SIAM J. SCI. COMPUT.
Vol. 41, No. 6, pp. A3703–A3727

ENERGY-DECAYING EXTRAPOLATED RK–SAV METHODS FOR THE ALLEN–CAHN AND CAHN–HILLIARD EQUATIONS*

GEORGIOS AKRIVIS†, BUYANG LI‡, AND DONGFANG LI§

Abstract. We construct and analyze a class of extrapolated and linearized Runge–Kutta (RK) methods, which can be of arbitrarily high order, for the time discretization of the Allen–Cahn and Cahn–Hilliard phase field equations, based on the scalar auxiliary variable (SAV) formulation. We prove that the proposed q-stage RK–SAV methods have qth-order convergence in time and satisfy a discrete version of the energy decay property. Numerical examples are provided to illustrate the discrete energy decay property and accuracy of the proposed methods.

Key words. Allen–Cahn equation, Cahn–Hilliard equation, energy decay, scalar auxiliary variable, Runge–Kutta methods, extrapolation, Gauss methods, Radau IIA methods, algebraic stability

AMS subject classifications. 65M12, 65L06

DOI. 10.1137/19M1264412

1. Introduction.

We consider the initial boundary value problems for the Allen–Cahn (AC) and Cahn–Hilliard (CH) phase field equations,

$$\begin{cases} \partial_t u - \Delta u + f(u) = 0 & \text{in } \Omega \times (0,T), \\ \partial_{\boldsymbol{n}} u = 0 & \text{on } \partial\Omega \times (0,T), \\ u(\cdot,0) = u_0 & \text{in } \Omega, \end{cases} \tag{1.1}$$

and

$$\begin{cases} \partial_t u - \Delta(-\Delta u + f(u)) = 0 & \text{in } \Omega \times (0,T), \\ \partial_{\boldsymbol{n}} u = \partial_{\boldsymbol{n}}(-\Delta u + f(u)) = 0 & \text{on } \partial\Omega \times (0,T), \\ u(\cdot,0) = u_0 & \text{in } \Omega, \end{cases} \tag{1.2}$$

respectively, in a smooth or a convex polygonal/polyhedral domain $\Omega \subset \mathbb{R}^d, 1 \leqslant d \leqslant 3$, with f being the derivative of a nonnegative potential function F, i.e., $f = F'$. For instance, the Ginzburg–Landau energy function

$$F(v) := \frac{1}{4\varepsilon^2}(v^2-1)^2 \quad \text{with a small parameter } \varepsilon$$

leads to $f(v) = (v^3 - v)/\varepsilon^2$, which is widely used in physics and engineering.

*Submitted to the journal's Methods and Algorithms for Scientific Computing section May 28, 2019; accepted for publication (in revised form) September 23, 2019; published electronically November 21, 2019.

https://doi.org/10.1137/19M1264412

Funding: The work of the second author was partially supported by Hong Kong RGC grant 15300519. The work of the third author was partially supported by NSFC grant 11771162.

†Department of Computer Science and Engineering, University of Ioannina, Ioannina, GR-45110, Greece, and Institute of Applied and Computational Mathematics, FORTH, Heraklion, Crete, GR-70013, Greece (akrivis@cse.uoi.gr).

‡Department of Applied Mathematics, Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong (buyang.li@polyu.edu.hk).

§School of Mathematics and Statistics, Huazhong University of Science and Technology, Wuhan 430074, China, and Hubei Key Laboratory of Engineering Modeling and Scientific Computing, Huazhong University of Science and Technology, Wuhan 430074, China (dfli@mail.hust.edu.cn).

The AC and CH equations are phase field models for phase separation in binary alloy systems, complex fluids, and soft matter [1, 5, 12, 45]. They are, respectively, the L^2 and H^{-1} gradient flows of the energy functional

$$E[u] = \int_\Omega \left(\frac{1}{2}|\nabla u|^2 + F(u) \right) \mathrm{d}x.$$

As a result, the solutions of the AC and CH equations have decaying energy. Indeed, testing (1.1) and (1.2) by $\partial_t u$ and $-\Delta u + f(u)$, respectively, yields

$$E[u(\cdot,t_2)] - E[u(\cdot,t_1)] = -\int_{t_1}^{t_2} \|\partial_t u\|^2_{L^2(\Omega)} \mathrm{d}t \qquad \text{for the AC equation,} \tag{1.3}$$

$$E[u(\cdot,t_2)] - E[u(\cdot,t_1)] = -\int_{t_1}^{t_2} \|\partial_t u\|^2_{H^{-1}(\Omega)} \mathrm{d}t \qquad \text{for the CH equation,} \tag{1.4}$$

where $0 \leqslant t_1 \leqslant t_2 \leqslant T$ and $\|\partial_t u\|_{H^{-1}(\Omega)} = \|\nabla w\|_{L^2(\Omega)}$ with $w = -\Delta u + f(u)$.

Correspondingly, great efforts have been devoted to the construction of efficient and accurate numerical methods preserving the energy decay property at the discrete level. Some popular energy-decaying nonlinearly implicit time stepping methods include convex splitting methods [14, 15, 25, 32, 36], discrete-gradient methods [20, 30, 11, 31], averaged collocation methods [22, 10], and Runge–Kutta methods [23]. More recently, fully implicit energy-decaying algebraically stable Runge–Kutta methods were proposed for the invariant energy quadratization (IEQ) and the scalar auxiliary variable (SAV) formulations (see [19]); they require solving a nonlinear system of equations at every time step. These methods are expected to be of higher order, like the implicit Runge–Kutta methods for the original formulation studied in [23]. Existence, uniqueness, and convergence of these methods require further analysis.

In contrast to fully implicit methods, linearly implicit methods only require solving a linear system of equations at each time step. Energy-decaying linearly implicit time stepping methods include stabilization methods [47, 35, 26, 27, 37, 38, 39, 46], Lagrange multiplier methods [2, 21], the IEQ approach [40, 41, 42, 44, 43], and the SAV approach [33, 34, 9]. By using the stabilization and Lagrange multiplier approaches, first-order time stepping methods can be easily constructed for phase field equations, while second-order methods have to be constructed by carefully choosing stabilization terms and Lagrange multipliers case by case. The IEQ and SAV approaches provide equivalent reformulations of general phase field equations. The main advantage of the reformulated equations is that the energy is expressed in terms of Hilbert space norms of the new variables, which considerably simplifies the construction of energy-decaying methods; in particular, the construction of energy-decaying second-order time stepping methods for the reformulated phase field equations is a relatively easy task. For example, applying the linearly implicit second-order backward difference formula (BDF) method to the reformulated equations, based either on the IEQ or on the SAV formulations, automatically yields an energy-decaying time discretization method, with the decay property properly interpreted since now two pairs of approximations, $(u_{m+1}, u_m), m = n, n-1$, are compared rather than two approximations u_{n+1} and u_n. However, high-order BDF methods for the reformulated equations, again based either on the IEQ or on the SAV formulations, do not immediately lead to energy-decaying numerical schemes. For all these approaches, the construction of general energy-decaying linearized time stepping methods of order higher than 2 has remained open.

We refer the reader to [3, 4, 6, 7, 8, 16, 17, 18, 28, 33, 39] for error analyses of numerical methods for the AC and CH equations.

In this article, we propose a class of extrapolated and linearized Runge–Kutta methods based on the SAV formulation (extrapolated RK–SAV methods) for the AC and CH phase field equations, with linearly implicit schemes for the linear part of the equations and linearized extrapolation for the nonlinear part. We prove that the proposed extrapolated q-stage RK–SAV methods can have qth-order convergence in time in approximating smooth solutions of the AC and CH equations and satisfy a discrete version of the energy decay, i.e.,

$$E_\tau[u_{n+1}, r_{n+1}] \leqslant E_\tau[u_n, r_n], \tag{1.5}$$

where $E_\tau[u_n, r_n]$ is the discrete (modified) energy of the numerical solution at time $t = t_n$; see section 3. Therefore, arbitrarily high order energy-decaying methods can be constructed by choosing large q. This again demonstrates the strength and suitability of the SAV and IEQ approaches for the construction not only of second-order but even of high-order energy-decaying linearly implicit methods.

The extrapolated RK–SAV methods proposed in this paper can be viewed as an extrapolated linearization of the Runge–Kutta methods studied in [19] for the IEQ and SAV formulations. The advantage of such extrapolation and linearization is a complete existence, uniqueness, and convergence analysis for a linearly implicit method.

In section 2, we recall the standard SAV reformulations of the AC and CH equations, then present our numerical methods based on the SAV formulation. The analogous numerical schemes based on IEQ are similar and therefore omitted in this article. In section 3, we prove the energy decay property of the proposed methods. Section 4 is devoted to the error analysis for the AC equation, while in section 5 we present results of our numerical experiments. Concluding remarks are presented in section 6.

2. Extrapolated RK–SAV methods. In this section, we present the extrapolated RK–SAV methods for the AC and CH equations. The AC and CH equations are first reformulated via the SAV, as in the literature, and then discretized in time by algebraically stable q-stage Runge–Kutta methods, with linearized extrapolation for the nonlinear terms in the reformulated equations.

2.1. SAV reformulations. The SAV approach introduces a scalar function of t,

$$r(t) := \sqrt{\textstyle\int_\Omega F(u(x,t))\mathrm{d}x + E_0}, \quad 0 \leqslant t \leqslant T, \tag{2.1}$$

with E_0 an arbitrary positive constant, for example, $E_0 := \int_\Omega (\frac{1}{2}|\nabla u_0|^2 + F(u_0))\mathrm{d}x$, and reformulates the AC and CH equations into

$$\begin{cases} \partial_t u - \Delta u + rW(u) = 0 & \text{in } \Omega \times (0,T), \\ \partial_{\boldsymbol{n}} u = 0 & \text{on } \partial\Omega \times (0,T), \\ r' = \dfrac{1}{2}(W(u), \partial_t u) & \text{in } (0,T), \\ u(\cdot,0) = u_0 & \text{in } \Omega, \\ r(0) = \sqrt{\textstyle\int_\Omega F(u_0)\mathrm{d}x + E_0}, \end{cases} \tag{2.2}$$

and

$$
\begin{cases}
\partial_t u - \Delta(-\Delta u + rW(u)) = 0 & \text{in } \Omega \times (0,T),\\
\partial_{\boldsymbol{n}} u = \partial_{\boldsymbol{n}}(-\Delta u + rW(u)) = 0 & \text{on } \partial\Omega \times (0,T),\\
r' = \dfrac{1}{2}(W(u), \partial_t u) & \text{in } (0,T),\\
u(\cdot,0) = u_0 & \text{in } \Omega,\\
r(0) = \sqrt{\int_\Omega F(u_0)\mathrm{d}x + E_0}, &
\end{cases} \tag{2.3}
$$

respectively, where $W(u) := f(u)/\sqrt{\int_\Omega F(u(x,t))\mathrm{d}x + E_0}$. First- and second-order linearly implicit schemes can be constructed by directly applying the implicit Euler and the second-order BDF methods to problems (2.2) and (2.3). The resulting methods automatically preserve the energy decay property; for instance, for the implicit Euler method, we have

$$
\int_\Omega \frac{1}{2}|\nabla u_{n+1}|^2\mathrm{d}x + r_{n+1}^2 \leqslant \int_\Omega \frac{1}{2}|\nabla u_n|^2\mathrm{d}x + r_n^2. \tag{2.4}
$$

However, as mentioned in the introduction, higher-order BDF methods for the SAV formulation do not automatically yield energy-decaying time discretization methods. Higher-order energy-decaying linear schemes have not been constructed so far. In sections 2.2–2.3, we present a class of extrapolated Runge–Kutta methods, which can be arbitrarily high order accurate, for the SAV formulations (2.2) and (2.3) that preserve the energy decay property at the discrete setting.

2.2. Algebraically stable Runge–Kutta methods. We consider a q-stage Runge–Kutta method, described by the Butcher tableau

$$
\begin{array}{ccc|c}
a_{11} & \dots & a_{1q} & c_1\\
\vdots & & \vdots & \vdots\\
a_{q1} & \dots & a_{qq} & c_q\\
\hline
b_1 & \dots & b_q &
\end{array} \tag{2.5}
$$

with $c_1, \dots, c_q \in (0,1]$. The Runge–Kutta method described by (2.5) discretizes an initial value problem for an autonomous ordinary differential equation

$$
v'(t) = f(v(t)), \quad t \in (0,T], \quad v(0) = v_0,
$$

in the following way. For a given approximation v_n of the nodal value $v(t_n)$, one computes v_{n+1} by

$$
\begin{cases}
v_{ni} = v_n + \tau \displaystyle\sum_{j=1}^{q} a_{ij}\dot{v}_{nj}, \quad i = 1, \dots, q,\\
v_{n+1} = v_n + \tau \displaystyle\sum_{i=1}^{q} b_i \dot{v}_{ni},
\end{cases}
$$

where $\dot{v}_{ni} = f(v_{ni}), i = 1, \dots, q$, v_{ni} are approximations to $v(t_{ni})$ for $i = 1, \dots, q$, with $t_{ni} = t_n + c_i\tau$ being the internal Runge–Kutta nodes and $\tau = t_{n+1} - t_n$ being the time stepsize.

We consider algebraically stable Runge–Kutta methods satisfying the following conditions:

$$\text{(2.6)}\qquad \begin{aligned} &(1)\ \text{The matrix } A=(a_{ij})_{i,j=1,\dots,q} \text{ is invertible,}\\ &(2)\ b_i>0, \quad i=1,\dots,q,\\ &(3)\ c_i\neq c_j \quad \text{for } i\neq j. \end{aligned}$$

Here, algebraic stability means, besides the positivity of the weights $b_1,\dots,b_q$, that the symmetric matrix $M\in\mathbb{R}^{q\times q}$ with entries $m_{ij}:=b_ia_{ij}+b_ja_{ji}-b_ib_j, i,j=1,\dots,q$, is positive semidefinite. This is our essential condition for the energy decay property.

Besides (2.6), we assume for the error analysis that the Runge–Kutta method is associated to a collocation method, that is, it has order $p\geqslant q$ and *stage order* at least q, i.e.,

$$\text{(B}(p)\text{)}\qquad \sum_{i=1}^{q} b_i c_i^{\ell-1}=\frac{1}{\ell},\qquad \ell=1,\dots,p,$$

$$\text{(C}(q)\text{)}\qquad \sum_{j=1}^{q} a_{ij} c_j^{\ell-1}=\frac{c_i^{\ell}}{\ell},\qquad \ell=1,\dots,q,\ i=1,\dots,q.$$

Two popular families of algebraically stable Runge–Kutta methods of collocation type satisfying (2.6), of orders $p=2q$ and $p=2q-1$, respectively, are the Gauss methods and the Radau IIA methods. For both families, arbitrarily high order methods can be constructed. The one-stage members of these families are the midpoint (or Crank–Nicolson) and implicit Euler methods, respectively. The tableaus of the two- and three-stage members of the Gauss and Radau IIA methods are given in [24, section IV, pp. 72, 74].

2.3. Extrapolated RK–SAV methods. Let N be a positive integer and $t_n:=n\tau, n=0,\dots,N$, be the uniform partition of the time interval $[0,T]$ with time stepsize $\tau:=T/N$. Furthermore, let $t_{ni}:=t_n+c_i\tau, i=1,\dots,q, n=0,\dots,N-1$, denote the internal Runge–Kutta nodes.

For given internal stages $u_{n-1,i}, i=1,\dots,q$, we denote by $u_{n-1}^{\tau}(t)$ the Lagrange interpolation polynomial of degree at most $q-1$ satisfying

$$u_{n-1}^{\tau}(t_{n-1,i})=u_{n-1,i},\quad i=1,\dots,q,$$

and use the abbreviation $I_{n-1}^{\tau}u_{ni}:=u_{n-1}^{\tau}(t_{ni})$, which approximates $u(t_{ni})$ by the extrapolation method using the values $u_{n-1,i}$, $i=1,\dots,q$.

Similarly, we denote by $I_{n-1}^{\tau}u(t)$ the Lagrange interpolation polynomial in t of degree at most $q-1$ interpolating the values of the exact solution u, i.e., satisfying

$$I_{n-1}^{\tau}u(t_{n-1,i})=u(t_{n-1,i}),\quad i=1,\dots,q.$$

Assuming that the nodal approximations u_n, r_n and the internal stages $u_{n-1,i}$, $i=1,\dots,q$, are known, we consider the following method for the reformulated AC equation (2.2):

$$\text{(2.8)}\qquad \begin{cases} \dot{u}_{ni}=\Delta u_{ni}-r_{ni}W(I_{n-1}^{\tau}u_{ni}) & \text{in } \Omega, \quad i=1,\dots,q,\\ u_{ni}=u_n+\tau\displaystyle\sum_{j=1}^{q}a_{ij}\dot{u}_{nj}, & \text{in } \Omega, \quad i=1,\dots,q,\\ \partial_{\boldsymbol{n}}u_{ni}=0 & \text{on } \partial\Omega, \quad i=1,\dots,q, \end{cases}$$

$$\text{(2.9)}\qquad \begin{cases} \dot{r}_{ni}=\dfrac{1}{2}(W(I_{n-1}^{\tau}u_{ni}),\dot{u}_{ni}), & i=1,\dots,q,\\ r_{ni}=r_n+\tau\displaystyle\sum_{j=1}^{q}a_{ij}\dot{r}_{nj}, & i=1,\dots,q. \end{cases}$$

Note that the quantities $\dot{u}_{ni}$ and $\dot{r}_{ni}$ have been introduced here for notational convenience only. In fact, substituting $\dot{u}_{ni}$ from the first relation of (2.8) into the second relation of (2.8) as well as into the first relation of (2.9), and subsequently substituting the new first relation of (2.9) into its second relation, we obtain a coupled system for the internal stages u_{ni} and r_{ni}, $i = 1, \dots, q$. Since the extrapolated quantities $W(I^{\tau}_{n-1}u_{ni})$ are known, the implementation of (2.8)–(2.9) requires only the solution of a coupled linear elliptic system for $(u_{ni}, r_{ni}) \in H^1(\Omega) \times \mathbb{R}$, $i = 1, \dots, q$. Once u_{ni} and r_{ni} have been determined, we obtain the quantities $\dot{u}_{ni} \in H^1(\Omega)$ and $\dot{r}_{ni} \in \mathbb{R}$ from the second relations of (2.8) and (2.9), respectively, using the invertibility of the matrix $A = (a_{ij})$. Using these values, one computes $(u_{n+1}, r_{n+1}) \in H^1(\Omega) \times \mathbb{R}$ through

$$\begin{cases} u_{n+1} := u_n + \tau \sum_{i=1}^{q} b_i \dot{u}_{ni}, \\ r_{n+1} := r_n + \tau \sum_{i=1}^{q} b_i \dot{r}_{ni}. \end{cases} \tag{2.10}$$

Since $I^{\tau}_{n-1}u(t)$ is the Lagrange interpolation polynomial of degree at most $q-1$, it follows that $I^{\tau}_{n-1}u(t_{ni})$ is a qth-order approximation of $u(t_{ni})$ for $i = 1, \dots, q$. Therefore, the method (2.8)–(2.10) has $(q+1)$th-order consistency error; see section 4.

For the CH equation, the extrapolated linearized Runge–Kutta method for (2.3) reads

$$\begin{cases} \dot{u}_{ni} = \Delta w_{ni} & \text{in } \Omega, \quad i = 1, \dots, q, \\ u_{ni} = u_n + \tau \sum_{j=1}^{q} a_{ij} \dot{u}_{nj} & \text{in } \Omega, \quad i = 1, \dots, q, \\ w_{ni} = -\Delta u_{ni} + r_{ni} W(I^{\tau}_{n-1}u_{ni}) & \text{in } \Omega, \quad i = 1, \dots, q, \\ \partial_{\boldsymbol{n}} u_{ni} = \partial_{\boldsymbol{n}} w_{ni} = 0 & \text{on } \partial\Omega, \quad i = 1, \dots, q, \end{cases} \tag{2.11}$$

and

$$\begin{cases} \dot{r}_{ni} = \dfrac{1}{2}(W(I^{\tau}_{n-1}u_{ni}), \dot{u}_{ni}), & i = 1, \dots, q, \\ r_{ni} = r_n + \tau \sum_{j=1}^{q} a_{ij} \dot{r}_{nj}, & i = 1, \dots, q, \end{cases} \tag{2.12}$$

which is an elliptic system of equations of the internal stages (u_{ni}, w_{ni}, r_{ni}), $i = 1, \dots, q$. Having solved the equations for these internal stages, we obtain $(\dot{u}_{ni}, \dot{r}_{ni}) \in H^1(\Omega) \times \mathbb{R}$, $i = 1, \dots, q$, from the second relations in (2.11)–(2.12) using the invertibility of the matrix $A = (a_{ij})$. Using these values, one computes $(u_{n+1}, r_{n+1}) \in H^1(\Omega) \times \mathbb{R}$ through

$$\begin{cases} u_{n+1} := u_n + \tau \sum_{i=1}^{q} b_i \dot{u}_{ni}, \\ r_{n+1} := r_n + \tau \sum_{i=1}^{q} b_i \dot{r}_{ni}. \end{cases} \tag{2.13}$$

As for method (2.8)–(2.10), the consistency error of method (2.11)–(2.13) is of $(q+1)$th-order.

3. Energy decay property of the extrapolated RK–SAV method. In this section, we use the algebraic stability of the Runge–Kutta methods to show that the extrapolated RK–SAV methods (2.8)–(2.10) and (2.11)–(2.13) satisfy discrete analogues of the energy decay. We denote by

$$E_\tau[u_n, r_n] := \frac{1}{2}\|\nabla u_n\|^2 + |r_n|^2 - E_0$$

the discrete energy (also referred to as modified energy since, in general, $|r_n|^2 - E_0$ does not coincide with $\int_\Omega F(u_n)\mathrm{d}x$) of the numerical solution at t_n.

THEOREM 3.1 (discrete energy decay for the AC equation). *Assume that the Runge–Kutta method* (2.5) *is algebraically stable and satisfies* (2.6)*, and assume that the values* $u_{n-1,i} \in H^1(\Omega)$*,* $i = 1, \dots, q$*, and* $(u_n, r_n) \in H^1(\Omega) \times \mathbb{R}$ *are given. Then, the extrapolated RK–SAV method* (2.8)–(2.9) *has a unique solution of internal stages* $(u_{ni}, r_{ni}) \in H^1(\Omega) \times \mathbb{R}$*,* $i = 1, \dots, q$*, and the nodal values* $(u_{n+1}, r_{n+1}) \in H^1(\Omega) \times \mathbb{R}$ *defined by* (2.10) *satisfy the energy decay property*

$$E_\tau[u_{n+1}, r_{n+1}] \leqslant E_\tau[u_n, r_n]. \tag{3.1}$$

Proof. We first prove (3.1) for any solution of internal stages $(u_{ni}, r_{ni}) \in H^1(\Omega) \times \mathbb{R}$, $i = 1, \dots, q$. In this case, the nodal value (u_{n+1}, r_{n+1}) defined by (2.10) is in $H^1(\Omega) \times \mathbb{R}$, as explained in section 2.3. Existence and uniqueness of solutions are proved afterward.

According to the first relation of (2.10), we have

$$\nabla u_{n+1} = \nabla u_n + \tau \sum_{i=1}^{q} b_i \nabla \dot{u}_{ni}.$$

Squaring the L^2-norms of both sides yields

$$\begin{aligned}\|\nabla u_{n+1}\|^2 &= \left(\nabla u_n + \tau \sum_{i=1}^{q} b_i \nabla \dot{u}_{ni}, \nabla u_n + \tau \sum_{j=1}^{q} b_j \nabla \dot{u}_{nj}\right) \\ &= \|\nabla u_n\|^2 + 2\tau \sum_{i=1}^{q} b_i (\nabla \dot{u}_{ni}, \nabla u_n) + \tau^2 \sum_{i,j=1}^{q} b_i b_j (\nabla \dot{u}_{ni}, \nabla \dot{u}_{nj}).\end{aligned}$$

Substituting $u_n = u_{ni} - \tau \sum_{j=1}^{q} a_{ij} \dot{u}_{nj}$ (the second relation in (2.8)) into the second term on the right-hand side of the last relation, we obtain

$$\|\nabla u_{n+1}\|^2 = \|\nabla u_n\|^2 + 2\tau \sum_{i=1}^{q} b_i \left(\nabla \dot{u}_{ni}, \nabla u_{ni} - \tau \sum_{j=1}^{q} a_{ij} \nabla \dot{u}_{nj}\right) + \tau^2 \sum_{i,j=1}^{q} b_i b_j (\nabla \dot{u}_{ni}, \nabla \dot{u}_{nj}),$$

whence

$$\|\nabla u_{n+1}\|^2 = \|\nabla u_n\|^2 + 2\tau \sum_{i=1}^{q} b_i \big(\nabla \dot{u}_{ni}, \nabla u_{ni}\big) - \tau^2 \sum_{i,j=1}^{q} m_{ij} (\nabla \dot{u}_{ni}, \nabla \dot{u}_{nj}),$$

with $m_{ij} = b_i a_{ij} + b_j a_{ji} - b_i b_j, i, j = 1, \dots, q$. Using here the positive semidefiniteness of the matrix $M = (m_{ij})$ we infer that

$$\|\nabla u_{n+1}\|^2 \leqslant \|\nabla u_n\|^2 + 2\tau \sum_{i=1}^{q} b_i \big(\nabla \dot{u}_{ni}, \nabla u_{ni}\big). \tag{3.2}$$

Testing the first relation of (2.8) by $\dot{u}_{ni}$ yields

$$\|\dot{u}_{ni}\|^2 = -(\nabla \dot{u}_{ni}, \nabla u_{ni}) - r_{ni}(\dot{u}_{ni}, W(I^\tau_{n-1}u_{ni})),$$

which implies $(\nabla \dot{u}_{ni}, \nabla u_{ni}) = -\|\dot{u}_{ni}\|^2 - r_{ni}(\dot{u}_{ni}, W(I^\tau_{n-1}u_{ni}))$. Then, substituting this into (3.2) yields

$$\|\nabla u_{n+1}\|^2 \leqslant \|\nabla u_n\|^2 - 2\tau \sum_{i=1}^{q} b_i \|\dot{u}_{ni}\|^2 - 2\tau \sum_{i=1}^{q} b_i r_{ni}\big(\dot{u}_{ni}, W(I^\tau_{n-1}u_{ni})\big). \tag{3.3}$$

Similarly, we can obtain

$$|r_{n+1}|^2 \leqslant |r_n|^2 + \tau \sum_{i=1}^{q} b_i r_{ni}(W(I^\tau_{n-1}u_{ni}), \dot{u}_{ni}). \tag{3.4}$$

Notice that up to this point we followed the proof for the algebraic stability of Runge–Kutta methods; of course, since we are interested in the energy decay property, we do not consider differences of approximations.

Now, multiplying (3.3) by $\frac{1}{2}$ and adding the result to (3.4), the last terms on their right-hand sides cancel and we obtain

$$\frac{1}{2}\|\nabla u_{n+1}\|^2 + |r_{n+1}|^2 \leqslant \frac{1}{2}\|\nabla u_n\|^2 + |r_n|^2 - \tau \sum_{i=1}^{q} b_i \|\dot{u}_{ni}\|^2. \tag{3.5}$$

Using here the nonnegativity of the weights $b_1, \dots, b_q$, we infer that the extrapolated RK–SAV method (2.8)–(2.10) has the desired energy decay property,

$$\frac{1}{2}\|\nabla u_{n+1}\|^2 + |r_{n+1}|^2 \leqslant \frac{1}{2}\|\nabla u_n\|^2 + |r_n|^2. \tag{3.6}$$

This proves (3.1).

Finally, we prove existence and uniqueness of solutions to system (2.8)–(2.9). To this end, we let $U_n := (u_{n1}, \dots, u_{nq})^T$, $R_n := (r_{n1}, \dots, r_{nq})^T$, and $\mathbb{1} = (1, \dots, 1)^T$ and denote by I the $q \times q$ identity matrix. Then, from (2.8) we obtain

$$(I - \Delta I)U_n = (I - \tau^{-1}A^{-1})U_n + u_n \tau^{-1}A^{-1}\mathbb{1} - B_1 R_n, \tag{3.7}$$

with the boundary condition $\partial_{\boldsymbol{n}} U_n = 0$ on $\partial\Omega$, where B_1 is the diagonal matrix-valued function

$$B_1 = \operatorname{diag}(w_{n1}, \dots, w_{nq}) \quad \text{with given functions } w_{ni} = W(I^\tau_{n-1}u_{ni}).$$

Multiplying both sides of (3.7) by $(I - \Delta I)^{-1}$, we further obtain

$$U_n = GU_n - B_2 R_n + g \tag{3.8}$$

with

$$G = (I - \Delta I)^{-1}(I - \tau^{-1}A^{-1}), \quad B_2 = (I - \Delta I)^{-1}B_1, \quad \text{and} \quad g = (I - \Delta I)^{-1}u_n\, \tau^{-1}A^{-1}\mathbb{1}.$$

Similarly, from (2.9) we obtain

$$R_n = r_n \mathbb{1} - \frac{1}{2}\tau A J \nabla W_n \cdot \nabla U_n - \frac{1}{2}\tau A B_3 R_n, \tag{3.9}$$

where $\nabla W_n \cdot \nabla U_n = (\nabla w_{n1} \cdot \nabla u_{n1}, \dots, \nabla w_{nq} \cdot \nabla u_{nq})^T$, B_3 is the diagonal matrix

$$B_3 = \operatorname{diag}\left(\|w_{n1}\|^2, \dots, \|w_{nq}\|^2\right),$$

and $J : L^2(\Omega)^q \to \mathbb{R}^q$ is defined by $Jv = \int_\Omega v \, \mathrm{d}x$ for $v \in L^2(\Omega)^q$.

Substituting (3.8) into (3.9) yields

$$R_n = -\frac{1}{2}\tau AJ\nabla W_n \cdot \nabla(GU_n) + \frac{1}{2}\tau AJ\nabla W_n \cdot \nabla(B_2 R_n) - \frac{1}{2}\tau AB_3 R_n + \tilde{g} \tag{3.10}$$

with $\tilde{g} = r_n \mathbb{1} - \frac{1}{2}\tau AJ\nabla W_n \cdot \nabla g$.

Equations (3.8) and (3.10) can be written abstractly as

$$\begin{pmatrix} U_n \\ R_n \end{pmatrix} = \tilde{G}\begin{pmatrix} U_n \\ R_n \end{pmatrix} + \begin{pmatrix} g \\ \tilde{g} \end{pmatrix} \tag{3.11}$$

with $\tilde{G} : (L^2(\Omega))^q \times \mathbb{R}^q \to (H^1(\Omega))^q \times \mathbb{R}^q \hookrightarrow\hookrightarrow (L^2(\Omega))^q \times \mathbb{R}^q$ being a compact operator on $(L^2(\Omega))^q \times \mathbb{R}^q$. Therefore, by the Fredholm theory of compact operators, (3.11) has a unique solution if and only if the corresponding homogeneous equation

$$\begin{pmatrix} U_n \\ R_n \end{pmatrix} = \tilde{G}\begin{pmatrix} U_n \\ R_n \end{pmatrix} \tag{3.12}$$

has only the trivial solution. Now, from (3.5) we obtain

$$\tau \sum_{i=1}^{q} b_i \|\dot{u}_{ni}\|^2 \leqslant \frac{1}{2}\|\nabla u_n\|^2 + |r_n|^2. \tag{3.13}$$

Since $b_i > 0$ for $i = 1, \dots, q$ (cf. condition (2.6)), estimate (3.13) implies that the homogeneous elliptic system corresponding to (2.8)–(2.9) has indeed only the trivial solution. In fact, for $u_n = 0$ and $r_n = 0$, (3.13) yields $\dot{u}_{ni} = 0, i = 1, \dots, q$, and we infer from (2.8)–(2.9) that also $u_{ni} = 0, r_{ni} = 0, i = 1, \dots, q$.

This proves that the elliptic system (2.8)–(2.9) has a unique solution $(u_{ni}, r_{ni}) \in H^1(\Omega) \times \mathbb{R}$, $i = 1, \dots, q$. ∎

Remark 3.1 (on the discrete energy decay property). It is evident from the proof of Theorem 3.1 that the discrete energy decay property (3.1) remains valid if we replace $I^\tau_{n-1} u_{ni}$ by arbitrary quantities $\tilde{u}_{ni}$. In particular, the base (nonlinear) version of the Runge–Kutta method, with $\tilde{u}_{ni} = u_{ni}$ (with unknown internal stages u_{ni}) as well as its fixed-point linearization, with given $\tilde{u}_{ni} = u_{ni}^{(\ell-1)}$ and unknown $u_{ni} = u_{ni}^{(\ell)}$, are also energy-decaying methods.

THEOREM 3.2 (discrete energy decay for the CH equation). *Let the Runge–Kutta method* (2.5) *be algebraically stable and satisfy* (2.6)*, and assume that the values* $u_{n-1,i} \in H^1(\Omega)$, $i = 1, \dots, q$, *and* $(u_n, r_n) \in H^1(\Omega) \times \mathbb{R}$ *are given. Then, the linearized Runge–Kutta method* (2.11)–(2.12) *has a unique solution of internal stages* $(u_{ni}, w_{ni}, r_{ni}) \in H^1(\Omega) \times H^1(\Omega) \times \mathbb{R}$ *for* $i = 1, \dots, q$*, and the nodal values* $(u_{n+1}, r_{n+1}) \in H^1(\Omega) \times \mathbb{R}$ *defined by* (2.13) *satisfy the following energy decay property:*

$$E_\tau[u_{n+1}, r_{n+1}] \leqslant E_\tau[u_n, r_n]. \tag{3.14}$$

Proof. We prove (3.14) for any solution of internal stages $(u_{ni}, w_{ni}, r_{ni}) \in H^1(\Omega) \times H^1(\Omega) \times \mathbb{R}$, $i = 1, \dots, q$. In this case, the nodal value (u_{n+1}, r_{n+1}) defined by (2.13)

is in $H^1(\Omega)\times\mathbb{R}$. Existence and uniqueness of a solution can be proved along the lines of the proof of Theorem 3.1 and are therefore omitted.

As in the proof of Theorem 3.1, both (3.2) and (3.4) hold for the numerical solution of the CH equation, i.e.,

$$\|\nabla u_{n+1}\|^2 \leqslant \|\nabla u_n\|^2 + 2\tau\sum_{i=1}^{q} b_i(\nabla \dot u_{ni}, \nabla u_{ni}), \tag{3.15}$$

$$|r_{n+1}|^2 \leqslant |r_n|^2 + \tau\sum_{i=1}^{q} b_i r_{ni}(W(I^\tau_{n-1}u_{ni}), \dot u_{ni}). \tag{3.16}$$

Testing the third relation of (2.11) by $\dot u_{ni}$ gives

$$(\nabla \dot u_{ni}, \nabla u_{ni}) = (\dot u_{ni}, w_{ni} - r_{ni}W(I^\tau_{n-1}u_{ni})).$$

Then, substituting this into (3.15) yields

$$\|\nabla u_{n+1}\|^2 \leqslant \|\nabla u_n\|^2 + 2\tau\sum_{i=1}^{q} b_i\big(\dot u_{ni}, w_{ni} - r_{ni}W(I^\tau_{n-1}u_{ni})\big). \tag{3.17}$$

Multiplying (3.17) by $\frac{1}{2}$ and adding the result to (3.16), we have

$$\frac{1}{2}\|\nabla u_{n+1}\|^2 + |r_{n+1}|^2 \leqslant \frac{1}{2}\|\nabla u_n\|^2 + |r_n|^2 + \tau\sum_{i=1}^{q} b_i\big(\dot u_{ni}, w_{ni}\big). \tag{3.18}$$

Substituting $\dot u_{ni} = \Delta w_{ni}$ (cf. the first relation of (2.11)) into the last estimate and using integration by parts, we obtain

$$\frac{1}{2}\|\nabla u_{n+1}\|^2 + |r_{n+1}|^2 \leqslant \frac{1}{2}\|\nabla u_n\|^2 + |r_n|^2 - \tau\sum_{i=1}^{q} b_i\|\nabla w_{ni}\|^2. \tag{3.19}$$

This proves the energy decay property (3.14) (since $b_i > 0$ for $i = 1,\dots,q$). ∎

4. Error analysis for the extrapolated RK–SAV method. In this section, we establish error estimates for the extrapolated RK–SAV method for the AC equation. The analysis for the CH equation is similar and is therefore omitted.

We assume that the values $u(t_{0i})$, $i = 1,\dots,q$, $u(t_1)$ and $r(t_1)$ are given or sufficiently accurate approximations thereof have been computed by other methods, and we examine the error of the numerical solutions given by the extrapolated RK–SAV method (2.8)–(2.10) for $n = 1,\dots,N-1$.

4.1. Consistency. We prove consistency of the extrapolated RK–SAV methods for the AC equation, assuming existence of a smooth solution. The consistency errors $\varepsilon_{ni}, \varepsilon_{n+1}, d_{ni}$, and d_{n+1} of the method are determined by

$$\begin{cases} \dot u^\star_{ni} = \Delta u^\star_{ni} - r^\star_{ni}W(I^\tau_{n-1}u^\star_{ni}) & \text{in } \Omega, \quad i = 1,\dots,q,\\ u^\star_{ni} = u^\star_n + \tau\displaystyle\sum_{j=1}^{q} a_{ij}\dot u^\star_{nj} + \varepsilon_{ni}, & \text{in } \Omega, \quad i = 1,\dots,q,\\ \partial_{\boldsymbol{n}} u^\star_{ni} = 0 & \text{on } \partial\Omega, \quad i = 1,\dots,q, \end{cases} \tag{4.1}$$

$$\begin{cases} \dot r^\star_{ni} = \dfrac{1}{2}(W(I^\tau_{n-1}u^\star_{ni}), \dot u^\star_{ni}), & i = 1,\dots,q,\\ r^\star_{ni} = r^\star_n + \tau\displaystyle\sum_{j=1}^{q} a_{ij}\dot r^\star_{nj} + d_{ni}, & i = 1,\dots,q, \end{cases} \tag{4.2}$$

and

$$\begin{cases} u^\star_{n+1} = u^\star_n + \tau \sum_{i=1}^{q} b_i \dot{u}^\star_{ni} + \varepsilon_{n+1}, \\ r^\star_{n+1} = r^\star_n + \tau \sum_{i=1}^{q} b_i \dot{r}^\star_{ni} + d_{n+1}, \end{cases} \tag{4.3}$$

with the notation

$$u^\star_n := u(t_n), \quad r^\star_n := r(t_n), \quad u^\star_{ni} := u(t_{ni}) = u(t_n + c_i\tau), \quad r^\star_{ni} := r(t_{ni}) = r(t_n + c_i\tau),$$

and $W(v) := f(v)/\sqrt{\int_\Omega F(v)\mathrm{d}x + E_0}$.

LEMMA 4.1 (consistency estimate). *If the solutions u and r of (2.2) are sufficiently smooth, then the following consistency estimate holds:*

$$\|\varepsilon_{n+1}\|_{H^1(\Omega)} + |d_{n+1}| + \tau \sum_{i=1}^{q} \big(\|\varepsilon_{ni}\|_{H^1(\Omega)} + |d_{ni}|\big) \leqslant c\tau^{q+1}. \tag{4.4}$$

Proof. We add and subtract $r^\star_{nj} W(u^\star_{nj})$ in the summation in the second relation of (4.1) and use the first differential equation in (2.2) to obtain

$$u^\star_{ni} - u^\star_n - \tau \sum_{j=1}^{q} a_{ij} u_t(t_{nj}) = \tau \sum_{j=1}^{q} a_{ij} r^\star_{nj} \big[W(u^\star_{nj}) - W(I^\tau_{n-1} u^\star_{nj})\big] + \varepsilon_{ni}. \tag{4.5}$$

Let us denote by $\tilde{\varepsilon}_{ni}, i = 1, \dots, q$, the quantity on the left-hand side of (4.5). Taylor expansion about t_n yields

$$\begin{aligned} \tilde{\varepsilon}_{ni} = {} & \sum_{\ell=1}^{q} \frac{\tau^\ell}{(\ell-1)!} \left(\frac{c_i^\ell}{\ell} - \sum_{j=1}^{q} a_{ij} c_j^{\ell-1} \right) u^{(\ell)}(t_n) + \frac{1}{q!} \int_{t_n}^{t_{ni}} (t_{ni} - s)^q u^{(q+1)}(s)\mathrm{d}s \\ & - \frac{\tau}{(q-1)!} \sum_{j=1}^{q} a_{ij} \int_{t_n}^{t_{nj}} (t_{nj} - s)^{q-1} u^{(q+1)}(s)\mathrm{d}s, \end{aligned}$$

with $u^{(\ell)} = \partial_t^\ell u$. In view of the stage order conditions (C(q)), leading terms of order up to q vanish, and $\tilde{\varepsilon}_{ni}$ can be represented in the form

$$\tilde{\varepsilon}_{ni} = \tau^q \int_{t_n}^{t_{n+1}} \kappa_i \left(\frac{s - t_n}{\tau} \right) u^{(q+1)}(s)\mathrm{d}s, \quad i = 1, \dots, q, \tag{4.6}$$

with the bounded Peano kernels

$$\kappa_i(t) := \frac{1}{q!} \big((c_i - t)_+\big)^q - \frac{1}{(q-1)!} \sum_{j=1}^{q} a_{ij} \big((c_j - t)_+\big)^{q-1}, \quad 0 \leqslant t \leqslant 1, \tag{4.7}$$

$i = 1, \dots, q$, where we used the standard notation $s_+ = s$ for $s \geqslant 0$ and $s_+ = 0$ for $s < 0$. We obtain the desired intermediate estimates for $\tilde{\varepsilon}_{ni}$,

$$\|\tilde{\varepsilon}_{ni}\|_{H^1(\Omega)} \leqslant C\tau^{q+1}, \quad i = 1, \dots, q. \tag{4.8}$$

Next, we note that the error $u^\star_{nj} - I^\tau_{n-1}u^\star_{nj}$ due to q-point extrapolation is $O(\tau^q)$ in $H^1(\Omega)$, i.e.,

$$\|u^\star_{nj} - I^\tau_{n-1}u^\star_{nj}\|_{H^1(\Omega)} \leqslant C\tau^q, \quad j = 1, \dots, q, \tag{4.9}$$

which also implies that

$$\|W(u^\star_{nj}) - W(I^\tau_{n-1}u^\star_{nj})\|_{H^1(\Omega)} \leqslant C\tau^q, \quad j = 1, \dots, q. \tag{4.10}$$

Combining (4.5) with (4.8) and (4.10), we obtain the desired final estimates for ε_{ni},

$$\|\varepsilon_{ni}\|_{H^1(\Omega)} \leqslant C\tau^q, \quad i = 1, \dots, q. \tag{4.11}$$

Analogously, we add and subtract $r^\star_{ni}W(u^\star_{ni})$ in the summation in the first relation of (4.3) and use the first differential equation in (2.2) to obtain

$$u^\star_{n+1} - u^\star_n - \tau\sum_{i=1}^q b_i\partial_t u(t_{ni}) = \tau\sum_{i=1}^q b_i r^\star_{ni}\big[W(u^\star_{ni}) - W(I^\tau_{n-1}u^\star_{ni})\big] + \varepsilon_{n+1}. \tag{4.12}$$

Let us denote by $\tilde\varepsilon_{n+1}$ the quantity on the left-hand side of (4.12). Taylor expansion about t_n yields

$$\begin{aligned}\tilde\varepsilon_{n+1} = {}& \sum_{\ell=1}^p \frac{\tau^\ell}{(\ell-1)!}\left(\frac{1}{\ell} - \sum_{i=1}^q b_i c_i^{\ell-1}\right)u^{(\ell)}(t_n) + \frac{1}{p!}\int_{t_n}^{t_{n+1}} (t_{n+1}-s)^p u^{(p+1)}(s)\mathrm{d}s \\ & - \frac{\tau}{(p-1)!}\sum_{i=1}^q b_i\int_{t_n}^{t_{ni}} (t_{ni}-s)^{p-1}u^{(p+1)}(s)\mathrm{d}s.\end{aligned}$$

Again, leading terms of order up to p vanish, this time in view of the order conditions (B(p)), and $\tilde\varepsilon_{n+1}$ can be represented in the form

$$\tilde\varepsilon_{n+1} = \tau^p\int_{t_n}^{t_{n+1}} \kappa\left(\frac{s-t_n}{\tau}\right)u^{(p+1)}(s)\mathrm{d}s \tag{4.13}$$

with the bounded Peano kernel

$$\kappa(t) := \frac{1}{p!}(1-t)^p - \frac{1}{(p-1)!}\sum_{i=1}^q b_i\big((c_i-t)_+\big)^{p-1}, \quad 0 \leqslant t \leqslant 1. \tag{4.14}$$

Relation (4.13) yields the desired intermediate estimate for $\tilde\varepsilon_{n+1}$,

$$\|\tilde\varepsilon_{n+1}\|_{H^1(\Omega)} \leqslant C\tau^{p+1}. \tag{4.15}$$

Combining (4.12) with (4.10) and (4.15), we obtain the desired final estimate for ε_{n+1},

$$\|\varepsilon_{n+1}\|_{H^1(\Omega)} \leqslant c\tau^{q+1}. \tag{4.16}$$

The consistency errors d_{ni} and d_{n+1} can be estimated similarly and the proofs are omitted. ∎

4.2. Error estimates. We have the following error estimates.

THEOREM 4.2 (error estimates). *We assume that $q \geqslant 2$ and that the following conditions hold:*

1. *The solution of* (2.2) *is sufficiently smooth.*
2. *The starting approximations (u_{0i}, r_{0i}) are sufficiently accurate such that*

$$\|u(t_1)-u_1\|^2 + |r(t_1) - r_1|^2+\tau \sum_{i=1}^{q} \left(\|u(t_{0i}) - u_{0i}\|^2 + |r(t_{0i}) - r_{0i}|^2\right) \leqslant C_0\tau^{2q},$$

$$\|u(t_{0i}) - u_{0i}\|_{L^\infty(\Omega)} \leqslant 1,$$

for some constant C_0 (independent of τ).

Then, the discrete solution given by (2.8)–(2.10) *satisfies the following error estimate:*

$$\max_{1\leqslant n\leqslant N-1}\left(\|u(t_{n+1}) - u_{n+1}\|^2 + |r(t_{n+1}) - r_{n+1}|^2 + \tau \sum_{i=1}^{q} \left(\|u(t_{ni}) - u_{ni}\|^2 + |r(t_{ni}) - r_{ni}|^2\right)\right) \leqslant C\tau^{2q}.$$

Proof. We subtract (2.8)–(2.10) from (4.1)–(4.3) and, with the notation

$$e_n := u_n^\star - u_n, \qquad e_{ni} := u_{ni}^\star - u_{ni}, \qquad \dot{e}_{ni} := \dot{u}_{ni}^\star - \dot{u}_{ni},$$
$$\eta_n := r_n^\star - r_n, \qquad \eta_{ni} := r_{ni}^\star - r_{ni}, \qquad \dot{\eta}_{ni} := \dot{r}_{ni}^\star - \dot{r}_{ni},$$

we obtain the error equations

(4.17)
$$\begin{cases} \dot{e}_{ni} = \Delta e_{ni} - \eta_{ni} W(I_{n-1}^\tau u_{ni}) - r_{ni}^\star (W(I_{n-1}^\tau u_{ni}^\star) - W(I_{n-1}^\tau u_{ni})), & i = 1, \dots, q, \\ e_{ni} = e_n + \tau \sum_{j=1}^{q} a_{ij} \dot{e}_{nj} + \varepsilon_{ni}, & i = 1, \dots, q, \\ \partial_{\boldsymbol{n}} e_{ni} = 0 \quad \text{on } \partial\Omega, & i = 1, \dots, q, \end{cases}$$

(4.18)
$$\begin{cases} \dot{\eta}_{ni}^\star = \dfrac{1}{2}(W(I_{n-1}^\tau u_{ni}^\star) - W(I_{n-1}^\tau u_{ni}), \dot{u}_{ni}^\star) + \dfrac{1}{2}(W(I_{n-1}^\tau u_{ni}), \dot{e}_{ni}), & i = 1, \dots, q, \\ \eta_{ni}^\star = \eta_n^\star + \tau \displaystyle\sum_{j=1}^{q} a_{ij} \dot{\eta}_{nj}^\star + d_{ni}, & i = 1, \dots, q, \end{cases}$$

and

(4.19)
$$\begin{cases} e_{n+1} = e_n + \tau \displaystyle\sum_{i=1}^{q} b_i \dot{e}_{ni} + \varepsilon_{n+1}, \\ \eta_{n+1} = \eta_n + \tau \displaystyle\sum_{i=1}^{q} b_i \dot{\eta}_{ni} + d_{n+1}. \end{cases}$$

Let $1 \leqslant m \leqslant N$. In the following, we assume that for $n \leqslant m$ the error functions satisfy

$$\|e_{n-1,i}\|_{L^\infty(\Omega)} \leqslant 1, \quad i = 1, \dots, q, \tag{4.20}$$

and prove that this inequality holds also for $n = m + 1$. Then, by mathematical induction, this inequality holds for all $1 \leqslant m \leqslant N$.

For $n \leqslant m$, taking the square of the L^2-norm of both sides of the first relation of (4.19), we have

$$\|e_{n+1}\|^2 = \left\|e_n + \tau \sum_{i=1}^{q} b_i \dot{e}_{ni}\right\|^2 + 2\left(\varepsilon_{n+1}, e_n + \tau \sum_{i=1}^{q} b_i \dot{e}_{ni}\right) + \|\varepsilon_{n+1}\|^2. \tag{4.21}$$

We next estimate the first two terms on the right-hand side of (4.21). For the first term, we have

$$\left\|e_n + \tau \sum_{i=1}^{q} b_i \dot{e}_{ni}\right\|^2 = \|e_n\|^2 + 2\tau \sum_{i=1}^{q} b_i(\dot{e}_{ni}, e_n) + \tau^2 \sum_{i,j=1}^{q} b_i b_j (\dot{e}_{ni}, \dot{e}_{nj}),$$

and, replacing e_n in the second term by $e_{ni} - \tau \sum_{j=1}^{q} a_{ij} \dot{e}_{nj} - \varepsilon_{ni}$ (see the second relation of the error equation (4.17)), we obtain

$$\left\|e_n + \tau \sum_{i=1}^{q} b_i \dot{e}_{ni}\right\|^2 = \|e_n\|^2 + 2\tau \sum_{i=1}^{q} b_i(\dot{e}_{ni}, e_{ni} - \varepsilon_{ni}) - \tau^2 \sum_{i,j=1}^{q} m_{ij}(\dot{e}_{ni}, \dot{e}_{nj}).$$

Using here the positive semidefiniteness of the matrix M, we infer that

$$\left\|e_n + \tau \sum_{i=1}^{q} b_i \dot{e}_{ni}\right\|^2 \leqslant \|e_n\|^2 + 2\tau \sum_{i=1}^{q} b_i(\dot{e}_{ni}, e_{ni} - \varepsilon_{ni}). \tag{4.22}$$

We also need to estimate the interpolation error $W(I_{n-1}^{\tau} u_{ni}^{\star}) - W(I_{n-1}^{\tau} u_{ni})$. With $\ell_1, \dots, \ell_q \in \mathbb{P}_{q-1}$ the Lagrange polynomials for the nodes $t_{n-1,1}, \dots, t_{n-1,q}$, i.e., such that

$$\ell_i(t_{n-1,k}) = \delta_{ik}, \quad i, k = 1, \dots, q,$$

we have the Lagrange representation of the interpolant,

$$(I_{n-1}^{\tau} \varphi)(t) = \sum_{i=1}^{q} \varphi(t_{n-1,i}) \ell_i(t),$$

and easily infer that

$$\max_{1 \leqslant i \leqslant q} \|I_{n-1}^{\tau} e_{ni}^{\star}\| \leqslant C \max_{1 \leqslant i \leqslant q} \|e_{n-1,i}^{\star}\|.$$

Since W is locally Lipschitz continuous, by the induction assumption (4.20) the following estimate holds for $n \leqslant m$:

$$\max_{1 \leqslant i \leqslant q} \|W(I_{n-1}^{\tau} u_{ni}^{\star}) - W(I_{n-1}^{\tau} u_{ni})\| \leqslant C \max_{1 \leqslant i \leqslant q} \|I_{n-1}^{\tau} e_{ni}\| \leqslant C \max_{1 \leqslant i \leqslant q} \|e_{n-1,i}\|. \tag{4.23}$$

The L^{∞}-boundedness of $I_{n-1}^{\tau} u_{ni}$, as implied by the induction assumption (4.20), and the inequality (4.23) imply

$$\|\eta_{ni} W(I_{n-1}^{\tau} u_{ni}) + r_{ni}^{\star}[W(I_{n-1}^{\tau} u_{ni}^{\star}) - W(I_{n-1}^{\tau} u_{ni})]\| \leqslant C|\eta_{ni}| + C \max_{1 \leqslant i \leqslant q} \|e_{n-1,i}\|. \tag{4.24}$$

Furthermore, for the second term on the right-hand side of (4.22), we have, in view of the first relation of (4.17),

$$\begin{aligned}&(\dot e_{ni}, e_{ni}-\varepsilon_{ni})\\ &=(\Delta e_{ni}-\eta_{ni}W(I^\tau_{n-1}u^\star_{ni})-r_{ni}(W(I^\tau_{n-1}u^\star_{ni})-W(I^\tau_{n-1}u_{ni})), e_{ni}-\varepsilon_{ni})\\ &=-\|\nabla e_{ni}\|^2+(\nabla e_{ni},\nabla\varepsilon_{ni})-(\eta_{ni}W(I^\tau_{n-1}u_{ni})+r^\star_{ni}(W(I^\tau_{n-1}u^\star_{ni})-W(I^\tau_{n-1}u_{ni})), e_{ni}-\varepsilon_{ni})\end{aligned}$$

and thus easily, using (4.24),

$$2(\dot e_{ni}, e_{ni}-\varepsilon_{ni})\leqslant -\|\nabla e_{ni}\|^2+\|\nabla\varepsilon_{ni}\|^2+C|\eta_{ni}|^2+C\max_{1\leqslant j\leqslant q}\|e_{n-1,j}\|^2+\|e_{ni}\|^2+\|\varepsilon_{ni}\|^2.$$

Therefore, (4.22) yields

$$\begin{aligned}\left\|e_n+\tau\sum_{i=1}^q b_i\dot e_{ni}\right\|^2 &\leqslant \|e_n\|^2-\tau\sum_{i=1}^q b_i\|\nabla e_{ni}\|^2+\tau\sum_{i=1}^q b_i\|\varepsilon_{ni}\|^2_{H^1(\Omega)}\\ &\quad+\tau\sum_{i=1}^q b_i(\|e_{ni}\|^2+C|\eta_{ni}|^2)+C\tau\max_{1\leqslant j\leqslant q}\|e_{n-1,j}\|^2.\end{aligned}\tag{4.25}$$

In view of the first relation of the error equation (4.17), we have the estimate

$$\|\dot e_{ni}\|_\star\leqslant C\bigg(\|\nabla e_{ni}\|+\sum_{j=1}^q\|e_{n-1,j}\|\bigg),\quad i=1,\dots,q,\tag{4.26}$$

with $\|\cdot\|_\star$ the norm of $\widetilde H^{-1}(\Omega)$, the dual space of $H^1(\Omega)$. Then, we can estimate the second term on the right-hand side of (4.21) as follows:

$$\begin{aligned}\bigg(\varepsilon_{n+1}, e_n+\tau\sum_{i=1}^q b_i\dot e_{ni}\bigg)&\leqslant\tau\|\varepsilon_{n+1}/\tau\|\,\|e_n\|+C\tau\|\varepsilon_{n+1}\|_{H^1(\Omega)}\sum_{i=1}^q b_i\|\nabla e_{ni}\|\\ &\quad+C\tau\|\varepsilon_{n+1}\|_{H^1(\Omega)}\sum_{j=1}^q\|e_{n-1,j}\|.\end{aligned}\tag{4.27}$$

Combining (4.21) with (4.25) and (4.27), we obtain

$$\begin{aligned}\|e_{n+1}\|^2+\frac{\tau}{2}\sum_{i=1}^q b_i\|\nabla e_{ni}\|^2&\leqslant(1+C\tau)\|e_n\|^2+\tau\sum_{i=1}^q b_i\|\varepsilon_{ni}\|^2_{H^1(\Omega)}\\ &\quad+C\tau\sum_{i=1}^q\left(\|e_{ni}\|^2+|\eta_{ni}|^2+\|e_{n-1,i}\|^2\right)\\ &\quad+C\tau\left(\|\varepsilon_{n+1}\|^2_{H^1(\Omega)}+\|\varepsilon_{n+1}/\tau\|^2\right)+\|\varepsilon_{n+1}\|^2,\end{aligned}$$

and thus, in view of the consistency estimate (4.4),

$$\begin{aligned}&\|e_{n+1}\|^2+\frac{\tau}{2}\sum_{i=1}^q b_i\|\nabla e_{ni}\|^2\\ &\quad\leqslant\|e_n\|^2+C\tau\sum_{i=1}^q(\|e_{ni}\|^2+|\eta_{ni}|^2+\|e_{n-1,i}\|^2)+C\tau^{2q+1}.\end{aligned}\tag{4.28}$$

Similarly, from (4.18) we can derive

$$|\eta_{n+1}|^2 \leqslant |\eta_n|^2 + C\tau \sum_{i=1}^{q} (\|e_{ni}\|^2 + |\eta_{ni}|^2 + \|e_{n-1,i}\|^2) + C\tau^{2q+1}. \tag{4.29}$$

Summing (4.28) and (4.29), we have

$$\begin{aligned} &\|e_{n+1}\|^2 + |\eta_{n+1}|^2 + \frac{\tau}{2} \sum_{i=1}^{q} b_i \|\nabla e_{ni}\|^2 \\ &\quad \leqslant \|e_n\|^2 + |\eta_n|^2 + C\tau \sum_{i=1}^{q} (\|e_{ni}\|^2 + |\eta_{ni}|^2 + \|e_{n-1,i}\|^2) + C\tau^{2q+1}. \end{aligned} \tag{4.30}$$

Now, we estimate the term $C\tau \sum_{i=1}^{q} (\|e_{ni}\|^2 + |\eta_{ni}|^2)$ on the right-hand side. To this end, we test the second relation of (4.17) by e_{ni}. This yields

$$\sum_{i=1}^{q} \|e_{ni}\|^2 \leqslant C\|e_n\|^2 + C\tau \sum_{i,j=1}^{q} a_{ij}(\dot{e}_{nj}, e_{ni}) + C \sum_{i=1}^{q} \|\varepsilon_{ni}\|^2.$$

Then, using the first relation of (4.17), we have

$$\begin{aligned} \sum_{i,j=1}^{q} a_{ij}(\dot{e}_{nj}, e_{ni}) &= - \sum_{i,j=1}^{q} a_{ij}(\nabla e_{nj}, \nabla e_{ni}) \\ &\quad - \sum_{i,j=1}^{q} a_{ij}(\eta_{nj} W(I_{n-1}^{\tau} u_{nj}) + r_{nj}^{\star}(W(I_{n-1}^{\tau} u_{nj}^{\star}) - W(I_{n-1}^{\tau} u_{nj})), e_{ni}) \\ &\leqslant C \sum_{i=1}^{q} (\|\nabla e_{ni}\|^2 + \|e_{ni}\|^2 + |\eta_{ni}|^2 + \|e_{n-1,i}\|^2) + C\|e_{n-1}\|^2. \end{aligned}$$

Combining the last two estimates, we obtain

$$\begin{aligned} \sum_{i=1}^{q} \|e_{ni}\|^2 \leqslant C\|e_n\|^2 &+ C\tau \sum_{i=1}^{q} (\|\nabla e_{ni}\|^2 + \|e_{ni}\|^2 + |\eta_{ni}|^2 + \|e_{n-1,i}\|^2) \\ &+ C\tau \|e_{n-1}\|^2 + C \sum_{i=1}^{q} \|\varepsilon_{ni}\|^2. \end{aligned}$$

Similarly, from (4.18) we can derive

$$\begin{aligned} \sum_{i=1}^{q} |\eta_{ni}|^2 \leqslant C|\eta_n|^2 &+ C\tau \sum_{i=1}^{q} (\|\nabla e_{ni}\|^2 + \|e_{ni}\|^2 + |\eta_{ni}|^2 + \|e_{n-1,i}\|^2) \\ &+ C\tau \|e_{n-1}\|^2 + C \sum_{i=1}^{q} \|d_{ni}\|^2. \end{aligned}$$

Sum up these two estimates and note that, for sufficiently small τ, the term $C\tau \sum_{i=1}^{q} (\|e_{ni}\|^2 + |\eta_{ni}|^2)$ on the right-hand side can be absorbed by the left-hand side. Then, we obtain

$$\begin{aligned} \sum_{i=1}^{q} (\|e_{ni}\|^2 + |\eta_{ni}|^2) \leqslant C(\|e_n\|^2 + |\eta_n|^2) &+ C\tau \sum_{i=1}^{q} b_i \|\nabla e_{ni}\|^2 \\ &+ C\tau \sum_{i=1}^{q} \|e_{n-1,i}\|^2 + C \sum_{i=1}^{q} (\|\varepsilon_{ni}\|^2 + \|d_{ni}\|^2), \end{aligned} \tag{4.31}$$

where we used the positivity of the weights $b_1, \dots, b_q$. Substituting this inequality into (4.30) yields

$$\begin{aligned}
&\|e_{n+1}\|^2 + |\eta_{n+1}|^2 + \frac{\tau}{2}\sum_{i=1}^{q} b_i \|\nabla e_{ni}\|^2 \\
&\leqslant (1 + C_1\tau)(\|e_n\|^2 + |\eta_n|^2) + C_1\tau^2 \sum_{i=1}^{q} b_i \|\nabla e_{ni}\|^2 \\
&\quad + C_1\tau \sum_{i=1}^{q} \|e_{n-1,i}\|^2 + C_1\tau \sum_{i=1}^{q} \|\varepsilon_{ni}\|^2 + C_1\tau^{2q+1},
\end{aligned} \tag{4.32}$$

with some constant C_1. Multiplying (4.31) by $2C_1\tau$ and adding to (4.32), we get

$$\begin{aligned}
&\left(\|e_{n+1}\|^2 + |\eta_{n+1}|^2 + 2C_1\tau \sum_{i=1}^{q} (\|e_{ni}\|^2 + |\eta_{ni}|^2)\right) + \frac{\tau}{2}\sum_{i=1}^{q} b_i \|\nabla e_{ni}\|^2 \\
&\leqslant (1 + C_2\tau)(\|e_n\|^2 + |\eta_n|^2) + C_2\tau^2 \sum_{i=1}^{q} b_i \|\nabla e_{ni}\|^2 \\
&\quad + (C_1 + C_2\tau)\tau \sum_{i=1}^{q} \|e_{n-1,i}\|^2 + C_2\tau^{2q+1},
\end{aligned}$$

with some constant C_2. For sufficiently small τ, the term $C_2\tau^2 \sum_{i=1}^{q} b_i \|\nabla e_{ni}\|^2$ can be absorbed by the left-hand side, and $C_1 + C_2\tau \leqslant 2C_1$. Therefore, the inequality above reduces to

$$\begin{aligned}
&\left(\|e_{n+1}\|^2 + |\eta_{n+1}|^2 + 2C_1\tau \sum_{i=1}^{q} (\|e_{ni}\|^2 + |\eta_{ni}|^2)\right) + \frac{\tau}{4}\sum_{i=1}^{q} b_i \|\nabla e_{ni}\|^2 \\
&\leqslant (1 + (2C_1 + C_2)\tau)\left[\|e_n\|^2 + |\eta_n|^2 + 2C_1\tau \sum_{i=1}^{q} (\|e_{n-1,i}\|^2 + |\eta_{n-1,i}|^2)\right] + C_2\tau^{2q+1}.
\end{aligned}$$

By using Gronwall's inequality, we obtain

$$\begin{aligned}
&\max_{1\leqslant n\leqslant m}\left(\|e_{n+1}\|^2 + |\eta_{n+1}|^2 + 2C_1\tau \sum_{i=1}^{q} (\|e_{ni}\|^2 + |\eta_{ni}|^2)\right) + \frac{\tau}{4}\sum_{i=1}^{q} b_i \|\nabla e_{ni}\|^2 \\
&\leqslant C\left[\|e_1\|^2 + |\eta_1|^2 + 2C_1\tau \sum_{i=1}^{q} (\|e_{0i}\|^2 + |\eta_{0i}|^2)\right] + C\tau^{2q}.
\end{aligned} \tag{4.33}$$

This estimate implies

$$\|e_{m+1}\| + \|e_m\| \leqslant C\tau^q \quad \text{and} \quad \|e_{mi}\| + |\eta_{mi}| \leqslant C\tau^{q-\frac{1}{2}}.$$

From the second relation of (4.17) we obtain

$$\|\dot{e}_{mi}\| \leqslant C\tau^{-1} \sum_{j=1}^{q} (\|e_{mj} - e_m\| + \|\varepsilon_{mj}\|) \leqslant C\tau^{q-\frac{3}{2}}. \tag{4.34}$$

Then, from the first relation of (4.17) we further derive

$$\|\Delta e_{mi}\| \leqslant C(\|\dot{e}_{mi}\| + \|\eta_{mi}\| + \|e_{m-1,i}\|) \leqslant C\tau^{q-\frac{3}{2}}, \tag{4.35}$$

and, therefore,

$$\|e_{mi}\|_{H^2(\Omega)} \leqslant C(\|e_{mi}\| + \|\Delta e_{mi}\|) \leqslant C\tau^{q-\frac{3}{2}}. \tag{4.36}$$

If $q \geqslant 2$, then for $d \in \{1, 2, 3\}$ the following Sobolev embedding inequality holds:

$$\|e_{mi}\|_{L^\infty(\Omega)} \leqslant C\|e_{mi}\|_{H^2(\Omega)} \leqslant 1$$

for sufficiently small τ (the smallness is independent of m). This completes the mathematical induction on (4.20). Therefore, the inequality (4.20) together with the error estimate (4.33) hold for all $1 \leqslant m \leqslant N$. ∎

5. Numerical examples. In this section, we present three numerical examples to illustrate the convergence and energy decay property of the extrapolated RK–SAV method.

Example 5.1. We consider the one-dimensional (1D) AC equation

$$\partial_t u - \partial_{xx} u + \frac{1}{\varepsilon^2}(u^3 - u) = 0, \quad (x, t) \in (0, \pi) \times (0, 1), \tag{5.1}$$

with the initial condition

$$u_0(x) = \sin^3 x.$$

The equation is discretized in time by the proposed extrapolated Gauss and Radau IIA RK–SAV methods and in space by the piecewise linear finite element method with mesh size $h = \pi/256$. First, we set $\varepsilon = 1$ and test the accuracy of the proposed method by choosing a reference solution with a very small time step, namely $\tau = 1/2560$.

The L^2-norms of the errors of numerical solutions at time $T = 1$ are presented in Table 5.1. The numerical results in Table 5.1 indicate that the extrapolated q-stage Gauss and Radau IIA methods actually have $(q+1)$th-order convergence, higher than the qth-order convergence proved in Theorem 4.2. Theoretical proof of the $(q+1)$th-order convergence remains open for the proposed method. However, instead of using q-point extrapolation, if we use $(q+1)$-point extrapolation in our numerical method, then we can actually prove $(q+1)$th-order convergence; see section 6.

TABLE 5.1
Numerical results for the 1D AC equation (Example 5.1).

		Gauss		Radau IIA	
q	N	Error	Order	Error	Order
2	160	2.32E-7	–	8.04E-8	–
	200	1.23E-7	2.86	4.16E-8	2.97
	240	7.25E-8	2.88	2.42E-8	2.97
	280	4.64E-8	2.90	1.53E-8	2.97
	320	3.14E-8	2.92	1.03E-8	2.98
3	160	5.45E-10	–	3.59E-10	–
	200	2.27E-10	3.92	1.50E-10	3.91
	240	1.11E-10	3.93	7.33E-11	3.92
	280	6.03E-11	3.94	3.99E-11	3.94
	320	3.56E-11	3.95	2.36E-11	3.94
4	160	1.09E-11	–	7.64E-12	–
	200	3.69E-12	4.87	2.58E-12	4.86
	240	1.51E-12	4.89	1.06E-12	4.89
	280	7.09E-13	4.91	4.97E-13	4.91
	320	3.68E-13	4.92	2.57E-13	4.93

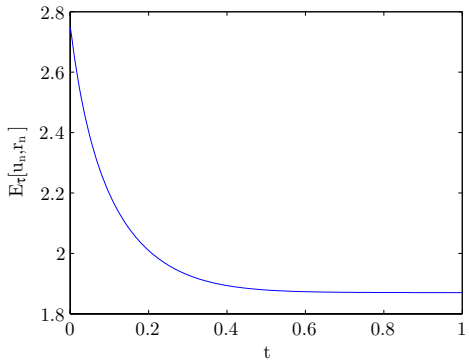

FIG. 5.1. *Discrete energy for the* 1*D AC equation with* $\varepsilon^2 = 0.25$ (*Example* 5.1).

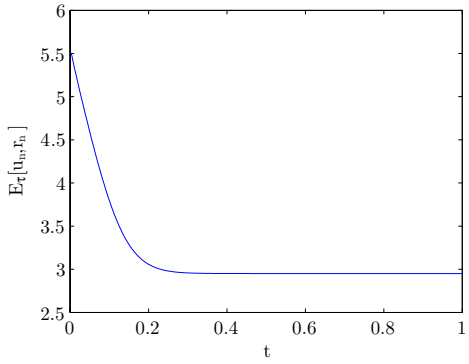

FIG. 5.2. *Discrete energy for the* 1*D AC equation with* $\varepsilon^2 = 0.1$ (*Example* 5.1).

Second, we illustrate the energy decay property of the proposed extrapolated RK–SAV method by presenting in Figures 5.1–5.2 the discrete energies of the numerical solutions given by the extrapolated 4-stage Radau IIA RK–SAV method with $\tau = 1/256$ and $h = \pi/256$, for $\varepsilon^2 = 0.25$ and $\varepsilon^2 = 0.1$, respectively. The numerical results show that the discrete energy decays, consistent with our theoretical result in Theorem 3.1.

Example 5.2. We consider the following 2D AC equation:

$$\partial_t u - \Delta u + \frac{1}{\varepsilon^2}(u^3 - u) = g(x, y, t), \quad x, y, t \in (0, 1). \tag{5.2}$$

First, we test the accuracy of the numerical methods by choosing the initial condition and the source term g such that the exact solution is

$$u(x, y, t) = \exp(-t)x^2(1 - x)^2y^2(1 - y)^2.$$

The L^2-norms of the errors of the extrapolated q-stage, $q = 2, 3, 4$, Gauss and Radau IIA RK–SAV methods at $T = 1$ are presented in Table 5.2. The corresponding spatial discretization is done by using piecewise quadratic finite elements in FreeFEM++ with $h = \tau$, $h = \tau^{4/3}$, and $h = \tau^{5/3}$, respectively. Again, the numerical results indicate that the extrapolated q-stage Gauss and Radau IIA RK–SAV methods are $(q + 1)$th-order convergent.

Second, we set $g = 0$ in (5.2) and consider the initial value $u_0(x, y) = x^2(1 - x)^2y^2(1 - y)^2$. We present the discrete energies for the numerical solutions given

TABLE 5.2
Numerical results for the 2D AC equation (Example 5.2).

		Gauss		Radau IIA	
q	N	Error	Order	Error	Order
2	10	1.08E-6	–	9.64E-7	–
	20	1.28E-7	3.07	1.22E-7	2.98
	30	3.73E-8	3.03	3.65E-8	2.99
	40	1.57E-8	3.02	1.54E-8	2.99
	80	7.99E-9	3.01	7.91E-9	2.99
3	12	5.23E-8	–	5.05E-8	–
	16	1.58E-8	4.15	1.55E-8	4.10
	20	6.40E-9	4.06	6.29E-9	4.04
	24	3.06E-9	4.05	3.01E-9	4.02
	28	1.63E-9	4.08	1.62E-9	4.05
4	8	3.09E-8	–	3.01E-8	–
	10	1.03E-8	4.92	1.02E-8	4.87
	12	4.20E-9	4.92	4.15E-9	4.93
	14	1.88E-9	5.21	1.86E-9	5.19
	16	9.67E-10	4.97	9.61E-10	4.96

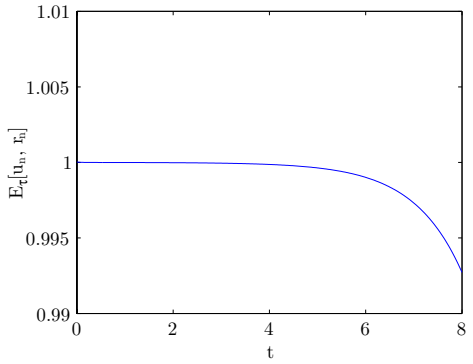

FIG. 5.3. *Discrete energy for the 2D AC equation with* $\varepsilon^2 = 0.25$ *(Example 5.2).*

by the extrapolated two-stage Gauss RK–SAV method in Figures 5.3 and 5.4 for $\varepsilon^2 = 0.25$ and $\varepsilon^2 = 0.1$, respectively, using $\tau = 1/40$ and $h = 1/40$. The numerical results show that the discrete energy decays, consistent with our theoretical result in Theorem 3.1.

Example 5.3. We consider the following 2D CH equation:

$$\partial_t u - \alpha \Delta \left(-\Delta u + \frac{1}{\varepsilon^2} f(u) \right) = 0, \quad x, y, t \in (0, 1). \tag{5.3}$$

We test the accuracy of the numerical methods by choosing $f(u) = u$, $\alpha = (4\pi^2 + \frac{2\pi^2}{\varepsilon^2})^{-1}$, and $\varepsilon^2 = 0.01$, which correspond to the exact solution

$$u(x, y, t) = \exp(-t) \cos(\pi x) \cos(\pi y).$$

The L^2-norms of the errors of the extrapolated q-stage, $q = 2, 3, 4$, Gauss and Radau IIA RK–SAV methods at $T = 1$ are presented in Table 5.3. The spatial discretization

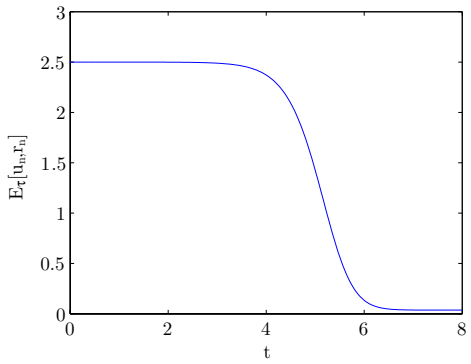

FIG. 5.4. *Discrete energy for the 2D AC equation with* $\varepsilon^2 = 0.1$ (*Example* 5.2).

TABLE 5.3
Numerical results for the 2D CH equation (*Example* 5.3).

		Gauss		Radau IIA	
q	N	Error	Order	Error	Order
2	10	1.05E-4	–	9.32E-5	–
	20	1.21E-5	3.12	1.13E-5	3.04
	30	3.48E-6	3.07	3.34E-6	3.01
	40	1.45E-6	3.04	1.41E-6	3.01
	80	7.37E-7	3.03	7.19E-7	3.00
3	12	4.81E-6	–	4.56E-6	–
	16	1.45E-6	4.17	1.40E-6	4.10
	20	5.84E-7	4.08	5.69E-7	4.03
	24	2.78E-7	4.06	2.73E-7	4.03
	28	1.48E-7	4.08	1.46E-7	4.06
4	8	2.86E-6	–	2.74E-6	–
	10	9.51E-7	4.95	9.22E-7	4.88
	12	3.85E-7	4.96	3.76E-7	4.91
	14	1.72E-7	5.22	1.69E-7	5.20
	16	8.83E-8	4.99	8.71E-8	4.96

is done by using piecewise quadratic finite elements in FreeFEM++ with $h = \tau$, $h = \tau^{4/3}$, and $h = \tau^{5/3}$, respectively. Again, the numerical results indicate that the extrapolated q-stage Gauss and Radau IIA RK–SAV methods are $(q+1)$th-order convergent.

For $f(u) = u^3 - u$, $\alpha = 1$, and $\varepsilon^2 = 0.001$, we solve (5.3) by the extrapolated two-stage Gauss RK–SAV finite element method (with $\tau = 10^{-6}$ and $h = 1/40$), with a random initial value

$$u(x, y, 0) = \mathrm{Rand}(x, y) - 0.5,$$

where $\mathrm{Rand}(x, y)$ stands for a uniform random variable in $[0, 1] \times [0, 1]$; therefore, the initial value takes on both positive and negative values. The discrete energy of the numerical solution is presented in Figure 5.5, and the evolution of the two phases is shown in Figure 5.6. The numerical results show that the discrete energy decays, which agrees with our theoretical analysis.

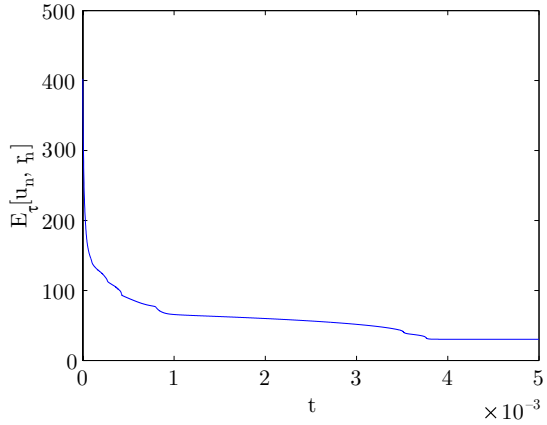

FIG. 5.5. *Discrete energy for the CH equation with* $\varepsilon^2 = 0.001$ (*Example* 5.3).

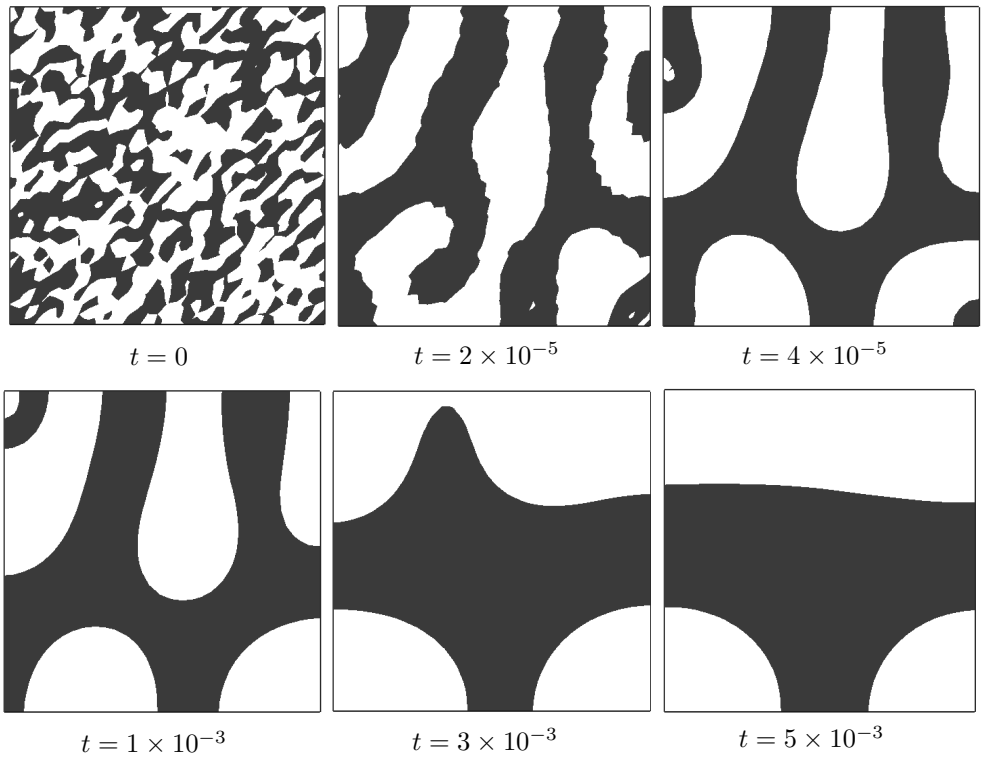

FIG. 5.6. *Numerical simulation of the CH equation with* $\varepsilon^2 = 0.001$ *(Example* 5.3*).*

6. Conclusion. We proposed a class of extrapolated RK–SAV methods for the time discretization of the AC and CH equations. This class of methods requires only the solution of a system of linear equations at each time level and preserves a discrete version of the energy decay property. By using q-point extrapolation in the nonlinear term, we proved that the extrapolated q-stage RK–SAV methods have qth-order convergence in time in approximating smooth solutions of the AC equation.

Since we used the SAV formulation, the energy stability of the proposed numerical methods is proved only for the modified energy $E_\tau[u, r]$ instead of the original energy $E[u]$. The construction of linearly implicit high-order methods preserving the energy decay for the original energy $E[u]$ remains still open.

In Theorem 4.2, we used a q-point extrapolation and proved qth-order convergence of the methods. A $(q+1)$-point extrapolation, with an *internal node* of the Runge–Kutta method as an additional interpolation point, at which we have control of the L^∞-norm by (4.36), would lead to order $q+1$ in Theorem 4.2. In particular, $(q+1)$th-order convergence can be established for q-stage Gauss and Radau IIA methods, provided we use $(q+1)$-point extrapolation.

We chose to present the details for the more general result of the q-point extrapolation in Theorem 4.2 for two reasons. First, for general Runge–Kutta methods, there are only q internal nodes in the time interval $[t_{n-1}, t_n]$ (except for the Radau IIA method, which has $q+1$ internal nodes, including both endpoints). But this is not an essential difficulty for practical computation, as it can be compensated by an internal node from the previous subinterval $[t_{n-2}, t_{n-1}]$, such as $t_{n-2,q}$. Second, q-point extrapolation also yields very good numerical results, i.e., $(q+1)$th-order convergence, in our numerical experiments. We believe that q-point extrapolation is practically a good and efficient choice for general Runge–Kutta methods, including Gauss methods.

In the convergence analysis, we proved the L^∞ bound by mathematical induction, i.e., assuming the a-priori L^∞ bound (4.20) at the previous time step, we derived the L^2 error estimate at the current time step with a bound (4.34) for its discrete temporal derivative. This in turn yields an H^2 error estimate with a reduced accuracy order. Subsequently, an application of Sobolev embedding implies the L^∞ bound of the numerical error function at the current time step. A similar technique was applied in [13, 33] to analyze incompressible Navier–Stokes equations and gradient flows.

An extension of our error analysis to fully discrete finite element methods requires a modification to yield the L^∞ bound of numerical solutions. For example, suppose that finite elements of polynomial degree $r \geqslant 1$ are used, with mesh size h. We denote $e_{ni}^{(h)} = R_h u_{ni}^\star - u_{ni}^{(h)}$, with R_h being the Ritz projection onto the finite element space. On the one hand, similarly as (4.34)–(4.35), one can derive an estimate

$$\left\| \Delta_h e_{mi}^{(h)} \right\| \leqslant C\tau^{-1} \left(\tau^{q-\frac{1}{2}} + h^{r+1} \right), \tag{6.1}$$

where Δ_h denotes the discrete Laplacian operator. On the other hand, one can use the inverse inequality of the finite element space, i.e.,

$$\left\| \Delta_h e_{mi}^{(h)} \right\| \leqslant Ch^{-2} \left(\tau^{q-\frac{1}{2}} + h^{r+1} \right). \tag{6.2}$$

Combining these two estimates, one can obtain

$$\left\| \Delta_h e_{mi}^{(h)} \right\| \leqslant C \min\left(\tau^{-1}, h^{-2}\right) \left(\tau^{q-\frac{1}{2}} + h^{r+1} \right) \leqslant C(\tau^{q-\frac{3}{2}} + h^{r-1}).$$

This can be further used to derive an L^∞ estimate of $e_{mi}^{(h)}$.

The extension of the convergence analysis to the CH equation can be done similarly for semidiscretization in time but needs to be modified for the standard C^0 finite element method for spatial discretization. The error estimation for the CH equation also requires L^∞ bounds of the numerical solutions. This can be done as for the AC equation combining two different types of estimates of the form (6.1) and (6.2).

REFERENCES

[1] S. M. Allen and J. W. Cahn, *A microscopic theory for antiphase boundary motion and its application to antiphase domain coarsening*, Acta Metall., 27 (1979), pp. 1085–1095.

[2] S. Badia, F. Guillén-González, and J. V. Gutiérrez-Santacreu, *Finite element approximation of nematic liquid crystal flows using a saddle-point structure*, J. Comput. Phys., 230 (2011), pp. 1686–1706.

[3] S. Bartels and R. Müller, *Quasi-optimal and robust a posteriori error estimates in $L^\infty(L^2)$ for the approximation of Allen–Cahn equations past singularities*, Math. Comp., 80 (2011), pp. 761–780.

[4] S. Bartels, R. Müller, and C. Ortner, *Robust a priori and a posteriori error analysis for the approximation of Allen–Cahn and Ginzburg–Landau equations past topological changes*, SIAM J. Numer. Anal., 49 (2011), pp. 110–134.

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

[6] Y. Cai, H. Choi, and J. Shen, *Error estimates for time discretizations of a phase-field model for two-phase incompressible flow*, Numer. Math., 137 (2017), pp. 417–449.

[7] Y. Cai and J. Shen, *Error estimates for a fully discretized scheme to a Cahn–Hilliard phase-field model for two-phase incompressible flows*, Math. Comp., 87 (2018), pp. 2057–2090.

[8] W. Chen, Y. Liu, C. Wang, and S. M. Wise, *Convergence analysis of a fully discrete finite difference scheme for the Cahn–Hilliard–Hele–Shaw equation*, Math. Comp., 85 (2016), pp. 2231–2257.

[9] Q. Cheng and J. Shen, *Multiple scalar auxiliary variable (MSAV) approach and its application to the phase-field vesicle membrane model*, SIAM J. Sci. Comput., 40 (2018), pp. A3982–A4006.

[10] D. Cohen and E. Hairer, *Linear energy-preserving integrators for Poisson systems*, BIT, 51 (2011), pp. 91–101.

[11] M. Dahlby, B. Owren, and T. Yaguchi, *Preserving multiple first integrals by discrete gradients*, J. Phys. A, 44 (2011) 305205.

[12] M. Doi and S. F. Edwards, *The Theory of Polymer Dynamics*, Oxford University Press, New York, 1988.

[13] W. E and J. G. Liu, *Projection method* I: *Convergence and numerical boundary layers*, SIAM J. Numer. Anal., 32 (1995), pp. 1017–1057.

[14] C. M. Elliott and A. M. Stuart, *The global dynamics of discrete semilinear parabolic equations,* SIAM J. Numer. Anal., 30 (1993), pp. 1622–1663.

[15] D. J. Eyre, *Unconditionally gradient stable time marching the Cahn–Hilliard equation,* Mater. Res. Soc. Symp. Proc., 529 (1998), pp. 39–46.

[16] X. Feng, Y. Li, and Y. Xing, *Analysis of mixed interior penalty discontinuous Galerkin method for the Cahn–Hilliard equation and the Hele–Shaw flow*, SIAM J. Numer. Anal., 54 (2016), pp. 825–847.

[17] X. Feng and A. Prohl, *Numerical analysis of the Allen–Cahn equation and approximation for mean curvature flows*, Numer. Math., 94 (2003), pp. 33–65.

[18] X. Feng and A. Prohl, *Numerical analysis of the Cahn–Hilliard equation and approximation for the Hele–Shaw problem*, Interfaces Free Bound., 7 (2005), pp. 1–28.

[19] Y. Gong and J. Zhao, *Energy-stable Runge–Kutta schemes for gradient flow models using the energy quadratization approach*, Appl. Math. Lett., 94 (2019), pp. 224–231.

[20] O. Gonzalez, *Time integration and discrete Hamiltonian systems*, J. Nonlinear Sci., 6 (1996), pp. 449–467.

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

[22] E. Hairer, *Energy-preserving variant of collocation methods*, J. Numer. Anal. Ind. Appl. Math., 5 (2010), pp. 73–84.

[23] E. Hairer and C. Lubich, *Energy-diminishing integration of gradient systems*, IMA J. Numer. Anal., 34 (2014), pp. 452–461.

[24] E. Hairer and G. Wanner, *Solving Ordinary Differential Equations* II: *Stiff and Differential–Algebraic Problems*, 2nd ed., Springer Ser. Comput. Math. 14, Springer, New York, 2002.

[25] Z. Hu, S. M. Wise, C. Wang, and J. S. Lowengrub, *Stable and efficient finite-difference nonlinear-multigrid schemes for the phase field crystal equation*, J. Comput. Phys., 228 (2009), pp. 5323–5339.

[26] D. Li and Z. Qiao, *On second order semi-implicit Fourier spectral methods for 2D Cahn–Hilliard equations*, J. Sci. Comput., 70 (2017), pp. 301–341.

[27] D. Li, Z. Qiao, and T. Tang, *Characterizing the stabilization size for semi-implicit Fourier-spectral method to phase field equations*, SIAM J. Numer. Anal., 54 (2016), pp. 1653–1681.
[28] X. Li, J. Shen, and H. Rui, *Energy stability and convergence of SAV block-centered finite difference method for gradient flows*, Math. Comp., 88 (2019), pp. 2047–2068.
[29] C. Lubich and A. Ostermann, *Runge–Kutta approximation of quasi-linear parabolic equations*, Math. Comp., 64 (1995), pp. 601–627.
[30] R. I. McLachlan, G. R. W. Quispel, and N. Robidoux, *Geometric integration using discrete gradients*, Philos. Trans. A, 357 (1999), pp. 1021–1045.
[31] R. Norton and G. Quispel, *Discrete gradient methods for preserving a first integral of an ordinary differential equation*, Discrete Contin. Dyn. Syst., 34 (2012), pp. 1147–1170.
[32] J. Shen, C. Wang, X. Wang, and S. M. Wise, *Second-order convex splitting schemes for gradient flows with Ehrlich–Schwoebel type energy: Application to thin film epitaxy*, SIAM J. Numer. Anal., 50 (2012), pp. 105–125.
[33] J. Shen and J. Xu, *Convergence and error analysis for the scalar auxiliary variable (SAV) schemes to gradient flows*, SIAM J. Numer. Anal., 56 (2018), pp. 2895–2912.
[34] J. Shen, J. Xu, and J. Yang, *The scalar auxiliary variable (SAV) approach for gradient flows*, J. Comput. Phys., 353 (2018), pp. 407–416.
[35] J. Shen and X. Yang, *Numerical approximations of Allen–Cahn and Cahn–Hilliard equations*, Discrete Contin. Dyn. Syst., 28 (2010), pp. 1669–1691.
[36] C. Wang and S. M. Wise, *An energy stable and convergent finite-difference scheme for the modified phase field crystal equation*, SIAM J. Numer. Anal., 49 (2011), pp. 945–969.
[37] L. Wang and H. Yu, *On efficient second order stabilized semi-implicit schemes for the Cahn–Hilliard phase-field equation*, J. Sci. Comput., 77 (2018), pp. 1185–1209.
[38] C. Xu and T. Tang, *Stability analysis of large time-stepping methods for epitaxial growth models*, SIAM J. Numer. Anal., 44 (2006), pp. 1759–1779.
[39] X. Yang, *Error analysis of stabilized semi-implicit method of Allen–Cahn equation*, Discrete Contin. Dyn. Syst. Ser. B, 11 (2009), pp. 1057–1070.
[40] 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), pp. 294–316.
[41] X. Yang and L. Ju, *Efficient linear schemes with unconditional energy stability for the phase field elastic bending energy model*, Comput. Methods Appl. Mech. Engrg., 315 (2017), pp. 691–712.
[42] X. Yang and L. Ju, *Linear and unconditionally energy stable schemes for the binary fluid-surfactant phase field model*, Comput. Methods Appl. Mech. Engrg., 318 (2017), pp. 1005–1029.
[43] X. Yang, J. Zhao, and Q. Wang, *Numerical approximations for the molecular beam epitaxial growth model based on the invariant energy quadratization method*, J. Comput. Phys., 333 (2017), pp. 104–127.
[44] X. Yang, J. Zhao, Q. Wang, and J. Shen, *Numerical approximations for a three components Cahn–Hilliard phase-field model based on the invariant energy quadratization method*, Math. Models Methods Appl. Sci., 27 (2017), pp. 1993–2030.
[45] P. Yue, J. J. Feng, C. Liu, and J. Shen, *A diffuse-interface method for simulating two-phase flows of complex fluids*, J. Fluid Mech., 515 (2004), pp. 293–317.
[46] J. Zhao, X. Yang, J. Li, and Q. Wang, *Energy stable numerical schemes for a hydrodynamic model of nematic liquid crystals*, SIAM J. Sci. Comput., 38 (2016), pp. A3264–A3290.
[47] J. Zhu, L. Chen, J. Shen, and V. Tikare, *Coarsening kinetics from a variable mobility Cahn–Hilliard equation—application of semi-implicit Fourier spectral method*, Phys. Rev. E, 60 (1999), pp. 3564–3572.