\mathbf{M}_h \end{pmatrix} \in \mathbb{R}^{2N_h\times 2N_h},$$

for $1\leqslant j<i\leqslant s$, and

$$\begin{aligned} \mathbf{D}_j\mathbf{F}_{h,s+1}(\mathbf{x}_h^m) &= (d_j\,\mathbf{I}_{N_h},\,\mathbf{0}) \in \mathbb{R}^{N_h\times 2N_h},\ 1\leqslant j<s+1, \\ \mathbf{D}_{s+1}\mathbf{F}_{h,s+1}(\mathbf{x}_h^m) &= -\mathbf{I}_{N_h}. \end{aligned}$$

5. Numerical solution of the fully discretized system

We first consider the application of Newton's method to (4.5) in subsection 5.1 and then elaborate on a predictor-corrector continuation strategy featuring an adaptive choice of the time step size in subsection 5.2.

5.1. Newtons's method

The nonlinear system (4.5) is solved by Newton's method

$$\mathbf{DF}_h(\mathbf{x}_h^{(m,\ell)},\tau_m)\Delta\mathbf{x}_h^{(m,\ell)} = -\mathbf{F}_h(\mathbf{x}_h^{(m,\ell)},\tau_m), \tag{5.1a}$$

$$\mathbf{x}_h^{(m,\ell+1)} = \mathbf{x}_h^{(m,\ell)} + \Delta\mathbf{x}_h^{(m,\ell)}, \tag{5.1b}$$

where $\mathbf{x}_h^{(m,\ell)}, \ell \in \mathbb{N}_0$, stands for the ℓ-th Newton iterate. As initial guess we choose

$$\mathbf{u}_h^{(m,i,0)} := \mathbf{0}, \quad \mathbf{w}_h^{(m,i,0)} := \mathbf{w}_h^{(m-1,i)}, \quad 1 \leqslant i \leqslant s, \qquad \mathbf{c}_h^{(m,0)} := \mathbf{c}_h^{m-1}. \tag{5.2}$$

The Jacobian $\mathbf{DF}_h(\mathbf{x}_h^{(m,\ell)},\tau_m)$ is regular, if this applies to each diagonal block $\mathbf{D}_i\mathbf{F}_{h,i}(\mathbf{x}_h^{(m,\ell)},\tau_m), 1 \leqslant i \leqslant s+1$. Obviously, $\mathbf{D}_{s+1}\mathbf{F}_{h,s+1}(\mathbf{x}_h^{(m,\ell)},\tau_m)$ is regular. The following result shows that $\mathbf{D}_i\mathbf{F}_{h,i}(\mathbf{x}_h^{(m,\ell)},\tau_m), 1 \leqslant i \leqslant s$, is invertible, if the time step size τ_m is appropriately chosen.

Theorem 5.1. *Let $\lambda_{\min}(\mathbf{M}_h)$ be the smallest eigenvalue of the symmetric positive definite matrix $\mathbf{M}_h$. If the time increment τ_m fulfills the inequalities*

$$0 < \tau_m \begin{cases} < \infty & \text{if } \mathbf{S}_h = \mathbf{A}_h + \mathbf{DG}_h(\mathbf{x}_h^{(m,i)}), \\ \leqslant \dfrac{\lambda_{\min}(\mathbf{M}_h)^2}{a_{ii}^2\,\|\mathbf{S}_h - (\mathbf{A}_h + \mathbf{DG}_h(\mathbf{x}_h^{(m,i)}))\|_2^2} & \text{else,} \end{cases} \tag{5.3}$$

where for simplicity we have used the notation $\mathbf{DG}_h(\mathbf{x}_h^{(m,i)}) := \mathbf{DG}_h(\mathbf{c}_h^{m-1} + \mathbf{u}_h^{(m,i)})$, then the matrix

$$\mathbf{B}_i(\tau_m) := \mathbf{D}_i\mathbf{F}_{h,i}(\mathbf{x}_h^m,\tau_m) = \begin{pmatrix} \mathbf{M}_h & \tau_m\, a_{ii}\, \mathbf{S}_h \\ a_{ii}\,(\mathbf{A}_h + \mathbf{DG}_h(\mathbf{x}_h^{(m,i)})) & -a_{ii}\,\mathbf{M}_h \end{pmatrix} \in \mathbb{R}^{2N_h \times 2N_h}$$

is regular and the spectral norm of its inverse can be estimated by

$$\left\|\mathbf{B}_i(\tau_m)^{-1}\right\|_2 \leqslant \begin{cases} \dfrac{2}{\lambda_{\min}(\mathbf{M}_h)}\,\dfrac{1}{\tau_m} & \text{for } 0 < \tau_m \leqslant 1 \\ \dfrac{2}{\lambda_{\min}(\mathbf{M}_h)}\,\tau_m & \text{for } \tau_m > 1. \end{cases} \tag{5.4}$$

Proof. We use a simplified version of the Banach-Nečas-Babuška theorem (see,e.g.,[8]): For $\mathbf{0} \neq \mathbf{v} \in \mathbb{R}^{2N_h}$ with components $\mathbf{v}_j \in \mathbb{R}^{N_h}$, $1 \leqslant j \leqslant 2$, we construct a vector $\mathbf{0} \neq \mathbf{w} = (\mathbf{w}_1, \mathbf{w}_2)^T \in \mathbb{R}^{2N_h}$ such that for some $\gamma_i > 0$ it holds

$$\mathbf{w}^T\mathbf{B}_i(\tau_m)\mathbf{v} \geqslant \gamma_i\, \|\mathbf{v}\|_2\, \|\mathbf{w}\|_2. \tag{5.5}$$

This implies $\|\mathbf{B}_i(\tau_m)\mathbf{v}\|_2 \geqslant \gamma_i \, \|\mathbf{v}\|_2$ which is equivalent to the injectivity (and therefore bijectivity) of the mapping $\mathbf{v} \mapsto \mathbf{B}_i(\tau_m)\mathbf{v}$.
We choose $\mathbf{w}_1 := \mathbf{v}_1$ and $\mathbf{w}_2 := -\tau_m \mathbf{v}_2$ so that $\mathbf{w} \neq \mathbf{0}$. Introducing for the moment the notation $\mu_h := \lambda_{\min}(\mathbf{M}_h)$, we obtain the estimate

$$\mathbf{w}^T \mathbf{B}_i(\tau_m)\mathbf{v} = \mathbf{v}_1^T \mathbf{M}_h \mathbf{v}_1 + \tau_m a_{ii} \, \mathbf{v}_2^T \left(\mathbf{S}_h - (\mathbf{A}_h + \mathbf{DG}_h(\mathbf{x}_h^{m,i})) \right) \mathbf{v}_1 + \tau_m a_{ii} \, \mathbf{v}_2^T \mathbf{M}_h \mathbf{v}_2$$

$$\geqslant \mu_h \, \|\mathbf{v}_1\|_2^2 - \tau_m a_{ii} \underbrace{\|\mathbf{S}_h - (\mathbf{A}_h + \mathbf{DG}_h(\mathbf{x}_h^{(m,i)}))\|_2}_{=: \, c_i(\mathbf{x}_h) =: c_i} \|\mathbf{v}_1\|_2 \|\mathbf{v}_2\|_2 + \tau_m a_{ii} \, \mu_h \|\mathbf{v}_2\|_2^2 \tag{5.6}$$

$$\geqslant (\mu_h - \tau_m a_{ii} \, c_i \, \varepsilon) \, \|\mathbf{v}_1\|_2^2 + \tau_m \left(\mu_h - a_{ii} \, c_i \, \frac{1}{4\,\varepsilon} \right) \|\mathbf{v}_2\|_2^2,$$

where in the last step we have used Young's inequality with some $\varepsilon > 0$. Let us first assume $c_i \neq 0$. We choose ε such that $\mu_h/(a_{ii} \, c_i) > \max(\tau_m \, \varepsilon, 1/(4\,\varepsilon))$. To this end, we set $\tau_m \, \varepsilon = 1/(4\,\varepsilon)$ and solve for $\varepsilon = 1/(2\sqrt{\tau_m})$. Then the above requirement reads $\mu_h/(a_{ii} \, c_i) > \sqrt{\tau_m}/2$. If we choose

$$0 < \frac{\sqrt{\tau_m}}{2} \leqslant \frac{1}{2} \frac{\mu_h}{a_{ii} \, c_i}$$

(which is equivalent to the second case of (5.3)), both factors in (5.6) are positive with value $\mu_h - \tau_m a_{ii} \, c_i \, \varepsilon = \mu_h - a_{ii} \, c_i/(4\,\varepsilon) \geqslant \mu_h/2$. Therefore,

$$\mathbf{w}^T \mathbf{B}_i(\tau_m)\mathbf{v} \geqslant \frac{\mu_h}{2} \|\mathbf{v}_1\|_2^2 + \tau_m \frac{\mu_h}{2} \, \|\mathbf{v}_2\|_2^2 \geqslant \begin{cases} \tau_m \dfrac{\mu_h}{2} \|\mathbf{v}\|_2^2 & \text{if } \tau_m \leqslant 1 \\ \dfrac{\mu_h}{2} \, \|\mathbf{v}\|_2^2 & \text{else} \end{cases} . \tag{5.7}$$

Furthermore, we have $\|\mathbf{w}\|_2 \leqslant \|\mathbf{v}\|_2$ for $\tau_m \leqslant 1$ and $\|\mathbf{w}\|_2 \leqslant \tau_m \|\mathbf{v}\|_2$ for $\tau_m > 1$. Combining this observation with (5.7) yields the desired estimate $\mathbf{w}^T \mathbf{B}_i(\tau_m)\mathbf{v} \geqslant \gamma_i \, \|\mathbf{v}\|_2 \, \|\mathbf{w}\|_2$ with inf-sup constant

$$\gamma_i := \begin{cases} \tau_m \dfrac{\mu_h}{2}, & \text{if } \tau_m \leqslant 1 \\ \dfrac{1}{\tau_m} \dfrac{\mu_h}{2}, & \text{else} \end{cases} .$$

In case of $\mathbf{S}_h = \mathbf{A}_h + \mathbf{DG}_h(\mathbf{x}_h^{m,i})$ (i.e., $c_i = 0$) there is no restriction on the time increment $\tau_m > 0$ as can be seen from the right-hand side of (5.6). As improved inf-sup constant we get $\gamma_i = \tau_m \, \mu_h$ for $\tau_m \leqslant 1$ and $\gamma_i = \mu_h/\tau_m$ for $\tau_m > 1$.
Setting $\mathbf{w} = \mathbf{B}_i(\tau_m)\mathbf{v} \Leftrightarrow \mathbf{v} = \mathbf{B}_i(\tau_m)^{-1}\mathbf{w}$, from the proven estimate $\|\mathbf{B}_i(\tau_m)\mathbf{v}\|_2 \geqslant \gamma_i \, \|\mathbf{v}\|_2$ we get $1/\gamma_i \geqslant \|\mathbf{B}_i(\tau_m)^{-1}\mathbf{w}\|_2/\|\mathbf{w}\|_2$ for $\mathbf{w} \neq \mathbf{0}$. Taking the maximum over all $\mathbf{w} \neq \mathbf{0}$ finally gives (5.4). □

As a side result of the previous theorem we obtain:

Corollary 5.2 *Under the assumptions of Theorem* 5.1 $\|\mathbf{B}_i(\tau_m)^{-1}\|_2$ *is bounded for all* $\tau_m \in [0, \tau_i(\mathbf{x}_h^m)]$, *where*

$$\tau_i(\mathbf{x}_h^m) := \begin{cases} T - t_{m-1}, & \text{if } \mathbf{S}_h = \mathbf{A}_h + \mathbf{DG}_h(\mathbf{x}_h^{(m,i)}), \\ \dfrac{\lambda_{\min}(\mathbf{M}_h)^2}{a_{ii}^2 \, \|\mathbf{S}_h - \left(\mathbf{A}_h + \mathbf{DG}_h(\mathbf{x}_h^{(m,i)})\right)\|_2^2}, & \text{else.} \end{cases}$$

Proof. The function $f(\tau_m) := \|\mathbf{B}_i(\tau_m)^{-1}\|_2$ is defined on $[0, \tau_i(\mathbf{x}_h^m)]$, since in addition to the result of the previous theorem

$$\mathbf{B}_i(0) = \begin{pmatrix} \mathbf{M}_h & \mathbf{0} \\ a_{ii} \, (\mathbf{A}_h + \mathbf{DG}_h(\mathbf{x}_h^{(m,i)})) & -a_{ii} \, \mathbf{M}_h \end{pmatrix}$$

is invertible with

$$\begin{aligned} \|\mathbf{B}_i(0)^{-1}\|_2 &\leqslant (1 + \frac{1}{a_{ii}}) \, \|\mathbf{M}_h^{-1}\|_2 + \|\mathbf{M}_h^{-1}\|_2^2 \, \|\mathbf{A}_h + \mathbf{DG}_h(\mathbf{x}_h^{(m,i)})\|_2 \\ &= \frac{1}{\lambda_{\min}(\mathbf{M}_h)} \left(\left(1 + \frac{1}{a_{ii}}\right) + \frac{\|\mathbf{A}_h + \mathbf{DG}_h(\mathbf{x}_h^{(m,i)})\|_2}{\lambda_{\min}(\mathbf{M}_h)} \right) =: C_i(\mathbf{x}_h). \end{aligned}$$

The function f is also continuous, since $\mathbf{B}_i(\tau_m)$ depends continuously on τ_m and $\mathbf{B}_i(\tau_m) \mapsto \mathbf{B}_i(\tau_m)^{-1}$ is continuous as well. Therefore, f attains its maximum over $[0, \tau_i(\mathbf{x}_h^m)]$ in some τ_0, and we get for all $\tau_m \in [0, \tau_i(\mathbf{x}_h^m)]$

$$\left\|\mathbf{B}_i(\tau_m)^{-1}\right\|_2 \leqslant \beta_i(\mathbf{x}_h^m) := \begin{cases} C_i(\mathbf{x}_h) & \text{for } \tau_0 = 0 \\ \dfrac{2}{\lambda_{\min}(\mathbf{M}_h)} \, \dfrac{1}{\tau_0} & \text{for } 0 < \tau_0 \leqslant 1 \\ \dfrac{2}{\lambda_{\min}(\mathbf{M}_h)} \, \tau_0 & \text{for } \tau_0 \geqslant 1 \end{cases} . \tag{5.8}$$

□

Remark 5.3 *As an alternative to the 'global' Newton method* (5.1a),(5.1b) *it is possible to solve*

$$\mathbf{F}_{h,i}(\mathbf{x}_h^m) = \mathbf{0}$$

with given $(\mathbf{u}_h^{(m,1)}, \mathbf{w}_h^{(m,1)}, \ldots, \mathbf{u}_h^{(m,i-1)}, \mathbf{w}_h^{(m,i-1)})$ *and unknowns* $(\mathbf{u}_h^{(m,i)}, \mathbf{w}_h^{(m,i)})$ *for* $i = 1, 2, \cdots, s$ *one after another by Newton's method, which also takes advantage of the staggered structure of* (4.5). *However, we have chosen the 'global' Newton formulation for a good reason: when it comes to the adaptive stepsize selection in the next subsection, we rely on the global information* $\Delta \mathbf{x}_h^{m,(0)}$.

5.2. Adaptive time step size selection

The system (4.5) can be seen as a parameter dependent nonlinear system with the time step size as the parameter. The vector

$$\mathbf{x}_h^{m-1}(\tau = 0) := (\mathbf{0}, \mathbf{w}_h^{(m-1,1)}, \ldots, \mathbf{0}, \mathbf{w}_h^{(m-1,s)}, \mathbf{c}_h^{m-1})$$

is the solution to $\mathbf{F}_h(\mathbf{x}_h, 0) = \mathbf{0}$. From Corollary 5.2 we know that $\mathbf{DF}_h(\mathbf{x}_h^{m-1}(0), 0)$ is invertible. Then the implicit function theorem guarantees the existence of a homotopy path $\mathbf{x}_h^{m-1} : [0, \tau^*] \to \Lambda \subset \mathbb{R}^{2N_h s + N_h}$ such that $\mathbf{F}_h(\mathbf{x}_h^{m-1}(\tau), \tau) = \mathbf{0}$ for all $\tau \in [0, \tau^*]$ for some $\tau^* = \tau^*(\mathbf{c}_h^{m-1}) > 0$. The parameter dependent nonlinear system can be solved by a predictor-corrector continuation strategy featuring an adaptive choice of the time step size. Having determined the solution at time instant t_{m-1}, one proceeds to the next time instant t_m by providing a predicted solution $\hat{\mathbf{x}}_h(\tau_m)$ that serves as an initial guess for Newton's method. The simplest possible choice for a continuation step is constant continuation, i.e., $\hat{\mathbf{x}}_h(\tau_m) := \mathbf{x}_h^{m-1}(0)$ which is in agreement with (5.2). Taylor expansion gives the approximation error

$$\|\hat{\mathbf{x}}_h(\tau_m) - \mathbf{x}_h^{m-1}(\tau_m)\| \leqslant \eta\ \tau_m, \quad \eta := \max_{\tau \in [0,\tau^*]} \left\| \frac{d\mathbf{x}_h^{m-1}}{d\tau}(\tau) \right\|_2 . \tag{5.9}$$

The predictor is based theoretically on the simplified Newton method. We recall that for a given start iterate $\mathbf{x}_h^{(m,0)}$ the simplified Newton method is of the form

$$\mathbf{DF}_h(\mathbf{x}_h^{(m,0)}, \tau_m)\, \overline{\Delta \mathbf{x}_h}^{(m,\ell)} = -\mathbf{F}_h(\mathbf{x}_h^{(m,\ell)}, \tau_m), \tag{5.10a}$$

$$\mathbf{x}_h^{(m,\ell+1)} = \mathbf{x}_h^{(m,\ell)} + \overline{\Delta \mathbf{x}_h}^{(m,\ell)}, \quad \ell = 0, 1, \ldots . \tag{5.10b}$$

Theorem 5.4. *Let* $\mathbf{DF}_h(\mathbf{x}_h, \tau)$ *be nonsingular for all* $(\mathbf{x}_h, \tau) \in \Lambda \times [0, \tau^*]$ *and assume that the homotopy path* $\mathbf{x}_h^{m-1} : [0, \tau^*] \to \Lambda$ *exists. Further, suppose that there exists a local Lipschitz constant* $\omega = \omega(\mathbf{c}_h^{m-1})$ *such that for all* $\mathbf{x}_h \in \Lambda$ *and for all* $0 \leqslant \tau \leqslant \tau^*$, *the affine covariant Lipschitz condition*

$$\|\mathbf{DF}_h(\hat{\mathbf{x}}_h(\tau), \tau)^{-1} \left(\mathbf{DF}_h(\mathbf{x}_h, \tau) - \mathbf{DF}_h(\hat{\mathbf{x}}_h(\tau), \tau)\right)\|_2 \leqslant \omega \, \|\mathbf{x}_h - \hat{\mathbf{x}}_h(\tau)\|_2 \tag{5.11}$$

holds true. Then, the simplified Newton method (5.10) *with initial guess* $\hat{\mathbf{x}}_h(\tau) := \mathbf{x}_h^{m-1}(0)$ *converges towards the solution* $\mathbf{x}_h^{m-1}(\tau)$ *for all stepsizes*

$$0 \leqslant \tau \leqslant \tau_m^{\max} := \frac{\sqrt{2}-1}{\omega\ \eta}. \tag{5.12}$$

Proof. We refer to Corollary 5.5 in [6]. □

The following result shows that for the nonlinear system (4.5) under consideration the affine covariant Lipschitz condition (5.11) from Theorem 5.4 is satisfied.

Theorem 5.5. *In addition to the conditions from Theorem* 5.1 *assume that* $\mathbf{G}_h$ *is Lipschitz-continuously differentiable with Lipschitz constant* $L_{\mathbf{DG}_h}$ *for* $\mathbf{DG}_h$. *Let*

$$0 \leqslant \tau \leqslant \tau(\hat{\mathbf{x}}_h^m) := \min_{1 \leqslant i \leqslant s} \tau_i(\hat{\mathbf{x}}_h^m).$$

Then, the nonlinear mapping $\mathbf{F}_h(\cdot,\tau)$ *from* (4.5) *satisfies the affine covariant Lipschitz condition* (5.11) *from Theorem* 5.4. *Further,* ω *can be estimated by the* τ*-independent upper bound*

$$\omega \leqslant \Big(1 + \varkappa_s\ (1 + s\ \|\mathbf{d}\|_\infty)\Big)\ L_{\mathbf{DG}_h}\ s\ \|\mathscr{A}\|_\infty, \tag{5.13}$$

where $\mathbf{d} = \mathscr{A}^{-T}\mathbf{b}$ *stems from section* 4. *Moreover,* $\varkappa_s$ *can be estimated recursively by*

$$\varkappa_i \leqslant \varkappa_{i-1}\ \Big(1 + \beta_i(\mathbf{x}_h^{m-1}(0))\ \Big(\tau(\mathbf{x}_h^{m-1}(0))\ \|\mathbf{S}_h\|_2 + \|\mathbf{A}_h + \mathbf{DG}(c_h^{m-1})\|_2 + \|\mathbf{M}_h\|_2\Big)\ \sum_{j=1}^{i-1} |a_{ij}|\Big) + \beta_i(\mathbf{x}_h^{m-1}(0)),\ 2 \leqslant i \leqslant s,$$

where $\varkappa_1 = \beta_1(\mathbf{x}_h^{m-1}(0))$ *(cf.* (5.8) *for the definition of* β_i*).*

Proof. Dropping the argument τ in $\mathbf{F}_h$ for notational simplicity, we split the left-hand side of (5.11) into

$$\begin{aligned} &\|\mathbf{DF}_h(\hat{\mathbf{x}}_h(\tau))^{-1}\ (\mathbf{DF}_h(\mathbf{x}_h) - \mathbf{DF}_h(\hat{\mathbf{x}}_h(\tau)))\|_2 \\ &\leqslant \|\mathbf{DF}_h(\hat{\mathbf{x}}_h(\tau))^{-1}\|_2\ \|\mathbf{DF}_h(\mathbf{x}_h) - \mathbf{DF}_h(\hat{\mathbf{x}}_h(\tau))\|_2. \end{aligned} \tag{5.14}$$

We note that $\mathbf{DF}_h(\mathbf{x}_h) - \mathbf{DF}_h(\hat{\mathbf{x}}_h(\tau))$ has a lower triangular block structure and its

norm can be estimated according to

$$\|\mathbf{DF}_h(\mathbf{x}_h)-\mathbf{DF}_h(\hat{\mathbf{x}}_h(\tau))\|_2 \leqslant \sum_{i=1}^{s}\sum_{j=1}^{i}|a_{ij}|\,\|\mathbf{DG}_h(\mathbf{c}_h^{m-1}+\mathbf{u}_h^{(\cdot,j)})-\mathbf{DG}_h(\mathbf{c}_h^{m-1}+\mathbf{0})\|_2$$

$$\leqslant\ L_{\mathbf{DG}_h}\sum_{i=1}^{s}\sum_{j=1}^{i}|a_{ij}|\,\|\mathbf{u}_h^{(\cdot,j)}-\mathbf{0}\|_2\ \leqslant L_{\mathbf{DG}_h}\,s\,\|\mathscr{A}\|_\infty\,\|\mathbf{x}_h-\hat{\mathbf{x}}_h(\tau)\|_2.$$

This gives the $\|\mathbf{x}_h-\hat{\mathbf{x}}_h(\tau)\|_2$-term in (5.11) and the last three factors in (5.13). It remains to estimate $\|\mathbf{DF}_h(\hat{\mathbf{x}}_h(\tau))^{-1}\|$. For a regular matrix of the form

$$\mathscr{A}^{(i)}=\begin{pmatrix}\mathscr{A}^{(i-1)} & \mathbf{0}\\ \mathscr{A}_{i,i-1} & \mathscr{A}_{ii}\end{pmatrix},\quad i\geqslant 2,$$

there holds

$$\mathscr{A}^{(i)^{-1}}=\begin{pmatrix}\mathscr{A}^{(i-1)^{-1}} & \mathbf{0}\\ -\mathscr{A}_{ii}^{-1}\,\mathscr{A}_{i,i-1}\,\mathscr{A}^{(i-1)^{-1}} & \mathscr{A}_{ii}^{-1}\end{pmatrix}. \tag{5.15}$$

Then one can estimate

$$\begin{aligned}\|\mathscr{A}^{(i)^{-1}}\|_2 &\leqslant \|\mathscr{A}^{(i-1)^{-1}}\|_2+\|\mathscr{A}_{ii}^{-1}\,\mathscr{A}_{i,i-1}\,\mathscr{A}^{(i-1)^{-1}}\|_2+\|\mathscr{A}_{ii}^{-1}\|_2\\ &\leqslant \|\mathscr{A}^{(i-1)^{-1}}\|_2\,\big(1+\|\mathscr{A}_{ii}^{-1}\|_2\,\|\mathscr{A}_{i,i-1}\|_2\big)+\|\mathscr{A}_{ii}^{-1}\|_2=:\varkappa_i.\end{aligned}$$

Setting $\mathscr{A}^{(1)}:=\mathscr{A}_{11}$ and $\varkappa_1:=\|\mathscr{A}^{(1)^{-1}}\|$, inductively one finds

$$\varkappa_i\leqslant\varkappa_{i-1}\,\big(1+\|\mathscr{A}_{ii}^{-1}\|_2\,\|\mathscr{A}_{i,i-1}\|_2\big)+\|\mathscr{A}_{ii}^{-1}\|_2.$$

Observing and combining

$$\begin{aligned}&\|\mathscr{A}_{ii}\|_2:=\|\mathbf{D}_i\mathbf{F}_{h,i}(\hat{\mathbf{x}}_h(\tau))\|_2\ \leqslant\ \beta_i(\mathbf{x}_h^{m-1}(0)),\ 1\leqslant i\leqslant s,\\ &\|\mathscr{A}_{s+1,s+1}\|_2:=\|-\mathbf{I}_{N_h}\|_2=1,\\ &\|\mathscr{A}_{i,i-1}\|_2:=\|(\mathbf{D}_1\mathbf{F}_{h,i}(\hat{\mathbf{x}}_h(\tau)),\cdots,\mathbf{D}_{i-1}\mathbf{F}_{h,i}(\hat{\mathbf{x}}_h(\tau)))\|_2\\ &\leqslant\ (\tau(\mathbf{x}_h^{m-1}(0))\,\|\mathbf{S}_h\|_2+\|\mathbf{A}_h+\mathbf{DG}_h(\mathbf{c}_h^{m-1})\|_2+\|\mathbf{M}_h\|_2)\sum_{j=1}^{i-1}|a_{ij}|,\ 2\leqslant i\leqslant s,\\ &\|\mathscr{A}_{s+1,s}\|_2:=\|(d_1\mathbf{I}_{N_h},\mathbf{0},\cdots,d_s\mathbf{I}_{N_h},\mathbf{0}))\|_2\leqslant s\,\|\mathbf{d}\|_\infty,\end{aligned}$$

we obtain the first factor of the upper bound in (5.13). □

From the simplified Newton method we compute the contraction factors

$$\Theta_\ell(\tau_m):=\frac{\|\overline{\Delta\mathbf{x}_h}^{(m,\ell+1)}\|_2}{\|\overline{\Delta\mathbf{x}_h}^{(m,\ell)}\|_2},\quad \ell\geqslant 0, \tag{5.16}$$

which will serve as convergence monitors for the convergence of the predictor-corrector continuation strategy. It can be shown (cf., for instance, [18]), that $\Theta_0(\tau_m)$ satisfies the estimate

$$\frac{2\,\Theta_0(\tau_m)}{\|\overline{\Delta \mathbf{x}_h}^{m,(0)}\|_2} \leqslant \omega, \tag{5.17}$$

which provides a lower bound for the critical local Lipschitz constant ω (cf. Theorem 5.5). Along with η from (5.9), this is the second key quantity for an adaptive time step size selection strategy. In order to exploit the steplength criterion (5.12) in an algorithmic realization, we compute estimates $[\cdot]$ of the a priori unknown constants and apply the steplength criterion with ω, η replaced by $[\omega]$, $[\eta]$.

5.2.1. τ-prediction strategy. We can use estimate (5.17) to obtain

$$\frac{2\,\Theta_0(\tau_m)}{\|\overline{\Delta \mathbf{x}_h}^{(m,0)}\|_2} =: [\omega] \leqslant \omega$$

as an estimate for ω and likewise

$$\frac{\|\mathbf{x}_h^{(m,0)} - \mathbf{x}_h^m\|_2}{\tau_m} =: [\eta] \leqslant \eta,$$

which obviously provides a lower bound of η due to (5.9). Inserting these computationally available quantities into (5.12) instead of ω and η results in

$$\tau_{m+1} := \tau_{m+1,0} := \frac{(\sqrt{2}-1)\,\|\overline{\Delta \mathbf{x}_h}^{(m,0)}\|_2}{2\,\Theta_0(\tau_m)\,\|\mathbf{x}_h^{(m,0)} - \mathbf{x}_h^m\|_2}\,\tau_m. \tag{5.18}$$

This formula predicts the next time increment τ_{m+1} adaptively based upon information about local and global constants of the homotopy gathered within the last Newton correction step. By nature of the constants ω, η, our approximations $[\omega]$, $[\eta]$ represent *lower* bounds. Thus (5.18) will in general overestimate the true maximal steplength $\tau_{m+1}^{\max}$. This explains why also a correction formula for τ_{m+1} is required.

5.2.2. τ-correction strategy. If convergence failure of the Newton correction step for $\tau_{m+1,j}$ occurs, we need to correct (decrease) the time step $\tau_{m+1,j}$ and repeat the last $\mathbf{x}_h$-prediction step with an adaptively reduced stepsize $\tau_{m+1,j+1}$. Here, the quantity $\Theta_0(\tau_{m+1,j})$ is available from the last unsuccessful Newton correction. It can be exploited to gain refined information about the crucial quantity $\omega\,\eta$ from (5.12). This leads to the τ-correction formula (for details see [18])

$$\tau_{m+1,j+1} := \frac{\sqrt{2}-1}{\sqrt{4\,\Theta_{\ell(j)}(\tau_{m+1,j})+1}-1}\,\tau_{m+1,j}, \quad j \geqslant 0.$$

6. Numerical results

We consider the initial-boundary value problem (2.1a)-(2.1c) in $Q := \Omega \times (0,T]$ with $\Omega := [0,L]^2, L := 1.0 \cdot 10^{-4} m$, and $T := 1.0 \cdot 10^{+1} s$. The physical parameters $\beta, \varkappa, \sigma$, and a_0, a_2, h_0, M are given in Table 2 in their physical units. We use the reference quantities

$$L_{ref} := 1.0 \cdot 10^{-5} m, \quad T_{ref} := 1.0 \cdot 10^{-2} s, \quad \sigma_{ref} := 1.0\, Jm^{-2} \tag{6.1}$$

and scale all parameters to dimensionless form. The values of the parameters in dimensionless form are also listed in Table 2. The initial concentration c_0 has been chosen randomly.

Table 2. Physical parameters in the sixth order Cahn Hilliard equation

Symbol	Value	Unit	Dimensionless Value
σ	1.0	Jm^{-2}	1.0
β	5.0	Jm^{-2}	5.0
h_0	$5.0 \cdot 10^{-1}$	1	$5.0 \cdot 10^{-1}$
M	$1.0 \cdot 10^{-13}$	$m^2 s^{-1}$	$1.0 \cdot 10^{-3}$
$\varkappa$	$1.0 \cdot 10^{-25}$	Jm^2	$1.0 \cdot 10^{-1}$
a_0	$-4.0 \cdot 10^{-12}$	J	-4.0
a_2	$1.0 \cdot 10^{-12}$	J	1.0

We have implemented the C^0IPDG method with $k = 2$ combined with a 2-stage SDIRK method of order 2 (cf. Table 3), the C^0IPDG method with $k = 3$ combined with a 3-stage SDIRK method of order 3 (cf. Table 4) and the C^0IPDG method with $k = 4$ combined with a 3-stage SDIRK method of order 4 (cf. Table 5).

Table 3. Butcher scheme of a 2-stage SDIRK method of order 2

$$\begin{array}{c|cc} \varkappa & \varkappa & 0 \\ 1 & 1-\varkappa & \varkappa \\ \hline & 1-\varkappa & \varkappa \end{array} \qquad \varkappa = 1 \pm \tfrac{1}{2}\sqrt{2}$$

Table 4. Butcher scheme of a 3-stage SDIRK method of order 3

α	α	0	0
$\frac{1+\alpha}{2}$	$\frac{1-\alpha}{2}$	α	0
1	b_0	b_1	α
	b_0	b_1	α

where $\alpha \approx 0.44$ is the root of $p(x) = x^3 - 3x^2 + \frac{3}{2}x - \frac{1}{6}$, $b_0 = -\frac{6\alpha^2-16\alpha+1}{4}$, and $b_1 = \frac{6\alpha^2-20\alpha+5}{4}$ (cf. [1]).

Table 5. Butcher scheme of a 3-stage SDIRK method of order 4

$(1+\varkappa)/2$	$(1+\varkappa)/2$	0	0	
$\frac{1}{2}$	$-\varkappa/2$	$(1+\varkappa)/2$	0	
$(1-\varkappa)/2$	$1+\varkappa$	$-(1+2\varkappa)$	$(1+\varkappa)/2$	$\varkappa = 2\cos(\pi/18)/\sqrt{3}$
	$1/(6\varkappa^2)$	$1-1/(3\varkappa^2)$	$1/(6\varkappa^2)$	

The associated spatial grid sizes are given in Table 6.

We note that the 2-stage SDIRK method of order $p = 2$ (cf. Table 3) and the 3-stage SDIRK method of order $p = 3$ (cf. Table 4) are both strongly stable, whereas the 3-stage SDIRK method of order $p = 4$ (cf. Table 5) is not strongly stable, but still A-stable (cf.,e.g.,[1]).

Figure 1 shows a visualization of the microemulsification process obtained by the numerical solution of the sixth order Cahn-Hilliard equation using C^0IPDG with $k = 2$ and 2-stage SDIRK with $p = 2$. The pure water phase ($c = 1$) is depicted in dark blue, the pure oil phase ($c = -1$) in dark red, and the microemulsion phase ($c = 0$) in light green. In Figure 1 (right), which represents the microemulsification process after $t = 3.86$, the formation of oil-in-water and water-in-oil droplets are clearly visible.

Table 6. Polynomial degree k , grid size, stage s and order p in the C^0IPD/SDIRK approach

C0IPDG		$(s,p)-$*SDIRK*	
Pol. Degr. k	Grid	Stage s	Order p
2	32×32	2	2
3	32×32	3	3
4	32×32	3	4

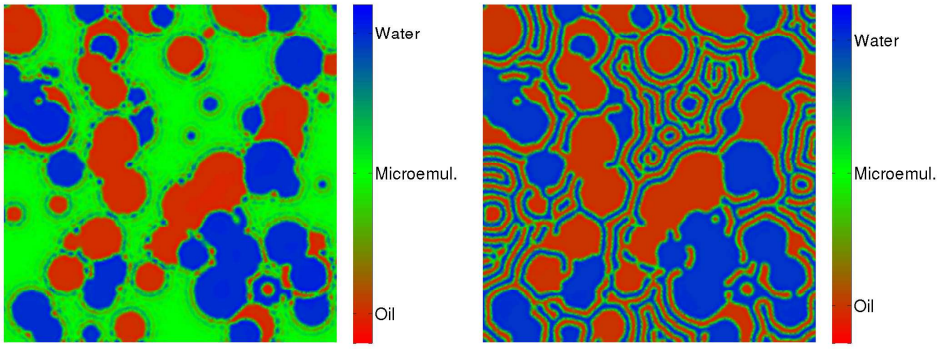

Figure 1. Formation of oil-in-water and water-in-oil droplets after $t = 0.60$ (left) and $t = 3.86$ (right). C^0IPDG with $k = 2$ on a 128×128 grid and 2-stage SDIRK with $p = 2$.

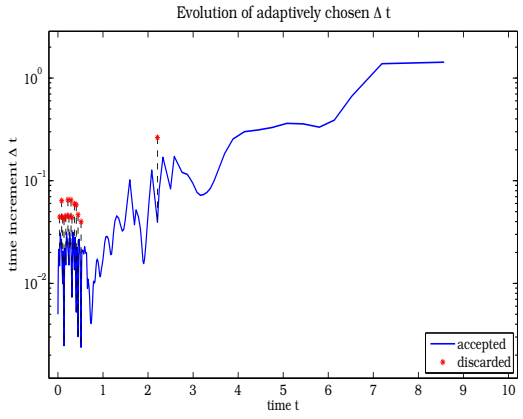

Figure 2. Evolution of the adaptively chosen time step sizes: C^0IPDG with $k = 2$ and 2-stage SDIRK with $p = 2$.

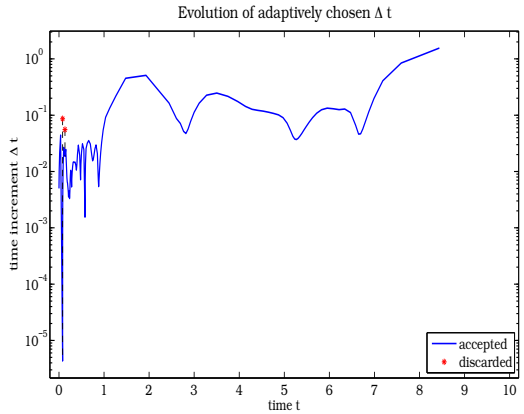

Figure 3. Evolution of the adaptively chosen time step sizes: C^{0}IPDG with $k = 3$ and 3-stage SDIRK with $p = 3$.

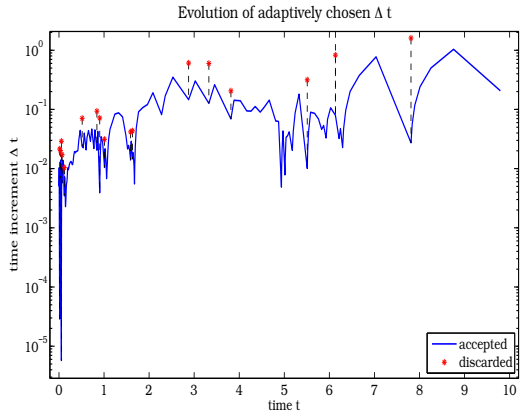

Figure 4. Evolution of the adaptively chosen time step sizes: C^{0}IPDG with $k = 4$ and 3-stage SDIRK with $p = 4$.

The performance of the predictor-corrector continuation strategy with an adaptive selection of the continuation parameter (time step size) is illustrated in Figure 2, Figure 3, and Figure 4. The figures display the evolution of the adaptively chosen time step sizes over the entire time interval $[0,10]$. Discarded time steps due to the τ-correction strategy as described at the end of section 5 are marked by stars. In particular, Figure 2 shows the result for C^{0}IPDG with $k = 2$ and 2-stage SDIRK with $p = 2$, Figure 3 the result for C^{0}IPDG with $k = 3$ and 3-stage SDIRK with $p = 3$, and Figure 4 the result for C^{0}IPDG with $k = 4$ and 3-stage SDIRK with $p = 4$. We see that the number of discarded time step sizes is significantly larger for C^{0}IPDG with $k = 4$ and 3-stage SDIRK with $p = 4$ than for the other two cases which is due to the strong stability of the 2-stage SDIRK method of order $p = 2$ and

the 3-stage SDIRK method of order $p = 2$, whereas the 3-stage SDIRK method of order $p = 4$ lacks strong stability.

7. Conclusions

Based on the mixed formulation of an initial-boundary value for a sixth order Cahn-Hilliard equation describing microemulsification processes we have considered high order space-time discretizations by C^0IPDG approximations in space with DG trial spaces composed of C^0 conforming Lagrangian finite elements of polynomial degree $k \geqslant 2$ and s-stage DIRK methods of order $p \geqslant 2$. The resulting parameter dependent nonlinear algebraic system has been solved by a predictor-corrector continuation strategy with an adaptive choice of the continuation parameter. For a ternary water-oil-microemulsion system, numerical results show the formation of water-in-oil and oil-in-water droplets and the performance of the approach by displaying the evolution of the adaptively chosen time step sizes over the entire time interval.

References

1. R. Alexander; Diagonally implicit Runge-Kutta methods for stiff O.D.E.'s. SIAM J. Numer. Anal. **14**, 1006–1021, 1977.
2. S.C. Brenner, T. Gudi, and L.-Y. Sung; An a posteriori error estimator for a quadratic C^0-interior penalty method for the biharmonic problem. IMA J. Numer. Anal., **30**, 777–798, 2010.
3. S.C. Brenner and L. Ridgway Scott; The Mathematical Theory of Finite Element Methods. 3rd Edition. Springer, New York, 2008
4. S.C. Brenner and L.-Y. Sung; C^0 interior penalty methods for fourth order elliptic boundary value problems on polygonal domains. J. Sci. Comput., **22/23**, 83–118, 2005.
5. J. Butcher; Numerical Methods for Ordinary Differential Equations. 2nd Edition. Wiley, Chichester-New York, 2008.
6. P. Deuflhard; Newton Methods for Nonlinear Problems - Affine Invariance and Adaptive Algorithms. Springer, Berlin-Heidelberg-New York, 2004.
7. G. Engel, K. Garikipati, T.J.R Hughes, M.G. Larson, L. Mazzei, and R.L. Taylor; Continuous/discontinuous finite element approximations of fourth order elliptic problems in structural and continuum mechanics with applications to thin beams and plates, and strain gradient elasticity. Comput. Methods Appl. Mech. Engrg., **191**, 3669–3750, 2002.
8. A. Ern and J.-L. Guermond; Theory and Practice of Finite Elements. Springer, Berlin-Heidelberg-New York, 2004.
9. T. Fraunholz, R.H.W. Hoppe, and M. Peter; Convergence analysis of an adaptive interior penalty discontinuous Galerkin method for the biharmonic problem. to appear in J. Numer. Math. **23**, Issue 4, 2015.
10. E.H. Georgoulis and P. Houston; Discontinuous Galerkin methods for the biharmonic problem. IMA J. Numer. Anal. **29**, 573–594, 2009.
11. E.H. Georgoulis, P. Houston, and J. Virtanen; An a posteriori error indicator for discontinuous Galerkin approximations of fourth order elliptic problems. IMA J. Numer. Anal. **31**, 281–298, 2011.

12. P.K. Ghosh and R.S.R. Murthy; Microemulsion potential drug delivery system. Current Drug Delivery **3**, 167–180, 2006.

13. G. Gompper and J. Goos; Fluctuating interfaces in microemulsion and sponge phases. Phys. Rev. E **50**, 1325–1335, 1994.

14. G. Gompper and M. Kraus; Ginzburg-Landau theory of ternary amphiphilic systems. I. Gaussian interface fluctuations. Phys. Rev. E **47**, 4289–4300, 1993.

15. G. Gompper and M. Kraus; Ginzburg-Landau theory of ternary amphiphilic systems. II. Monte Carlo simulations. Phys. Rev. E **47**, 4301–4312, 1993.

16. G. Gompper and S. Zschocke; Ginzburg-Landau theory of oil-water mixtures. Phys. Rev. A **46**, 4836–4851, 1992.

17. E. Hairer and G. Wanner; Solving Ordinary Differential Equations. II: Stiff and Differential-Algebraic Equations. 2nd Revised Edition. Springer, Berlin-Heidelberg-New York, 1996.

18. R.H.W. Hoppe and C. Linsenmann; An adaptive Newton continuation strategy for the fully implicit finite element immersed boundary method. J. Comp. Phys. **231**, 4676–4693, 2012.

19. S.K. Jha, R. Karki, D.P. Venkatesh, and A. Geethalakshami; Formulation development & characterization of microemulsion drug delivery systems containing antiulcer drug. Int. Journal of Drug Development & Research **3**, 336–343, 2011.

20. M.D. Korzec, P.L. Evans, A. Münch, and B. Wagner; Stationary solutions of driven fourth- and sixth-order Cahn-Hilliard type equations. SIAM J. Appl. Math. **69**, 348–374, 2008.

21. S.K. Mehta and G. Kaur; Microemulsions: Thermodynamics and Dynamics. In: Thermodynamics (M. Tadashi; ed.), pp. 381–406, InTech, 2011 (available at: http://www.intechopen/com/books/thermodynamics/microemulsions-thermodynamics-and-dynamics).

22. S.P. Moulik and A.K. Rashit; Physiochemistry and applications of microemulsions. J. Surf. Sci. Technol. 22, 159-186, 2006.

23. I. Pawlow and W. Zajaczkowski; A sixth order Cahn-Hilliard type equation arising in oil-water surfactant mixtures. Commun. Pure Appl. Anal. **10**, 1823–1847, 2011.

24. I. Pawlow and W. Zajaczkowski; On a class of sixth order viscous Cahn-Hilliard type equations. Discrete and Continuous Dynamical Systems, Series S, 6, 517-546, 2013.

25. L.M. Prince; Microemulsions in Theory and Practice. Academic Press, New York, 1977.

26. H.L. Rosano and M. Clausse; Microemulsion Systems. Marcel Dekker, New York, 1987.

27. G. Schimperna and I. Pawlow; On a class of Cahn-Hilliard models with nonlinear diffusion. SIAM J. Math. Anal. **45**, 31–63, 2013.

28. G.N. Wells, E. Kuhl, and K. Garikipati; A discontinuous Galerkin method for the Cahn-Hilliard equation. J. Comp. Phys. **218**, 860–877, 2006.