Journal of Computational Physics 383 (2019) 24–54

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

A family of second-order energy-stable schemes for Cahn–Hilliard type equations

Zhiguo Yang [a], Lianlei Lin [a,b], Suchuan Dong [a,*]

[a] *Center for Computational and Applied Mathematics, Department of Mathematics Purdue University, USA*
[b] *School of Electrical Engineering and Automation, Harbin Institute of Technology, China*

ARTICLE INFO

Article history:
Received 16 March 2018
Received in revised form 7 November 2018
Accepted 13 January 2019
Available online 25 January 2019

Keywords:
Cahn–Hilliard equation
Energy stability
Unconditional stability
Free energy
Phase field
Two-phase flow

ABSTRACT

We focus on the numerical approximation of the Cahn–Hilliard type equations, and present a family of second-order unconditionally energy-stable schemes. By reformulating the equation into an equivalent system employing a scalar auxiliary variable, we approximate the system at the time step $(n+\theta)$ (n denoting the time step index and θ is a real-valued parameter), and devise a family of corresponding approximations that are second-order accurate and unconditionally energy stable. This family of approximations contains the often-used Crank–Nicolson scheme and the second-order backward differentiation formula as particular cases. We further develop two efficient solution algorithms for the resultant discrete system of equations to overcome the difficulty caused by the unknown scalar auxiliary variable. Within each time step, our method requires only the solution of either four de-coupled individual Helmholtz type equations, or two separate individual systems with each system consisting of two coupled Helmholtz type equations. All the resultant linear algebraic systems involve only constant and time-independent coefficient matrices that can be pre-computed. A number of numerical examples are presented to demonstrate the performance of the family of schemes developed herein. We note that this family of second-order approximations can be readily applied to devise energy-stable schemes for other types of gradient flows when combined with the auxiliary variable approaches.

1. Introduction

Diffuse interface or phase field approach [2,30,29,42,36,31] has become one of the main techniques for modeling and simulating two-phase and multiphase problems involving fluid interfaces and the effect of surface tensions. Cahn–Hilliard type equations [6] are one of the commonly encountered equations in such models for describing the evolution of the phase field functions. Indeed, with an appropriate free energy density form, the mass balance equations for the individual fluid components in a multicomponent mixture will reduce to the Cahn–Hilliard type equations [23,22,1,8,11]. Devising efficient numerical schemes for Cahn–Hilliard equations therefore has crucial implications to two-phase and multiphase problems, and this has attracted a sustained interest from the community [37,40,5].

Nonlinearity and high spatial order (fourth order) are the main issues encountered when numerically solving the Cahn–Hilliard equations. The interfacial thickness scale parameter, when small, also exacts high mesh resolutions in numerical simulations. The energy stability property of a numerical scheme, when the computational cost is manageable, is a desir-

[*] Corresponding author.
E-mail address: sdong@purdue.edu (S. Dong).

https://doi.org/10.1016/j.jcp.2019.01.014
0021-9991/© 2019 Elsevier Inc. All rights reserved.

able feature for solving the Cahn–Hilliard equations. While other types (e.g. conditionally stable, semi-implicit) of schemes for the Cahn–Hilliard equation also exist in the literature (see e.g. [3,42,12,9,21,10], among others), in what follows we will focus on the energy-stable schemes in the review of literature.

The nonlinearity of Cahn–Hilliard equation is induced by the potential free energy density function. Ensuring discrete energy stability in a numerical scheme hinges on the treatment of the nonlinear term. Based on the strategies for treating the nonlinear term, existing energy stable schemes for the Cahn–Hilliard equations can be broadly classified into two categories: nonlinear schemes and linear schemes. Nonlinear schemes (see e.g. [14,19,18,27,32,17,39,24]) entail the solution of a system of nonlinear algebraic equations within a time step after discretization. Convex splitting of the potential energy term and its variants are a popular approach to treat the nonlinear term in this category [15,16]. Other treatments include the midpoint approximation [14,28], specially designed quadrature formulas [20], and Taylor expansion approximations [27] of the potential term, among others. On the other hand, linear schemes (see e.g. [34,21,41,33]) involve only the solution of a system of linear algebraic equations after discretization, thanks to certain special treatment of the nonlinear term or the introduction of certain auxiliary variables. Adding a stabilization term that is equivalent to zero while using a potential energy with bounded second derivative and treating the nonlinear term explicitly [34,21] is a widely used technique in this category. Using a Lagrange multiplier [4,21] is another technique to derive unconditionally energy-stable schemes for the Cahn–Hilliard equation. The invariant energy quadratization (IEQ) [41] is a general technique that generalizes the Lagrange multiplier approach and can be applied to a large class of free energy forms. IEQ introduces an auxiliary field function related to the square root of the potential free energy density function together with a dynamic equation for this auxiliary variable, and allows one to devise schemes to ensure the energy stability relatively easily. The IEQ method gives rise to a system of linear algebraic equations involving time-dependent coefficient matrices after discretization. A further development of the auxiliary variable strategy is introduced in [33] very recently, in which an auxiliary variable, which is a scalar value rather than a field function, related to the square root of the total potential energy integral has been employed. The scalar auxiliary variable (SAV) method retains the main advantage of IEQ, and further can lead to a constant coefficient matrix for the resultant linear algebraic system of equations after discretization.

In the current work we focus on the numerical approximation of the Cahn–Hilliard equation, reformulated using the scalar auxiliary variable approach. We present a family of second-order accurate linear schemes for the system, and show that this family of schemes is unconditionally energy-stable. This family of approximations contains the Crank–Nicolson scheme (or trapezoidal rule) and the second-order backward differentiation formula (BDF2) as particular cases. The key idea of the schemes lies in enforcing the system of equations at the time step $(n+\theta)$, where n is the time step index and θ is a real-valued parameter ($\frac{1}{2} \leqslant \theta \leqslant \frac{3}{2}$), and then devising appropriate corresponding approximations at $(n+\theta)$ with second-order accuracy that guarantee the energy stability of the system. We further present two efficient solution algorithms for the discretized system of equations to overcome the numerical difficulty induced by the unknown scalar auxiliary variable. Within each time step, the overall method requires the computation of four de-coupled individual Helmholtz type equations with one solution algorithm, and the computation of two separate individual systems with each consisting of two coupled Helmholtz type equations for the other solution procedure. All equations with these solution algorithms involve only constant coefficient matrices that can be pre-computed.

The new aspects of this paper include the following: (i) the family of second-order accurate energy-stable schemes, and (ii) the efficient solution algorithms for the discrete system resulting from the Cahn–Hilliard type equations.

While we only consider the numerical approximation of the Cahn–Hilliard equation in this work, we would like to point out that the family of second-order energy-stable approximations herein are general, and are readily applicable to other types of equations resulting from gradient flows for designing energy-stable schemes when combined with the IEQ or SAV strategy.

The rest of this paper is structured as follows. In Section 2 we discuss the SAV reformulation of the Cahn–Hilliard equation, and present the family of second-order energy-stable approximations of the reformulated system of equations. We will also present two solution algorithms for the discretized equations and the implementation of these algorithms based on the C^0-continuous spectral element method for spatial discretizations. In Section 3 we test the performance of the algorithms using several representative numerical examples, and demonstrate numerically the stability of computations with large time step sizes. Section 4 concludes the presentation with some closing remarks. The Appendices A and B provide proofs for the energy stability and another property of the presented family of schemes.

2. A family of second-order energy-stable schemes

2.1. Cahn–Hilliard equation, boundary conditions, and transformed system

Consider a domain Ω in two or three dimensions, whose boundary is denoted by $\partial\Omega$, and the Cahn–Hilliard equation [6] with a source term within this domain:

$$\frac{\partial \phi}{\partial t} = m\nabla^2\left[-\lambda\nabla^2\phi + h(\phi)\right] + g(\mathbf{x}, t) \tag{1}$$

where $\phi(\mathbf{x}, t)$ ($\phi \in [-1, 1]$) is the phase field function, $m > 0$ is the mobility and assumed to be a constant in this work, and $\mathbf{x}$ and t denote the spatial coordinates and time. $g(\mathbf{x}, t)$ is a prescribed source term for the purpose of numerical testing (for convergence rates) only, and will be set to $g(\mathbf{x},t)=0$ in actual simulations. λ is the mixing energy density coefficient and is related to other physical parameters, e.g. for two-phase flow problems it is given by (see [42]) $\lambda=\frac{3}{2\sqrt{2}}\sigma\eta$, where σ is the surface tension and η is the characteristic interfacial thickness. The nonlinear term $h(\phi)$ is given by $h(\phi)=\frac{dF(\phi)}{d\phi}$, where $F(\phi)$ is the potential free energy density function of the system. A double-well potential is often used for $F(\phi)$, in the form $F(\phi)=\frac{\lambda}{4\eta^2}(\phi^2-1)^2$, and this form will be used for the numerical tests in Section 3.

We consider the wall-type boundary with a neutral wettability (i.e. 90^0 contact angle) for $\partial\Omega$, characterized by the following boundary conditions [25,43,7]

$$\mathbf{n}\cdot\nabla\left[-\lambda\nabla^2\phi+h(\phi)\right]=g_a(\mathbf{x},t),\quad \text{on } \partial\Omega \tag{2a}$$

$$\mathbf{n}\cdot\nabla\phi=g_b(\mathbf{x},t),\quad \text{on } \partial\Omega \tag{2b}$$

where $\mathbf{n}$ is the outward-pointing unit vector normal to $\partial\Omega$. $g_a(\mathbf{x},t)$ and $g_b(\mathbf{x},t)$ are prescribed source terms for the purpose of numerical testing only, and will be set to $g_a=0$ and $g_b=0$ in actual simulations.

The system consisting of the equation (1) and the boundary conditions (2a) and (2b), with $g=0$, $g_a=0$ and $g_b=0$, satisfies the following energy balance equation:

$$\frac{\partial}{\partial t}\int_\Omega\left(\frac{\lambda}{2}\nabla\phi\cdot\nabla\phi+F(\phi)\right)=-m\int_\Omega\left|\nabla\left[-\lambda\nabla^2\phi+h(\phi)\right]\right|^2. \tag{3}$$

This system is to be supplemented by the following initial condition

$$\phi(\mathbf{x},0)=\phi_{in}(\mathbf{x}) \tag{4}$$

where $\phi_{in}(\mathbf{x})$ denotes the initial phase field distribution.

We next reformulate this system of equations and boundary conditions into a modified equivalent system, by introducing an auxiliary variable associated with the total potential energy in a way similar to [33]. Define the total potential free energy $E(t)$ by

$$E(t)=E[\phi]=C_0+\int_\Omega F(\phi) \tag{5}$$

where C_0 is a chosen constant such that $E(t)>0$ for all $0\leqslant t\leqslant T$ (T denoting the period of time to find the solution for). For all the numerical experiments presented in the current work we employ $C_0=0$ in the simulations. It is important to note that $E[\phi]$ as defined above is a scalar value, not a field function. We define an auxiliary variable $r(t)$ by

$$r(t)=\sqrt{E(t)}. \tag{6}$$

Then $r(t)$ satisfies the dynamic equation

$$\frac{dr}{dt}=\frac{1}{2\sqrt{E[\phi]}}\int_\Omega h(\phi)\frac{\partial\phi}{\partial t}. \tag{7}$$

We re-write the Cahn–Hilliard equation (1) into an equivalent form

$$\frac{\partial\phi}{\partial t}=m\nabla^2\mathcal{H}+g, \tag{8a}$$

$$\mathcal{H}=-\lambda\nabla^2\phi+\frac{r(t)}{\sqrt{E[\phi]}}h(\phi), \tag{8b}$$

where the definition (6) has been used. The boundary conditions (2a) is also re-written into an equivalent form

$$\mathbf{n}\cdot\nabla\left[-\lambda\nabla^2\phi+\frac{r(t)}{\sqrt{E[\phi]}}h(\phi)\right]=g_a,\quad \text{on } \partial\Omega, \tag{9a}$$

The system consisting of equations (8a)–(8b) and (7), the boundary conditions (9a) and (2b), and the initial conditions of (4) and the following

$$r(0)=\sqrt{E[\phi_{in}]}=\left[\int_\Omega F(\phi_{in})+C_0\right]^{1/2}, \tag{10}$$

is equivalent to the original system consisting of equations (1), (2a)–(2b) and (4).

2.2. A family of second-order energy-stable approximations

We focus on the numerical approximation of the reformulated equivalent system consisting of equations (8a)–(8b) and (7), the boundary conditions (9a) and (2b), and the initial conditions (4) and (10). We present a family of second-order energy-stable schemes for this system that allows for very efficient implementations.

Let $n \geqslant 0$ denote the time step index, and $(\cdot)^n$ represent the variable $(\cdot)$ at time step n, corresponding to the time $t = n\Delta t$, where Δt is the time step size.

Let θ ($\frac{1}{2} \leqslant \theta \leqslant \frac{3}{2}$) denote a real-valued parameter. We approximate the variables at time step $(n+\theta)$ (corresponding to time $(n+\theta)\Delta t$) as follows with a scheme of second-order accuracy in time (χ denoting a generic variable below),

$$\left.\frac{\partial \chi}{\partial t}\right|^{n+\theta} = \frac{1}{\Delta t}\left[\left(\theta+\frac{1}{2}\right)\chi^{n+1} - 2\theta\chi^n + \left(\theta-\frac{1}{2}\right)\chi^{n-1}\right], \tag{11a}$$

$$\chi^{n+\theta} = \left[\theta\left(\frac{5}{2}-\theta\right)-\frac{1}{2}\right]\chi^{n+1} + 2(1-\theta)^2\chi^n + \left(\theta-\frac{1}{2}\right)(1-\theta)\chi^{n-1}, \tag{11b}$$

$$\bar{\chi}^{n+\theta} = (1+\theta)\chi^n - \theta\chi^{n-1}. \tag{11c}$$

In the above $\chi^{n+\theta}$ and $\bar{\chi}^{n+\theta}$ are respectively an implicit and an explicit approximation of χ at time step $(n+\theta)$. The 2nd-order temporal accuracy of these approximations can be verified by Taylor expansions in a straightforward fashion. The approximation (11a) is not new, and it has been studied in previous works (see e.g. [38]). The approximation (11b) seems to have not been explored before to the best of our knowledge. This approximation is critical to the energy stability of the current family of schemes, due to the following property:

$$\begin{aligned}
&\chi^{n+\theta}\left[\left(\theta+\frac{1}{2}\right)\chi^{n+1} - 2\theta\chi^n + \left(\theta-\frac{1}{2}\right)\chi^{n-1}\right] \\
&= \frac{1}{2}\theta\left(\theta-\frac{1}{2}\right)(3-2\theta)\left|\chi^{n+1}-2\chi^n+\chi^{n-1}\right|^2 + \frac{1}{2}\left(\frac{3}{2}-\theta\right)\left(\left|\chi^{n+1}\right|^2 - \left|\chi^n\right|^2\right) \\
&\quad + \frac{1}{2}\left(\theta-\frac{1}{2}\right)\left(\left|2\chi^{n+1}-\chi^n\right|^2 - \left|2\chi^n-\chi^{n-1}\right|^2\right).
\end{aligned} \tag{12}$$

This property can be verified by elementary manipulations. The family of approximations given by (11a)–(11c) contains the often-used 2nd-order backward differentiation formula (or BDF2, corresponding to $\theta = 1$) and the Crank–Nicolson approximation (corresponding to $\theta = 1/2$).

Given (ϕ^n, r^n), we compute (ϕ^{n+1}, r^{n+1}) by enforcing the system consisting of (7)–(9a) and (2b) at time step $(n+\theta)$ and using the above approximations (11a)–(11c), as follows,

$$\frac{\left(\theta+\frac{1}{2}\right)\phi^{n+1} - 2\theta\phi^n + \left(\theta-\frac{1}{2}\right)\phi^{n-1}}{\Delta t} = m\nabla^2\mathcal{H}^{n+\theta} + g^{n+\theta}, \tag{13a}$$

$$\mathcal{H}^{n+\theta} = -\lambda\nabla^2\phi^{n+\theta} + S(\phi^{n+1} - 2\phi^n + \phi^{n-1}) + \frac{r^{n+\theta}}{\sqrt{E[\bar{\phi}^{n+\theta}]}}h(\bar{\phi}^{n+\theta}), \tag{13b}$$

$$\frac{\left(\theta+\frac{1}{2}\right)r^{n+1} - 2\theta r^n + \left(\theta-\frac{1}{2}\right)r^{n-1}}{\Delta t} = \int_\Omega \frac{h(\bar{\phi}^{n+\theta})}{2\sqrt{E[\bar{\phi}^{n+\theta}]}}\,\frac{\left(\theta+\frac{1}{2}\right)\phi^{n+1} - 2\theta\phi^n + \left(\theta-\frac{1}{2}\right)\phi^{n-1}}{\Delta t}, \tag{13c}$$

$$\mathbf{n}\cdot\nabla\mathcal{H}^{n+\theta} = g_a^{n+\theta}, \quad \text{on } \partial\Omega, \tag{13d}$$

$$\mathbf{n}\cdot\nabla\phi^{n+\theta} = g_b^{n+\theta}, \quad \text{on } \partial\Omega. \tag{13e}$$

In the above discrete equation (13b), note that we have included an extra term, $S(\phi^{n+1} - 2\phi^n + \phi^{n-1})$, where $S \geqslant 0$ is a chosen constant, which, depending on the solution algorithm in the implementation, may need to satisfy a condition to be specified later in Section 2.3.1. Because $\phi^{n+1} - 2\phi^n + \phi^{n-1} = \mathcal{O}(\Delta t)^2$, this extra term does not spoil the second-order accuracy of the overall scheme. In the above equations the variables $\phi^{n+\theta}$, $\bar{\phi}^{n+\theta}$, and $r^{n+\theta}$ are given by the equations (11b) and (11c). $g^{n+\theta}$, $g_a^{n+\theta}$ and $g_b^{n+\theta}$ are the prescribed functions $g(\mathbf{x}, t)$, $g_a(\mathbf{x}, t)$ and $g_b(\mathbf{x}, t)$ evaluated at time $t = (n+\theta)\Delta t$, respectively.

The above scheme has the following property:

Theorem 2.1. *The scheme consisting of equations (13a)–(13e), in the absence of the source terms (i.e. $g = 0$, $g_a = 0$ and $g_b = 0$), satisfies the discrete energy balance equation for $\frac{1}{2} \leqslant \theta \leqslant \frac{3}{2}$ and $S \geqslant 0$,*

$$
\begin{aligned}
& W^{n+1} - W^n + \theta\left(\theta - \frac{1}{2}\right)(3-2\theta)\left(|r^{n+1} - 2r^n + r^{n-1}|^2\right. \\
& \quad \left. + \int_\Omega \frac{\lambda}{2}\left|\nabla\phi^{n+1} - 2\nabla\phi^n + \nabla\phi^{n-1}\right|^2\right) + \theta S \int_\Omega \left|\phi^{n+1} - 2\phi^n + \phi^{n-1}\right|^2 \\
& = -m\Delta t \int_\Omega |\nabla\mathcal{H}^{n+\theta}|^2
\end{aligned}
\tag{14}
$$

where

$$
\begin{aligned}
W^n = & \left(\frac{3}{2} - \theta\right)\left(|r^n|^2 + \int_\Omega \frac{\lambda}{2}\left|\nabla\phi^n\right|^2\right) + \left(\theta - \frac{1}{2}\right)\left(\left|2r^n - r^{n-1}\right|^2 + \int_\Omega \frac{\lambda}{2}\left|2\nabla\phi^n - \nabla\phi^{n-1}\right|^2\right) \\
& + \frac{S}{2}\int_\Omega \left|\phi^n - \phi^{n-1}\right|^2 .
\end{aligned}
\tag{15}
$$

A proof of this theorem is provided in the Appendix A.

Remark 1. Based on this theorem the scheme given by equations (13a)–(13e) constitutes a family of unconditionally stable algorithms for $\frac{1}{2} \leqslant \theta \leqslant \frac{3}{2}$. Note that this energy stability property holds regardless of the specific form of the potential free energy density $F(\phi)$, as long as it is such that an appropriate constant C_0 in (5) can be chosen to ensure $E(t) > 0$ for $0 \leqslant t \leqslant T$.

Remark 2. The first term on the right hand side of (12) determines the dissipativeness of the approximations (11a)–(11b). The algorithm with $\theta = \frac{2}{3} + \frac{\sqrt{7}}{6} \approx 1.11$ is the most dissipative among this family of approximations, while both $\theta = \frac{1}{2}$ (Crank–Nicolson) and $\theta = \frac{3}{2}$ are non-dissipative (energy-conserving). In terms of numerical dissipation, BDF2 ($\theta = 1$) is close to, but not as dissipative as, the scheme with $\theta = \frac{2}{3} + \frac{\sqrt{7}}{6}$.

Remark 3. The incorporation of the extra zero term, $S(\phi^{n+1} - 2\phi^n + \phi^{n-1})$, is not necessary for the proof of unconditional energy stability of the current family of schemes; see equations (14) and (15). But inclusion of such a term enables a very efficient solution procedure through a reformulation of the fourth-order equation into two de-coupled 2nd-order equations; see Section 2.3.1 below. Similar terms have been employed in previous works for approximating the phase field Allen–Cahn or Cahn–Hilliard equations (see e.g. [34,35], among others), where the magnitude of the second derivative of the double-well potential energy is assumed to be bounded from above and this extra term involves a coefficient proportional to this upper bound.

Remark 4. While we consider only the Cahn–Hilliard type equations in this work, the application of the family of 2nd-order approximations (11a)–(11c) is not limited to this class of equations. They can be readily applied to other types of equations describing gradient flows. For example, by combining these approximations and the auxiliary variable approaches of [41,33], one can derive a family of energy-stable schemes for a large class of gradient flows.

2.3. Efficient solution algorithm

The scheme represented by equations (13a)–(13e) involves integrals of the unknown field variable ϕ^{n+1} over the domain Ω. Despite this apparent complication, the formulation allows for simple and very efficient solution algorithms. We present two such algorithms below.

2.3.1. De-coupled algorithm

To facilitate subsequent discussions and make the representation more compact, we first introduce several notations (χ denoting a generic variable):

$$
\gamma_0 = \gamma_0(\theta) = \theta + \frac{1}{2}, \qquad \omega_0 = \omega_0(\theta) = \theta\left(\frac{5}{2} - \theta\right) - \frac{1}{2}, \tag{16a}
$$

$$
\hat{\chi} = 2\theta\chi^n - \left(\theta - \frac{1}{2}\right)\chi^{n-1}, \tag{16b}
$$

$$\tilde{\chi} = 2(1-\theta)^2\chi^n + \left(\theta - \frac{1}{2}\right)(1-\theta)\chi^{n-1}. \tag{16c}$$

Then the approximations in (11a) and (11b) can be written as

$$\left.\frac{\partial \chi}{\partial t}\right|^{n+\theta} = \frac{\gamma_0\chi^{n+1} - \hat{\chi}}{\Delta t}, \tag{17a}$$

$$\chi^{n+\theta} = \omega_0\chi^{n+1} + \tilde{\chi}. \tag{17b}$$

Combining equations (13a) and (13b) leads to

$$\frac{\gamma_0\phi^{n+1} - \hat{\phi}}{\Delta t} = m\nabla^2\left[-\lambda\nabla^2\left(\omega_0\phi^{n+1} + \tilde{\phi}\right) + S(\phi^{n+1} - \bar{\phi}^{n+1}) + \left(\omega_0 r^{n+1} + \tilde{r}\right)\frac{h(\bar{\phi}^{n+\theta})}{\sqrt{E[\bar{\phi}^{n+\theta}]}}\right] + g^{n+\theta} \tag{18}$$

where γ_0 and ω_0 are given by (16a), $\hat{\phi}$ is defined by (16b), $\tilde{\phi}$ and $\tilde{r}$ are defined by (16c), $\bar{\phi}^{n+1} = 2\phi^n - \phi^{n-1}$ based on equation (11c), and we have used equations (17a) and (17b). It is important to note that ϕ^{n+1}, $\hat{\phi}$, $\tilde{\phi}$, $\bar{\phi}^{n+\theta}$, $\bar{\phi}^{n+1}$ are field functions, while r^{n+1}, $\tilde{r}$ and $E[\bar{\phi}^{n+\theta}]$ are scalar-valued variables. Define

$$b^n = \frac{h(\bar{\phi}^{n+\theta})}{\sqrt{E[\bar{\phi}^{n+\theta}]}}, \qquad z^{n+1} = \int_\Omega b^n\phi^{n+1}. \tag{19}$$

Note that b^n is a field function, and z^{n+1} is a scalar-valued variable. Equation (18) is then transformed into

$$\begin{aligned}&\nabla^2(\nabla^2\phi^{n+1}) - \frac{S}{\lambda\omega_0}\nabla^2\phi^{n+1} + \frac{\gamma_0}{\lambda\omega_0 m\Delta t}\phi^{n+1}\\ &= \frac{1}{\lambda\omega_0 m}\left[g^{n+\theta} + \frac{\hat{\phi}}{\Delta t}\right] - \frac{S}{\lambda\omega_0}\nabla^2\bar{\phi}^{n+1} - \frac{1}{\omega_0}\nabla^2(\nabla^2\tilde{\phi}) + \frac{\omega_0 r^{n+1} + \tilde{r}}{\lambda\omega_0}\nabla^2 b^n.\end{aligned} \tag{20}$$

Equation (13c) can be written as $\gamma_0 r^{n+1} - \hat{r} = \frac{1}{2}\int_\Omega b^n(\gamma_0\phi^{n+1} - \hat{\phi})$, from which we get

$$r^{n+1} = \frac{1}{\gamma_0}\left(\hat{r} - \frac{1}{2}\int_\Omega b^n\hat{\phi}\right) + \frac{z^{n+1}}{2} \tag{21}$$

where z^{n+1} is defined in (19) and involves the unknown field function ϕ^{n+1}. In light of this expression, equation (20) is transformed into

$$\nabla^2(\nabla^2\phi^{n+1}) - \frac{S}{\lambda\omega_0}\nabla^2\phi^{n+1} + \frac{\gamma_0}{\lambda\omega_0 m\Delta t}\phi^{n+1} = Q^n + \frac{z^{n+1}}{2\lambda}\nabla^2 b^n, \tag{22}$$

where

$$Q^n = \frac{1}{\lambda\omega_0 m}\left[g^{n+\theta} + \frac{\hat{\phi}}{\Delta t}\right] - \frac{S}{\lambda\omega_0}\nabla^2\bar{\phi}^{n+1} - \frac{1}{\omega_0}\nabla^2(\nabla^2\tilde{\phi}) + \left[\frac{1}{\lambda\gamma_0}\left(\hat{r} - \frac{1}{2}\int_\Omega b^n\hat{\phi}\right) + \frac{\tilde{r}}{\lambda\omega_0}\right]\nabla^2 b^n. \tag{23}$$

Barring the unknown scalar variable z^{n+1}, equation (22) is a fourth-order equation about ϕ^{n+1}. The left-hand-side (LHS) of this equation can be reformulated into two de-coupled Helmholtz type equations (see e.g. [42,12,7]). By adding/subtracting a term $\alpha\nabla^2\phi^{n+1}$ (α denoting a constant to be determined) on the LHS of (22), we get

$$\nabla^2\left[\nabla^2\phi^{n+1} + \alpha\phi^{n+1}\right] - \left(\alpha + \frac{S}{\lambda\omega_0}\right)\left[\nabla^2\phi^{n+1} - \frac{\gamma_0}{\lambda\omega_0 m\Delta t\left(\alpha + \frac{S}{\lambda\omega_0}\right)}\phi^{n+1}\right] = Q^n + \frac{z^{n+1}}{2\lambda}\nabla^2 b^n. \tag{24}$$

By requiring $\alpha = -\frac{\gamma_0}{\lambda\omega_0 m\Delta t\left(\alpha + \frac{S}{\lambda\omega_0}\right)}$, we can determine the constant α,

$$\alpha = -\frac{S}{2\lambda\omega_0}\left[1 - \sqrt{1 - \frac{4\gamma_0}{\lambda\omega_0 m\Delta t}\left(\frac{\lambda\omega_0}{S}\right)^2}\right] \tag{25}$$

with the requirement

$$S \geqslant \lambda\omega_0\sqrt{\frac{4\gamma_0}{\lambda\omega_0 m\Delta t}}=\sqrt{\frac{4\gamma_0\lambda\omega_0}{m\Delta t}}. \tag{26}$$

This is the condition the chosen constant S should satisfy for the solution algorithm in this subsection.

Therefore, equation (24) can be transformed into the following equivalent form

$$\nabla^2\psi^{n+1}-\left(\alpha+\frac{S}{\lambda\omega_0}\right)\psi^{n+1}=Q^n+\frac{z^{n+1}}{2\lambda}\nabla^2 b^n, \tag{27a}$$

$$\nabla^2\phi^{n+1}+\alpha\phi^{n+1}=\psi^{n+1}, \tag{27b}$$

where ψ^{n+1} is an auxiliary field variable defined by (27b). Note that $\alpha<0$ and $\alpha+\frac{S}{\lambda\omega_0}>0$ under the condition (26) for S. It can also be noted that, if z^{n+1} is known, then the two equations (27a) and (27b) are not coupled. One can first solve (27a) for ψ^{n+1}, and then solve (27b) for ϕ^{n+1}. The unknown variable z^{n+1}, which depends on ϕ^{n+1}, causes a complication in the solution of this system.

Let us now turn the attention to the boundary conditions. Equation (13e) can be written as

$$\mathbf{n}\cdot\nabla\phi^{n+1}=\frac{1}{\omega_0}\left(g_b^{n+\theta}-\mathbf{n}\cdot\nabla\tilde{\phi}\right) \quad \text{on } \partial\Omega. \tag{28}$$

Equation (13d) can be re-written as

$$-\lambda\mathbf{n}\cdot\nabla\left(\omega_0\nabla^2\phi^{n+1}+\nabla^2\tilde{\phi}\right)+S\mathbf{n}\cdot\nabla\left(\phi^{n+1}-\bar{\phi}^{n+1}\right)+\left(\omega_0 r^{n+1}+\tilde{r}\right)\mathbf{n}\cdot\nabla b^n=g_a^{n+\theta}, \quad \text{on } \partial\Omega. \tag{29}$$

In light of the equations (27b) and (21), we can transform this equation into

$$\begin{aligned}\mathbf{n}\cdot\nabla\psi^{n+1}=&\left(\alpha+\frac{S}{\lambda\omega_0}\right)\mathbf{n}\cdot\nabla\phi^{n+1}-\frac{1}{\lambda\omega_0}\left(g_a^{n+\theta}+S\mathbf{n}\cdot\nabla\bar{\phi}^{n+1}\right)-\frac{1}{\omega_0}\mathbf{n}\cdot\nabla(\nabla^2\tilde{\phi})\\&+\left[\frac{\tilde{r}}{\lambda\omega_0}+\frac{1}{\lambda\gamma_0}\left(\hat{r}-\frac{1}{2}\int_\Omega b^n\hat{\phi}\right)+\frac{z^{n+1}}{2\lambda}\right]\mathbf{n}\cdot\nabla b^n, \quad \text{on } \partial\Omega.\end{aligned} \tag{30}$$

Substitute equation (28) into the above equation, and we have the boundary condition about ψ^{n+1}:

$$\mathbf{n}\cdot\nabla\psi^{n+1}=T^n+\frac{z^{n+1}}{2\lambda}\mathbf{n}\cdot\nabla b^n, \quad \text{on } \partial\Omega, \tag{31}$$

where

$$\begin{aligned}T^n=&\left(\alpha+\frac{S}{\lambda\omega_0}\right)\frac{1}{\omega_0}\left(g_b^{n+\theta}-\mathbf{n}\cdot\nabla\tilde{\phi}\right)-\frac{1}{\lambda\omega_0}\left(g_a^{n+\theta}+S\mathbf{n}\cdot\nabla\bar{\phi}^{n+1}\right)\\&-\frac{1}{\omega_0}\mathbf{n}\cdot\nabla(\nabla^2\tilde{\phi})+\left[\frac{\tilde{r}}{\lambda\omega_0}+\frac{1}{\lambda\gamma_0}\left(\hat{r}-\frac{1}{2}\int_\Omega b^n\hat{\phi}\right)\right]\mathbf{n}\cdot\nabla b^n.\end{aligned} \tag{32}$$

Therefore, the original system of equations (13a)–(13e) has been reduced to the system consisting of equations (27a)–(27b), (31) and (28). After ϕ^{n+1} is solved from this system, r^{n+1} can be computed based on equation (21).

To solve the system consisting of equations (27a)–(27b), (31) and (28), it is critical to note that the unknown variable z^{n+1} involved therein is a scalar number, not a field function, and that the equations are linear with respect to ϕ^{n+1}, ψ^{n+1}, and z^{n+1}. In what follows we present an efficient algorithm for solving this system. We define two sets of field variables, $(\psi_i^{n+1},\phi_i^{n+1})$ $(i=1,2)$, as follows:

For ψ_1^{n+1}:

$$\nabla^2\psi_1^{n+1}-\left(\alpha+\frac{S}{\lambda\omega_0}\right)\psi_1^{n+1}=Q^n, \tag{33a}$$

$$\mathbf{n}\cdot\nabla\psi_1^{n+1}=T^n, \quad \text{on } \partial\Omega. \tag{33b}$$

For ϕ_1^{n+1}:

$$\nabla^2\phi_1^{n+1}+\alpha\phi_1^{n+1}=\psi_1^{n+1}, \tag{34a}$$

$$\mathbf{n}\cdot\nabla\phi_1^{n+1}=\frac{1}{\omega_0}\left(g_b^{n+\theta}-\mathbf{n}\cdot\nabla\tilde{\phi}\right), \quad \text{on } \partial\Omega. \tag{34b}$$

For ψ_2^{n+1}:

$$\nabla^2\psi_2^{n+1} - \left(\alpha + \frac{S}{\lambda\omega_0}\right)\psi_2^{n+1} = \frac{1}{2\lambda}\nabla^2 b^n, \tag{35a}$$

$$\mathbf{n}\cdot\nabla\psi_2^{n+1} = \frac{1}{2\lambda}\mathbf{n}\cdot\nabla b^n, \quad \text{on } \partial\Omega. \tag{35b}$$

For ϕ_2^{n+1}:

$$\nabla^2\phi_2^{n+1} + \alpha\phi_2^{n+1} = \psi_2^{n+1}, \tag{36a}$$

$$\mathbf{n}\cdot\nabla\phi_2^{n+1} = 0, \quad \text{on } \partial\Omega. \tag{36b}$$

Then we have the following result.

Theorem 2.2. *For given scalar value z^{n+1}, the following field functions solve the system consisting of equations (27a)–(27b), (31) and (28):*

$$\psi^{n+1} = \psi_1^{n+1} + z^{n+1}\psi_2^{n+1}, \tag{37a}$$

$$\phi^{n+1} = \phi_1^{n+1} + z^{n+1}\phi_2^{n+1}, \tag{37b}$$

where $(\psi_i^{n+1}, \phi_i^{n+1})$ $(i = 1, 2)$ are given by equations (33a)–(36b).

This theorem can be proved by straightforward substitutions and verifications.

We still need to determine the value for z^{n+1}. Substituting the expression ϕ^{n+1} in (37b) into the definition for z^{n+1} in (19) results in

$$z^{n+1} = \frac{\int_\Omega b^n\phi_1^{n+1}}{1 - \int_\Omega b^n\phi_2^{n+1}}. \tag{38}$$

We have the following result:

Theorem 2.3. *The ϕ_2^{n+1} defined by (35a)–(36b) satisfies the property,*

$$\int_\Omega b^n\phi_2^{n+1} \leqslant 0. \tag{39}$$

A proof of this theorem is provided in Appendix B. Based on this theorem, z^{n+1} given by (38) is well defined for any Δt.

Combining the above discussions, we arrive at the de-coupled solution algorithm for solving the system consisting of equations (13a)–(13e). Given (ϕ^n, r^n), we compute $(\phi^{n+1}, \nabla^2\phi^{n+1}, r^{n+1})$ through the following steps:

(S-1): Solve equations (33a)–(33b) for ψ_1^{n+1};
Solve equations (34a)–(34b) for ϕ_1^{n+1}.
(S-2): Solve equations (35a)–(35b) for ψ_2^{n+1};
Solve equations (36a)–(36b) for ϕ_2^{n+1}.
(S-3): Solve equations (38) for z^{n+1}.
(S-4): Compute ψ^{n+1} and ϕ^{n+1} based on equations (37a) and (37b), respectively.
Compute $\nabla^2\phi^{n+1}$ as follows in light of equation (27b)

$$\nabla^2\phi^{n+1} = \psi^{n+1} - \alpha\phi^{n+1}. \tag{40}$$

(S-5): Compute r^{n+1} based on equation (21).

Remark 5. This algorithm involves only the solution of four de-coupled Helmholtz type equations within a time step, hence the name de-coupled solution algorithm. These Helmholtz equations involve only two distinct coefficient matrices after spatial discretization, and they are both constant coefficient matrices and can be pre-computed. No nonlinear algebraic solver is involved in this algorithm. Thanks to these characteristics, the family of second-order energy-stable schemes represented by (13a)–(13e) can be implemented in a very efficient fashion.

Remark 6. This de-coupled solution algorithm requires the chosen constant S to be sufficiently large (satisfying condition (26)). The method breaks down if this condition is not satisfied, because in this case the decomposition given by equations (33a)–(36b) is no longer valid. This is a restriction of the de-coupled solution algorithm. For a given Δt one can always choose a constant S satisfying the condition (26). For a given constant S, if one reduces Δt successively, the de-coupled algorithm will no longer work when Δt reaches a small enough value $\Delta t_c = \frac{4\gamma_0\lambda\omega_0}{mS^2}$. In this case, one can employ a larger constant S in order to perform simulations using the de-coupled solution algorithm for smaller time step sizes. With a range of time step size (Δt) values one plans to cover in the simulations, it is straightforward to estimate an appropriate constant S for these simulations. So this restriction on S is not an issue in practical simulations. Alternatively, one can allow S to vary with respect to Δt instead of being a constant in order to use the de-coupled algorithm, e.g. by choosing S as the lower bound in condition (26). With this mode of usage, however, note that the temporal convergence rate will not be of second order, because S is not a constant. For example, it will be of 3/2-th order if S is always chosen to be the lower bound in condition (26). The algorithm presented in the next sub-section (Section 2.3.2) removes this restriction on the constant S.

2.3.2. Coupled algorithm

The system (13a)–(13e) can also be solved without the decomposition represented by (27a)–(27b). In this case, $\mathcal{H}^{n+1}$ and ϕ^{n+1} are the dynamic variables to solve, and they need to be solved in a coupled fashion. The complication caused by the integral of the unknown field function ϕ^{n+1} over the domain can be dealt with in a way analogous to in the previous subsection. The upside is that the constant S can now take any value $S \geqslant 0$. We will refer to this method as the coupled algorithm.

More specifically, equation (13a) can be rewritten as

$$\nabla^2\mathcal{H}^{n+1} - \frac{\gamma_0}{m\omega_0\Delta t}\phi^{n+1} = -\frac{1}{\omega_0}\nabla^2\tilde{\mathcal{H}} - \frac{1}{m\omega_0}\left(\frac{\hat{\phi}}{\Delta t} + g^{n+\theta}\right), \tag{41}$$

where $\tilde{\mathcal{H}}$ is defined by (16c). In light of equation (21), equation (13b) is transformed into

$$\nabla^2\phi^{n+1} - \frac{S}{\lambda\omega_0}\phi^{n+1} + \frac{1}{\lambda}\mathcal{H}^{n+1} = -\frac{1}{\omega_0}\nabla^2\tilde{\phi} + R^n + \frac{z^{n+1}}{2\lambda}b^n, \tag{42}$$

where

$$R^n = -\frac{S}{\lambda\omega_0}\bar{\phi} - \frac{1}{\lambda\omega_0}\tilde{\mathcal{H}} + \frac{1}{\lambda\gamma_0}\left(\hat{r} - \frac{1}{2}\int_\Omega b^n\hat{\phi}\right)b^n + \frac{\tilde{r}}{\lambda\omega_0}b^n. \tag{43}$$

The boundary condition (13d) can be rewritten as

$$\mathbf{n}\cdot\nabla\mathcal{H}^{n+1} = -\frac{1}{\omega_0}\mathbf{n}\cdot\nabla\tilde{\mathcal{H}} + \frac{1}{\omega_0}g_a^{n+\theta}. \tag{44}$$

Analogous to the decoupled algorithm, to solve the system consisting of equations (41)–(42), (21), (44) and (28), we define two sets of field variables, $(\mathcal{H}_i^{n+1}, \phi_i^{n+1})$ $(i = 1, 2)$, as follows:

For $(\mathcal{H}_1^{n+1}, \phi_1^{n+1})$:

$$\nabla^2\mathcal{H}_1^{n+1} - \frac{\gamma_0}{m\omega_0\Delta t}\phi_1^{n+1} = -\frac{1}{\omega_0}\nabla^2\tilde{\mathcal{H}} - \frac{1}{m\omega_0}\left(\frac{\hat{\phi}}{\Delta t} + g^{n+\theta}\right), \tag{45a}$$

$$\nabla^2\phi_1^{n+1} - \frac{S}{\lambda\omega_0}\phi_1^{n+1} + \frac{1}{\lambda}\mathcal{H}_1^{n+1} = -\frac{1}{\omega_0}\nabla^2\tilde{\phi} + R^n, \tag{45b}$$

$$\mathbf{n}\cdot\nabla\mathcal{H}_1^{n+1} = \frac{1}{\omega_0}\left(g_a^{n+\theta} - \mathbf{n}\cdot\nabla\tilde{\mathcal{H}}\right), \quad \text{on } \partial\Omega, \tag{45c}$$

$$\mathbf{n}\cdot\nabla\phi_1^{n+1} = \frac{1}{\omega_0}\left(g_b^{n+\theta} - \mathbf{n}\cdot\nabla\tilde{\phi}\right), \quad \text{on } \partial\Omega. \tag{45d}$$

For $(\mathcal{H}_2^{n+1}, \phi_2^{n+1})$:

$$\nabla^2\mathcal{H}_2^{n+1} - \frac{\gamma_0}{m\omega_0\Delta t}\phi_2^{n+1} = 0, \tag{46a}$$

$$\nabla^2\phi_2^{n+1} - \frac{S}{\lambda\omega_0}\phi_2^{n+1} + \frac{1}{\lambda}\mathcal{H}_2^{n+1} = \frac{1}{2\lambda}b^n, \tag{46b}$$

$$\mathbf{n}\cdot\nabla\mathcal{H}_2^{n+1} = 0, \quad \text{on } \partial\Omega, \tag{46c}$$

$$\mathbf{n}\cdot\nabla\phi_2^{n+1} = 0, \quad \text{on } \partial\Omega. \tag{46d}$$

Then similar to Theorem 2.2, we have the following result:

Theorem 2.4. *For given scalar value z^{n+1}, the following field functions solve the system consisting of equations (41)–(42), (21), (44) and (28):*

$$\mathcal{H}^{n+1} = \mathcal{H}_1^{n+1} + z^{n+1}\mathcal{H}_2^{n+1}, \tag{47a}$$

$$\phi^{n+1} = \phi_1^{n+1} + z^{n+1}\phi_2^{n+1}, \tag{47b}$$

where $(\mathcal{H}_i^{n+1}, \phi_i^{n+1})$ $(i = 1, 2)$ are given by equations (45a)–(46d).

It is straightforward to verify that the combination of equations (46a) and (46b) leads to equation (75a) in the Appendix B, and that the combination of equations (46c) and (46b) gives rise to equation (75b). Therefore, the system consisting of equations (46a)–(46d) is equivalent to that consisting of equations (75a), (75b) and (46d). All the results from Appendix B apply to the system of (46a)–(46d). We therefore have the following result:

Theorem 2.5. *The field function ϕ_2^{n+1} defined by (46a)–(46d) satisfies the property*

$$\int_\Omega b^n\phi_2^{n+1} \leqslant 0. \tag{48}$$

In summary, after $(\mathcal{H}_i^{n+1}, \phi_i^{n+1})$ $(i = 1, 2)$ are obtained based on equations (45a)–(46d), z^{n+1} can be computed by equation (38). Then $\mathcal{H}^{n+1}$, ϕ^{n+1} and r^{n+1} are given by equations (47a), (47b) and (21), respectively.

Remark 7. With the coupled algorithm the chosen constant S can be any value with $S \geqslant 0$. If the number of degrees of freedom in the system is characterized by N, then roughly speaking, with the coupled algorithm one needs to solve two separate linear systems within a time step, each involving a $2N \times 2N$ coefficient matrix; see equations (45a)–(46d). On the other hand, with the de-coupled algorithm one needs to solve four separate linear systems within a time step, each involving an $N \times N$ coefficient matrix. The coefficient matrix resulting from the coupled algorithm is non-symmetric, while the coefficient matrix from the de-coupled algorithm is symmetric positive definite.

2.4. Implementation with C^0 spectral elements

The solution algorithms presented in the previous subsection can be implemented using any commonly-used method for the spatial discretization. In the current work we will employ C^0 continuous high-order spectral elements [26,44] for spatial discretizations. We next consider the implementation of the algorithm using C^0 spectral elements. The discussions in this subsection without change also apply to implementations using low-order finite elements.

De-coupled algorithm We will first derive the weak forms about $(\psi_i^{n+1}, \phi_i^{n+1})$ $(i = 1, 2)$ in the continuous space, and then specify the discrete function space for the spectral element approximations. In the process, certain terms involving derivatives of order two or higher will be dealt with appropriately so that all quantities involved in the weak formulation can be computed directly using C^0 elements in the discrete function space.

Let $\varphi(\mathbf{x})$ denote a test function. Taking the L^2 inner product between φ and the equation (33a) leads to

$$\int_\Omega \nabla\psi_1^{n+1}\cdot\nabla\varphi + \left(\alpha + \frac{S}{\lambda\omega_0}\right)\int_\Omega \psi_1^{n+1}\varphi = -\int_\Omega Q^n\varphi + \int_{\partial\Omega} T^n\varphi, \quad \forall\varphi, \tag{49}$$

where we have used integration by part and the equation (33b). By substituting the Q^n expression from (23) and T^n expression from (32) into the RHS of the above equation and integration by part, we obtain the weak form about ψ_1^{n+1},

$$\begin{aligned}
&\int_\Omega \nabla\psi_1^{n+1}\cdot\nabla\varphi + \left(\alpha + \frac{S}{\lambda\omega_0}\right)\int_\Omega \psi_1^{n+1}\varphi \\
&= -\frac{1}{\lambda\omega_0 m}\int_\Omega \left(g^{n+\theta} + \frac{\hat{\phi}}{\Delta t}\right)\varphi - \frac{S}{\lambda\omega_0}\int_\Omega \nabla\bar{\phi}^{n+1}\cdot\nabla\varphi - \frac{1}{\omega_0}\int_\Omega \nabla\left(\nabla^2\tilde{\phi}\right)\cdot\nabla\varphi - \frac{1}{\lambda\omega_0}\int_{\partial\Omega} g_a^{n+\theta}\varphi \\
&\quad + \left[\frac{1}{\lambda\gamma_0}\left(\hat{r} - \frac{1}{2}\int_\Omega b^n\hat{\phi}\right) + \frac{\tilde{r}}{\lambda\omega_0}\right]\int_\Omega \nabla b^n\cdot\nabla\varphi + \left(\alpha + \frac{S}{\lambda\omega_0}\right)\frac{1}{\omega_0}\int_{\partial\Omega}\left(g_b^{n+\theta} - \mathbf{n}\cdot\nabla\tilde{\phi}\right)\varphi.
\end{aligned} \tag{50}$$

In the above equation note that the term $\nabla^2\tilde{\phi}$ is to be computed by, in light of equation (16c),

$$\begin{aligned}\nabla^2\tilde{\phi} &= 2(1-\theta)^2(\nabla^2\phi^n) + \left(\theta - \frac{1}{2}\right)(1-\theta)(\nabla^2\phi^{n-1}) \\ &= 2(1-\theta)^2(\psi^n - \alpha\phi^n) + \left(\theta - \frac{1}{2}\right)(1-\theta)(\psi^{n-1} - \alpha\phi^{n-1})\end{aligned} \tag{51}$$

where $\nabla^2\phi^n$ has been computed based on equation (40).

Taking the L^2 inner product between φ and the equation (34a) and integration by part results in the weak form about ϕ_1^{n+1}:

$$\int_\Omega \nabla\phi_1^{n+1}\cdot\nabla\varphi - \alpha\int_\Omega \phi_1^{n+1}\varphi = -\int_\Omega \psi_1^{n+1}\varphi + \frac{1}{\omega_0}\int_{\partial\Omega}\left(g_b^{n+\theta} - \mathbf{n}\cdot\nabla\tilde{\phi}\right)\varphi, \quad \forall\varphi, \tag{52}$$

where we have use equations (34b).

By taking the L^2 inner product between φ and the equation (35a) and integration by part, we arrive at the weak form about ψ_2^{n+1},

$$\int_\Omega \nabla\psi_2^{n+1}\cdot\nabla\varphi + \left(\alpha + \frac{S}{\lambda\omega_0}\right)\int_\Omega \psi_2^{n+1}\varphi = \frac{1}{2\lambda}\int_\Omega \nabla b^n\cdot\nabla\varphi, \quad \forall\varphi, \tag{53}$$

where we have used equation (35b). Taking the L^2 inner product between φ and equation (36a) and integration by part leads to the weak form about ϕ_2^{n+1},

$$\int_\Omega \nabla\phi_2^{n+1}\cdot\nabla\varphi - \alpha\int_\Omega \phi_2^{n+1}\varphi = -\int_\Omega \psi_2^{n+1}\varphi, \quad \forall\varphi, \tag{54}$$

where the equation (36b) has been used.

We discretize the domain Ω using a mesh of N_{el} spectral elements. Let K (positive integer) denote the element order, which is a measure of the highest polynomial degree in field expansions within an element. Let Ω_h denote the discretized domain, and Ω_h^e ($1 \leqslant e \leqslant N_{el}$) denote the element e, $\Omega_h = \cup_{e=1}^{N_{el}}\Omega_h^e$. Define function space

$$H_\phi = \left\{\, v \in H^1(\Omega_h) \ : \ v \text{ is a polynomial of degree characterized by } K \text{ on } \Omega_h^e, \text{ for } 1 \leqslant e \leqslant N_{el} \,\right\}.$$

Let $(\cdot)_h$ denote the discretized version of the variable $(\cdot)$ below. The fully discretized equations for $(\psi_i^{n+1}, \phi_i^{n+1})$ $(i = 1, 2)$ are the following.

For ψ_{1h}^{n+1}: find $\psi_{1h}^{n+1} \in H_\phi$ such that

$$\begin{aligned}&\int_{\Omega_h} \nabla\psi_{1h}^{n+1}\cdot\nabla\varphi_h + \left(\alpha + \frac{S}{\lambda\omega_0}\right)\int_{\Omega_h}\psi_{1h}^{n+1}\varphi_h \\ &= -\frac{1}{\lambda\omega_0 m}\int_{\Omega_h}\left(g_h^{n+\theta} + \frac{\hat{\phi}_h}{\Delta t}\right)\varphi_h - \frac{S}{\lambda\omega_0}\int_{\Omega_h}\nabla\bar{\phi}_h^{n+1}\cdot\nabla\varphi_h - \frac{1}{\omega_0}\int_{\Omega_h}\nabla\left(\nabla^2\tilde{\phi}_h\right)\cdot\nabla\varphi_h \\ &\quad + \left[\frac{1}{\lambda\gamma_0}\left(\hat{r} - \frac{1}{2}\int_{\Omega_h} b_h^n\hat{\phi}_h\right) + \frac{\tilde{r}}{\lambda\omega_0}\right]\int_{\Omega_h}\nabla b_h^n\cdot\nabla\varphi_h \\ &\quad + \left(\alpha + \frac{S}{\lambda\omega_0}\right)\frac{1}{\omega_0}\int_{\partial\Omega_h}\left(g_{bh}^{n+\theta} - \mathbf{n}_h\cdot\nabla\tilde{\phi}_h\right)\varphi_h - \frac{1}{\lambda\omega_0}\int_{\partial\Omega_h} g_{ah}^{n+\theta}\varphi_h, \qquad \forall\varphi_h \in H_\phi,\end{aligned} \tag{55}$$

where note that $\nabla^2\tilde{\phi}_h$ is to be computed according to equation (51).

For ϕ_{1h}^{n+1}: find $\phi_{1h}^{n+1} \in H_\phi$ such that

$$\int_{\Omega_h} \nabla\phi_{1h}^{n+1}\cdot\nabla\varphi_h - \alpha\int_{\Omega_h}\phi_{1h}^{n+1}\varphi_h = -\int_{\Omega_h}\psi_{1h}^{n+1}\varphi_h + \frac{1}{\omega_0}\int_{\partial\Omega_h}\left(g_{bh}^{n+\theta} - \mathbf{n}_h\cdot\nabla\tilde{\phi}_h\right)\varphi_h, \quad \forall\varphi_h \in H_\phi, \tag{56}$$

For ψ_{2h}^{n+1}: find $\psi_{2h}^{n+1} \in H_\phi$ such that

$$\int_{\Omega_h} \nabla\psi_{2h}^{n+1} \cdot \nabla\varphi_h + \left(\alpha + \frac{S}{\lambda\omega_0}\right) \int_{\Omega_h} \psi_{2h}^{n+1}\varphi_h = \frac{1}{2\lambda} \int_{\Omega_h} \nabla b_h^n \cdot \nabla\varphi_h, \qquad \forall\varphi_h \in H_\phi. \tag{57}$$

For ϕ_{2h}^{n+1}: find $\phi_{2h}^{n+1} \in H_\phi$ such that

$$\int_{\Omega_h} \nabla\phi_{2h}^{n+1} \cdot \nabla\varphi_h - \alpha \int_{\Omega_h} \phi_{2h}^{n+1}\varphi_h = -\int_{\Omega_h} \psi_{2h}^{n+1}\varphi_h, \qquad \forall\varphi_h \in H_\phi. \tag{58}$$

The final discretized de-coupled algorithm consists of the following within each time step: (i) Solve equations (55)–(58) for $(\psi_i^{n+1}, \phi_i^{n+1})$ $(i = 1, 2)$, respectively. (ii) Compute z^{n+1} from equations (38). (iii) Compute ψ^{n+1}, ϕ^{n+1}, and r^{n+1} based on equations (37a), (37b), and (21), respectively.

Coupled algorithm Similarly, the weak forms about $(\mathcal{H}_i^{n+1}, \phi_i^{n+1})$ $(i = 1, 2)$ can be obtained by taking the L^2 inner product between a test function and the equations (45a)–(45b) or (46a)–(46b), respectively, and with integration by part. Correspondingly, the fully discretized equations for $(\mathcal{H}_i^{n+1}, \phi_i^{n+1})$ $(i = 1, 2)$ are given by the following.

For $(\mathcal{H}_{1h}^{n+1}, \phi_{1h}^{n+1})$: find $\mathcal{H}_{1h}^{n+1}, \phi_{1h}^{n+1} \in H_\phi$ such that

$$\begin{aligned} &\int_{\Omega_h} \nabla\mathcal{H}_{1h}^{n+1} \cdot \nabla\varphi_h + \frac{\gamma_0}{m\omega_0\Delta t} \int_{\Omega_h} \phi_{1h}^{n+1}\varphi_h \\ &\quad = \frac{1}{\omega_0 m} \int_{\Omega_h} \left(g_h^{n+\theta} + \frac{\hat{\phi}_h}{\Delta t}\right)\varphi_h - \frac{1}{\omega_0} \int_{\Omega_h} \nabla\tilde{\mathcal{H}}_h \cdot \nabla\varphi_h + \frac{1}{\omega_0} \int_{\partial\Omega_h} g_{ah}^{n+\theta}\varphi_h, \quad \forall\varphi_h \in H_\phi, \end{aligned} \tag{59}$$

$$\begin{aligned} &\int_{\Omega_h} \nabla\phi_{1h}^{n+1} \cdot \nabla\varphi_h + \frac{S}{\lambda\omega_0} \int_{\Omega_h} \phi_{1h}^{n+1}\varphi_h - \frac{1}{\lambda} \int_{\Omega_h} \mathcal{H}_{1h}^{n+1}\varphi_h = -\frac{1}{\omega_0} \int_{\Omega_h} \nabla\tilde{\phi}_h \cdot \nabla\varphi_h + \frac{S}{\lambda\omega_0} \int_{\Omega_h} \bar{\phi}_h\varphi_h \\ &\quad + \frac{1}{\lambda\omega_0} \int_{\Omega_h} \tilde{\mathcal{H}}_h\varphi_h - \left[\frac{1}{\lambda\gamma_0}\left(\hat{r} - \frac{1}{2}\int_{\Omega_h} b_h^n\hat{\phi}_h\right) + \frac{\tilde{r}}{\lambda\omega_0}\right] \int_{\Omega_h} b_h^n\varphi_h + \frac{1}{\omega_0} \int_{\partial\Omega_h} g_{bh}^{n+\theta}\varphi_h, \quad \forall\varphi_h \in H_\phi. \end{aligned} \tag{60}$$

For $(\mathcal{H}_{2h}^{n+1}, \phi_{2h}^{n+1})$: find $\mathcal{H}_{2h}^{n+1}, \phi_{2h}^{n+1} \in H_\phi$ such that

$$\int_{\Omega_h} \nabla\mathcal{H}_{2h}^{n+1} \cdot \nabla\varphi_h + \frac{\gamma_0}{m\omega_0\Delta t} \int_{\Omega_h} \phi_{2h}^{n+1}\varphi_h = 0, \quad \forall\varphi_h \in H_\phi. \tag{61}$$

$$\int_{\Omega_h} \nabla\phi_{2h}^{n+1} \cdot \nabla\varphi_h + \frac{S}{\lambda\omega_0} \int_{\Omega_h} \phi_{2h}^{n+1}\varphi_h - \frac{1}{\lambda} \int_{\Omega_h} \mathcal{H}_{2h}^{n+1}\varphi_h = -\frac{1}{2\lambda} \int_{\Omega} b_h^n\varphi_h, \quad \forall\varphi_h \in H_\phi. \tag{62}$$

The final discretized coupled algorithm consists of the following within each time step: (i) Solve equations (59)–(62) for $(\mathcal{H}_i^{n+1}, \phi_i^{n+1})$ $(i = 1, 2)$, respectively. (ii) Compute z^{n+1} from equation (38). (iii) Compute $\mathcal{H}^{n+1}$, ϕ^{n+1}, and r^{n+1} based on equations (47a), (47b), and (21), respectively.

3. Representative numerical examples

In this section we present several example problems in two dimensions to test the performance of the family of energy-stable schemes developed in the previous section. In the numerical simulations we have non-dimensionalized the physical variables, the governing equations, and the boundary conditions. As detailed in previous works (see e.g. [8]), the non-dimensional form of the governing equations and boundary conditions will remain the same, if the physical variables are normalized consistently. Let L denote a length scale, U_0 a velocity scale, ϱ_0 a density scale, and d ($d = 2$ or 3) the spatial dimension. The normalization constants for consistent non-dimensionalization of different physical variables involved in the current work are listed in Table 1.

Table 1
Normalization constants for consistent non-dimensionalization of physical variables.

Variable	Normalization constant	Variable	Normalization constant
$\mathbf{x}$, η	L	t, Δt	L/U_0
ϕ, ϕ_1, ϕ_2, ϕ_{in}, θ, γ_0, ω_0, z^{n+1}	1	λ	$\varrho_0 U_0^2 L^2$
m	$L/(\varrho_0 U_0)$	$h(\phi)$, $F(\phi)$, $\mathcal{H}$, S	$\varrho_0 U_0^2$
$g(\mathbf{x})$	U_0/L	$g_a(\mathbf{x})$	$\varrho_0 U_0/L$
$g_b(\mathbf{x})$	$1/L$	$E[\phi]$, C_0	$\varrho_0 U_0^2 L^d$
$r(t)$	$\sqrt{\varrho_0 U_0^2 L^d}$	α, ψ_1, ψ_2, ψ	$1/L^2$

Table 2
Simulation parameter values for convergence tests.

Parameter	Value	Parameter	Value
C_0	0	λ	0.01
m	0.01	η	0.1
θ	$1/2 \sim 3/2$	Δt	(varied)
$\Delta t_{\min}$	1e−4	S	$\sqrt{\frac{4\gamma_0\lambda\omega_0}{m\Delta t}}$ or $\sqrt{\frac{4\gamma_0\lambda\omega_0}{m\Delta t_{\min}}}$ (decoupled solver) or 1 (coupled solver)
t_0	0.1	t_f	0.2 (spatial tests) or 0.3 (temporal tests)
Element order	(varied)	Elements	2

3.1. Convergence rates

We first use a manufactured analytic solution to the Cahn–Hilliard equation to numerically demonstrate the rates of convergence in space and time of the algorithms presented in Section 2.

We consider the domain $0 \leqslant x \leqslant 2$ and $-1 \leqslant y \leqslant 1$, and the following solution to the Cahn–Hilliard equation (1) on this domain,

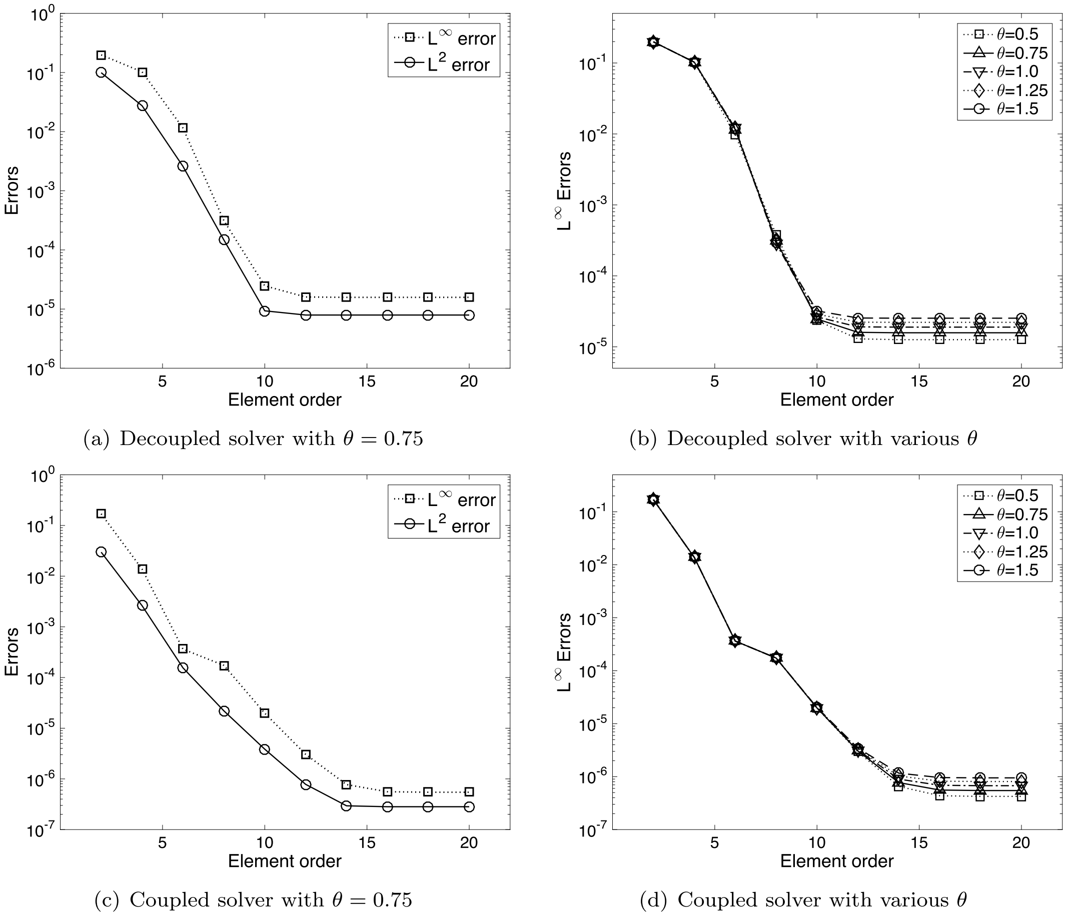
$$\phi = \cos(\pi x)\cos(\pi y)\sin(t). \tag{63}$$

The source term $g(\mathbf{x}, t)$ in (1) is chosen such that the analytic expression of (63) satisfies (1). The conditions (2a) and (2b) are imposed on the domain boundary, and the source terms $g_a(\mathbf{x}, t)$ and $g_b(\mathbf{x}, t)$ are chosen such that the analytic expression of (63) satisfies (2a) and (2b) on the domain boundary.

We discretize the domain using two equal-sized quadrilateral elements, with the element order and the time step size Δt varied systematically in the spatial and temporal convergence tests. The algorithms from Section 2 are employed to numerically integrate the Cahn–Hilliard equation in time from $t = t_0$ to $t = t_f$. The initial phase field function ϕ_{in} is obtained by setting $t = t_0$ in the analytic expression (63). The numerical solution at $t = t_f$ is then compared with the analytic solution, and various norms of the errors are computed. The values for the simulation parameters are summarized in Table 2 for this problem.

To test the spatial convergence rate, we employ a fixed $\Delta t = 0.001$, $t_0 = 0.1$ and $t_f = 0.2$, and vary the element order systematically between 2 and 20. Then for each element order, the numerical errors of ϕ in L^∞ and L^2 norms at $t = t_f$ are obtained. Figs. 1(a) and (c) show these errors as a function of the element order obtained using $\theta = 3/4$ with the de-coupled and coupled algorithms, respectively. In this group of tests $S = \sqrt{\frac{4\gamma_0\lambda\omega_0}{m\Delta t}}$ for the de-coupled algorithm and $S = 1$ for the coupled algorithm. We observe an exponential decrease of the numerical errors with increasing element order, and a level-off of the error curves beyond element orders 10 and 14, respectively for the decoupled and coupled algorithms, due to the saturation of temporal errors. Figs. 1(b) and (d) show the L^∞ errors versus the element order obtained using the de-coupled and coupled algorithms corresponding several θ values ranging from 1/2 to 3/2. Some differences in the saturation error level can be observed with different algorithms. The saturation error is larger for algorithms with larger θ values.

To test the temporal convergence rate, we employ a fixed element order 20, $t_0 = 0.1$ and $t_f = 0.3$, and vary the time step size Δt systematically between 0.1 and 1.953125×10^{-4}. For each Δt the numerical errors in different norms are computed at $t = t_f$. Figs. 2(a) and (c) show the numerical errors as a function of Δt obtained using the de-coupled and coupled algorithms with $\theta = 3/4$, respectively. A second-order convergence rate is observed. In these tests the constant S is chosen as $S = S_0 = \sqrt{\frac{4\gamma_0\lambda\omega_0}{m\Delta t_{\min}}}$, where $\Delta t_{\min} = 10^{-4}$, with the de-coupled algorithm and $S = 1$ with the coupled algorithm. Figs. 2(b) and (d) show the L^∞ errors versus Δt obtained using the two algorithms corresponding to several θ values. The errors are generally smaller with a smaller θ value in the algorithm. However, with $\theta = 1/2$ (the lower boundary for the θ range, corresponding to the Crank–Nicolson scheme) we observe a weak instability when Δt becomes small for the decoupled algorithm, which results in larger errors with the two smallest Δt values in the test. Note that t_f is fixed in these tests, and a larger number of time steps are computed with a smaller Δt.



(a) Decoupled solver with $\theta = 0.75$

(b) Decoupled solver with various θ

(c) Coupled solver with $\theta = 0.75$

(d) Coupled solver with various θ

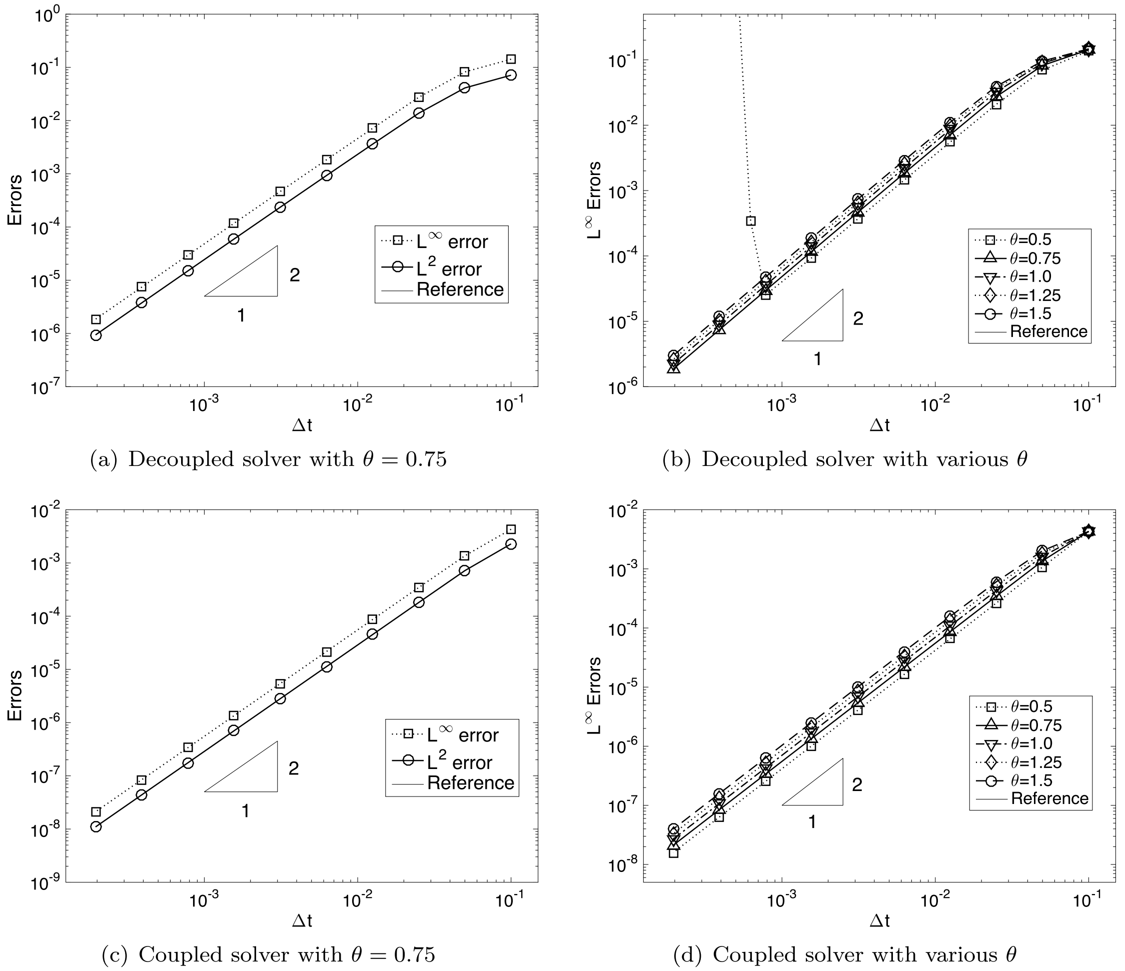
Fig. 1. Spatial convergence tests: (a) L^∞, L^2 errors as a function of the element order computed using decoupled algorithm with $\theta = 0.75$. (b) L^∞ errors as a function of the element order computed using decoupled algorithm corresponding to several θ values. (c) L^∞, L^2 errors as a function of the element order computed using coupled algorithm with $\theta = 0.75$. (d) L^∞ errors as a function of the element order computed using coupled algorithm corresponding to several θ values. These results are obtained with a fixed $\Delta t = 0.001$.

Fig. 3 demonstrates the effect of the constant S in the algorithms on the numerical errors. Here we plot the L^∞ and L^2 errors as a function of S/S_0, where S_0 is given in the previous paragraph, obtained with $\theta = \frac{3}{4}$ for both the de-coupled and coupled algorithms. In this group of tests, we have a fixed $t_0 = 0.1$, $t_f = 0.2$, $\Delta t = 0.001$, and the element order 20. The two algorithms result in essentially identical error values. The error data are observed to fall onto a straight line, with a unit slope. This indicates that as the value for the constant S increases, the numerical errors increase proportionally.

As shown by Fig. 2(b), the scheme with $\theta = \frac{1}{2}$ (Crank–Nicolson) exhibits a weak instability and is unstable for the de-coupled algorithm in long-time simulations. We observe that this instability is also encountered with the other θ values in a small neighborhood of $\theta = \frac{1}{2}$, approximately $\theta \in [0.5, 0.512]$. This is demonstrated by Fig. 4. Fig. 4(a) shows time histories of the L^∞ error obtained using the de-coupled algorithm with $\theta = 0.5$, 0.51, and 0.75 in a long-time simulation, with $t_0 = 0.1$ and $t_f = 100$. Only a window of the first 20 time units is shown here for clarity. The other crucial parameters in this group of tests are the time step size $\Delta t = 0.001$, element order 20, and the constant $S = \sqrt{\frac{4\gamma_0\lambda\omega_0}{m\Delta t}}$. It is observed that the computations with both $\theta = 0.5$ and 0.51 blow up not long into the simulation, while the simulation with $\theta = 0.75$ is long-term stable. We have varied θ systematically and studied the stability of long-time simulations. Fig. 4(b) shows the blow-up time, i.e. the time when the computation becomes unstable, as a function of θ. The blow-up time increases as θ increases from 0.5. Beyond $\theta = 0.512$ the simulation is observed to be long-term stable (With $\theta = 0.513$ e.g. the L^∞ error is approximately 7.11e−4 at $t = 100$).

We generally observe a poor performance in the scheme $\theta = \frac{1}{2}$ (lower boundary of the θ range) in actual simulations. While the results produced by the scheme $\theta = \frac{3}{2}$ (upper boundary of θ) seem fine with the manufactured solution here, numerical tests in later sections reveal that its performance is also poor (but slightly better than $\theta = \frac{1}{2}$) in actual long-time simulations. We will further investigate the performance of these two boundary schemes in subsequent sections.

The test results of this section indicate that the family of algorithms presented in Section 2 exhibits a spatial exponential convergence rate and a temporal second-order convergence rate.



(a) Decoupled solver with $\theta = 0.75$

(b) Decoupled solver with various θ

(c) Coupled solver with $\theta = 0.75$

(d) Coupled solver with various θ

Fig. 2. Temporal convergence tests: (a) L^∞, L^2 errors as a function of time step size Δt computed using the decoupled algorithm with $\theta = 0.75$. (b) L^∞ errors as a function of Δt computed using the decoupled algorithm with various θ values. (c) L^∞, L^2 errors as a function of Δt computed using the coupled algorithm with $\theta = 0.75$. (d) L^∞ errors as a function of Δt computed using the coupled algorithm with various θ values. Results are obtained with a fixed element order 20.

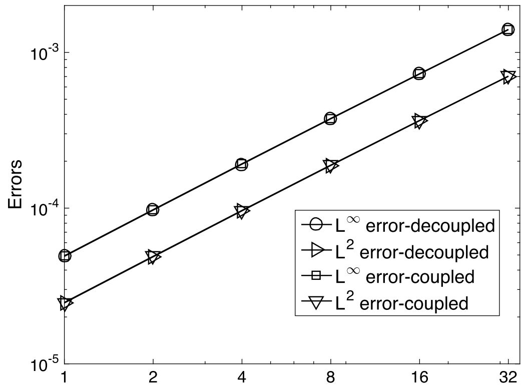
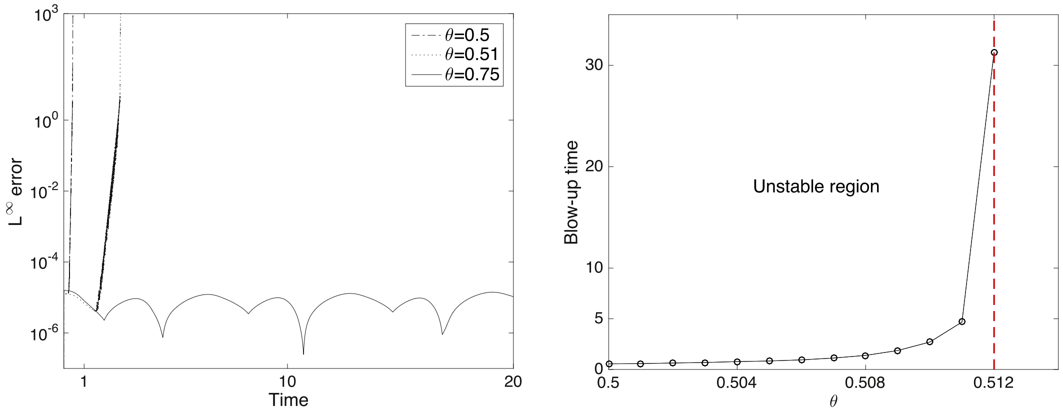


Fig. 3. L^∞ and L^2 errors as a function of $S/S_0 = (1, 2, 4, 8, 16, 32)$ with an element order 20 and $\theta = 0.75$ for the de-coupled and coupled algorithms. The constant $S_0 = \sqrt{\frac{4\gamma_0\lambda\omega_0}{m\Delta t_{\min}}}$, where $\Delta t_{\min} = 1\text{e}{-4}$.



(a) Error history of decoupled algorithm (b) Unstable θ region

Fig. 4. (a) Time histories of L^∞ error obtained with $\theta = 0.5, 0.51, 0.75$. (b) The blow-up time as a function of θ. Results are obtained using the decoupled algorithm with an element order 20, $\Delta t = 0.001$, and $S = \sqrt{\frac{4\gamma_0\lambda\omega_0}{m\Delta t}}$.

3.2. Evolution and coalescence of drops

In this section we consider the evolution of a drop and the coalescence of two drops to illustrate the dynamics of the Cahn–Hilliard equation and the energy stability of the family of algorithms from Section 2.

Evolution of a drop We first look into the evolution of a material drop under the Cahn–Hilliard dynamics. Consider a square domain $\Omega = \{\ (x, y)\ |\ 0 \leqslant x, y \leqslant 1\ \}$, and two materials contained in this domain. It is assumed that the evolution of the material regions is described by the Cahn–Hilliard equation, and that $\phi = 1$ and -1 correspond to the bulk of the first and second materials, respectively. At $t = 0$, the first material is located in a smaller square region with dimension $2h_0 \times 2h_0$ (where $h_0 = 0.2$) around the center of the domain, and the second material fills the rest of the domain. The goal here is to study the evolution dynamics of these material regions.

We employ the algorithms from Section 2 to numerically solve the Cahn–Hilliard equation with $g(\mathbf{x}, t) = 0$ in this domain. We discretize the domain using 100 quadrilateral elements, with 10 equal-sized elements along both x and y directions. The boundary conditions (2a) and (2b), with $g_a = 0$ and $g_b = 0$, are imposed on the domain boundaries. The initial distribution of the materials is given by

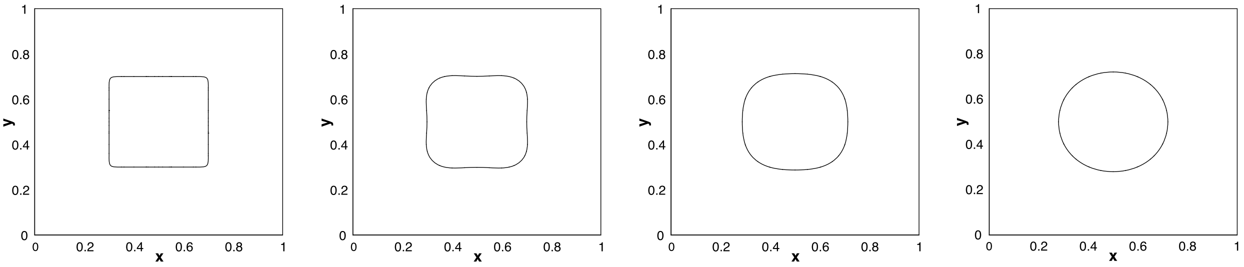
$$\phi_{in}(\mathbf{x}) = \frac{1}{2}\left[\tanh\frac{x - x_0 + h_0}{\sqrt{2}\eta} - \tanh\frac{x - x_0 - h_0}{\sqrt{2}\eta}\right]\cdot\left[\tanh\frac{y - y_0 + h_0}{\sqrt{2}\eta} - \tanh\frac{y - y_0 - h_0}{\sqrt{2}\eta}\right] - 1, \tag{64}$$

where $(x_0, y_0) = (0.5, 0.5)$ is the center of the domain. We employ the following (non-dimensional) parameter values for this problem:

$$\begin{cases} \eta = 0.01, \quad \sigma = 151.15, \quad C_0 = 0, \quad \lambda = \dfrac{3}{2\sqrt{2}}\sigma\eta, \quad m = \dfrac{10^{-6}}{\lambda}, \\ S = \sqrt{\dfrac{4\gamma_0\lambda\omega_0}{m\Delta t_{\min}}} \ \text{ or } \ \sqrt{\dfrac{4\gamma_0\lambda\omega_0}{m\Delta t}} \text{ (decoupled algorithm); and } S = 1 \text{ (coupled algorithm)}, \\ \text{element order: 12,} \quad \text{number of elements: 100,} \\ \Delta t_{\min} = 10^{-5}, \quad \Delta t \text{ varied,} \quad \theta \text{ varied.} \end{cases} \tag{65}$$

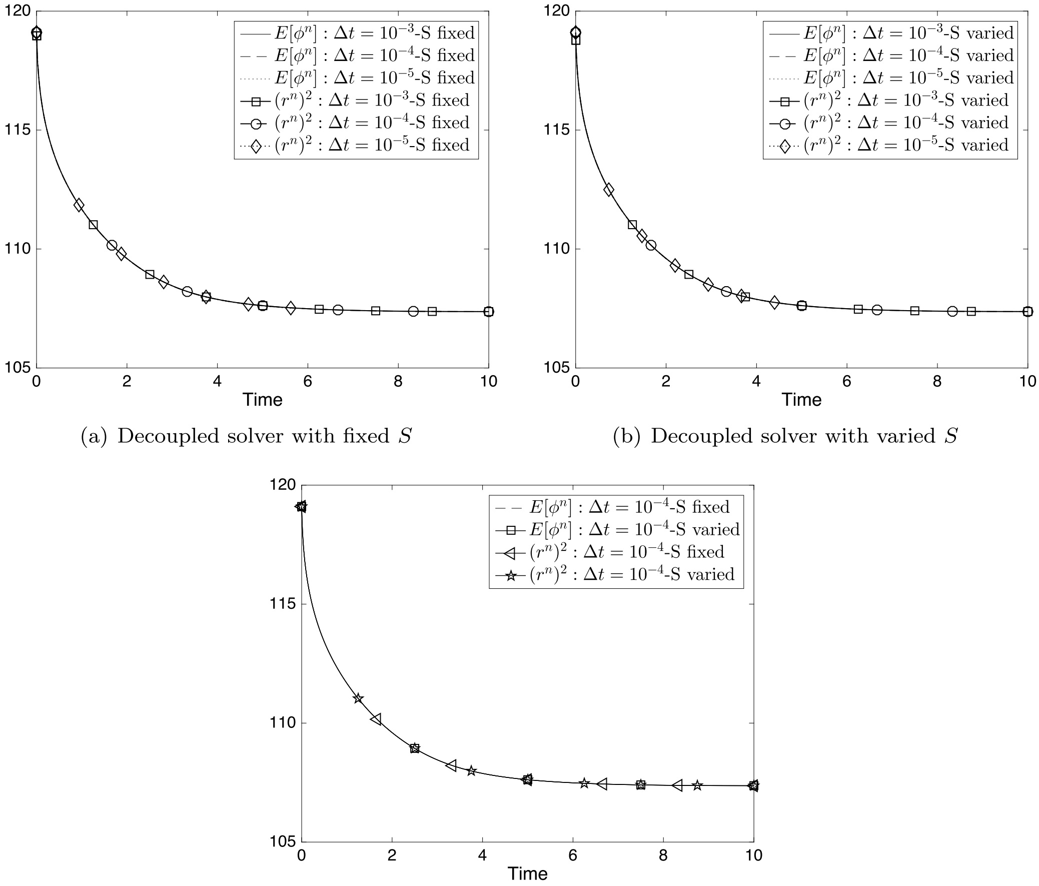
Fig. 5 shows the evolution of the material regions with a temporal sequence of snapshots of the interface (visualized by the contour level $\phi = 0$) between the two materials. These results are computed with a time step size $\Delta t = 0.001$ and the constant $S = \sqrt{\frac{4\gamma_0\lambda\omega_0}{m\Delta t}}$ using the decoupled algorithm corresponding to $\theta = 0.75$. The initial square region of the first material evolves gradually into a circular region under the Cahn–Hilliard dynamics.

We now look into the performance of the de-coupled algorithm. Fig. 6(a) shows time histories of the potential energy $E[\phi]$ defined in (5) and the quantity $[r(t)]^2$ computed from equation (7) obtained using the de-coupled algorithm with $\theta = 0.75$ and several small time step sizes (Δt ranging from 10^{-5} to 10^{-3}). In this group of tests we have employed a fixed $S = \sqrt{\frac{4\gamma_0\lambda\omega_0}{m\Delta t_{\min}}}$, where $\Delta t_{\min} = 10^{-5}$. It can be observed that both $E[\phi]$ and $[r(t)]^2$ decrease over time, and they gradually level off at a certain level. In particular, we observe that the history curves for $E[\phi]$ and $[r(t)]^2$ overlap with one another,



(a) $t = 0.001$ (b) $t = 1$ (c) $t = 4$ (d) $t = 10$

Fig. 5. Temporal sequence of snapshots showing the evolution of a drop visualized by the contour level $\phi = 0$. Results are obtained with $\theta = 0.75$ and $\Delta t = 10^{-3}$ using the decoupled algorithm.



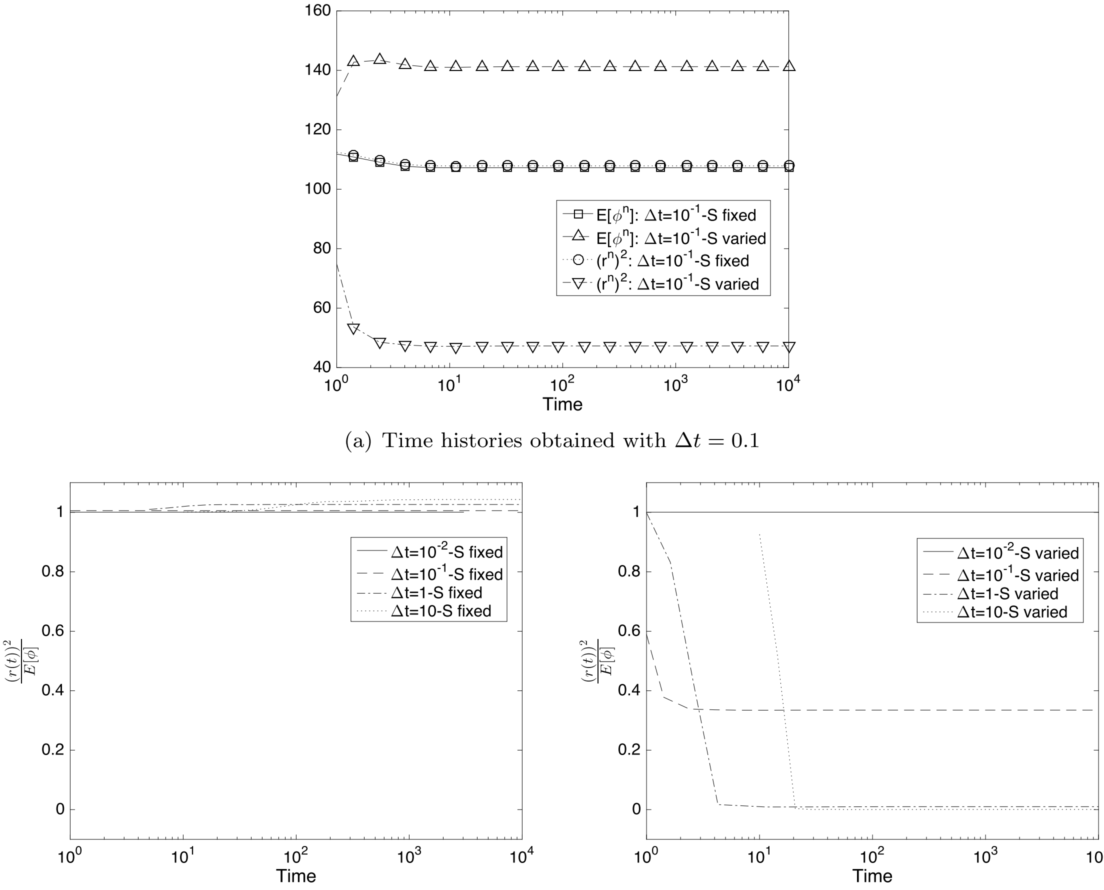
(a) Decoupled solver with fixed S

(b) Decoupled solver with varied S

(c) Comparison with S fixed and varied

Fig. 6. Time histories of $E[\phi]$ and $[r(t)]^2$ computed using the decoupled algorithm ($\theta = 0.75$) with (a) $S = \sqrt{\frac{4\gamma_0\lambda\omega_0}{m\Delta t_{\min}}}$ (fixed), (b) $S = \sqrt{\frac{4\gamma_0\lambda\omega_0}{m\Delta t}}$ (varied with Δt), at several small Δt values. Plot (c) compares the time histories obtained with S fixed and varied at $\Delta t = 10^{-4}$. In these tests $\Delta t_{\min} = 1\text{e}{-5}$.

which indicates that $r(t)$ computed based on equation (7) approximates well the quantity $\sqrt{E[\phi]}$ (see equation (6)). Fig. 6(b) shows the corresponding time histories for $E[\phi]$ and $r^2(t)$ obtained using the de-coupled algorithm with the same time step size values, but with S taking the lower bound in the condition (26), i.e. $S = \sqrt{\frac{4\gamma_0\lambda\omega_0}{m\Delta t}}$. Since S varies with Δt, theoretically the temporal rate of convergence would be of order $3/2$ in this group of tests. It is observed that these time history curves overlap with one another, suggesting that the computed $r^2(t)$ faithfully reflects the evolution of the potential energy of the system. We further observe that the simulation results obtained using a fixed S and with S varying with respect to Δt are the same, which is demonstrated by a comparison of the $E[\phi]$ and $r^2(t)$ time histories in Fig. 6(c) obtained with $\Delta t = 10^{-4}$ for these two cases. This indicates that with the de-coupled algorithm and small time step sizes, when S is allowed to vary with respect to Δt, there is little or no difference in the simulation results when compared with the case with S fixed.



(a) Time histories obtained with $\Delta t = 0.1$

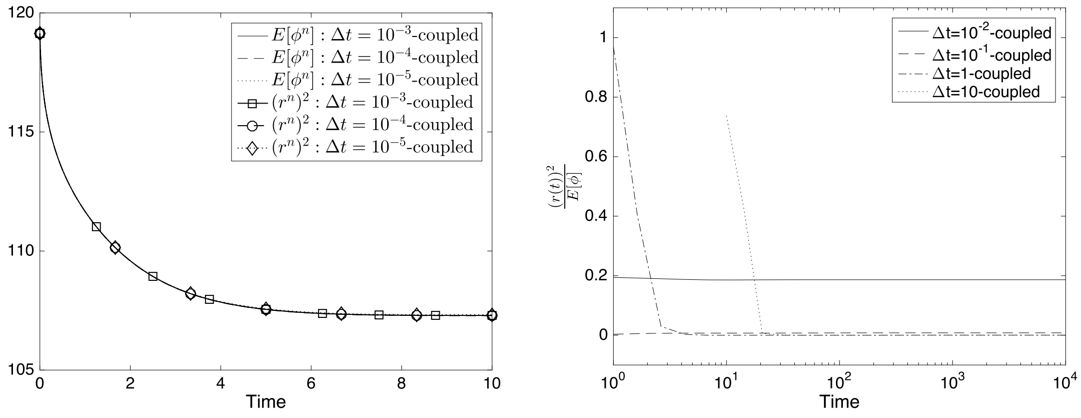
(b) Decoupled solver with fixed S

(c) Decoupled solver with varied S

Fig. 7. Time histories of $E[\phi]$, $r^2(t)$ and $[r(t)]^2/E[\phi]$ computed using the decoupled algorithm ($\theta = 0.75$) with (a) $S = \sqrt{\frac{4\gamma_0\lambda\omega_0}{m\Delta t_{\min}}}$ (fixed), and (b) $S = \sqrt{\frac{4\gamma_0\lambda\omega_0}{m\Delta t}}$ (varying with Δt), at several large Δt values.

The energy stable nature of the current algorithms allows the use of large or fairly large time step sizes in the computations. The behavior of the de-coupled algorithm at several fairly large time step sizes is illustrated in Fig. 7 with $\theta = 0.75$. The time steps size ranges from $\Delta t = 0.01$ to $\Delta t = 10$ in this group of tests. The de-coupled algorithm with S fixed at $S = \sqrt{\frac{4\gamma_0\lambda\omega_0}{m\Delta t_{\min}}}$ ($\Delta t_{\min} = 1\text{e}{-2}$) and with S varied $\left(S = \sqrt{\frac{4\gamma_0\lambda\omega_0}{m\Delta t}}\right)$ with respect to Δt have been tested. Fig. 7(a) compares the long-time histories of $E[\phi]$ and $r^2(t)$ (up to $t = 10^4$) obtained with $\Delta t = 0.1$ using the de-coupled algorithm with S fixed and with S varied. With the fixed S, the histories of $E[\phi]$ and $r^2(t)$ obtained with $\Delta t = 0.1$ agree with each other quite well (with slight difference), and both are in good agreement with those time histories obtained with smaller Δt values (see Fig. 6(a)). On the other hand, with the S varied with respect to Δt, the time histories for $E[\phi]$ and $r^2(t)$ computed with $\Delta t = 0.1$ exhibit large discrepancies. $E[\phi]$ increases initially, and levels off over time around $E[\phi] \approx 140$. In contrast, $r^2(t)$ decreases sharply initially and then levels off gradually over time around $r^2(t) \approx 47$. These time histories are quite different, both qualitatively and quantitatively, from those obtained using smaller Δt values (see Fig. 6(b)). The discrepancies between $E[\phi]$ and $r^2(t)$ are understandable, because at very large time step sizes we generally cannot expect that the computation results will be accurate.

The large difference between $r^2(t)$ and $E[\phi]$ (with S varied) suggests that in this case $r(t)$ is no longer an accurate approximation of $\sqrt{E[\phi]}$. We observe that the ratio $\frac{r^2(t)}{E[\phi]}$ indeed seems to be a good indicator of the accuracy. It should physically take the unit value. When this ratio deviates from the unit value substantially, it suggests that the simulation is no longer accurate and the result likely contains significant errors. In Figs. 7(b) and (c) we show long-time histories of the ratio $r^2(t)/E[\phi]$ corresponding to several large Δt values ($\Delta t = 0.01 \sim 10$), obtained with S fixed and with S varied with respect to Δt in the de-coupled algorithm. We observe that all these simulations are stable, signifying that our algorithms can indeed produce stable results at large time step sizes. With S fixed, the ratio $r^2(t)/E[\phi]$ corresponding to $\Delta t = 0.01$ and 0.1 is very close to 1, while with the even larger time step sizes $\Delta t = 1.0$ and $\Delta t = 10$ this ratio becomes slightly larger than 1 over time. On the other hand, with S varied with Δt, the ratio $r^2(t)/E[\phi]$ deviates significantly from the unit value



(a) Coupled solver using small Δt with $\theta = 0.75$ (b) Coupled solver using large Δt with $\theta = 0.75$

Fig. 8. Time histories of $E[\phi]$ and $[r(t)]^2$ computed with (a) several small time step sizes, and (b) several large time step sizes. Results are obtained using the coupled algorithm with $\theta = 0.75$ and $S = 1$.

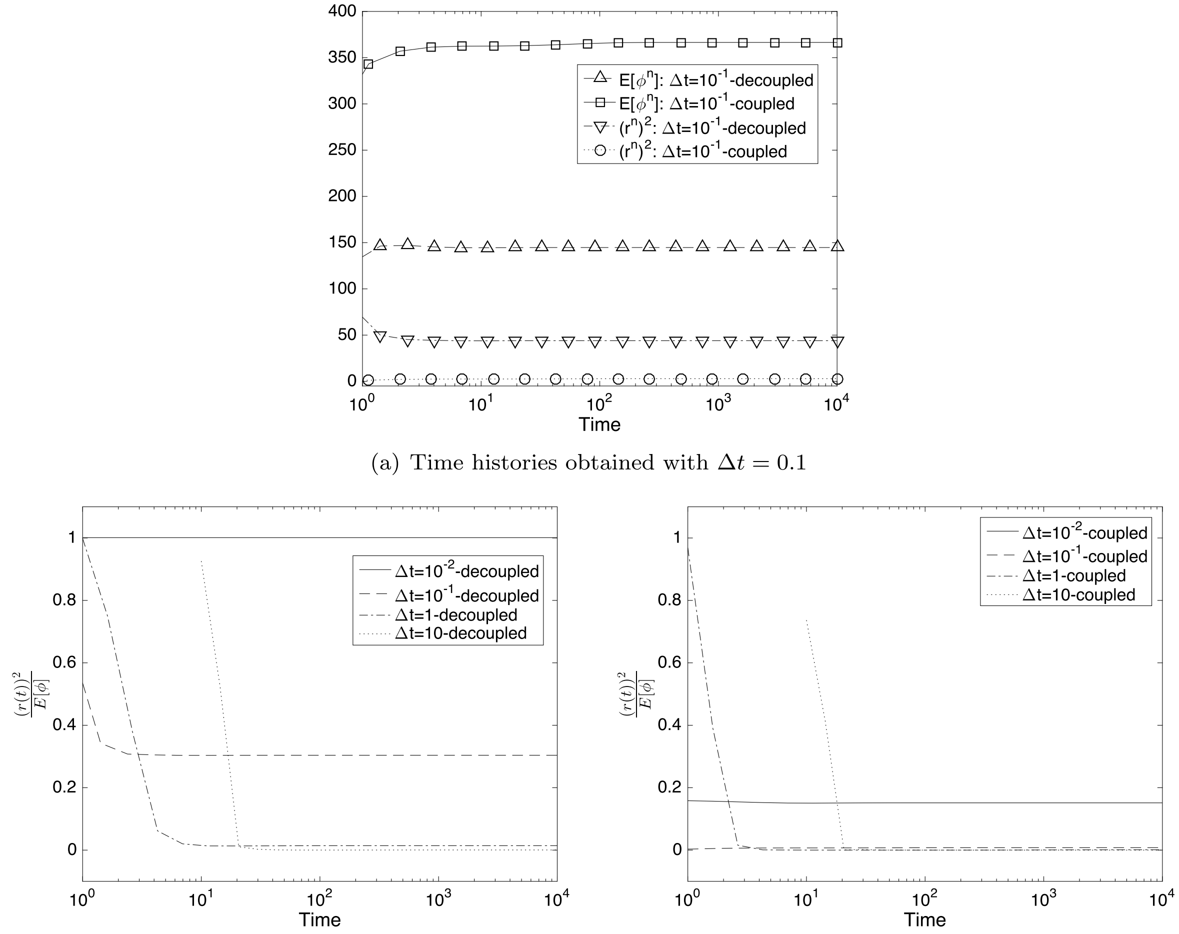
at these time step sizes, except $\Delta t = 0.01$. With $\Delta t = 0.1$ this ratio attains a value around 0.33, and with $\Delta t = 1.0$ and 10 it essentially becomes zero over time. These results suggest that the current schemes are indeed stable at large time step sizes. But their accuracy deteriorates with increasing Δt, and with large time step sizes the computed results are generally no longer accurate. These results also suggest that the de-coupled algorithm with a fixed S seems to result in more accurate simulation results than with S varied with respect to Δt at large time step sizes.

The behavior of the coupled algorithm is illustrated by Fig. 8 with $\theta = 0.75$. Fig. 8(a) shows time histories of $E[\phi]$ and $r^2(t)$ obtained with a range of small time steps sizes ($\Delta t = 10^{-5} \sim 10^{-3}$), and Fig. 8(b) shows long-time histories (up to $t = 10^4$) of the ratio $r^2(t)/E[\phi]$ computed with several large time step sizes ($\Delta t = 0.01 \sim 10$). These results are computed using the coupled algorithm with a constant $S = 1$. For the range of small Δt, the histories for $r^2(t)$ and $E[\phi]$ overlap with one another, and they are the same as those obtained using the de-coupled algorithm (see Fig. 6(a)). With the range of large Δt, the coupled algorithm produces stable results, as is evident from Fig. 8(b). But the accuracy seems to be generally worse than that with the de-coupled algorithm. For example, with $\Delta t = 0.01$ the coupled algorithm gives rise to a ratio $r^2(t)/E[\phi] \approx 0.2$, indicating that the simulation is already inaccurate. In contrast, with the de-coupled algorithm this ratio is the unit value (see Fig. 7), indicating that the computed $r(t)$ is an accurate approximation of $\sqrt{E[\phi]}$ in those simulations.

Similar behaviors have been observed with the other members of the family of schemes for large Δt values. Fig. 9 demonstrates the performance of the de-coupled and coupled algorithms at large time step sizes using $\theta = 1.25$. Fig. 9(a) is a long-time history of $E[\phi]$ and $r^2(t)$ obtained with $\Delta t = 0.1$ using the de-coupled algorithm (with $S = \sqrt{\frac{4\gamma_0\lambda\omega_0}{m\Delta t}}$) and the coupled algorithm (with $S = 1$). The stability of these simulations is evident. There exists a substantial difference between the computed $r^2(t)$ and $E[\phi]$ values, and this difference is significantly larger with the coupled algorithm. This suggests that the simulation results contains significant errors at this Δt and that the errors with the coupled algorithm is much larger. Figs. 9(b) and (c) are time histories of the ratio $r^2(t)/E[\phi]$ obtained with several large time step sizes ($\Delta t = 0.01, 0.1, 1, 10$) using the de-coupled and coupled algorithms with $\theta = 1.25$. We observe similar characteristics in these results to those in Figs. 7(b)–(c) and 8(b) corresponding to $\theta = 0.75$. The methods are evidently stable with these large time step sizes.

The value of constant S in the coupled algorithm is observed to have little or no effect on the simulation results. This point is demonstrated by Fig. 10, in which the time histories of $E[\phi]$ and $r^2(t)$ are depicted corresponding to several S values for two different time step sizes $\Delta t = 10^{-4}$ and $\Delta t = 0.01$ in the coupled algorithm with $\theta = 1.25$. The difference of $E[\phi]$ and $[r(t)]^2$ obtained by different S are negligible for both small and large time step sizes. In the following simulations we will fix $S = 1$ for the coupled algorithm.

Among the family of algorithms ($\frac{1}{2} \leqslant \theta \leqslant \frac{3}{2}$) presented in Section 2, we observe from numerical simulations that the two members on the borders, $\theta = \frac{1}{2}$ and $\theta = \frac{3}{2}$, exhibit a performance inferior to the rest of this family. With the manufactured solution in Section 3.1 we have already observed that the scheme with $\theta = \frac{1}{2}$ (and the θ values in a neighborhood of $\frac{1}{2}$) encounters difficulties for the de-coupled algorithm in long-time simulations. Numerical tests with the current problem reinforce this observation, and additionally reveal difficulties also for the coupled algorithm. Fig. 11 shows the time histories of $E[\phi]$ corresponding to several small Δt values, ranging from 10^{-3} to 10^{-5}, computed using the de-coupled and coupled algorithms with $\theta = \frac{1}{2}$ and $\theta = \frac{3}{2}$. We observe that for the de-coupled algorithm the computation using $\theta = \frac{1}{2}$ becomes unstable with $\Delta t = 10^{-3}$, 10^{-4} and 10^{-5}, and that using $\theta = \frac{3}{2}$ it becomes unstable with $\Delta t = 10^{-3}$. For the coupled algorithm, the computation does not blow up (generating exponentially increasing values). But the computed potential energy $E[\phi]$ is completely wrong with $\Delta t = 10^{-4}$ and 10^{-3} for both $\theta = \frac{1}{2}$ and $\theta = \frac{3}{2}$. In contrast, with the other members of this family of schemes accurate results have been attained with this range of small Δt values; see e.g. Fig. 8(a) for $\theta = 0.75$ and Fig. 10(a) for $\theta = 1.25$. There is a possibility that the spatial discretization might be a contributing factor to the

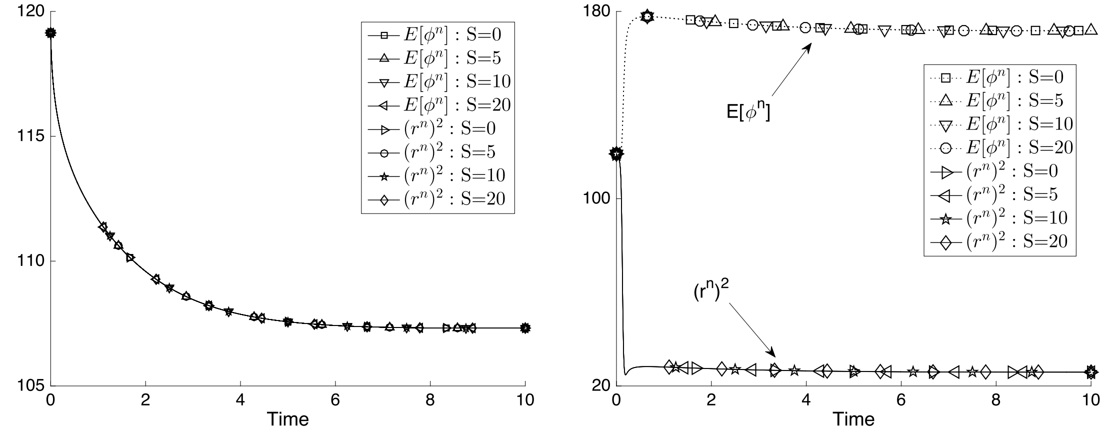


(a) Time histories obtained with $\Delta t = 0.1$

(b) Decoupled solver with $\theta = 1.25$

(c) Coupled solver with $\theta = 1.25$

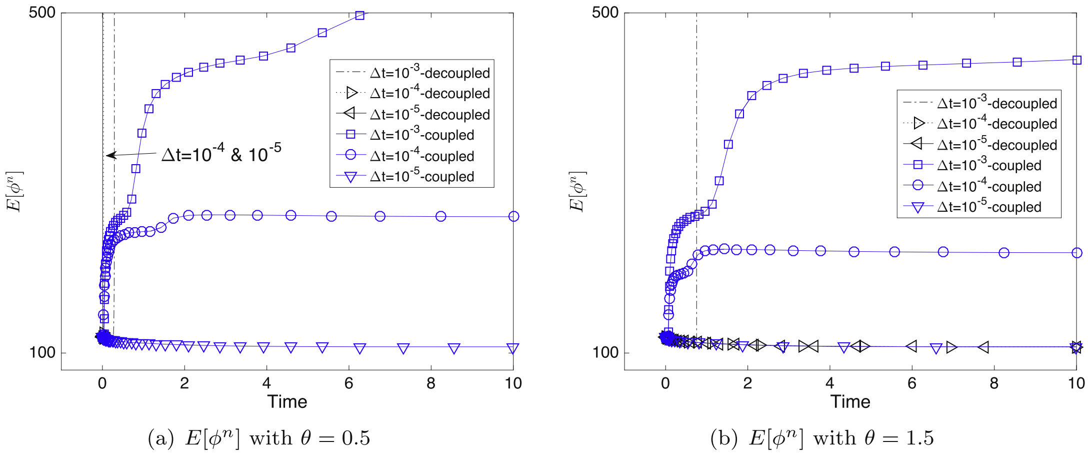
Fig. 9. Simulation results obtained with $\theta = 1.25$ and large Δt: (a) Time histories of $E[\phi]$ and $r^2(t)$ computed with $\Delta t = 0.1$. Time histories of $[r(t)]^2/E[\phi]$ computed using the de-coupled algorithm (b) and the coupled algorithm (c) with several large time step sizes.



(a) coupled solver with $\Delta t = 1e-4$

(b) coupled solver with $\Delta t = 1e-2$

Fig. 10. Time histories of $E[\phi]$ and $[r(t)]^2$ computed using the coupled algorithm with $\theta = 1.25$ for several S values: (a) $\Delta t = 1e{-}4$, (b) $\Delta t = 1e{-}2$.



(a) $E[\phi^n]$ with $\theta = 0.5$

(b) $E[\phi^n]$ with $\theta = 1.5$

Fig. 11. Time histories of $E[\phi]$ computed using the decoupled and coupled algorithms with (a) $\theta = 0.5$ and (b) $\theta = 1.5$, and various Δt values.

poor performance of $\theta = \frac{1}{2}$ and $\frac{3}{2}$. Note that the Theorem 2.1 ensures the energy stability of the semi-discretized algorithm (discrete in time, continuous in space) given by (13a)–(13e) for $\theta = \frac{1}{2}$ and $\theta = \frac{3}{2}$. It is possible that the fully discretized algorithm (in both space and time) may not preserve this energy stability, which may play a role in the instability we have observed with the algorithms $\theta = \frac{1}{2}$ and $\theta = \frac{3}{2}$.

Coalescence of two drops We next consider the coalescence of two material drops governed by the Cahn–Hilliard equation. The computational domain and the setting follow those for the single-drop case discussed above. The difference lies in the initial distribution of the materials. Here we assume that at $t = 0$ the first material occupies two circular regions that are right next to each other and the second material fills the rest of the domain. The two regions of the first material then coalesce with each other to form a single drop under the Cahn–Hilliard dynamics. The goal is to illustrate this process using the current algorithms.

More specifically, we employ the following initial distribution for the materials

$$\phi_{in}(\boldsymbol{x}) = 1 - \tanh \frac{|\boldsymbol{x} - \boldsymbol{x}_0| - R_0}{\sqrt{2}\eta} - \tanh \frac{|\boldsymbol{x} - \boldsymbol{x}_1| - R_0}{\sqrt{2}\eta}, \tag{66}$$

where $\boldsymbol{x}_0 = (x_0, y_0) = (0.3, 0.5)$ and $\boldsymbol{x}_1 = (x_1, y_1) = (0.7, 0.5)$ are the centers of the circular regions for the first material, and $R_0 = 0.19$ is the radius of these circles. In the Cahn–Hilliard equation (1) we set $g = 0$. The boundary conditions (2a) and (2b) with $g_a = 0$ and $g_b = 0$ are imposed on the domain boundaries. The other simulation parameters follow those given in (65).

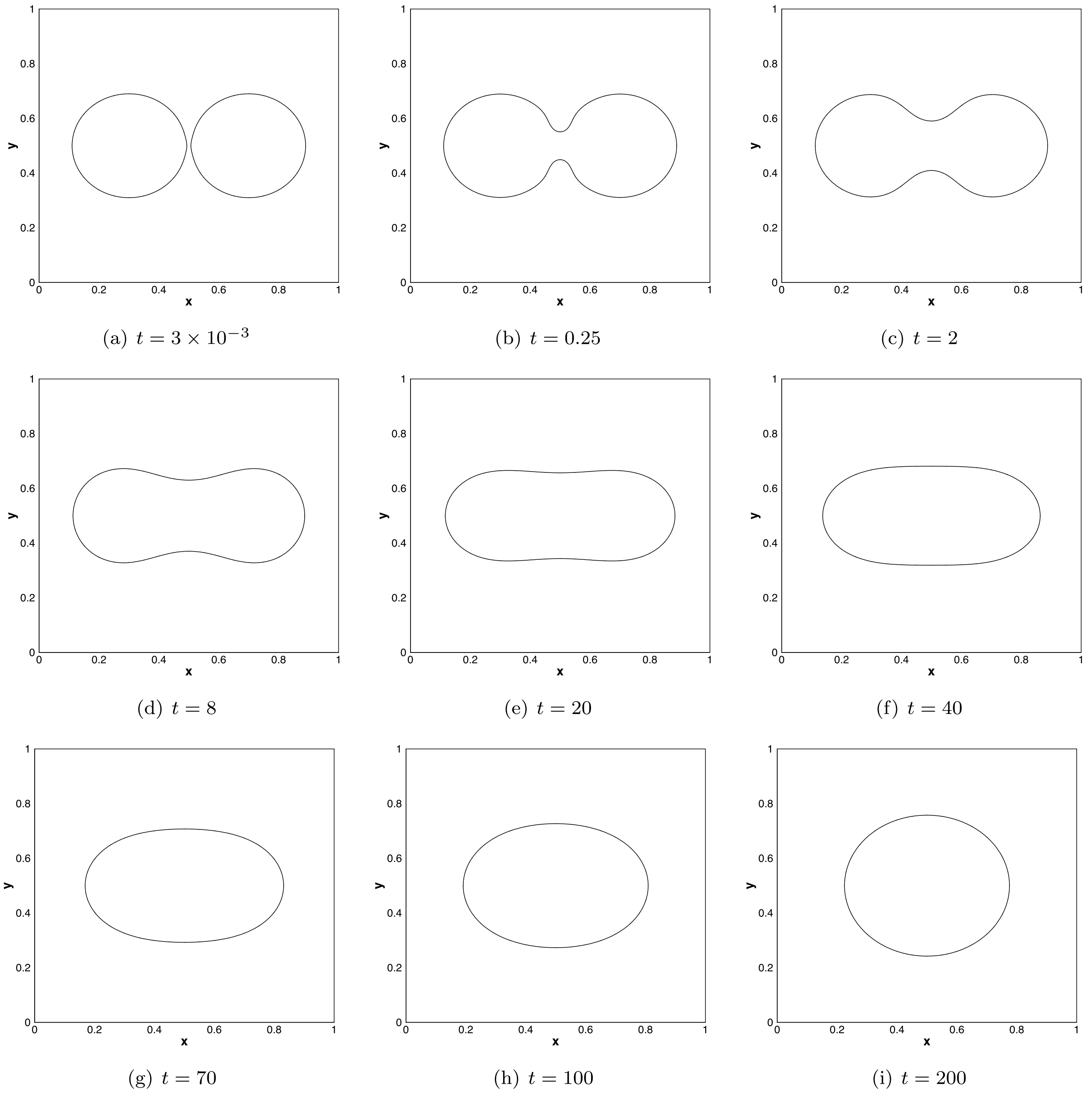
The process of coalescence of the two drops is demonstrated in Fig. 12 with a temporal sequence of snapshots of the interfaces between the materials (visualized by the contour level $\phi = 0$). These results are computed using the de-coupled algorithm with $\theta = 0.52$ and a time step size $\Delta t = 10^{-3}$. Fig. 13 shows the time histories of $E[\phi]$ and $[r(t)]^2$ computed using both the de-coupled and coupled algorithms corresponding to $\theta = 0.52$ and $\theta = 1.48$. The general characteristics for these history signals are similar to those for the case with a single drop. Both $E[\phi]$ and $[r(t)]^2$ are observed to decrease over time, and their history curves essentially overlap with each other.

3.3. Spinodal decomposition

In this section we consider the spinodal decomposition of a homogeneous mixture into two coexisting phases governed by the Cahn–Hilliard equation as another test of the algorithms developed herein.

Consider the domain $\Omega = \{\ (x, y)\ |\ 0 \leqslant x, y \leqslant 1\ \}$, and a homogeneous mixture of two materials with a random initial distribution (see Fig. 14(a)). The evolution of the materials is assumed be described by the Cahn–Hilliard equation (1) with $g(\mathbf{x}, t) = 0$, and the goal is to simulate this evolution process.

We simulate this problem using the algorithms from Section 2. The domain is discretized using 100 quadrilateral elements, with 10 uniform elements along both x and y directions. The boundary conditions (2a) and (2b) with $g_a = 0$ and $g_b = 0$ are imposed on the domain boundaries. The initial random distribution of $\phi_{in}(\mathbf{x})$ is generated using a random number generator from the standard library of C language (see Fig. 14(a)). The following simulation parameter values are employed for this problem:



(a) $t = 3 \times 10^{-3}$

(b) $t = 0.25$

(c) $t = 2$

(d) $t = 8$

(e) $t = 20$

(f) $t = 40$

(g) $t = 70$

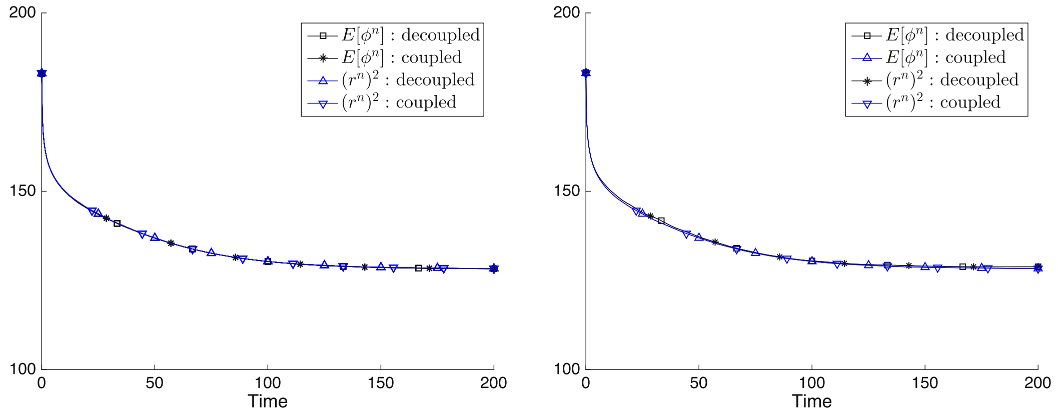
(h) $t = 100$

(i) $t = 200$

Fig. 12. Temporal sequence of snapshots showing the coalescence of two circular drops. Results are obtained with the decoupled algorithm $\theta = 0.52$.

$$\begin{cases} \eta = 0.01, \quad \lambda = 0.001, \quad m = 0.001, \\ C_0 = 0, \quad \Delta t = 0.001 \text{ (or varied)}, \\ S = \sqrt{\dfrac{4\gamma_0 \lambda \omega_0}{m \Delta t}} \text{(decoupled algorithm) or 1 (coupled algorithm)}, \\ \text{element order: 12,} \quad \text{number of elements: 100,} \\ \theta = \ 0.5, \ 0.75, \ 1.0, \ 1.25, \ 1.5. \end{cases} \tag{67}$$

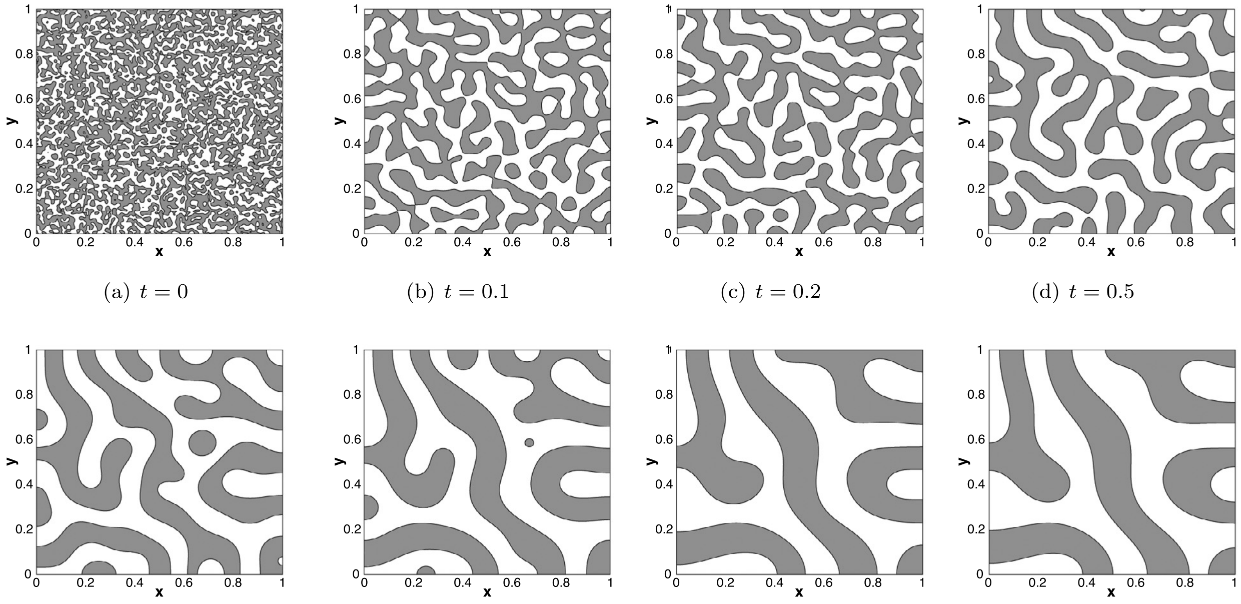
Fig. 14 shows the typical evolution process of the mixture with a temporal sequence of snapshots of the interfaces formed between the two phases. The interface is visualized by the contour level $\phi(\mathbf{x}, t) = 0$. The lighter regions represent the first phase and the darker region represents the second phase. These results are obtained using the de-coupled algorithm with $\theta = 1.25$. It can be observed that two phases emerge from the initially homogeneous distribution of the mixture. Over time the grains of the two phases become increasingly coarser, and a certain pattern can be observed from the two regions.



(a) $\theta = 0.52$

(b) $\theta = 1.48$

Fig. 13. Time histories of $E[\phi]$ and $[r(t)]^2$ for the coalescence of two drops, computed using the de-coupled and coupled algorithms with (a) $\theta = 0.52$ and (b) $\theta = 1.48$.



(a) $t = 0$

(b) $t = 0.1$

(c) $t = 0.2$

(d) $t = 0.5$

(e) $t = 2$

(f) $t = 4$

(g) $t = 7$

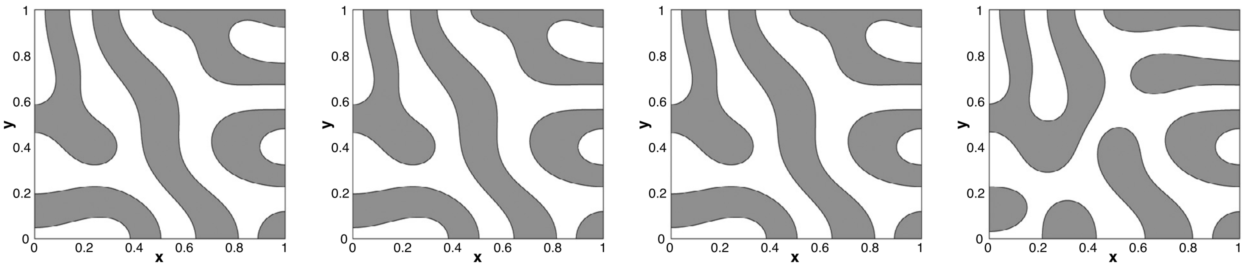
(h) $t = 10$

Fig. 14. Spinodal decomposition: time sequence of snapshots of $\phi = 0$ at (a) $t = 0$, (b) $t = 0.1$, (c) $t = 0.2$, (d) $t = 0.5$, (e) $t = 2$, (f) $t = 4$, (g) $t = 7$, (h) $t = 10$. Results correspond to $\theta = 1.25$ and $\Delta t = 10^{-3}$.

The distributions of the material interfaces at $t = 10$ obtained with several time step sizes, ranging from $\Delta t = 10^{-5}$ to $\Delta t = 10^{-2}$, using the de-couple algorithm with $\theta = 1.25$ are shown in Fig. 15. The distributions computed with $\Delta t = 10^{-5}$, 10^{-4} and 10^{-3} are essentially the same. With the larger time step size $\Delta t = 10^{-2}$, we can observe some notable difference in the material distribution from those obtained using smaller Δt values, suggesting that the simulation is starting to lose accuracy notably with this time step size (or larger).

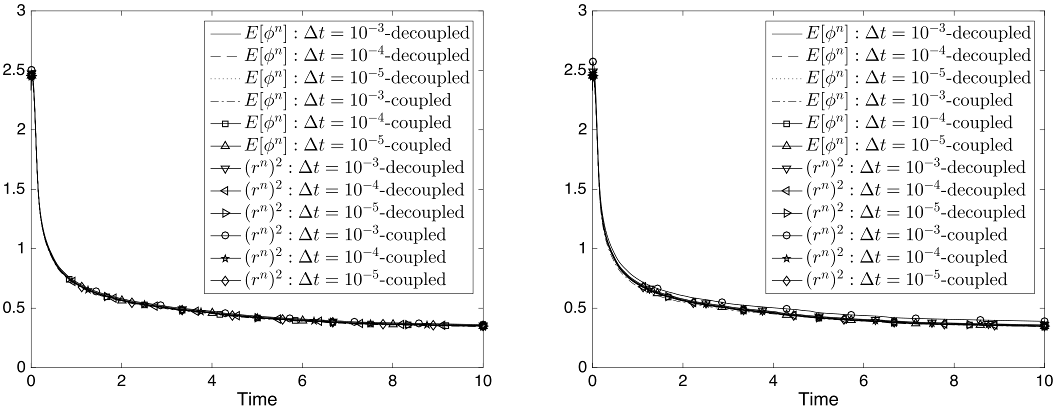
Fig. 16 shows the time histories of the potential free energy $E[\phi]$ and the variable $[r(t)]^2$ of the system obtained using both the decoupled and coupled algorithms, corresponding to $\theta = 0.75$ and $\theta = 1.25$ with several small time step sizes $\Delta t = 10^{-5}, 10^{-4}, 10^{-3}$. We note that the history curves for $E[\phi]$ and $[r(t)]^2$ all overlap with one another with $\theta = 0.75$. Those time histories corresponding to $\theta = 1.25$ also overlap with one another essentially, and some difference can be discerned in the curve obtained using the coupled algorithm with $\Delta t = 10^{-3}$ when compared with the rest.

Similar to the test problems in previous subsections, stable simulation results have been obtained with the current method for the spinodal decomposition using fairly large and large time step sizes. Fig. 17 is a demonstration of this point using the time histories of $E[\phi]$ and $r^2(t)$ computed with the de-coupled and coupled algorithms ($\theta = 0.75$) at several large or fairly large time step sizes: $\Delta t = 0.01$, 0.1, 1 and 10. The long time histories (up to $t = 1000$) signifies the stability of



(a) $\Delta t = 10^{-5}$ (b) $\Delta t = 10^{-4}$ (c) $\Delta t = 10^{-3}$ (d) $\Delta t = 10^{-2}$

Fig. 15. Spinodal decomposition: snapshots of material interfaces at $t = 10$ computed with: (a) $\Delta t = 10^{-5}$, (b) $\Delta t = 10^{-4}$, (c) $\Delta t = 10^{-3}$, (d) $\Delta t = 10^{-2}$. Results are obtained using the decoupled algorithm with $\theta = 1.25$.



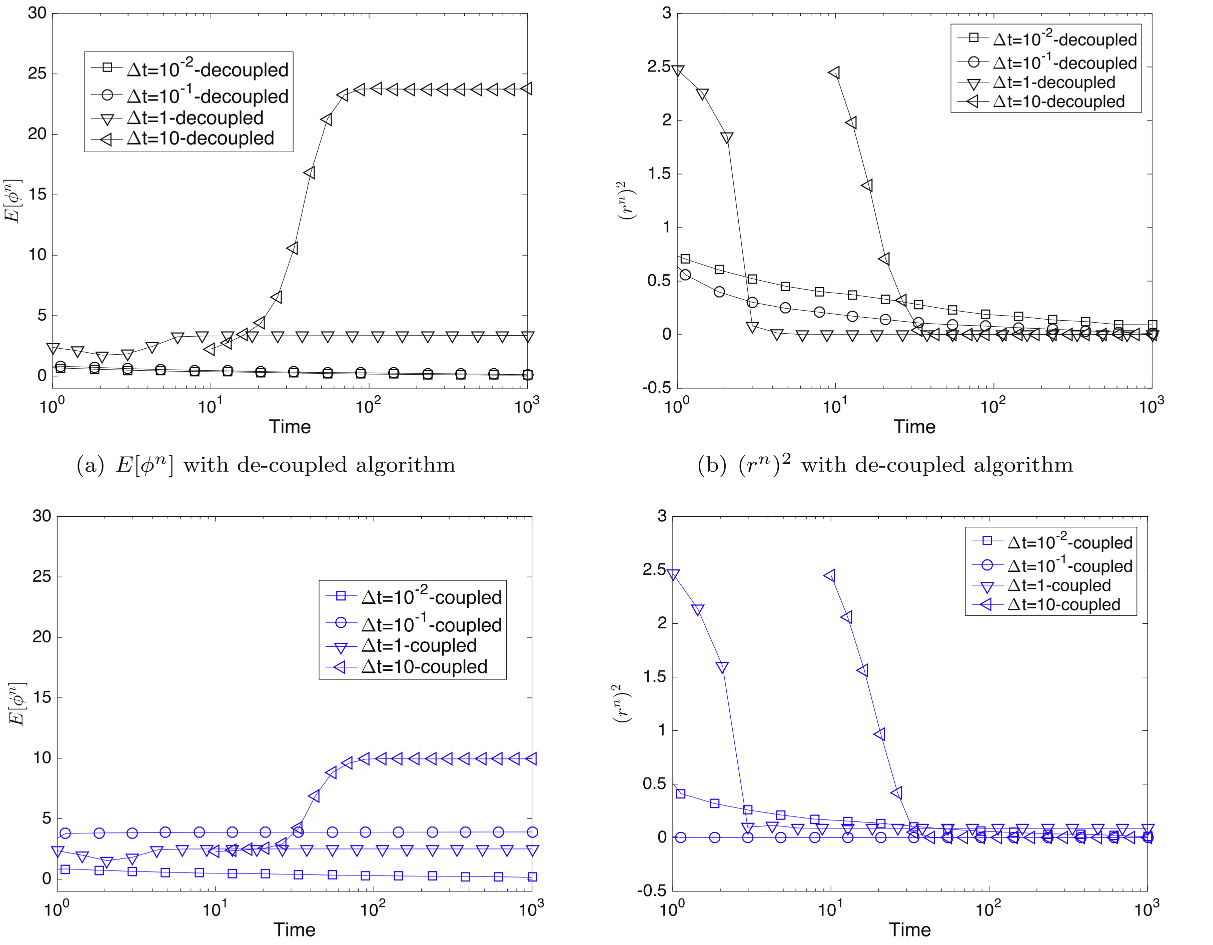
(a) $\theta = 0.75$ (b) $\theta = 1.25$

Fig. 16. Time histories of $E[\phi]$ and $[r(t)]^2$ for spinoidal decomposition obtained with the schemes (a) $\theta = 0.75$, (b) $\theta = 1.25$ and a range of small time step sizes.

simulations. At the large Δt values we can no longer expect the results to be accurate. Indeed, significant differences can be noticed between the $E[\phi]$ and $r^2(t)$ histories corresponding to $\Delta t = 0.1$ and larger time step sizes.

The performance of the two members on the boundary of this family of schemes ($\theta = \frac{1}{2}$ and $\frac{3}{2}$) is demonstrated by Fig. 18 for the spinodal decomposition problem. Here we have plotted the time histories of $E[\phi]$ corresponding to $\theta = \frac{1}{2}$ in (a) and $\theta = \frac{3}{2}$ in (b) obtained with a range of small time step sizes ($\Delta t = 10^{-5}$, 10^{-4} and 10^{-3}) using the de-coupled and coupled algorithms. We observe similar behaviors with these two schemes to those in previous sections. For the scheme $\theta = \frac{1}{2}$, simulations using the de-coupled algorithm blow up shortly into the run with all three time step sizes. The coupled algorithm does not blow up. But it produces erroneous $E[\phi]$ histories with time step sizes $\Delta t = 10^{-3}$ and 10^{-4} (Fig. 18(a)). We observe that, shortly into the simulation, the computed $E[\phi]$ value with the coupled algorithm abruptly jumps to a new higher level and remains a constant at that level throughout the rest of computation. With the smallest $\Delta t = 10^{-5}$, the de-coupled algorithm produces a reasonable $E[\phi]$ signal. The scheme $\theta = \frac{3}{2}$ exhibits behaviors similar to, and perhaps slightly better than, the scheme $\theta = \frac{1}{2}$. For $\theta = \frac{3}{2}$, the de-coupled algorithm produces reasonable $E[\phi]$ histories with $\Delta t = 10^{-5}$ and 10^{-4}, and it blows up with $\Delta t = 10^{-3}$. On the other hand, the coupled algorithm produces a reasonable $E[\phi]$ history with $\Delta t = 10^{-5}$, but gives rise to erroneous $E[\phi]$ signals with the larger $\Delta t = 10^{-4}$ and 10^{-3}. These results can be compared with those produced by other members of this family of schemes (see e.g. Fig. 16). The two boundary schemes ($\theta = \frac{1}{2}$ and $\frac{3}{2}$), while apparently "energy-conserving" according to equation (12), give rise to simulation results inferior to the rest of this family of schemes in practice.

Fig. 19 is a comparison of the accuracy corresponding to several representative θ values among the family of schemes developed herein. Here we plot the ratio $r^2(t)/E[\phi]$ as a function of time obtained with the time step size $\Delta t = 0.02$ corresponding to $\theta = 0.52$, 0.75, 1.0, 1.25 and 1.48 in the de-coupled algorithm for the spinodal decomposition problem. This ratio should physically be the unit value, and a deviation from the unit value indicates errors in the simulation results. The result shows that a larger θ in the scheme results in a consistently larger deviation of the ratio $r^2(t)/E[\phi]$ from the unit value. This suggests that the smaller the θ is the more accurate the scheme tends to be. However, the scheme at the lower boundary $\theta = \frac{1}{2}$ (and those θ values in a small neighborhood of $\frac{1}{2}$) tends to produce poor results in terms of stability and accuracy, as is demonstrated by the test results in this and previous sections. The numerical results from this



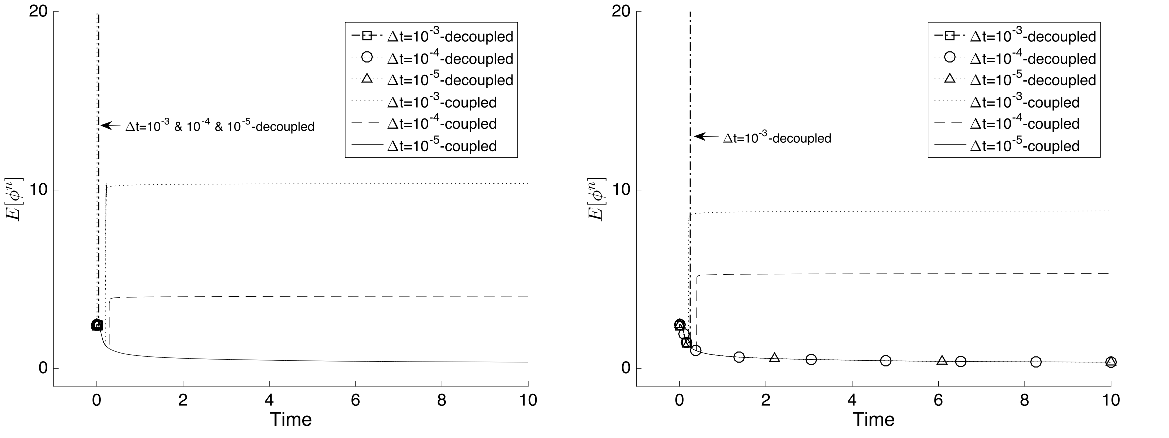
(a) $E[\phi^n]$ with de-coupled algorithm

(b) $(r^n)^2$ with de-coupled algorithm

(c) $E[\phi^n]$ with coupled algorithm

(d) $(r^n)^2$ with coupled algorithm

Fig. 17. Time histories of $E[\phi]$ ((a) and (c)) and $[r(t)]^2$ ((b) and (d)) for spinoidal decomposition computed with $\theta = 0.75$ using various time step sizes. Results are obtained by the de-coupled algorithm ((a) and (b)) and coupled algorithm ((c) and (d)).



(a) $E[\phi^n]$ with $\theta = 0.5$

(b) $E[\phi^n]$ with $\theta = 1.5$

Fig. 18. Time histories of $E[\phi]$ for spinoidal decomposition obtained using the de-coupled and coupled algorithms with (a) $\theta = 0.5$, and (b) $\theta = 1.5$.

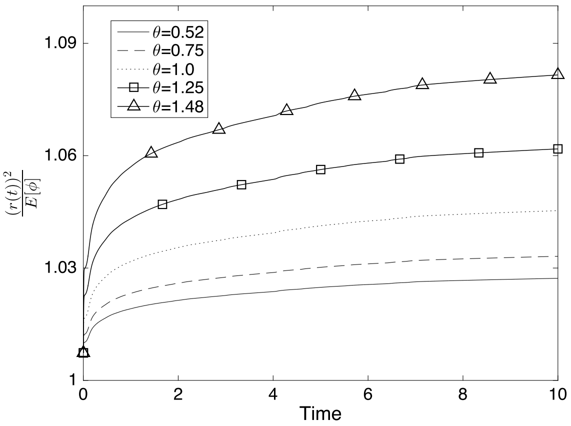


Fig. 19. Time histories of $[r(t)]^2/E[\phi]$ for spinodal decomposition computed using the de-coupled algorithm with various θ values and $\Delta t = 0.02$.

and previous sections suggest that in terms of accuracy and stability the best members among this family of energy-stable schemes appears to correspond to those θ values close to $\theta \approx 0.52$.

3.4. Two-phase flow: rising air bubble in water

As another test case, in this section we demonstrate the performance of the method developed herein in the context of a two-phase flow solver, and simulate the two-phase flow of an air bubble rising through the water.

Following [12,7], we consider a system consisting of two immiscible incompressible fluids, and combine the Cahn–Hilliard equation and the incompressible Navier–Stokes equations with variable density and variable viscosity to model such a system; see [12,13] for details. We then combine the family of schemes presented in Section 2 for the Cahn–Hilliard equation and the algorithm from [12] for the momentum equations to form an overall method for numerically solving the coupled system of Cahn–Hilliard and Navier–Stokes equations. Note that the algorithm for the momentum equation employed here is a semi-implicit type scheme and is only conditionally stable [12]. So the overall algorithm for the two-phase flows is not energy stable. Only the de-coupled algorithm for the Cahn–Hilliard equation (see Section 2.3.1) has been implemented in the two-phase flow solver.

We consider a solid container occupying the domain $-\frac{L}{2} \leqslant x \leqslant \frac{L}{2}$ and $0 \leqslant y \leqslant \frac{3}{2}L$, where $L = 1$ cm (see Fig. 20(a)). The container is filled with water, and an air bubble is trapped inside the water. The air bubble is initially circular with a diameter $2R_0 = 0.5L = 0.5$ cm and its center located at $\mathbf{x}_0 = (x_0, y_0) = (0, 0.5L)$, and it is held at rest. The container walls are assumed to have a neutral wettability (90-degree contact-angle), and the gravity is assumed to point downward. At $t = 0$, the air bubble is released, and starts to rise through the water due to buoyancy. The bubble eventually breaks up on the upper wall and forms an air cavity at the top of the container. The goal is to simulate this process.

The physical parameters for the air and water are summarized in Table 3. We choose L as the characteristic length scale, the air density as the characteristic density scale ϱ_0, and $U_0 = \sqrt{g_{r0}L}$ ($g_{r0} = 1$ m/s^2) as the characteristic velocity scale. All the physical parameters are then normalized according to Table 1.

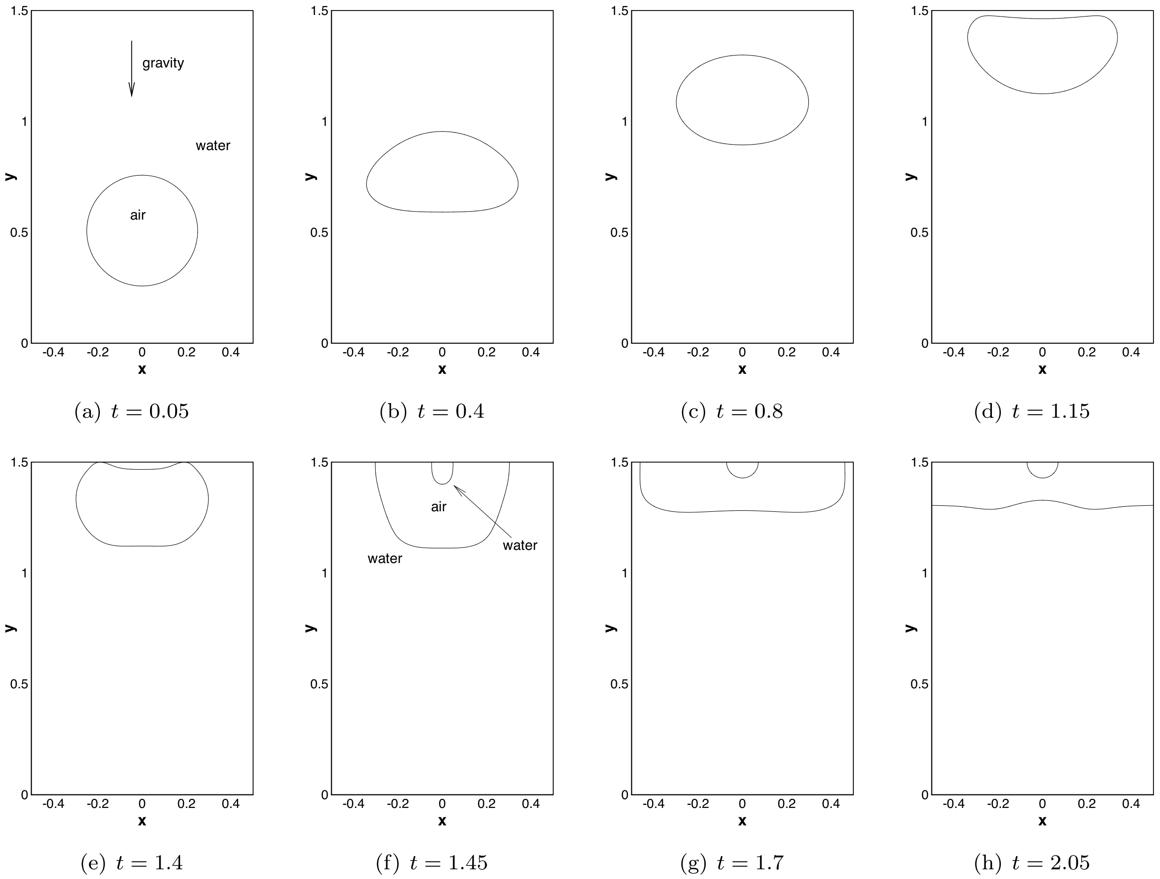
In the simulations the domain is discretized using 600 quadrilateral elements, with 20 and 30 uniform elements in x and y directions, respectively. An element order 8 is employed in the simulations. We impose the boundary conditions (2a) and (2b) with $g_a = 0$ and $g_b = 0$ for the phase field function ϕ and the no-slip condition for the velocity on the container walls. The initial velocity is assumed to be zero, and the initial distribution of the phase field function is given by

$$\phi_{in}(\boldsymbol{x}) = -\tanh \frac{|\boldsymbol{x} - \boldsymbol{x}_0| - R_0}{\sqrt{2}\eta}. \tag{68}$$

The values for the simulation parameters in this problem are given by

$$\begin{cases} \dfrac{\eta}{L} = 0.01, \quad \dfrac{\sigma}{\varrho_0 U_0^2 L} = 604.6 \text{ (surface tension)}, \quad \lambda = \dfrac{3}{2\sqrt{2}}\sigma\eta, \quad \dfrac{m}{L/(\varrho_0 U_0)} = \dfrac{10^{-7}}{\lambda/(\varrho_0 U_0^2 L^2)}, \\ C_0 = 0, \quad \dfrac{\Delta t}{L/U_0} = 2.5\text{e}{-5}, \quad S = \sqrt{\dfrac{4\gamma_0 \lambda \omega_0}{m\Delta t}}, \\ \text{element order: 8,} \quad \text{number of elements: 600,} \\ \theta = 0.75. \end{cases} \tag{69}$$

The dynamics of this two-phase flow is illustrated by Fig. 20, in which we have shown a temporal sequence of snapshots of the air-water interface in this system. The interface is visualized by the contour level $\phi(\mathbf{x}, t) = 0$. As the system is released



(a) $t = 0.05$ (b) $t = 0.4$ (c) $t = 0.8$ (d) $t = 1.15$

(e) $t = 1.4$ (f) $t = 1.45$ (g) $t = 1.7$ (h) $t = 2.05$

Fig. 20. Temporal sequence of snapshots of an air bubble (initially circular) rising in water within a solid container, showing the bubble breakup on the upper wall and the formation of an air cavity.

Table 3
Physical property values of air and water.

Density [kg/m^3]:	air – $\rho_1 = 1.204$	water – $\rho_2 = 998.207$
Dynamic viscosity [kg/(m s)]:	air – $\mu_1 = 1.78 \times 10^{-5}$	water – $\mu_2 = 1.002 \times 10^{-3}$
Surface tension [kg/s^2]:	air/water – $\sigma = 7.28 \times 10^{-2}$	
Gravity [m/s^2]:	$g_r = 9.8$	

the air bubble rises through the water and experiences significant deformations (Fig. 20(b)–(d)). The bubble exhibits the typical shape of a circular “cap” (Fig. 20(b)) when it is still far away from the upper wall. But as the bubble approaches the upper wall, its shape is affected by the presence of the wall significantly (Figs. 20(c)–(d)). After the bubble touches the upper wall, it traps a layer of water between the wall and its upper side (Fig. 20(e)). The trapped water becomes a water drop sitting on the upper wall over time (Fig. 20(g)). The interface formed between the bulk of air and the bulk of water moves sideways on the wall, and the air forms a cavity at the top of the container (Figs. 20(f)–(h)). Our method has captured this process and the interaction between the air-water interface and the wall.

4. Concluding remarks

In this paper we have developed a family of energy-stable schemes for the Cahn–Hilliard type equations. We start with the reformulated system of equations based on the scalar auxiliary variable approach, and approximate this system at time step $(n+\theta)$, and then develop corresponding numerical approximations that are second-order accurate and unconditionally energy stable. This family of approximations contains the often-used Crank–Nicolson scheme and the second-order backward differentiation formula as particular cases. We have also developed efficient solution algorithms to overcome the difficulty caused by the unknown scalar auxiliary variable in the discrete system of equations resulting from this family of schemes. Our overall method only requires the solution of four de-coupled individual Helmholtz type equations with the de-coupled algorithm, and the solution of two de-coupled individual systems of equations (with each system consisting of two coupled

Helmholtz type equations) with the coupled algorithm within each time step. All the equations involved in the solution algorithms only concern constant and time-independent coefficient matrices that can be pre-computed.

Among the family of approximations ($\frac{1}{2} \leqslant \theta \leqslant \frac{3}{2}$) considered in the current work, the schemes with $\theta = \frac{1}{2}$ and $\theta = \frac{3}{2}$ involve zero numerical dissipation, and the rest of the family are numerically dissipative, with $\theta = \frac{2}{3} + \frac{\sqrt{7}}{6}$ being the most dissipative. The two boundary schemes $\theta = \frac{1}{2}$ and $\frac{3}{2}$, while energy-conserving, exhibit consistently inferior performance to the rest of the family of schemes in actual simulations. Both the de-coupled and the coupled algorithms encounter difficulties (in terms of stability or accuracy) with these two schemes in actual long-time simulations. Poor performance is also observed for schemes with θ in a small neighborhood of $\frac{1}{2}$ in long-time simulations. For the rest of this family of schemes, stable computations can be achieved with time step sizes ranging from very small to very large values in actual long-time simulations, verifying their unconditional energy stability. In terms of their accuracy, numerical tests show that the scheme with a smaller θ produces more accurate simulation results. All these results suggest that the best members among this family of schemes, in terms of both stability and accuracy, are those with a θ value toward the lower end (close to $\theta \approx 0.52$), but not the lower boundary ($\theta = \frac{1}{2}$).

In the discrete formulation for the Cahn–Hilliard type equations, the incorporation of the extra term $S(\phi^{n+1} - 2\phi^n + \phi^{n-1})$, which is equivalent to zero to the second order accuracy in time, is not essential to the unconditional energy stability of the current family of schemes. To the coupled solution algorithm for the resultant equations, it is also not essential. But to the de-coupled solution algorithm this term is critical, and it enables the reformulation of the fourth-order equation into two de-coupled Helmholtz type equations. Tests with the manufactured analytical solution indicate that the numerical error associated with this term is proportional to the value of the chosen constant S. Tests with actual simulations, however, show that the value of the constant S apparently has little or basically no effect on the actual simulation results.

While the current paper is only concerned with the numerical approximation of the Cahn–Hilliard equation, the family of energy-stable approximations herein is readily applicable to other types of equations resulting from gradient flows. When combined with the invariant energy quadratization or scalar auxiliary variable approach, they can be readily used to design energy-stable schemes for other gradient flows.

Acknowledgement

This work was partially supported by NSF (DMS-1318820, DMS-1522537).

Appendix A. Proof of Theorem 2.1

We note first the following useful relation (χ denoting a generic scalar variable):

$$\begin{aligned}&(\chi^{n+1} - 2\chi^n + \chi^{n-1})\left[\left(\theta + \frac{1}{2}\right)\chi^{n+1} - 2\theta\chi^n + \left(\theta - \frac{1}{2}\right)\chi^{n-1}\right] \\ &\quad = \theta\left|\chi^{n+1} - 2\chi^n + \chi^{n-1}\right|^2 + \frac{1}{2}\left(\left|\chi^{n+1} - \chi^n\right|^2 - \left|\chi^n - \chi^{n-1}\right|^2\right). \end{aligned} \tag{70}$$

This relation can be verified by elementary operations.

Assume that $g = 0$, $g_a = 0$, and $g_b = 0$. Multiply $\Delta t \mathcal{H}^{n+\theta}$ to equation (13a) and integrate over Ω, and we have

$$\int_\Omega \left[\left(\theta + \frac{1}{2}\right)\phi^{n+1} - 2\theta\phi^n + \left(\theta - \frac{1}{2}\right)\phi^{n-1}\right]\mathcal{H}^{n+\theta} = m\Delta t \int_\Omega \mathcal{H}^{n+\theta}\nabla^2\mathcal{H}^{n+\theta}. \tag{71}$$

Multiplying $\Delta t \left.\frac{\partial\phi}{\partial t}\right|^{n+\theta}$ to equation (13b) and integrating over Ω leads to

$$\begin{aligned}&\int_\Omega \mathcal{H}^{n+\theta}\left[\left(\theta + \frac{1}{2}\right)\phi^{n+1} - 2\theta\phi^n + \left(\theta - \frac{1}{2}\right)\phi^{n-1}\right] \\ &= -\lambda\int_\Omega \nabla^2\phi^{n+\theta}\left[\left(\theta + \frac{1}{2}\right)\phi^{n+1} - 2\theta\phi^n + \left(\theta - \frac{1}{2}\right)\phi^{n-1}\right] \\ &\quad + S\int_\Omega (\phi^{n+1} - 2\phi^n + \phi^{n-1})\left[\left(\theta + \frac{1}{2}\right)\phi^{n+1} - 2\theta\phi^n + \left(\theta - \frac{1}{2}\right)\phi^{n-1}\right] \\ &\quad + \frac{r^{n+\theta}}{\sqrt{E[\bar{\phi}^{n+\theta}]}}\int_\Omega h(\bar{\phi}^{n+\theta})\left[\left(\theta + \frac{1}{2}\right)\phi^{n+1} - 2\theta\phi^n + \left(\theta - \frac{1}{2}\right)\phi^{n-1}\right]. \end{aligned} \tag{72}$$

Multiplying $2r^{n+\theta}\Delta t$ to equation (13c) leads to

$$
\begin{aligned}
&2r^{n+\theta}\left[\left(\theta+\frac{1}{2}\right)r^{n+1}-2\theta r^{n}+\left(\theta-\frac{1}{2}\right)r^{n-1}\right]\\
&=\frac{r^{n+\theta}}{\sqrt{E[\bar{\phi}^{n+\theta}]}}\int_{\Omega} h(\bar{\phi}^{n+\theta})\left[\left(\theta+\frac{1}{2}\right)\phi^{n+1}-2\theta\phi^{n}+\left(\theta-\frac{1}{2}\right)\phi^{n-1}\right].
\end{aligned}
\tag{73}
$$

Summing up equations (71) and (73), and then subtracting equation (72), results in

$$
\begin{aligned}
&2r^{n+\theta}\left[\left(\theta+\frac{1}{2}\right)r^{n+1}-2\theta r^{n}+\left(\theta-\frac{1}{2}\right)r^{n-1}\right]\\
&=-m\Delta t\int_{\Omega}\nabla\mathcal{H}^{n+\theta}\cdot\nabla\mathcal{H}^{n+\theta}-\lambda\int_{\Omega}\nabla\phi^{n+\theta}\cdot\left[\left(\theta+\frac{1}{2}\right)\nabla\phi^{n+1}-2\theta\nabla\phi^{n}+\left(\theta-\frac{1}{2}\right)\nabla\phi^{n-1}\right]\\
&\quad-S\int_{\Omega}\left(\phi^{n+1}-2\phi^{n}+\phi^{n-1}\right)\left[\left(\theta+\frac{1}{2}\right)\phi^{n+1}-2\theta\phi^{n}+\left(\theta-\frac{1}{2}\right)\phi^{n-1}\right]
\end{aligned}
\tag{74}
$$

where we have performed integration by part on the right-hand-side (RHS) of equation (71) and the first term on the RHS of equation (72), and used equations (13b), (13d) and (13e).

Use the relations (70) and (12) to transform the corresponding terms in (74), and then collect related terms, and one would arrive at the discrete energy balance equation (14).

Appendix B. Proof of Theorem 2.3

Substituting ψ_2^{n+1} from (36a) into equations (35a) and (35b) leads to

$$
\nabla^2\left(\nabla^2\phi_2^{n+1}\right)-\frac{S}{\lambda\omega_0}\nabla^2\phi_2^{n+1}+\frac{\gamma_0}{\lambda\omega_0 m\Delta t}\phi_2^{n+1}=\frac{1}{2\lambda}\nabla^2 b^n,
\tag{75a}
$$

$$
\mathbf{n}\cdot\nabla\left(\nabla^2\phi_2^{n+1}\right)=\frac{1}{2\lambda}\mathbf{n}\cdot\nabla b^n,\quad \text{on } \partial\Omega,
\tag{75b}
$$

where we have used equation (36b) and the relation $\alpha\left(\alpha+\frac{S}{\lambda\omega_0}\right)=-\frac{\gamma_0}{\lambda\omega_0 m\Delta t}$. The system of equations (75a), (75b) and (36b) is equivalent to the system consisting of equations (35a)–(36b). By integrating equation (75a) over the domain Ω, we conclude that ϕ_2^{n+1} has the property

$$
\int_{\Omega}\phi_2^{n+1}=0,
\tag{76}
$$

where we have used the divergence theorem and the equations (75b) and (36b).

Define function $\xi(\mathbf{x})$ by

$$
\begin{cases}
\nabla^2\xi=\phi_2^{n+1},\\
\mathbf{n}\cdot\nabla\xi=0,\quad \text{on } \partial\Omega,\\
\displaystyle\int_{\Omega}\xi=\frac{\omega_0 m\Delta t}{2\gamma_0}\int_{\Omega}b^n.
\end{cases}
\tag{77}
$$

Let

$$
\Phi=\nabla^2\phi_2^{n+1}-\frac{S}{\lambda\omega_0}\phi_2^{n+1}+\frac{\gamma_0}{\lambda\omega_0 m\Delta t}\xi-\frac{1}{2\lambda}b^n.
\tag{78}
$$

Then equations (75a), (75b) and (76) are transformed into

$$
\nabla^2\Phi=0,
\tag{79a}
$$

$$
\mathbf{n}\cdot\nabla\Phi=0,\quad \text{on } \partial\Omega,
\tag{79b}
$$

$$
\int_{\Omega}\Phi=0,
\tag{79c}
$$

where we have used (77) and (36b). So we conclude that $\Phi=0$ and

$$\nabla^2\phi_2^{n+1} - \frac{S}{\lambda\omega_0}\phi_2^{n+1} + \frac{\gamma_0}{\lambda\omega_0 m\Delta t}\xi = \frac{1}{2\lambda}b^n. \tag{80}$$

Taking the L^2 inner product between this equation and ϕ_2^{n+1} and integrating by part, we get

$$\frac{1}{2\lambda}\int_\Omega b^n\phi_2^{n+1} = -\int_\Omega \left|\nabla\phi_2^{n+1}\right|^2 - \frac{S}{\lambda\omega_0}\int_\Omega \left|\phi_2^{n+1}\right|^2 - \frac{\gamma_0}{\lambda\omega_0 m\Delta t}\int_\Omega |\nabla\xi|^2 \leqslant 0, \tag{81}$$

where we have used the divergence theorem and equation (77).

References

[1] H. Abels, H. Garcke, G. Grün, Thermodynamically consistent, frame indifferent diffuse interface models for incompressible two-phase flows with different densities, Int. J. Math. Models Methods Appl. Sci. 22 (2012) 1150013.

[2] D.M. Anderson, G.B. McFadden, A.A. Wheeler, Diffuse-interface methods in fluid mechanics, Annu. Rev. Fluid Mech. 30 (1998) 139–165.

[3] V.E. Badalassi, H.D. Ceniceros, S. Banerjee, Computation of multiphase systems with phase field models, J. Comput. Phys. 190 (2003) 371–397.

[4] S. Badia, F. Guillen-Gonzalez, J.V. Gutierrez-Santacreu, Finite element approximation of nematic liquid crystal flows using a saddle-point structure, J. Comput. Phys. 230 (2011) 1686–1706.

[5] J. Bosch, C. Kahle, M. Stoll, Preconditioning of a coupled Cahn–Hilliard Navier–Stokes system, Commun. Comput. Phys. 23 (2018) 603–628.

[6] J.W. Cahn, J.E. Hilliard, Free energy of a nonuniform system, I: interfacial free energy, J. Chem. Phys. 28 (1958) 258–267.

[7] S. Dong, On imposing dynamic contact-angle boundary conditions for wall-bounded liquid–gas flows, Comput. Methods Appl. Mech. Eng. 247–248 (2012) 179–200.

[8] S. Dong, An efficient algorithm for incompressible N-phase flows, J. Comput. Phys. 276 (2014) 691–728.

[9] S. Dong, An outflow boundary condition and algorithm for incompressible two-phase flows with phase field approach, J. Comput. Phys. 266 (2014) 47–73.

[10] S. Dong, Wall-bounded multiphase flows of N immiscible incompressible fluids: consistency and contact-angle boundary condition, J. Comput. Phys. 338 (2017) 21–67.

[11] S. Dong, Multiphase flows of N immiscible incompressible fluids: a reduction-consistent and thermodynamically-consistent formulation and associated algorithm, J. Comput. Phys. 361 (2018) 1–49.

[12] S. Dong, J. Shen, A time-stepping scheme involving constant coefficient matrices for phase field simulations of two-phase incompressible flows with large density ratios, J. Comput. Phys. 231 (2012) 5788–5804.

[13] S. Dong, X. Wang, A rotational pressure-correction scheme for incompressible two-phase flows with open boundaries, PLoS ONE 11 (5) (2016) e0154565.

[14] C.M. Elliot, D.A. French, F.A. Milner, A second order splitting method for the Cahn–Hilliard equation, Numer. Math. 54 (1989) 575–590.

[15] C.M. Elliot, A.M. Stuart, The global dynamics of discrete semilinear parabolic equations, SIAM J. Numer. Anal. 30 (1993) 1622–1663.

[16] J.D. Eyre, An unconditionally stable one-step scheme for gradient system, unpublished, www.math.utah.edu/~eyre/research/methods/stable.ps.

[17] X. Feng, Fully discrete finite element approximations of the Navier–Stokes–Cahn–Hilliard diffuse interface model for two-phase fluid flows, SIAM J. Numer. Anal. 44 (2006) 1049–1072.

[18] X. Feng, A. Prohl, Error analysis of a mixed finite element method for the Cahn–Hilliard equation, Numer. Math. 99 (2004) 47–84.

[19] D. Furihata, A stable and conservative finite difference scheme for the Cahn–Hilliard equation, Numer. Math. 87 (2001) 675–699.

[20] H. Gomez, T.J.R. Hughes, Provably unconditionally stable, second-order time accurate, mixed variational methods for phase-field models, J. Comput. Phys. 230 (2011) 5310–5327.

[21] F. Guillen-Gonzalez, G. Tierra, On linear schemes for a Cahn–Hilliard diffuse interface model, J. Comput. Phys. 234 (2013) 140–171.

[22] D. Gurtin, D. Polignone, J. Vinals, Two-phase binary fluids and immiscible fluids described by an order parameter, Int. J. Math. Models Methods Appl. Sci. 6 (1996) 815–831.

[23] P.C. Hohenberg, B.I. Halperin, Theory of dynamic critical phenomena, Rev. Mod. Phys. 49 (1977) 435–479.

[24] J. Hua, P. Lin, C. Liu, Q. Wang, Energy law preserving C^0 finite element schemes for phase field models in two-phase flow computations, J. Comput. Phys. 230 (2011) 7115–7131.

[25] D. Jacqmin, Calculation of two-phase Navier–Stokes flows using phase-field modeling, J. Comput. Phys. 155 (1999) 96–127.

[26] G.E. Karniadakis, S.J. Sherwin, Spectral/Hp Element Methods for Computational Fluid Dynamics, 2nd edn., Oxford University Press, 2005.

[27] J. Kim, K. Kang, J. Lowengrub, Conservative multigrid methods for Cahn–Hilliard fluids, J. Comput. Phys. 193 (2004) 511–543.

[28] P. Lin, C. Liu, H. Zhang, An energy law preserving c^0 finite element scheme for simulating the kinematic effects in liquid crystal dynamics, J. Comput. Phys. 227 (2007) 1411–1427.

[29] C. Liu, J. Shen, A phase field model for the mixture of two incompressible fluids and its approximation by a Fourier-spectral method, Physica D 179 (2003) 211–228.

[30] J. Lowengrub, L. Truskinovsky, Quasi-incompressible Cahn–Hilliard fluids and topological transitions, Proc. R. Soc. Lond. A 454 (1998) 2617–2654.

[31] L. Ma, R. Chen, X. Yang, H. Zhang, Numerical approximations for Allen–Cahn type phase field model of two-phase incompressible fluids with moving contact lines, Commun. Comput. Phys. 21 (2017) 867–889.

[32] E.V.L. Mello, O.T.S. Filho, Numerical study of the Cahn–Hilliard equation in one, two and three dimensions, Physica A 347 (2005) 429–443.

[33] J. Shen, J. Xu, J. Yang, The scalar auxiliary variable (SAV) approach for gradient flows, J. Comput. Phys. 353 (2018) 407–416.

[34] J. Shen, X. Yang, A phase-field model and its numerical approximation for two-phase incompressible flows with different densities and viscosities, SIAM J. Sci. Comput. 32 (2010) 1159–1179.

[35] J. Shen, X. Yang, Decoupled energy stable schemes for phase-field models of two-phase incompressible flows, SIAM J. Numer. Anal. 53 (2015) 279–296.

[36] P. Song, T. Yang, Y. Ji, Z. Wang, Z. Yang, L. Chen, L. Chen, A comparison of Fourier spectral iterative perturbation method and finite element method in solving phase-field equilibrium equations, Commun. Comput. Phys. 22 (2017) 1325–1349.

[37] G. Tierra, F. Guillen-Gonzalez, Numerical methods for solving the Cahn–Hilliard equation and its applicability to related energy-based models, Arch. Comput. Methods Eng. 22 (2015) 269–289.

[38] J.M. Varah, Stability restrictions on second order, three level finite difference schemes for parabolic equations, SIAM J. Numer. Anal. 17 (1980) 300–309.

[39] S.M. Wise, Unconditionally stable finite difference, nonlinear multigrid simulation of the Cahn–Hilliard–Hele–Shaw system of equations, J. Sci. Comput. 44 (1) (2010).

[40] Y. Yan, W. Chen, C. Wang, S.M. Wise, A second-order energy stable BDF numerical scheme for the Cahn–Hilliard equation, Commun. Comput. Phys. 23 (2018) 572–602.

[41] X. Yang, Linear, first and second-order, unconditionally energy stable numerical schemes for the phase field model of homopolymer blends, J. Comput. Phys. 327 (2016) 294–316.

[42] P. Yue, J.J. Feng, C. Liu, J. Shen, A diffuse-interface method for simulating two-phase flows of complex fluids, J. Fluid Mech. 515 (2004) 293–317.

[43] P. Yue, C. Zhou, J.J. Feng, Sharp-interface limit of the Cahn–Hilliard model for moving contact lines, J. Fluid Mech. 645 (2010) 279–294.

[44] X. Zheng, S. Dong, An eigen-based high-order expansion basis for structured spectral elements, J. Comput. Phys. 230 (2011) 8573–8602.