Arbitrarily High Order and Fully Discrete Extrapolated RK–SAV/DG Schemes for Phase-field Gradient Flows

Article *in* Journal of Scientific Computing · September 2022

DOI: 10.1007/s10915-022-01995-5

CITATION
1

READS
114

3 authors, including:

Xu Wu
Southern University of Science and Technology
6 PUBLICATIONS **4** CITATIONS

SEE PROFILE

Journal of Scientific Computing (2022) 93:38
https://doi.org/10.1007/s10915-022-01995-5

Arbitrarily High Order and Fully Discrete Extrapolated RK–SAV/DG Schemes for Phase-field Gradient Flows

Tao Tang[1,2] · **Xu Wu**[3,4] · **Jiang Yang**[4,5]

Received: 29 May 2022 / Revised: 28 July 2022 / Accepted: 19 August 2022 /
Published online: 16 September 2022

Abstract
In this paper, we construct and analyze a fully discrete method for phase-field gradient flows, which uses extrapolated Runge–Kutta with scalar auxiliary variable (RK–SAV) method in time and discontinuous Galerkin (DG) method in space. We propose a novel technique to decouple the system, after which only several elliptic scalar problems with constant coefficients need to be solved independently. Discrete energy decay property of the method is proved for gradient flows. The scheme can be of arbitrarily high order both in time and space, which is demonstrated rigorously for the Allen–Cahn equation and the Cahn–Hilliard equation. More precisely, optimal L^2-error bound in space and qth-order convergence rate in time are obtained for q-stage extrapolated RK–SAV/DG method. Several numerical experiments are carried out to verify the theoretical results.

Keywords Phase-field models · Gradient flows · Energy stability · Convergence and error analysis · Allen–Cahn equation · Cahn–Hilliard equation

✉ Xu Wu
11849596@mail.sustech.edu.cn

Tao Tang
tangt@sustech.edu.cn

Jiang Yang
yangj7@sustech.edu.cn

1 Division of Science and Technology, BNU-HKBU United International College, Zhuhai, Guangdong, China

2 SUSTech International Center for Mathematics, Southern University of Science and Technology, Shenzhen, China

3 School of Mathematics, Harbin Institute of Technology, Harbin, China

4 Department of Mathematics, Southern University of Science and Technology, Shenzhen, China

5 Guangdong Provincial Key Laboratory of Computational Science and Material Design, Southern University of Science and Technology, Shenzhen, China

1 Introduction

Gradient flows become more and more significant in science and engineering. A large class of mathematical models can be read as PDEs in the form of gradient flows, for instance crystal growth, liquid crystals, thin films, tumor growth, solidification, interface dynamics, see, e.g., [3, 5, 6, 19, 36, 37]. Generally, we consider a system with the total free energy in the form:

$$\mathcal{E}[u] = \frac{1}{2}(u, \mathcal{L}u) + \mathcal{E}_1[u], \tag{1.1}$$

where $(\cdot, \cdot)$ is the standard L^2 inner product, $\mathcal{L}$ is a symmetric non-negative linear operator, and $\mathcal{E}_1[u]$ is nonlinear but with only lower-order derivatives than $\mathcal{L}$, and $\mathcal{E}_1[u]$ is bounded from below. A general form of the gradient flow associating with the free energy (1.1) can be written as

$$\frac{\partial u}{\partial t} = \mathcal{G}\mu, \quad \text{and} \quad \mu = \frac{\delta \mathcal{E}}{\delta u}, \tag{1.2}$$

supplemented with suitable boundary conditions and initial data. Here, a non-positive symmetric operator $\mathcal{G}$ determines the dissipation law of the system, e.g., $\mathcal{G} = -\mathcal{I}$ and $\mathcal{G} = \Delta$ leading to the L^2 gradient flow and the H^{-1} gradient flow, respectively. Since $\mathcal{G}$ is non-positive, the free energy is non-increasing due to the following energy dissipation law

$$\frac{\mathrm{d}\mathcal{E}[u]}{\mathrm{d}t} = \left(\frac{\delta \mathcal{E}}{\delta u}, \frac{\partial u}{\partial t}\right) = (\mu, \mathcal{G}\mu) \leq 0.$$

Various gradient flows are listed in Table 1. Without loss of the generality and to keep the presentation short, we limit our concentration on the AC and CH equations in this work. The AC equation was originally introduced by Allen and Cahn in [2] to describe the motion of anti-phase boundaries in crystalline solids. and the CH equation was introduced by Cahn and Hilliard in [8] to describe the complicated phase separation and coarsening phenomena in a solid. In recent decades, the AC and CH equations have become two commonly used phase-field equations, which have been widely applied to many complicated moving interface problems in materials science and fluid dynamics through a phase-field approach coupled with other models, see, e.g., [3, 10, 44]. Efficient and energy stable numerical schemes for phase-field gradient flows are very prevalent in the last few decades [16, 18, 20, 25, 30, 33, 40], and we refer to [17] for an up-to-date extensive review on this subject. In particular, inspired by the invariant energy quadratization (IEQ) approach [42, 43], Shen et al. [12, 31, 32] propose the scalar auxiliary variable (SAV) method, which can be implemented efficiently for a large class of gradient flows. And [11] proposes a new Lagrange multiplier approach to design unconditional energy stable schemes for gradient flows, which can keep the original energy dissipating. However, high order (higher than second order) backward difference formula (BDF) methods based either on the IEQ or the SAV formulations do not immediately lead to energy-decaying numerical schemes theoretically, even though they perform very well in numerical simulations. Recently the authors in [1] have break this barrier. They construct a class of extrapolated and linearized Runge–Kutta (RK) methods based on the SAV formulation which can be of arbitrarily high order in time for the AC and CH phase field equations. Also, they have proved that the schemes satisfy a discrete version of the energy decay property. More recently, highly efficient and accurate scheme just by BDFs based on a new SAV formulation is proposed in [23].

Moving the concentration from time discretizations to space discretizations, it is well-known that discontinuous Galerkin (DG) methods are very powerful. The DG method is a

Table 1 Example of the gradient flows

	Type of $\mathcal{G}$	$\mathcal{E}[u]$
Allen–Cahn (AC) equation	$\mathcal{G} = -\mathcal{I}$	$\int_\Omega \left(\frac{1}{2}\lvert\nabla u\rvert^2 + \frac{1}{4\varepsilon^2}(u^2-1)^2\right) \mathrm{d}x$
Cahn–Hilliard (CH) equation	$\mathcal{G} = \Delta$	$\int_\Omega \left(\frac{1}{2}\lvert\nabla u\rvert^2 + \frac{1}{4\varepsilon^2}(u^2-1)^2\right) \mathrm{d}x$
Thin film model	$\mathcal{G} = -\mathcal{I}$	$\int_\Omega \left(\frac{1}{2}\lvert\Delta u\rvert^2 + \frac{1}{2\varepsilon^2}\ln(1+\lvert\nabla u\rvert^2)\right) \mathrm{d}x$
Phase field crystals	$\mathcal{G} = \Delta$	$\int_\Omega \left(\frac{1}{2}\lvert\Delta u\rvert^2 - \lvert\nabla u\rvert^2 + \frac{1}{4}u^4 + \frac{1-\varepsilon}{2}u^2\right) \mathrm{d}x$

class of finite element methods, which uses completely discontinuous piecewise polynomials as basis to approximate solutions. The first DG methods [28] for hyperbolic equations were introduced by Reed and Hill in 1973, and the local discontinuous Galerkin (LDG) methods were introduced by Cockburn and Shu [15]. More references on theoretical analysis of DG/LDG can be found e.g. in [4, 7, 14].

In particular, the DG methods have been also successfully applied for the phase field problems, see, e.g., [21, 34, 39].

In this paper, we present a fully discrete scheme for a large class of phase-field gradient flows, which can be of arbitrarily high order both in time and space. The time discretization is based on the extrapolated RK–SAV method as in [1], and the space discretization is based on the DG method. We call the scheme as extrapolated RK–SAV/DG method for short. We show that the scheme satisfies a discrete energy decay property as

$$\mathcal{E}[u_{n+1}^h, r_{n+1}] \le \mathcal{E}[u_n^h, r_n],$$

where $\mathcal{E}[u_n^h, r_n]$ is the discrete (modified) energy of the numerical solution. Moreover, we show the optimal error estimates in space as well as the qth-order convergence rate in time of the proposed q-stage extrapolated RK–SAV/DG method for the AC and CH equations. In addition, we propose a technique to decouple the system, after which only several elliptic scalar problems with constant coefficients need to be solved independently.

The paper is organized as follows. In Sect. 2, we briefly recall the SAV reformulation, then we combine the extrapolated RK–SAV method and the DG method to obtain the fully discrete scheme for gradient flows and present the implementation process. The discrete energy decay property is proved in Sect. 3. Then we show the optimal error estimate for the AC and CH equations in Sect. 4. In Sect. 5, several numerical experiments are carried out to validate the theoretical results. Some concluding remarks are given in the finial section.

2 Extrapolated RK–SAV/DG Method

In this section, we present the extrapolated RK–SAV/DG method for gradient flows. It is started with rewriting (1.2) into SAV reformulation as in [31, 32]. Then we combine the extrapolated RK–SAV method [1] in temporal discretization and DG method in space discretization to obtain the fully discrete scheme for gradient flows, called as extrapolated RK–SAV/DG method. We provide the detailed implementation process in the last. Periodic boundary conditions, homogeneous Neumann boundary conditions or homogeneous Dirich-

let boundary conditions can be applied. Without loss of generality, we only consider the homogeneous Dirichlet boundary conditions in the analysis.

2.1 Extrapolated RK–SAV Method

The SAV approach first introduces a scalar function only depending on t as

$$r(t) := \sqrt{\mathcal{E}_1[u] + \mathcal{E}_0}, \quad 0 \le t \le T,$$

where $\mathcal{E}_0$ is a positive constant to keep $r(t)$ being real. Thus, the gradient flow (1.2) is rewritten as

$$\begin{cases} \dfrac{\partial u}{\partial t} = \mathcal{G}\mu, \\ \mu = \mathcal{L}u + rH(u), \\ r_t = \dfrac{1}{2}(H(u), u_t), \end{cases} \tag{2.1}$$

where

$$H(u) = \frac{1}{\sqrt{\mathcal{E}_1[u] + \mathcal{E}_0}} \frac{\delta \mathcal{E}_1}{\delta u}. \tag{2.2}$$

Note that the first- and second-order semi-implicit BDF methods schemes can be applied directly for the above system [32]. More importantly, it can be demonstrated that the energy-decay property is also preserved for the BDF methods. To obtain higher order temporal discretizations to preserve energy-decay property, we will use the extrapolated RK–SAV method. Let N be a positive integer and $t_n := n\tau$, $n = 0, \ldots, N$, be the uniform partition of the time interval $[0, T]$ with time stepsize $\tau := T/N$. Furthermore, let $t_{ni} := t_n + c_i\tau$, $i = 1, \ldots, q$, $n = 0, \ldots, N-1$, denote the inner Runge–Kutta nodes. Thus, the q-stage Runge–Kutta method, described by the Butcher tableau

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

Definition 1 We call a RK method is algebraically stable if for $i, j = 1, \ldots, q$

- The matrix $A = (a_{ij})$ is invertible,
- $b_i > 0$, $c_i \neq c_j$ for $i \neq j$,
- The symmetric matrix $M \in R^{q\times q}$ with $m_{ij} = b_i a_{ij} + b_j a_{ji} - b_i b_j$ is positive semidefinite.

Both Gauss type methods and Radau type IIA methods are algebraically stable Runge–Kutta methods, for further details of the tableau for Gauss and Radau IIA type methods, we refer to Chapter IV, section IV.5 in [38].

Given internal stages $u_{n-1,i}$, $i = 1, \ldots, q$, we denote by $u^\tau_{n-1}(t)$ the Lagrange interpolation polynomial of degree at most $q-1$ satisfying

$$u^\tau_{n-1}(t_{n-1,i}) = u_{n-1,i}, \quad i = 1, \ldots, q,$$

and use the abbreviation $I^\tau_{n-1}u_{ni} := u^\tau_{n-1}(t_{ni})$, which approximates $u(t_{ni})$ by the extrapolation method using the values $u_{n-1,i}$, $i = 1, \ldots, q$. Similarly, we denote by $I^\tau_{n-1}u(t)$ the

Lagrange interpolation polynomial in t of degree at most $q-1$ interpolating the exact solution u satisfying

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

Providing that the nodal approximations u_n, r_n and the internal stages $u_{n-1,i}$, $i = 1, \ldots, q$, are given, applying above q-stage RK to the SAV formulation (2.1) immediately yields

$$\begin{cases} \dot{u}_{ni} = \mathcal{G}\mu_{ni}, & \text{in } \Omega, \\ \mu_{ni} = \mathcal{L}u_{ni} + r_{ni} H(I_{n-1}^{\tau} u_{ni}), & \text{in } \Omega, \\ u_{ni} = u_n + \tau \sum_{j=1}^{q} a_{ij} \dot{u}_{nj}, & \text{in } \Omega, \end{cases} \tag{2.3}$$

$$\begin{cases} \dot{r}_{ni} = \frac{1}{2}(H(I_{n-1}^{\tau} u_{ni}), \dot{u}_{ni}), \\ r_{ni} = r_n + \tau \sum_{j=1}^{q} a_{ij} \dot{r}_{nj}, \end{cases} \tag{2.4}$$

and the numerical solution at time level t_{n+1} is given by

$$\begin{cases} u_{n+1} := u_n + \tau \sum_{i=1}^{q} b_i \dot{u}_{ni}, \\ r_{n+1} := r_n + \tau \sum_{i=1}^{q} b_i \dot{r}_{ni}. \end{cases} \tag{2.5}$$

Homogeneous Dirichlet boundary conditions are imposed for u_{ni} and μ_{ni}.

2.2 DG Methods

Next we briefly introduce the DG method. Let $\mathcal{T}_h = \{K\}$ be the triangulation of the domain Ω, and we assume the triangles K to be shape-regular. Thus, the discontinuous finite element space is defined by

$$V_h := \{v \in L^2(\Omega) : v|_K \in P(K), \quad \forall K \in \mathcal{T}_h\},$$
$$\Sigma_h := \{\boldsymbol{\sigma} \in [L^2(\Omega)]^d : \boldsymbol{\sigma}|_K \in \Sigma(K), \quad \forall K \in \mathcal{T}_h\},$$

where $P(K) = P_k(K)$ is the space of polynomial functions of degree at most $k \geq 1$ on K and $\Sigma(K) = [P_k(K)]^d$. Furthermore, we define the inner product notation as

$$(w, v)_K = \int_K w\, v \, \mathrm{d}K, \quad \langle w, v \rangle_{\partial K} = \int_{\partial K} w\, v \, \mathrm{d}s,$$
$$(\boldsymbol{q}, \boldsymbol{p})_K = \int_K \boldsymbol{q} \cdot \boldsymbol{p} \, \mathrm{d}K, \quad \langle \boldsymbol{q}, \boldsymbol{p} \rangle_{\partial K} = \int_{\partial K} \boldsymbol{q} \cdot \boldsymbol{p} \, \mathrm{d}s,$$

for scalar variables w, v and vector variables $\boldsymbol{q}$, $\boldsymbol{p}$, respectively. Naturally, the inner products on Ω are defined as

$$(w, v) := (w, v)_\Omega = \sum_K (w, v)_K, \quad (\boldsymbol{q}, \boldsymbol{p}) := (\boldsymbol{q}, \boldsymbol{p})_\Omega = \sum_K (\boldsymbol{q}, \boldsymbol{p})_K.$$

And the L^2 norm, L^∞ norm on the domain Ω and the boundary Γ are given by the standard definitions:

$$\|u\| := \|u\|_\Omega = \sqrt{(u, u)_\Omega}, \quad \|u\|_{L^\infty(\Omega)} = \operatorname{ess\,sup}_{x \in \Omega} |u|, \quad \|u\|_\Gamma = \sqrt{(u, u)_\Gamma}.$$

Then we consider the model problem:

$$\begin{cases} -\Delta u = f, & \text{in } \Omega, \\ u = 0, & \text{on } \partial\Omega. \end{cases} \tag{2.6}$$

Table 2 Overview of numerical flux choices

	u_h^*	$\boldsymbol{\pi}_h^*$
Central flux	$\{\{u_h\}\}$	$\{\{\boldsymbol{\pi}_h\}\} - \gamma \left[\left[u_h\right]\right]$
Local DG flux (LDG)	$\{\{u_h\}\} + \boldsymbol{\beta} \cdot \left[\left[u_h\right]\right]$	$\{\{\boldsymbol{\pi}_h\}\} - \boldsymbol{\beta} \left[\left[\boldsymbol{\pi}_h\right]\right] - \gamma \left[\left[u_h\right]\right]$
Interior penalty flux (IP)	$\{\{u_h\}\}$	$\{\{\nabla u_h\}\} - \gamma \left[\left[u_h\right]\right]$

Let us rewrite the problem as a first-order system:

$$\begin{cases} \boldsymbol{\pi} = \nabla u, & \text{in } \Omega, \\ -\nabla \cdot \boldsymbol{p}\boldsymbol{\pi} = f, & \text{in } \Omega, \\ u = 0, & \text{on } \partial\Omega. \end{cases}$$

Following Cockburn and Shu [15], we consider the following general formulation: finding $u_h \in V_h$ and $\boldsymbol{\pi}_h \in \Sigma_h$ such that for all $K \in \mathcal{T}_h$

$$\begin{cases} (\boldsymbol{\pi}_h, \boldsymbol{\sigma})_K = -(u_h, \nabla \cdot \boldsymbol{p}\boldsymbol{\sigma})_K + \left\langle u_K^*, \boldsymbol{n}_K \cdot \boldsymbol{p}\boldsymbol{\sigma} \right\rangle_{\partial K} \quad \forall \boldsymbol{\sigma} \in \Sigma(K), \\ (\boldsymbol{\pi}_h, \nabla \phi_h)_K = (f, \phi_h)_K + \left\langle \boldsymbol{\pi}_K^* \cdot \boldsymbol{p}\boldsymbol{n}_K, \phi_h \right\rangle_{\partial K} \quad \forall \phi_h \in P(K), \end{cases} \tag{2.7}$$

where the $\boldsymbol{n}_K$ is the outward normal unit vector to ∂K and the numerical fluxes u_K^* and $\boldsymbol{\pi}_K^*$ are approximations to u and $\boldsymbol{\pi}$, respectively, on the boundary of K. Some choices of the flux is present in Table 2. The standard notation is taken as

$$\{\{u\}\} = \frac{u^- + u^+}{2},$$

where u can be either a scalar or a vector. The jumps along a normal $\boldsymbol{n}$ is defined as

$$[[u]] = \boldsymbol{n}^- u^- + \boldsymbol{n}^+ u^+, \ [[\boldsymbol{u}]] = \boldsymbol{n}^- \cdot \boldsymbol{u}^- + \boldsymbol{n}^+ \cdot \boldsymbol{u}^+.$$

Note that there are several alternatives to these three options for the numerical fluxes and we refer to [4] for a complete discussion of these.

In (2.7), summing up all the elements and eliminating the auxiliary variable $\boldsymbol{\pi}_h$, we derive the following formulation:

$$\mathcal{B}_h(u_h, \phi_h) = (f, \phi_h), \qquad \forall \phi_h \in V_h, \tag{2.8}$$

where

$$\begin{aligned} \mathcal{B}_h(u_h, \phi_h) = (\nabla u_h, \nabla \phi_h) &- \oint_\Gamma \left(\left[\left[u_h - u_h^*\right]\right] \cdot \{\{\nabla \phi_h\}\} + \left\{\left\{\boldsymbol{\pi}_h^*\right\}\right\} \cdot [[\phi_h]] \right) \mathrm{d}x \\ &- \oint_{\Gamma_i} \left([[\nabla \phi_h]] \left\{\left\{u_h - u_h^*\right\}\right\} + \left[\left[\boldsymbol{\pi}_h^*\right]\right] \{\{\phi_h\}\} \right) \mathrm{d}x, \end{aligned}$$

with Γ and Γ_i represents the set of unique edges and the set of unique purely internal edges, respectively. We call (2.8) as the primal formulation of the method with the bilinear form $\mathcal{B}_h(\cdot \boldsymbol{p}, \cdot \boldsymbol{p})$. For details of the formulation we refer to [4, 22].

With the above definitions, we can simplify the notations by defining the discrete Laplacian $\Delta_h : V_h \to V_h$ as

$$(\Delta_h u_h, \phi_h) = -\mathcal{B}_h(u_h, \phi_h), \quad \forall \phi_h \in V_h,$$

and the L^2 projection $\Pi_h : L^2 \to V_h$ as

$$(\Pi_h H(I_{n-1}^\tau u_{ni}^h), \phi_h) = (H(I_{n-1}^\tau u_{ni}^h), \phi_h), \quad \forall \phi_h \in V_h,$$

and let $H_{ni} := \Pi_h H(I_{n-1}^{\tau} u_{ni}^h)$. Hence, $\mathcal{G}$ and $\mathcal{L}$ can be defined in the discrete case, corresponding to the non-positive symmetric operate $\mathcal{G}_h$ and the non-negative symmetric operate $\mathcal{L}_h$, respectively.

2.3 Fully Discrete RK–SAV/DG Scheme and Decoupled Technique

Now it is ready to introduce the fully discrete extrapolated RK–SAV/DG scheme. For the system (2.3)–(2.5), the fully discrete scheme reads as for $i = 1, \ldots, q$

$$\begin{cases} \dot{u}_{ni}^h = \mathcal{G}_h \mu_{ni}^h, \\ \mu_{ni}^h = \mathcal{L}_h u_{ni}^h + r_{ni} H_{ni}, \\ u_{ni}^h = u_n^h + \tau \sum_{j=1}^q a_{ij} \dot{u}_{nj}^h, \end{cases} \tag{2.9}$$

$$\begin{cases} \dot{r}_{ni} = \frac{1}{2} \left(H_{ni}, \dot{u}_{ni}^h \right), \\ r_{ni} = r_n + \tau \sum_{j=1}^q a_{ij} \dot{r}_{nj}, \end{cases} \tag{2.10}$$

and updating u_{n+1}^h and r_{n+1} by

$$\begin{cases} u_{n+1}^h := u_n^h + \tau \sum_{i=1}^q b_i \dot{u}_{ni}^h, \\ r_{n+1} := r_n + \tau \sum_{i=1}^q b_i \dot{r}_{ni}. \end{cases} \tag{2.11}$$

Homogeneous Dirichlet boundary conditions are imposed for u_{ni}^h and μ_{ni}^h.

Note that we need to solve coupled the system (2.9)–(2.11). To enhance the efficiency, in the following, we intend to decouple the system. Denote $U_n := (u_{n1}^h, \ldots, u_{nq}^h)^T$, $R_n := (r_{n1}, \ldots, r_{nq})^T$, $\mathbf{1} := (1, \ldots, 1)^T$, and I as $q \times q$ identity matrix. From (2.9) we have

$$(\tau^{-1} A^{-1} - \mathcal{G}_h \mathcal{L}_h I) U_n = u_n^h \tau^{-1} A^{-1} \mathbf{1} + B_1 R_n, \tag{2.12}$$

where B_1 is the diagonal matrix-valued function

$$B_1 = \mathcal{G}_h \mathrm{diag}(H_{n1}, \ldots, H_{nq}).$$

Now we decompose $\tau^{-1} A^{-1} = T^{-1} \Lambda T$, where Λ is a $q \times q$ diagonal matrix, and T is nonsingular. The above equation (2.12) can be represented as

$$\begin{aligned} T^{-1}(\Lambda - \mathcal{G}_h \mathcal{L}_h I) T U_n &= u_n^h \tau^{-1} A^{-1} \mathbf{1} + B_1 R_n, \\ (\Lambda - \mathcal{G}_h \mathcal{L}_h I) T U_n &= T u_n^h \tau^{-1} A^{-1} \mathbf{1} + T B_1 R_n. \end{aligned}$$

Then we compute $T U_n$ by

$$T U_n = (\Lambda - \mathcal{G}_h \mathcal{L}_h I)^{-1} T u_n^h \tau^{-1} A^{-1} \mathbf{1} + (\Lambda - \mathcal{G}_h \mathcal{L}_h I)^{-1} T B_1 R_n,$$

and get U_n

$$U_n = T^{-1}(\Lambda - \mathcal{G}_h \mathcal{L}_h I)^{-1} T u_n^h \tau^{-1} A^{-1} \mathbf{1} + T^{-1}(\Lambda - \mathcal{G}_h \mathcal{L}_h I)^{-1} T B_1 R_n. \tag{2.13}$$

For simplicity, we denote $g_n = T^{-1}(\Lambda - \mathcal{G}_h \mathcal{L}_h I)^{-1} T u_n^h \tau^{-1} A^{-1} \mathbf{1}$, $B_n = T^{-1}(\Lambda - \mathcal{G}_h \mathcal{L}_h I)^{-1} T B_1$, and rewrite (2.13) as

$$U_n = g_n + B_n R_n. \tag{2.14}$$

Using the third relation of (2.9) gives

$$\dot{U}_n = \tau^{-1} A^{-1}(g_n - u_n^h \mathbf{1}) + \tau^{-1} A^{-1} B_n R_n. \tag{2.15}$$

Substituting (2.15) to the first relation of (2.10) yields

$$\dot{R}_n = \tau^{-1} f_n + \tau^{-1} F_n R_n, \tag{2.16}$$

where

$$f_n = \frac{1}{2}\left(\left(H_{n1}, A^{-1}(g_n - u_n^h \mathbf{1})(1)\right), \ldots, \left(H_{nq}, A^{-1}(g_n - u_n^h \mathbf{1})(q)\right) \right)^T$$

and

$$F_n = \frac{1}{2}(F_{ij}), \quad F_{ij} = \left(H_{ni}, (A^{-1}B_n)_{ij}\right).$$

Substituting (2.16) into the second relation of (2.10), it follows

$$R_n = r_n \mathbf{1} + A f_n + A F_n R_n. \tag{2.17}$$

Consequently, we finally derive

$$R_n = (I - A F_n)^{-1}(r_n \mathbf{1} + A f_n).$$

Now we have R_n, then from (2.14), (2.15) and (2.16), we can get U_n, $\dot{U}_n$ and $\dot{R}_n$ instantly, and update u_{n+1}^h and r_{n+1} by (2.11).

In summary, we can decouple the scheme (2.9)–(2.11) as the following way:

1. Compute g_n and B_n from (2.12)–(2.13).
2. Compute f_n and F_n from (2.16).
3. Compute R_n from (2.17) and get U_n, $\dot{U}_n$ and $\dot{R}_n$ from (2.14), (2.15) and (2.16). Then update u_{n+1}^h and r_{n+1} by (2.11).

Remark 1 In the above, we treat $\mathcal{G}_h$ and $\mathcal{L}_h$ as operators. When computing g_n and B_n from (2.12)–(2.13), we only need to solve a few equations like $(\lambda - \mathcal{G}_h\mathcal{L}_h)x = b$, which can be solved efficiently, instead of solving the system $(\tau^{-1}A^{-1} - \mathcal{G}_h\mathcal{L}_h I)\mathbf{x} = \mathbf{b}$. And the coupled system (2.9)–(2.11) has a unique solution (U_n, R_n) which is obviously seen by these decoupled elliptic equations.

3 Energy Decay of the Extrapolated RK–SAV/DG Method

In this section we will show that the extrapolated RK–SAV/DG method of the system (2.9)–(2.11) preserves the discrete energy decay property. We denote the discrete energy of the numerical solution at t_n by

$$\mathcal{E}[u_n^h, r_n] = \frac{1}{2}(u_n^h, \mathcal{L}_h u_n^h) + r_n^2 - \mathcal{E}_0,$$

also referred as the modified energy, since in general, $r_n^2 - \mathcal{E}_0$ does not coincide with $\mathcal{E}_1(u_n^h)$.

Remark 2 $(u_n^h, \mathcal{L}_h u_n^h)$ is defined by $\mathcal{B}_h$, for example if $\mathcal{L} = -\Delta + \beta I$, with β being a positive constant, then $(u_n^h, \mathcal{L}_h u_n^h) = \mathcal{B}_h(u_n^h, u_n^h) + \beta(u_n^h, u_n^h)$. Thus we have $(u_n^h, \mathcal{L}_h u_n^h) \geq 0$. Here, the constant β can be viewed as a stabilization term to produce more regular solutions. More discussions about β can be found in [32].

Theorem 1 (Discrete energy decay property) *Let the Runge–Kutta method be algebraically stable, i.e., satisfy Definition 1, and assume that the values $u_{n-1,i}$, $i = 1, \dots q$, and (u_n, r_n) are given. Then the extrapolated RK–SAV/DG method (2.9)–(2.11) preserves the energy decay property in sense of*

$$\mathcal{E}[u_{n+1}^h, r_{n+1}] \le \mathcal{E}[u_n^h, r_n]. \tag{3.1}$$

Proof According to the first relation of (2.11), we have

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

since $(u_n^h, \mathcal{L}_h \dot{u}_{ni}^h) = (\mathcal{L}_h u_n^h, \dot{u}_{ni}^h)$. Substituting $u_n^h = u_{ni}^h - \tau \sum_{j=1}^q a_{ij} \dot{u}_{nj}^h$ (the third relation in (2.9)) into the second term on the right-hand side of the last relation gives

$$\begin{aligned}\left(u_{n+1}^h, \mathcal{L}_h u_{n+1}^h\right) = \left(u_n^h, u_n^h\right) &+ 2\tau \sum_{i=1}^q b_i \left(u_{ni}^h - \tau \sum_{j=1}^q a_{ij} \dot{u}_{nj}^h, \mathcal{L}_h \dot{u}_{ni}^h\right) \\ &+ \tau^2 \sum_{i,j=1}^q b_i b_j \left(\dot{u}_{ni}^h, \mathcal{L}_h \dot{u}_{nj}^h\right).\end{aligned}$$

Hence,

$$\left(u_{n+1}^h, \mathcal{L}_h u_{n+1}^h\right) = \left(u_n^h, \mathcal{L}_h u_n^h\right) + 2\tau \sum_{i=1}^q b_i \left(u_{ni}^h, \mathcal{L}_h \dot{u}_{ni}^h\right) - \tau^2 \sum_{i,j=1}^q m_{ij} \left(\dot{u}_{ni}^h, \mathcal{L}_h \dot{u}_{nj}^h\right),$$

with $m_{ij} = b_i a_{ij} + b_i a_{ji} - b_i b_j,\ i, j = 1, \dots, q$. Using the positive semi-definiteness of matrix $M = (m_{ij})$ and $(\dot{u}_{ni}^h, \mathcal{L}_h \dot{u}_{nj}^h) \ge 0$ yields

$$\left(u_{n+1}^h, \mathcal{L}_h u_{n+1}^h\right) \le \left(u_n^h, \mathcal{L}_h u_n^h\right) + 2\tau \sum_{i=1}^q b_i \left(u_{ni}^h, \mathcal{L}_h \dot{u}_{ni}^h\right). \tag{3.2}$$

Similarly, we can obtain

$$r_{n+1}^2 \le r_n^2 + \tau \sum_{i=1}^q b_i r_{ni} \left(H_{ni}, \dot{u}_{ni}^h\right). \tag{3.3}$$

From the first and second relation in (2.9), we can derive that

$$(\mu_{ni}^h, \dot{u}_{ni}^h) = \left(\mu_{ni}^h, \mathcal{G}_h \mu_{ni}^h\right), \tag{3.4}$$

$$(\mu_{ni}^h, \dot{u}_{ni}^h) = \left(u_n^h, \mathcal{L}_h \dot{u}_{ni}^h\right) + r_{ni} \left(H_{ni}, \dot{u}_{ni}^h\right). \tag{3.5}$$

Combining with (3.2)–(3.5) yields

$$\frac{1}{2}\left(u_{n+1}^h, \mathcal{L}_h u_{n+1}^h\right) + r_{n+1}^2 \le \frac{1}{2}\left(u_n^h, \mathcal{L}_h u_n^h\right) + r_n^2 + \tau \sum_{i=1}^q b_i \left(\mu_{ni}^h, \mathcal{G}_h \mu_{ni}^h\right).$$

This completes the proof of (3.1), since $b_i > 0$ for $i = 1, \dots, q$, and $(\mu_{ni}^h, \mathcal{G}_h \mu_{ni}^h) \le 0$. □

4 Error Analysis for the Extrapolated RK–SAV/DG Method

In this section, we follow [1] with the elliptic projection technique [26, 29, 35] to establish the optimal error estimates of the extrapolated RK–SAV/DG method for the AC and CH equations. We assume that the values $u^h(t_{0i}), i = 1, \ldots, q$, $u^h(t_1)$ and $r(t_1)$ have been given or approximated accurately.

4.1 Preliminaries

First, the local truncation errors ε_{ni}, ε_{n+1}, d_{ni} and d_{n+1} for the semi-discrete extrapolated RK/SAV scheme are defined by for $i = 1, \ldots, q$,

$$\begin{cases} \dot{u}^\star_{ni} = \mathcal{G}\mu^\star_{ni}, \\ \mu^\star_{ni} = -\Delta u^\star_{ni} + r^\star_{ni} H\left(I^\tau_{n-1} u^\star_{ni}\right), \\ u^\star_{ni} = u^\star_n + \tau \sum_{j=1}^q a_{ij} \dot{u}^\star_{nj} + \varepsilon_{ni}, \end{cases} \tag{4.1}$$

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

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

where $H(u)$ is defined by (2.2), and

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

Lemma 1 (Consistency estimate, [1]) *If the exact solutions u and r of AC equation are sufficiently smooth, then the following consistency estimate holds*

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

In space discretization, following [4, 9, 22], we have some knowledge about the bilinear form $\mathcal{B}_h(u_h, \phi_h)$ and the error estimates of the elliptic problem (2.6):

- Coercivity:

$$\mathcal{B}_h(\phi_h, \phi_h) \ge C_c \|\phi_h\|^2_{\mathrm{DG}}, \quad \forall \phi_h \in V_h,$$

with the natural DG norm :

$$\|\phi\|^2_{\mathrm{DG}} = \|\nabla\phi\|^2 + \|h^{-1/2}[[\phi]]\|^2_{\Gamma_i} + \|h^{-1/2}\phi\|^2_{\Gamma_b}, \quad \Gamma_b = \Gamma/\Gamma_i.$$

- Continuity:

$$\mathcal{B}_h(u_h, \phi_h) \le C_k \|u_h\|_{\mathrm{DG}} \|\phi_h\|_{\mathrm{DG}}, \quad \forall \phi_h \in V_h.$$

- Galerkin orthogonality: let u be the smooth solution satisfies (2.6), and all numerical fluxes in Table 2 are consistent, from which

$$\mathcal{B}_h(u, \phi_h) = (f, \phi_h), \quad \forall \phi_h \in V_h,$$

then we have the Galerkin orthogonality

$$\mathcal{B}_h(u - u_h, \phi_h) = 0, \quad \forall \phi_h \in V_h.$$

- Error estimate of the elliptic problem (2.6): supposing the solution u of (2.6) is sufficient smooth, u_h is the numerical solution of (2.8), then there are the error estimates as follows

$$\|u - u_h\| \le Ch^{k+1}, \quad \|u - u_h\|_{\mathrm{DG}} \le Ch^k, \tag{4.5}$$

where C only depends on $\|u\|_{H^{k+1}(\Omega)}$, and $\|\cdot \boldsymbol{p}\|_{H^n(\Omega)}$ is the Sobolev norm [13].

4.2 Error Estimate

Theorem 2 (Error estimate for the AC equation) *We assume that Runge–Kutta method is algebraically stable with $q \ge 2$, and that the following conditions hold*:

- *the exact solution of the AC equation is sufficiently smooth*;
- *the starting approximations (u_{0i}^h, r_{0i}) are sufficiently accurate such that*

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

for some constant C_0 independent of τ, h;
- *for $i = 1, \ldots, q$, starting approximations satisfy $\left\|u(t_{0i}) - u_{0i}^h\right\|_{L^\infty(\Omega)} \le 1$.*

Then the discrete solution to (2.9)–(2.11) for the AC equation satisfies the following error estimate with sufficiently small time step τ and mesh grid size h:

$$\max_{1\le n\le N-1} \left(\left\|u(t_{n+1}) - u_{n+1}^h\right\|^2 + |r(t_{n+1}) - r_{n+1}|^2 + \tau \sum_{i=1}^{q} \left(\left\|u(t_{ni}) - u_{ni}^h\right\|^2 + |r(t_{ni}) - r_{ni})|^2 \right) \right) \le C(\tau^{2q} + h^{2k+2}). \tag{4.6}$$

Proof We first define the elliptic projection $\mathrm{P}u \in V_h$ of a smooth solution u:

$$\mathcal{B}_h\left(u - \mathrm{P}u, \phi_h\right) = 0, \quad \forall \phi_h \in V_h,$$

and from (4.5) we have $\|u - \mathrm{P}u\| \le Ch^{k+1}$.

For the AC equation $\mathcal{G} = -\mathcal{I}$, we have from (4.1)–(4.3) that for any $\phi_n^h \in V_h$, and $i = 1, \ldots, q$,

$$\begin{cases} \left(\dot{e}_{ni}, \phi_n^h\right) = -\mathcal{B}_h\left(e_{ni}, \phi_n^h\right) - \left(\eta_{ni} H\left(I_{n-1}^\tau u_{ni}^h\right) + r_{ni}^\star \left(H(I_{n-1}^\tau u_{ni}^\star) - H(I_{n-1}^\tau u_{ni}^h)\right), \phi_n^h\right), \\ e_{ni} = e_n + \tau \sum_{j=1}^q a_{ij} \dot{e}_{nj} + \varepsilon_{ni}, \end{cases} \tag{4.7}$$

$$\begin{cases} \dot{\eta}_{ni} = \dfrac{1}{2}\left(H(I_{n-1}^\tau u_{ni}^\star) - H(I_{n-1}^\tau u_{ni}^h), \dot{u}_{ni}^\star\right) + \dfrac{1}{2}\left(H(I_{n-1}^\tau u_{ni}^h), \dot{e}_{ni}\right), \\ \eta_{ni} = \eta_n + \tau \sum_{j=1}^q a_{ij} \dot{\eta}_{nj} + d_{ni}, \end{cases} \tag{4.8}$$

$$\begin{cases} e_{n+1} := e_n + \tau \sum_{i=1}^q b_i \dot{e}_{ni} + \varepsilon_{n+1}, \\ \eta_{n+1} := e_n + \tau \sum_{i=1}^q b_i \dot{\eta}_{ni} + d_{n+1}, \end{cases} \tag{4.9}$$

with the boundary conditions $e_{ni} = 0$, and the following notations:

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

Furthermore, we have

$$e_n = u_n^\star - \mathrm{P}u_n^\star + \mathrm{P}e_n, \quad e_{ni} = u_{ni}^\star - \mathrm{P}u_{ni}^\star + \mathrm{P}e_{ni}, \quad \dot{e}_{ni} = \dot{u}_{ni}^\star - \mathrm{P}\dot{u}_{ni}^\star + \mathrm{P}e_{ni}.$$

Let $1 \le m \le N$. In the following, we assume that for $n \le m$ the error function satisfies

$$\|e_{n-1,i}\|_{L^\infty(\Omega)} \le 1, \quad i = 1, \ldots, q. \tag{4.10}$$

We will prove that above inequality holds also for $n = m + 1$ by mathematical induction.

From the first relation of (4.9) we have

$$\mathrm{P}e_{n+1} = \mathrm{P}e_n + \tau \sum_{i=1}^{q} b_i \mathrm{P}\dot{e}_{ni} + \mathrm{P}\varepsilon_{n+1}. \tag{4.11}$$

Taking the square of L^2-norm of both sides of (4.11) gives

$$\|\mathrm{P}e_{n+1}\|^2 = \left\| \mathrm{P}e_n + \tau \sum_{i=1}^{q} b_i \mathrm{P}\dot{e}_{ni} \right\|^2 + 2\left(\mathrm{P}\varepsilon_{n+1}, \mathrm{P}e_n + \tau \sum_{i=1}^{q} b_i \mathrm{P}\dot{e}_{ni} \right) + \|\mathrm{P}\varepsilon_{n+1}\|^2. \tag{4.12}$$

Next we approximate the first two terms on the right-hand side of (4.12). For the first term, we find

$$\left\| \mathrm{P}e_n + \tau \sum_{i=1}^{q} b_i \mathrm{P}\dot{e}_{ni} \right\|^2 = \|\mathrm{P}e_n\|^2 + 2\tau \sum_{i=1}^{q} b_i (\mathrm{P}\dot{e}_{ni}, \mathrm{P}e_n) + \tau^2 \sum_{i,j=1}^{q} b_i b_j (\mathrm{P}\dot{e}_{ni}, \mathrm{P}\dot{e}_{nj}). \tag{4.13}$$

Noticing that from the second relation of (4.7), there is

$$\mathrm{P}e_{ni} = \mathrm{P}e_n + \tau \sum_{j=1}^{q} a_{ij} \mathrm{P}\dot{e}_{nj} + \mathrm{P}\varepsilon_{ni}. \tag{4.14}$$

Replacing $\mathrm{P}e_n$ in the second term of (4.13) by $\mathrm{P}e_{ni} - \tau \sum_{j=1}^{q} a_{ij} \mathrm{P}\dot{e}_{nj} - \mathrm{P}\varepsilon_{ni}$ gives

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

Using the positive semidefinite of the matrix M yields

$$\left\| \mathrm{P}e_n + \tau \sum_{i=1}^{q} b_i \mathrm{P}\dot{e}_{ni} \right\|^2 \le \|\mathrm{P}e_n\|^2 + 2\tau \sum_{i=1}^{q} b_i (\mathrm{P}\dot{e}_{ni}, \mathrm{P}e_{ni} - \mathrm{P}\varepsilon_{ni}). \tag{4.15}$$

To estimate $(\mathrm{P}\dot{e}_{ni}, \mathrm{P}e_{ni})$, inserting $\phi_n^h = \mathrm{P}e_{ni}$ into the first relation of (4.7) gives

$$(\dot{e}_{ni}, \mathrm{P}e_{ni}) = -\mathcal{B}_h(e_{ni}, \mathrm{P}e_{ni}) - \left(\eta_{ni} H\left(I_{n-1}^\tau u_{ni}^h\right) + r_{ni}^\star \left(H(I_{n-1}^\tau u_{ni}^\star) - H(I_{n-1}^\tau u_{ni}^h) \right), \mathrm{P}e_{ni} \right).$$

Applying Cauchy-Schwarz inequality, there is

$$\begin{aligned}
(\mathrm{P}\dot{e}_{ni}, \mathrm{P}e_{ni}) \le & -(\dot{u}_{ni}^\star - \mathrm{P}\dot{u}_{ni}^\star, \mathrm{P}e_{ni}) - \mathcal{B}_h(\mathrm{P}e_{ni}, \mathrm{P}e_{ni}) \\
& + C|\eta_{ni}|^2 + C \max_{1\le i\le q} \|e_{n-1,i}\|^2 + \frac{1}{4}\|\mathrm{P}e_{ni}\|^2 \\
\le & Ch^{2k+2} - \mathcal{B}_h(\mathrm{P}e_{ni}, \mathrm{P}e_{ni}) + C|\eta_{ni}|^2 + C \max_{1\le i\le q} \|e_{n-1,i}\|^2 + \frac{1}{2}\|\mathrm{P}e_{ni}\|^2,
\end{aligned}$$

which gives

$$\begin{aligned}&(\mathrm{P}\dot{e}_{ni}, \mathrm{P}e_{ni}) + \mathcal{B}_h(\mathrm{P}e_{ni}, \mathrm{P}e_{ni})\\ &\le \frac{1}{2}\|\mathrm{P}e_{ni}\|^2 + C|\eta_{ni}|^2 + C\max_{1\le i\le q}\|e_{n-1,i}\|^2 + \frac{1}{2}\|\mathrm{P}e_{ni}\|^2 + Ch^{2k+2},\end{aligned}$$

where we have used the following inequality from [1], when (4.10) holds

$$\left\|\eta_{ni}H(I_{n-1}^{\tau}u_{ni}^{h}) + r_{ni}^{\star}\left(H(I_{n-1}^{\tau}u_{ni}^{\star}) - H(I_{n-1}^{\tau}u_{ni}^{h})\right)\right\| \le C|\eta_{ni}| + C\max_{1\le i\le q}\|e_{n-1,i}\|,$$

and the Galerkin orthogonality and elliptic projection:

$$\mathcal{B}_h(\mathrm{P}e_{ni}, \mathrm{P}e_{ni}) = \mathcal{B}_h(e_{ni}, \mathrm{P}e_{ni}), \qquad \|\dot{u}_{ni}^{\star} - \mathrm{P}\dot{u}_{ni}^{\star}\| \le Ch^{k+1}.$$

Applying Cauchy-Schwarz inequality for $\tau\sum_{i=1}^{q} b_i\,(\mathrm{P}\dot{e}_{ni}, -\mathrm{P}\varepsilon_{ni})$ yields

$$-\tau\sum_{i=1}^{q} b_i\,(\mathrm{P}\dot{e}_{ni}, \mathrm{P}\varepsilon_{ni}) \le C\tau(\|\tau\mathrm{P}\dot{e}_{ni}\|^2 + \|\mathrm{P}\varepsilon_{ni}/\tau\|^2).$$

It is apparent from Lemma 1 that

$$\|\mathrm{P}\varepsilon_{ni}/\tau\| = \tau^q\|\mathrm{P}\varepsilon_{ni}/\tau^{q+1}\|^2 \le \tau^q(\|\varepsilon_{ni}/\tau^{q+1}\| + \|\varepsilon_{ni}/\tau^{q+1} - \mathrm{P}\varepsilon_{ni}/\tau^{q+1}\|) \le C\tau^q.$$

And from (4.11), there is

$$\|\tau\mathrm{P}\dot{e}_{ni}\|^2 \le C\|\mathrm{P}e_n\|^2 + C\|\mathrm{P}e_{n+1}\|^2 + C\tau^{2q+2} \tag{4.16}$$

so we can derive

$$-\tau\sum_{i=1}^{q} b_i\,(\mathrm{P}\dot{e}_{ni}, \mathrm{P}\varepsilon_{ni}) \le C\tau\|\mathrm{P}e_{n+1}\|^2 + C\tau\|\mathrm{P}e_n\|^2 + C\tau^{2q+1}.$$

Using the estimates of $(\mathrm{P}\dot{e}_{ni}, \mathrm{P}e_{ni})$ and $\tau\sum_{i=1}^{q} b_i\,(\mathrm{P}\dot{e}_{ni}, -\mathrm{P}\varepsilon_{ni})$, we can rewrite (4.15) into

$$\begin{aligned}&\left\|\mathrm{P}e_n + \tau\sum_{i=1}^{q} b_i\mathrm{P}\dot{e}_{ni}\right\|^2 + \tau\sum_{i=1}^{q} b_i\mathcal{B}_h(\mathrm{P}e_{ni}, \mathrm{P}e_{ni}) - C\tau\|\mathrm{P}e_{n+1}\|^2\\ &\le (1 + C\tau)\|\mathrm{P}e_n\|^2 + \tau\sum_{i=1}^{q} b_i(\|\mathrm{P}e_{ni}\|^2 + C|\eta_{ni}|^2)\\ &\quad + C\tau\max_{1\le i\le q}\|e_{n-1,i}\|^2 + C\tau(h^{2k+2} + \tau^{2q}).\end{aligned} \tag{4.17}$$

Now we estimate $2\left(\mathrm{P}\varepsilon_{n+1}, \mathrm{P}e_n + \tau\sum_{i=1}^{q} b_i\mathrm{P}\dot{e}_{ni}\right) + \|\mathrm{P}\varepsilon_{n+1}\|^2$, which is the remaining term of the right hand side (4.12). From (4.11) we can find

$$2(\mathrm{P}\varepsilon_{n+1}, \mathrm{P}e_n + \tau\sum_{i=1}^{q} b_i\mathrm{P}\dot{e}_{ni}) + \|\mathrm{P}\varepsilon_{n+1}\|^2 = 2(\mathrm{P}e_{n+1}, \mathrm{P}\varepsilon_{n+1}) - \|\mathrm{P}\varepsilon_{n+1}\|^2.$$

By Cauchy-Schwarz inequality, we find

$$(\mathrm{P}e_{n+1}, \mathrm{P}\varepsilon_{n+1}) \le \tau\|\mathrm{P}e_{n+1}\|\|\mathrm{P}\varepsilon_{n+1}/\tau\| \le C\tau(\|\mathrm{P}e_{n+1}\|^2 + \|\mathrm{P}\varepsilon_{n+1}/\tau\|^2).$$

Thus there is the following estimate

$$2(\mathrm{P}\varepsilon_{n+1}, \mathrm{P}e_n + \tau \sum_{i=1}^{q} b_i \mathrm{P}\dot{e}_{ni}) + \|\mathrm{P}\varepsilon_{n+1}\|^2 \le C\tau \|\mathrm{P}e_{n+1}\|^2 + C\tau^{2q+1}. \tag{4.18}$$

Combining (4.12) with (4.17) and (4.18) gives

$$\begin{aligned} &\|\mathrm{P}e_{n+1}\|^2 + \frac{\tau}{2} \sum_{i=1}^{q} b_i \mathcal{B}_h(\mathrm{P}e_{ni}, \mathrm{P}e_{ni}) \\ &\le (1 + C\tau)\|\mathrm{P}e_n\|^2 + C\tau \sum_{i=1}^{q} \left(\|\mathrm{P}e_{ni}\|^2 + |\eta_{ni}|^2 + \|e_{n-1,i}\|^2\right) \\ &+ C\tau(h^{2k+2} + \tau^{2q}) \end{aligned} \tag{4.19}$$

Similarly, from (4.8) we can infer that

$$\begin{aligned} |\eta_{n+1}|^2 \le &(1 + C\tau)|\eta_n|^2 + \frac{\tau}{4} \sum_{i=1}^{q} b_i \mathcal{B}_h(\mathrm{P}e_{ni}, \mathrm{P}e_{ni}) \\ &+ C\tau \sum_{i=1}^{q} \left(\|\mathrm{P}e_{ni}\|^2 + |\eta_{ni}|^2 + \|e_{n-1,i}\|^2\right) + C\tau(h^{2k+2} + \tau^{2q}). \end{aligned} \tag{4.20}$$

It is apparent from (4.19) and (4.20) that

$$\begin{aligned} &\|\mathrm{P}e_{n+1}\|^2 + |\eta_{n+1}|^2 + \frac{\tau}{4} \sum_{i=1}^{q} b_i \mathcal{B}_h(\mathrm{P}e_{ni}, \mathrm{P}e_{ni}) \\ &\le (1 + C\tau)\left(\|\mathrm{P}e_n\|^2 + |\eta_n|^2\right) + C\tau \sum_{i=1}^{q} \left(\|\mathrm{P}e_{ni}\|^2 + |\eta_{ni}|^2 + \|e_{n-1,i}\|^2\right) \\ &+ C\tau(h^{2k+2} + \tau^{2q}). \end{aligned} \tag{4.21}$$

We now approximate the term $C\tau \sum_{i=1}^{q} (\|\mathrm{P}e_{ni}\|^2 + |\eta_{ni}|^2)$ on the right-hand side. To this end, multiplying (4.14) by $\mathrm{P}e_{ni}$ and using the Cauchy-Schwarz inequality yield

$$\sum_{i=1}^{q} \|\mathrm{P}e_{ni}\|^2 \le C\|\mathrm{P}e_n\|^2 + C\tau \sum_{i,j=1}^{q} a_{ij}(\mathrm{P}\dot{e}_{nj}, \mathrm{P}e_{ni}) + C \sum_{i=1}^{q} \|\mathrm{P}\varepsilon_{ni}\|^2. \tag{4.22}$$

Similarly, substituting $\phi_n^h = \mathrm{P}e_{nj}$ in the first relation of (4.7) gives

$$\begin{aligned} (\mathrm{P}\dot{e}_{ni}, \mathrm{P}e_{nj}) \le &\frac{1}{2}\mathcal{B}_h(\mathrm{P}e_{ni}, \mathrm{P}e_{ni}) + \frac{1}{2}\mathcal{B}_h(\mathrm{P}e_{nj}, \mathrm{P}e_{nj}) + \frac{1}{2}\|\mathrm{P}e_{ni}\|^2 + C|\eta_{ni}|^2 \\ &+ C \max_{1 \le i \le q} \|e_{n-1,i}\|^2 + \frac{1}{2}\|\mathrm{P}e_{nj}\|^2 + Ch^{2k+2}, \end{aligned}$$

which leads to

$$\begin{aligned} &\sum_{i,j=1}^{q} a_{ij}(\mathrm{P}\dot{e}_{nj}, \mathrm{P}e_{ni}) \\ &\le C \sum_{i=1}^{q} \left(\mathcal{B}_h(\mathrm{P}e_{ni}, \mathrm{P}e_{ni}) + \|\mathrm{P}e_{ni}\|^2 + |\eta_{ni}|^2 + \|e_{n-1,i}\|^2\right) + Ch^{2k+2}. \end{aligned} \tag{4.23}$$

Combining (4.23) with (4.22) and (4.16) leads to

$$\sum_{i=1}^{q}\|\mathrm{P}e_{ni}\| \le C\|\mathrm{P}e_n\|^2 + C\tau\sum_{i=1}^{q}\left(\mathcal{B}_h(\mathrm{P}e_{ni},\mathrm{P}e_{ni}) + \|\mathrm{P}e_{ni}\|^2 + |\eta_{ni}|^2 + \|e_{n-1,i}\|^2\right) + C\tau(h^{2k+2}+\tau^{2q}). \tag{4.24}$$

Similarly, from (4.8) we can prove

$$\sum_{i=1}^{q}|\eta_{ni}|^2 \le C|\eta_n|^2 + C\tau\sum_{i=1}^{q}\left(\mathcal{B}_h(\mathrm{P}e_{ni},\mathrm{P}e_{ni}) + \|\mathrm{P}e_{ni}\|^2 + |\eta_{ni}|^2 + \|e_{n-1,i}\|^2\right) + C\tau(h^{2k+2}+\tau^{2q}). \tag{4.25}$$

Summing up (4.24)–(4.25) and noting that the term $C\tau\sum_{i=1}^{q}(\|\mathrm{P}e_{ni}\|^2+|\eta_{ni}|^2)$ on the right-hand side can be absorbed by the left-hand side, for sufficiently small τ. It follows that

$$\sum_{i=1}^{q}\left(\|\mathrm{P}e_{ni}\|^2+|\eta_{ni}|^2\right) \le C\left(\|\mathrm{P}e_n\|^2+|\eta_n|^2\right) + C\tau\sum_{i=1}^{q}b_i\mathcal{B}_h(\mathrm{P}e_{ni},\mathrm{P}e_{ni}) + C\tau\sum_{i=1}^{q}\|e_{n-1,i}\|^2 + C\tau(h^{2k+2}+\tau^{2q}), \tag{4.26}$$

where we use the positivity of the weights $b_1,\ldots,b_q$. Substituting this inequality into (4.21) yields

$$\begin{aligned}&\|\mathrm{P}e_{n+1}\|^2+|\eta_{n+1}|^2+\frac{\tau}{4}\sum_{i=1}^{q}b_i\mathcal{B}_h(\mathrm{P}e_{ni},\mathrm{P}e_{ni})\\ &\le (1+C_1\tau)\left(\|\mathrm{P}e_n\|^2+|\eta_n|^2\right)+C_1\tau^2\sum_{i=1}^{q}b_i\mathcal{B}_h(\mathrm{P}e_{ni},\mathrm{P}e_{ni})\\ &\quad+C_1\tau\sum_{i=1}^{q}\|e_{n-1,i}\|^2+C_1\tau(h^{2k+2}+\tau^{2q}),\end{aligned}$$

with some constant C_1. Multiplying (4.26) by $2C_1\tau$ and adding to the above inequality gives

$$\begin{aligned}&\left(\|\mathrm{P}e_{n+1}\|^2+|\eta_{n+1}|^2+2C_1\tau\sum_{i=1}^{q}(\|\mathrm{P}e_{ni}\|^2+|\eta_{ni}|^2)\right)+\frac{\tau}{4}\sum_{i=1}^{q}b_i\mathcal{B}_h(\mathrm{P}e_{ni},\mathrm{P}e_{ni})\\ &\le(1+C_2\tau)\left(\|\mathrm{P}e_n\|^2+|\eta_n|^2\right)+C_2\tau^2\sum_{i=1}^{q}b_i\mathcal{B}_h(\mathrm{P}e_{ni},\mathrm{P}e_{ni})\\ &\quad+(C_1+C_2\tau)\tau\sum_{i=1}^{q}\|\mathrm{P}e_{n-1,i}\|^2+C_2\tau(h^{2k+2}+\tau^{2q}),\end{aligned}$$

where the elliptic projection property $\|e_{n-1,i}\| \le \|\mathrm{P}e_{n-1,i}\| + Ch^{k+1}$ is applied and the constant C_2 is independent of τ. The term $C_2\tau^2\sum_{i=1}^{q}b_i\mathcal{B}_h(\mathrm{P}e_{ni},\mathrm{P}e_{ni})$ can be absorbed by the left-hand side with sufficiently small τ if $C_1+C_2\tau \le 2C_1$. Therefore, the inequality above reduces to

$$\left(\|\mathrm{P}e_{n+1}\|^2+|\eta_{n+1}|^2+2C_1\tau\sum_{i=1}^{q}(\|\mathrm{P}e_{ni}\|^2+|\eta_{ni}|^2)\right)+\frac{\tau}{8}\sum_{i=1}^{q}b_i\mathcal{B}_h(\mathrm{P}e_{ni},\mathrm{P}e_{ni})$$

$$\leq (1+(2C_1+C_2)\tau)\left[\|\mathrm{P}e_n\|^2+|\eta_n|^2+2C_1\tau\sum_{i=1}^{q}(\|\mathrm{P}e_{n-1,i}\|^2+|\eta_{n-1,i}|^2)\right]$$
$$+C_2\tau(h^{2k+2}+\tau^{2q}).$$

By Gronwall's inequality, we obtain

$$\max_{1\leq n\leq m}\left(\|\mathrm{P}e_{n+1}\|^2+|\eta_{n+1}|^2+2C_1\tau\sum_{i=1}^{q}(\|\mathrm{P}e_{ni}\|^2+|\eta_{ni}|^2)\right)+\frac{\tau}{8}\sum_{i=1}^{q}b_i\mathcal{B}_h(\mathrm{P}e_{ni},\mathrm{P}e_{ni})$$
$$\leq C\left[\|\mathrm{P}e_1\|^2+|\eta_1|^2+2C_1\tau\sum_{i=1}^{q}(\|\mathrm{P}e_{0i}\|^2+|\eta_{0i}|^2)\right]+C(h^{2k+2}+\tau^{2q}).$$

Using the elliptic projection property of P yields

$$\|e_{n+1}\|\leq\|u^{\star}_{n+1}-\mathrm{P}u^{\star}_{n+1}\|+\|\mathrm{P}e_{n+1}\|\leq Ch^{k+1}+\|\mathrm{P}e_{n+1}\|.$$

It is easy to see that

$$\|e_{m+1}\|+|\eta_{m+1}|\leq C(h^{k+1}+\tau^q),$$
$$\|e_{mi}\|+|\eta_{mi}|\leq C\tau^{-\frac{1}{2}}(\tau^q+h^{k+1}),$$

Moreover, using inverse inequality gives

$$\|e_{mi}\|_{L^\infty(\Omega)}\leq Ch^{-\frac{d}{2}}\|e_{mi}\|\leq Ch^{-\frac{d}{2}}\tau^{-\frac{1}{2}}(\tau^q+h^{k+1}).$$

By noting that $1\leq d\leq 3$, $q\geq 2$ and $k\geq 1$, we then obtain

$$\|e_{mi}\|_{L^\infty(\Omega)}\leq 1,$$

for sufficiently small τ and h (the smallness is independent of m). This completes the mathematical induction on (4.10). Consequently, the error estimate (4.6) follows as (4.51) and (4.51) hold for all $1\leq m\leq N-1$. □

Remark 3 One can also use the LDG method for space discretization and the optimal error estimate in space can be achieved with Cartesian meshes similar as in [21, 27]. Again, the energy decay property of can be preserved by using the LDG method for both Cartesian meshes and triangle mesh. One major advantage of the LDG method is that it can be implemented easily for the case with variable coefficient like $\mathcal{G}\mu=\nabla\cdot(\phi\nabla\mu)$.

Theorem 3 (Error estimate for the CH equation) *We assume that Runge–Kutta method is algebraically stable with $q\geq 2$, and that the following conditions hold*:

- *the exact solution of the CH equation is sufficiently smooth*;
- *the starting approximations (u^h_{0i}, r_{0i}) are sufficiently accurate such that*

$$\left\|u(t_1)-u^h_1\right\|^2_{\mathrm{B}}+|r(t_1)-r_1|^2+\tau\sum_{i=1}^{q}\left(\left\|u(t_{0i})-u^h_{0i}\right\|^2_{\mathrm{B}}+|r(t_{0i})-r_{0i})|^2\right)$$
$$\leq C_0(\tau^{2q}+h^{2k}),$$

  for some constant C_0 independent of τ, h;
- *for $i=1,\ldots,q$, starting approximations satisfy $\left\|u(t_{0i})-u^h_{0i}\right\|_{L^\infty(\Omega)}\leq 1$.*

Then the discrete solution to (2.9)–(2.11) for the AC equation satisfies the following error estimate with sufficiently small time step τ and mesh grid size h:

$$\max_{1\le n\le N-1}\left(\left\|u(t_{n+1})-u^h_{n+1}\right\|^2_{\mathrm{B}}+|r(t_{n+1})-r_{n+1}|^2 +\tau\sum_{i=1}^{q}\left(\left\|u(t_{ni})-u^h_{ni}\right\|^2_{\mathrm{B}}+|r(t_{ni})-r_{ni})|^2\right)\right)\le C(\tau^{2q}+h^{2k}),$$

where the energy-norm $\|\cdot\|_{\mathrm{B}}:=\sqrt{\mathcal{B}_h(\cdot,\cdot)}$ is equivalent to the semi-H^1 norm for smooth functions.

The basic idea and technique involved for the proof of the above theorem are very similar as ones in Theorem 2. The main difference is to replace the $\|\cdot\|$ by the energy norm $\|\cdot\|_{\mathrm{B}}$. Thus, the optimal L^2 estimate for the CH equation can be obtained by Sobolev inequality. A specific point is given below: as the error function of the CH equation is similar to (4.7)–(4.9), we just need to rewrite the first relation of (4.7) in the following form

$$(\dot e_{ni},\psi^h_n)=-\mathcal{B}_h(e^\mu_{ni},\psi^h_n),$$
$$(e^\mu_{ni},\phi^h_n)=\mathcal{B}_h\left(e_{ni},\phi^h_n\right)+\left(\eta_{ni}H\left(I^\tau_{n-1}u^h_{ni}\right)+r^\star_{ni}\left(H(I^\tau_{n-1}u^\star_{ni})-H(I^\tau_{n-1}u^h_{ni})\right),\phi^h_n\right),$$

where the $e^\mu_{ni}:=\mu^\star_{ni}-\mu^h_{ni}$. Comparing to Theorem 2, the critical difficulty in the proof of Theorem 3 is to bound the term $\mathcal{B}_h(\mathrm{P}\dot e_{ni},\mathrm{P}e_{ni})$. Similar as in [26], we use Lax-Milgram theorem to define an invertible operator $\mathcal{J}:V_h\to V_h$ as

$$\mathcal{B}_h(\phi_h,\mathcal{J}\psi_h)=(\phi_h,\psi_h),\ \ \forall\phi_h\in V_h.$$

Then we take test function $\psi^h_n=\mathcal{J}\mathrm{P}\dot e_{ni}$ and $\phi^h_n=\mathrm{P}\dot e_{ni}$ to get the estimate of $\mathcal{B}_h(\mathrm{P}\dot e_{ni},\mathrm{P}e_{ni})$.

We close this section by pointing out that in Theorem 3 the error estimate is optimal in sense of the energy norm.

5 Numerical Examples

In this section, we present several numerical examples to illustrate the accuracy and the energy decay property of the extrapolated RK–SAV/DG method for the AC and the CH equations. In the following, we choose Gauss type RK methods for temporal discretization and the interior penalty (IP) flux, which leads to IPDG for the space discretization.

We first take $\mathcal{L}=-\Delta+\dfrac{\beta}{\varepsilon^2}$ and

$$\mathcal{E}_1[u]=-\frac{\beta}{2\varepsilon^2}(u,u)+\int_\Omega\frac{1}{4\varepsilon^2}(u^2-1)^2\,\mathrm{d}x,$$

in the original energy (1.1). Note that this modification does not affect the implementation process and the error analysis described above.

Example 1 We consider the following one-dimensional AC equation

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

with two types of initial values and boundary conditions as follows

$$\text{Case I}:\ \ u_0(x)=\sin(x)\ \text{with}\ \ u=0\ \text{on}\ \partial\Omega,\tag{5.1}$$

Table 3 (Example 1) $\max_{1\le k\le n}\|u(t_n)-u^n\|_{L^2(\Omega)}$ for the Gauss type RK3 by 4-point extrapolation for Dirichlet boundary condition Case I and Neumann boundary condition Case II, respectively

	$\tau = 1/40$	$\tau = 1/60$	$\tau = 1/80$	$\tau = 1/100$	$\tau = 1/120$
Case I	3.8841e-08	6.6875e-09	1.9359e-09	7.6193e-10	3.5904e-10
	–	4.3388	4.3090	4.1790	4.1268
Case II	3.8835e-08	6.6879e-09	1.9372e-09	7.5448e-10	3.5387e-10
	–	4.3383	4.3069	4.2260	4.1524

$$\text{Case II:}\quad u_0(x)=\cos(x)\ \text{with}\ \partial_{\boldsymbol{n}}u=0\ \text{on}\ \partial\Omega. \tag{5.2}$$

We first choose $\varepsilon^2=1$ and test the time accuracy for the numerical methods. The source terms $g(x,t)$ are chosen such that the exact solutions for (5.1)–(5.2) are respectively

$$\text{Case I: } u(x,t)=e^{-t}\sin(x),\quad \text{and}\quad \text{Case II: } u(x,t)=e^{-t}\cos(x).$$

Here we take $\beta=0,\mathcal{E}_0=0$ and implement the DG method with polynomial degree $k=3$ and mesh size $h=2\pi/360$. The L^2 errors of the numerical solutions at time $T=1$ are demonstrated in Fig. 1. Optimal convergence rates in time are observed. As mentioned in [1, 41], if we use $(q+1)$-point extrapolation, with an internal node of the Runge–Kutta method as an additional interpolation point for the nonlinear term, then the scheme is of $(q+1)$th order in the time for AC equation. And we give a numerical test in Table 3.

Secondly, we set $g(x,t)=0$ and consider the same initial value

$$u_0(x)=0.5\sin(x)+0.1*\mathrm{Rand}(x), \tag{5.3}$$

with homogeneous Dirichlet and Neumann boundary conditions respectively. We illustrate the energy decay property of the 3-stage extrapolated RK–SAV/DG method in Fig. 1. Here we take $\beta=0,\mathcal{E}_0=0,\varepsilon^2=0.01$, and implement the DG method with polynomial degree $k=2$ and mesh size $h=\pi/20$. The first stages $u_{0i}^h, i=1,\dots,q$ and u_1^h are derived by ETDRK4 method [24]. The time step is $\tau=1/400$ and the final time is $T=0.1$. The numerical results show that the discrete energy decays, which are consistent with our theoretical result in Theorem 1.

Example 2 We consider the following one-dimensional CH equation

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

with periodic boundary condition and the initial value as follows

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

We first choose $\varepsilon^2=1$ and test the accuracy of the numerical methods by choosing the source term $g(x,t)$ such that the exact solution is

$$u(x,t)=e^{-t}\sin(x).$$

We test the 2-stage extrapolated RK–SAV/DG method for time accuracy. Here, we take $\beta=0,\mathcal{E}_0=100$ and implement the DG method with polynomial degree $k=3$ and mesh size $h=2\pi/80$. The L^2 errors of the numerical solutions at time $T=1$ are presented in Fig. 2. To test the space accuracy, we choose $\beta=2,\mathcal{E}_0=1000$ and use the time step $\tau=10^{-4}$

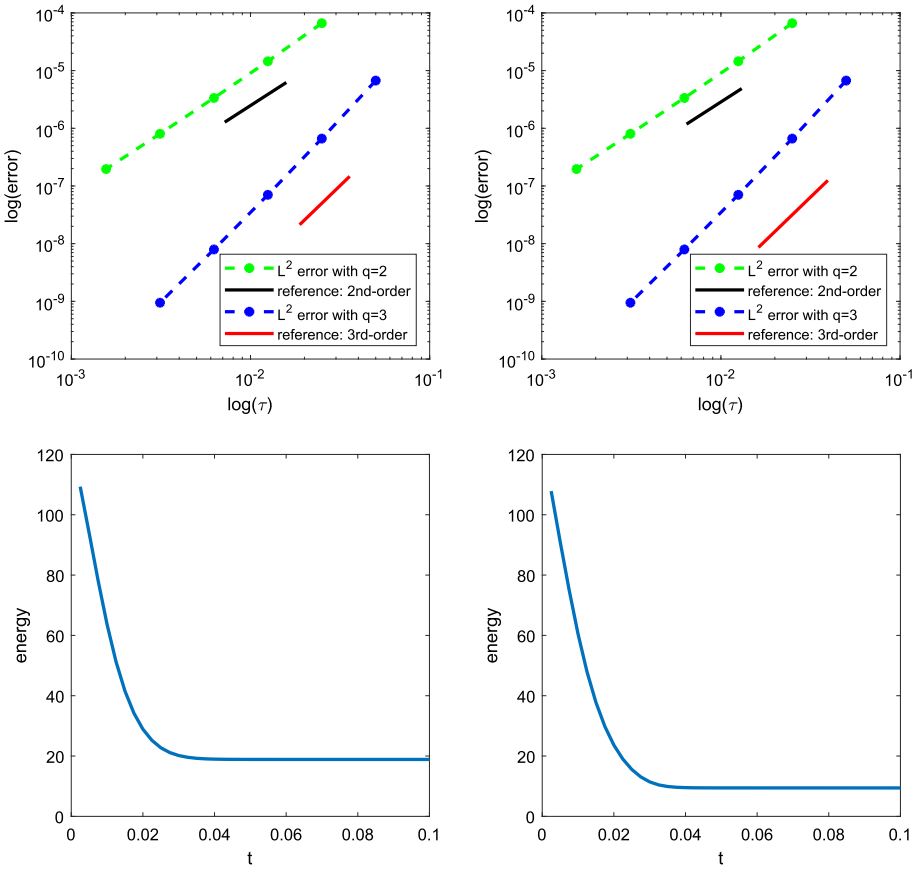

Fig. 1 (Example 1) 1D AC equation. Upper: error of time discretization, (left) Dirichlet boundary condition; (right) Neumann boundary condition. Bottom: discrete energy evolution, (left) Dirichlet boundary condition; (right) Neumann boundary condition

with 3-stage extrapolated RK–SAV/DG method and apply the DG method with polynomial degree $k = 2$. The L^2 errors of the numerical solutions at time $T = 0.3$ are presented in Fig. 2. Again, optimal convergence rates in both time and space are observed.

In the second part, we set $g(x, t) = 0$ and consider the initial value as in (5.3)

with periodic boundary conditions. The energy decay property of the 2-stage extrapolated RK–SAV/DG method is demonstrated in Fig. 2. The first stages u_{0i}^h $i = 1, \ldots, q$ and u_1^h are derived by ETDRK4 method. Here we set $\varepsilon^2 = 0.01$, $\beta = 2$, $\mathcal{E}_0 = 1000$ and implement the DG method with polynomial degree $k = 2$ and mesh size $h = \pi/20$. The time step is $\tau = 1/500$ and the final time is $T = 0.1$.

Example 3 We consider the following two-dimensional AC equation

$$\partial_t u - \Delta u + \frac{1}{\varepsilon^2}(u^3 - u) = g(x, y, t), \quad x, y, t \in (0, 2\pi)^2 \times (0, 1), \tag{5.4}$$

here we just consider Dirichlet boundary condition and the initial value give by

$$u_0(x, y) = \sin(x)\sin(y), \quad \text{with } u = 0 \quad \text{on } \partial\Omega.$$

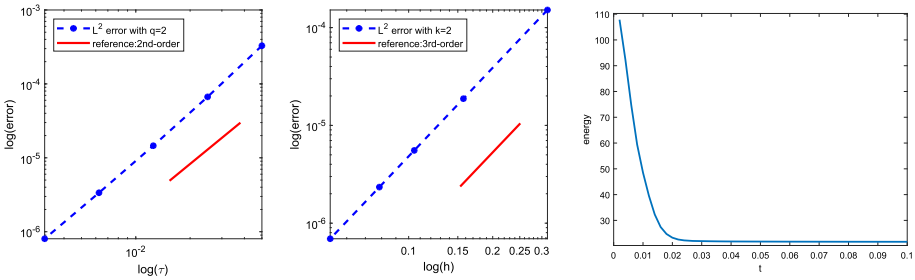

Fig. 2 (Example 2) 1D CH equation: left: time accuracy; middle: space accuracy; right: discrete energy evolution

Table 4 (Example 3) space accuracy for 2D AC equation with $k = 3$

h	$\pi/10$	$\pi/15$	$\pi/20$	$\pi/30$	$\pi/40$
L^2- error	2.4858e-05	4.8768e-06	1.5389e-06	3.0330e-07	9.5880e-08
order	–	4.0169	4.0094	4.0054	4.0031

We first set $\varepsilon^2 = 1$ and test the accuracy of the numerical method by choosing the source term $g(x, y, t)$ such that the exact solution is

$$u(x, y, t) = e^{-t} \sin(x) \sin(y).$$

We set $\beta = 0, \mathcal{E}_0 = 0$ and use the time step $\tau = 10^{-3}$ with 3-stage extrapolated RK–SAV/DG method and test space accuracy by applying the DG method with polynomial degree $k = 3$. The L^2 errors of the numerical solution at time $T = 0.5$ are presented in Table 4, which are consistent with the analysis. In this example, we omit the time accuracy test.

To show the energy decay property, we set $g = 0$ in Eq. (5.4) and consider the initial value and boundary condition given by

$$u = 0.4 * \text{Rand}(x, y) + 0.25, \quad \text{with } u = 0 \quad \text{on } \partial\Omega.$$

We present the discrete energy decay property for the 2-stage extrapolated RK–SAV/DG method in Fig. 3 . Here we set $\varepsilon^2 = 0.01$, $\beta = 0$, $\mathcal{E}_0 = 0$ and implement the DG method with polynomial degree $k = 2$ and mesh size $h = \pi/10$. The time step is $\tau = 1/500$ and the final time is $T = 0.1$.

We close this section by pointing out that although all numerical examples are carried out by the Gauss type RK methods other relevant RK methods such as the Radau IIA type methods also work well.

6 Conclusion

In this work, we developed an extrapolated RK–SAV/DG method for solving phase-field problems. The novelty of this method is that the corresponding fully discrete system requires only solving a system of linear equations at each time level. A strategy is proposed to decouple the system, which results in solving several scalar elliptic problems independently. We have proved the optimal error bound of the fully discrete scheme for the AC and CH equations.

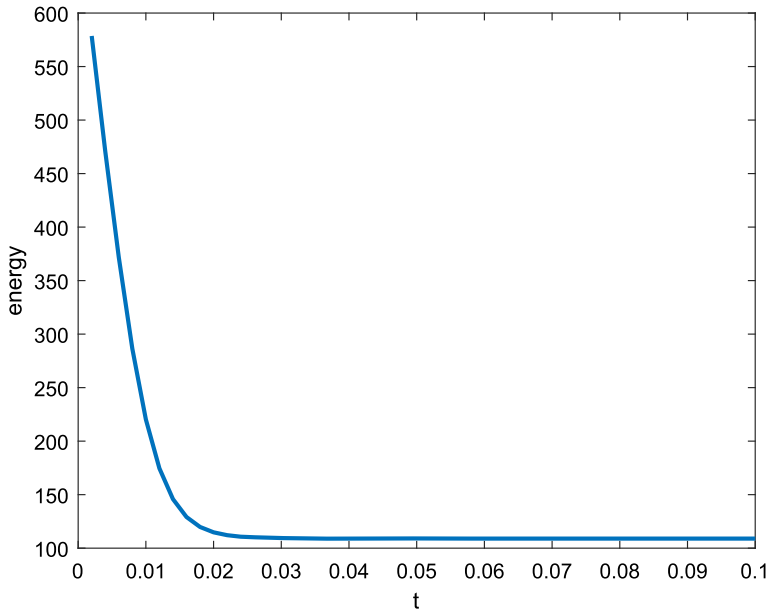

Fig. 3 (Example 3) Discrete energy for 2D AC equation with $\varepsilon^2 = 0.01,\ \tau = 1/500$

Moreover, it can be shown that the scheme preserves the discrete energy decay property at all time levels.

On the other hand, as mentioned in [1], if we use $(q+1)$-point extrapolation, with an internal node of the Runge–Kutta method as an additional interpolation point for the nonlinear term, then the scheme is of $(q+1)$th order in the time for AC equation. We test the 4th-order by using Gauss type RK3 with the 4-point extrapolation.

Acknowledgements This work is partially supported by the National Science Foundation of China and Hong Kong RGC Joint Research Scheme (NSFC/RGC 11961160718), and the fund of the Guangdong Provincial Key Laboratory of Computational Science and Material Design (No. 2019B030301001). The work of J. Yang is supported by the National Science Foundation of China (NSFC-11871264) and the Shenzhen Natural Science Fund (RCJC20210609103819018).

References

1. Akrivis, G., Li, B., Li, D.: Energy-decaying extrapolated RK–SAV methods for the Allen–Cahn and Cahn–Hilliard equations. SIAM J. Sci. Comput. **41**(6), A3703–A3727 (2019)
2. Allen, S.M., Cahn, J.W.: A microscopic theory for antiphase boundary motion and its application to antiphase domain coarsening. Acta Metall. **27**(6), 1085–1095 (1979)
3. Anderson, D.M., McFadden, G.B., Wheeler, A.A.: Diffuse-interface methods in fluid mechanics. Ann. Rev. Fluid Mech. **30**(1), 139–165 (1998)
4. Arnold, D.N., Brezzi, F., Cockburn, B., Marini, L.D.: Unified analysis of discontinuous Galerkin methods for elliptic problems. SIAM J. Numer. Anal. **39**(5), 1749–1779 (2002)
5. Boettinger, W.J., Warren, J.A., Beckermann, C., Karma, A.: Phase-field simulation of solidification. Ann. Rev. Mater. Res. **32**(1), 163–194 (2002)
6. Braun, R.J., Murray, B.T.: Adaptive phase-field computations of dendritic crystal growth. J. Cryst. Growth **174**(1–4), 41–53 (1997)

7. Brezzi, F., Manzini, G., Marini, D., Pietra, P., Russo, A.: Discontinuous Galerkin approximations for elliptic problems. Numer. Methods Partial Differ. Equ. Int. J. **16**(4), 365–378 (2000)
8. Cahn, J.W., Hilliard, J.E.: Free energy of a nonuniform system. I. Interfacial free energy. J. Chem. Phys. **28**(2), 258–267 (1958)
9. Castillo, P., Cockburn, B., Perugia, I., Schötzau, D.: An a priori error analysis of the local discontinuous Galerkin method for elliptic problems. SIAM J. Numer. Anal. **38**(5), 1676–1706 (2000)
10. Chen, L.-Q.: Phase-field models for microstructure evolution. Ann. Rev. Mater. Res. **32**(1), 113–140 (2002)
11. Cheng, Q., Liu, C., Shen, J.: A new Lagrange Multiplier approach for gradient flows. Comput. Methods Appl. Mech. Eng. **367**, 113070 (2020)
12. Cheng, Q., Shen, J.: Multiple scalar auxiliary variable (MSAV) approach and its application to the phase-field vesicle membrane model. SIAM J. Sci. Comput. **40**(6), A3982–A4006 (2018)
13. Ciarlet, P.G.: The Finite Element Method for Elliptic Problems. SIAM, Philadelphia (2002)
14. Cockburn, B., Kanschat, G., Perugia, I., Schötzau, D.: Superconvergence of the local discontinuous Galerkin method for elliptic problems on cartesian grids. SIAM J. Numer. Anal. **39**(1), 264–285 (2001)
15. Cockburn, B., Shu, C.-W.: The local discontinuous Galerkin method for time-dependent convection-diffusion systems. SIAM J. Numer. Anal. **35**(6), 2440–2463 (1998)
16. Du, Q., Ju, L., Li, X., Qiao, Z.: Maximum bound principles for a class of semilinear parabolic equations and exponential time-differencing schemes. SIAM Rev. **63**(2), 317–359 (2021)
17. Du, Q., Feng, X.: The Phase Field Method for Geometric Moving Interfaces and Their Numerical Approximations. Handbook of Numerical Analysis, vol. 21, pp. 425–508. Elsevier, Amsterdam (2020)
18. Feng, X., Tang, T., Yang, J.: Long time numerical simulations for phase-field problems using p-adaptive spectral deferred correction methods. SIAM J. Sci. Comput. **37**(1), A271–A294 (2015)
19. Frieboes, H.B., Jin, F., Chuang, Y.-L., Wise, S.M., Lowengrub, J.S., Cristini, V.: Three-dimensional multispecies nonlinear tumor growth-II: tumor invasion and angiogenesis. J. Theor. Biol. **264**(4), 1254–1278 (2010)
20. Fu, Z., Yang, J.: Energy-decreasing exponential time differencing Runge–Kutta methods for phase-field models. J. Comput. Phys. **454**, 110943 (2022)
21. Guo, R., Ji, L., Xu, Y.: High order local discontinuous Galerkin methods for the Allen–Cahn equation: analysis and simulation. J. Comput. Math. **34**(2), 135–158 (2016)
22. Hesthaven, J.S., Warburton, T.: Nodal Discontinuous Galerkin Methods: Algorithms, Analysis, and Applications. Springer, Berlin (2007)
23. Huang, F., Shen, J., Yang, Z.: A highly efficient and accurate new scalar auxiliary variable approach for gradient flows. SIAM J. Sci. Comput. **42**(4), A2514–A2536 (2020)
24. Kassam, A.-K., Trefethen, L.N.: Fourth-order time-stepping for stiff PDEs. SIAM J. Sci. Comput. **26**(4), 1214–1233 (2005)
25. Li, D., Quan, C., Xu, J.: Stability and convergence of Strang splitting. Part I: Scalar Allen–Cahn equation. J. Comput. Phys. **458**(6), 111087 (2022)
26. Liu, C., Frank, F., Rivière, B.M.: Numerical error analysis for nonsymmetric interior penalty discontinuous Galerkin method of Cahn–Hilliard equation. Numer. Methods Partial Differ. Equ. **35**(4), 1509–1537 (2019)
27. Liu, H., Yan, J.: A local discontinuous Galerkin method for the Korteweg–de Vries equation with boundary effect. J. Comput. Phys. **215**(1), 197–218 (2006)
28. Reed, W.H., Hill, T.R.: Triangular Mesh Methods for the Neutron Transport Equation, Tech. report, Los Alamos Scientific Lab., N. Mex. (USA), (1973)
29. Riviere, B.: Discontinuous Galerkin Methods for Solving Elliptic and Parabolic Equations: Theory and Implementation. SIAM, Philadelphia (2008)
30. Shen, J., Wang, C., Wang, X., Wise, S.M.: Second-order convex splitting schemes for gradient flows with Ehrlich–Schwoebel type energy: application to thin film epitaxy. SIAM J. Numer. Anal. **50**, 105–125 (2012)
31. Shen, J., Xu, J., Yang, J.: The scalar auxiliary variable (SAV) approach for gradient flows. J. Comput. Phys. **353**, 407–416 (2018)
32. Shen, J., Xu, J., Yang, J.: A new class of efficient and robust energy stable schemes for gradient flows. SIAM Rev. **61**(3), 474–506 (2019)
33. Shen, J., Yang, X.: Numerical approximations of Allen–Cahn and Cahn–Hilliard equations. Discret. Contin. Dyn. Syst. **28**, 1669–1691 (2010)
34. Song, H., Shu, C.-W.: Unconditional energy stability analysis of a second order implicit-explicit local discontinuous Galerkin method for the Cahn-Hilliard equation. J. Sci. Comput. **73**(2–3), 1178–1203 (2017)
35. Thomée, V.: Galerkin Finite Element Methods for Parabolic Problems, vol. 1054. Springer, Berlin (1984)

36. Villain, J.: Continuum models of crystal growth from atomic beams with and without desorption. Journal de Physique I **1**(1), 19–42 (1991)
37. Wang, Y.U., Jin, Y.M., Khachaturyan, A.G.: Phase field microelasticity modeling of dislocation dynamics near free surface and in heteroepitaxial thin films. Acta Materialia **51**(14), 4209–4223 (2003)
38. Wanner, G., Hairer, E.: Solving Ordinary Differential Equations II. Springer, Berlin (1996)
39. Xia, Y., Yan, X., Shu, C.-W.: Application of the local discontinuous Galerkin method for the Allen–Cahn/Cahn–Hilliard system. Commun. Comput. Phys. **5**, 821–835 (2009)
40. Xu, C., Tang, T.: Stability analysis of large time-stepping methods for epitaxial growth models. SIAM J. Numer. Anal. **44**, 1759–1779 (2006)
41. Yang, J., Yuan, Z., Zhou, Z.: Arbitrarily high-order maximum bound preserving schemes with cut-off postprocessing for Allen–Cahn equations. J. Sci. Comput. **90**(2), 1–36 (2022)
42. Yang, X.: Linear, first and second-order, unconditionally energy stable numerical schemes for the phase field model of homopolymer blends. J. Comput. Phys. **327**, 294–316 (2016)
43. Yang, X., Zhao, J., Wang, Q., Shen, J.: Numerical approximations for a three-component Cahn–Hilliard phase-field model based on the invariant energy quadratization method. Math. Models Methods Appl. Sci. **27**(11), 1993–2030 (2017)
44. Yue, P., Feng, J.J., Liu, C., Shen, J.: A diffuse-interface method for simulating two-phase flows of complex fluids. J. Fluid Mech. **515**, 293 (2004)