SIAM J. NUMER. ANAL.
Vol. 56, No. 5, pp. 2895–2912

CONVERGENCE AND ERROR ANALYSIS FOR THE SCALAR AUXILIARY VARIABLE (SAV) SCHEMES TO GRADIENT FLOWS*

JIE SHEN† AND JIE XU‡

Abstract. We carry out convergence and error analysis of the scalar auxiliary variable (SAV) methods for L^2 and H^{-1} gradient flows with a typical form of free energy. We first derive H^2 bounds, under certain assumptions suitable for both the gradient flows and the SAV schemes, which allow us to establish the convergence of the SAV schemes under mild conditions. We then derive error estimates with further regularity assumptions. We also discuss several other gradient flows, which cannot be cast in the general framework used in this paper, but for which convergence and error analysis can still be established using a similar procedure.

Key words. convergence and error analysis, gradient flows, energy stability, Allen–Cahn, Cahn–Hilliard

AMS subject classifications. 65M12, 35K20, 35K35, 35K55, 65Z05

DOI. 10.1137/17M1159968

1. Introduction.

The main goal of this paper is to conduct convergence and error analysis for the recently proposed scalar auxiliary variable (SAV) approach [17, 18]. The SAV approach is proposed for a large class of gradient flows that describe energy dissipative physical systems [1, 4, 8, 14, 6, 3]. The schemes for gradient flows that introduce auxiliary variables (or Lagrange multipliers) are probably first proposed in [12] for fourth-order polynomial double-well free energies, and then generalized to other free energies and known as the invariant energy quadratization (IEQ) approach (cf. [22] and many works afterwards). Instead of introducing an auxiliary function in the IEQ approach, the SAV approach introduces an auxiliary scalar, which leads to numerical schemes that enjoy the following remarkable properties:

- second-order unconditionally energy stable, and easily extendable to higher-order (though not unconditionally stable);
- only requires solving linear, decoupled systems with constant coefficients at each time step so it is easy to implement and extremely efficient;
- only requires that the free energy functional is bounded from below so it applies to a large class of gradient flows (cf. [18] for some examples).

Ample numerical evidences presented in [17, 18] have shown that the SAV schemes are superior to the commonly used schemes for gradient flows such as convex splitting, stabilized semi-implicit, and IEQ methods. However, only the energy stability, which is usually not sufficient for the convergence, has been proved for the SAV schemes. In particular, since the stability is proved only for a modified energy, it is essential to carry out a convergence and error analysis for the SAV approach to ensure that the SAV schemes do converge to the correct solutions at the expected rates.

Although some convergence and error analyses are available for fully implicit

*Received by the editors December 5, 2017; accepted for publication (in revised form) June 27, 2018; published electronically September 25, 2018.
http://www.siam.org/journals/sinum/56-5/M115996.html
Funding: The work of the authors was supported in part by NSF DMS-1620262, DMS- 1720442, and AFOSR FA9550-16-1-0102.
†Department of Mathematics, Purdue University, West Lafayette, IN 47907 (shen7@purdue.edu).
‡Corresponding author. Department of Mathematics, Purdue University, West Lafayette, IN 47907 (xu924@purdue.edu).

(such as backward Euler) [7, 10, 23, 11, 5] or nonlinearly implicit (such as convex splitting) [9, 2] schemes without restrictive assumptions on the free energy, most of the convergence and error analyses for linearly implicit (such as semi-implicit or stabilized semi-implicit) schemes [9, 13, 19] are based on the so-called Lipschitz assumption, i.e.,

$$|F'(x) - F'(y)| \le L|x-y| \quad \forall x, y \in \mathbb{R}, \tag{1.1}$$

where $F(u)$ is the nonlinear free energy density. However, this assumption greatly limits its range of applicability. Moreover, most of these analyses are for simple free energies, and no analysis is available for the IEQ approach. We aim to establish the convergence and error estimates of the SAV approach with minimum assumptions, in particular without the Lipschitz assumption.

To be specific, we consider the gradient flow on a bounded domain $\Omega \in \mathbb{R}^n$ ($n = 1, 2, 3$) with smooth boundary. Let $F(u)$ be a nonlinear free energy density. We focus on a typical energy functional $E[u(\boldsymbol{x})]$ given by

$$E[u] = \int_\Omega \left(\frac{\lambda}{2}u^2 + \frac{1}{2}|\nabla u|^2\right) d\boldsymbol{x} + E_1[u], \tag{1.2}$$

where $\lambda \ge 0$ [1] and $E_1[u] = \int_\Omega F(u) d\boldsymbol{x} \ge -c_0$ for some $c_0 > 0$, i.e., it is bounded from below, and consider the gradient flow

$$\frac{\partial u}{\partial t} = \mathcal{G}\mu = \mathcal{G}(-\Delta u + \lambda u + g(u)), \tag{1.3}$$

where $\mathcal{G} = -I$ for the L^2 gradient flow, $\mathcal{G} = \Delta$ for the H^{-1} gradient flow, and $g(u) = F'(u)$. As an example, when $E_1[u] = \int_\Omega \alpha(1-u^2)^2 d\boldsymbol{x}$, the two gradient flows are the celebrated Allen–Cahn and Cahn–Hilliard equations [1, 4]. The equation is supplemented with the initial condition $u(\boldsymbol{x}, 0) = u^0(\boldsymbol{x})$ and the boundary conditions

$$\begin{aligned} &\text{periodic, or} \quad u|_{\partial\Omega} = 0, \quad \text{or} \quad \frac{\partial u}{\partial \boldsymbol{n}} = 0 \quad \text{if } \mathcal{G} = -I; \\ &\text{periodic, or} \quad \frac{\partial u}{\partial \boldsymbol{n}} = \frac{\partial \mu}{\partial \boldsymbol{n}} = 0 \quad \text{if } \mathcal{G} = \Delta, \end{aligned} \tag{1.4}$$

where $\mu = \frac{\delta E}{\delta u}$. The equation satisfies the energy dissipation law,

$$\frac{dE}{dt} = \int_\Omega \frac{\partial u}{\partial t}\mu d\boldsymbol{x} = \int_\Omega \mu \mathcal{G}\mu d\boldsymbol{x} \le 0. \tag{1.5}$$

Let $C_0 > c_0$ so that $E_1[u] + C_0 > 0$. Without loss of generality, we substitute E_1 with $E_1 + C_0$ without changing the gradient flow. In this setting, E_1 has a positive lower bound $C_0' = C_0 - c_0$, which we still denote as C_0.

In the SAV approach, we introduce a scalar variable $r(t) = \sqrt{E_1[u]}$ and rewrite (1.3) as

$$\frac{\partial u}{\partial t} = \mathcal{G}\mu, \tag{1.6a}$$

[1] We need $\lambda > 0$ to ensure that $\frac{\lambda}{2}\|u\|^2 + \frac{1}{2}\|\nabla u\|^2$ is a norm in H^1, only in the rare case of L^2 gradient flows for which maximum principle is not satisfied, e.g., with nonpolynomial nonlinear potential subjected to periodic or homogeneous Neumann boundary conditions. This will be assumed throughout the paper. In all other cases we can take $\lambda = 0$.

$$\mu = -\Delta u + \lambda u + \frac{r}{\sqrt{E_1[u]}} g(u), \tag{1.6b}$$

$$r_t = \frac{1}{2\sqrt{E_1[u]}} \int_\Omega g(u) u_t d\boldsymbol{x}. \tag{1.6c}$$

To fix the idea, we shall concentrate our analysis on the following first-order SAV scheme:

$$\frac{u^{n+1} - u^n}{\Delta t} = \mathcal{G}\mu^{n+1}, \tag{1.7a}$$

$$\mu^{n+1} = -\Delta u^{n+1} + \lambda u^{n+1} + \frac{r^{n+1}}{\sqrt{E_1^n}} g(u^n), \tag{1.7b}$$

$$r^{n+1} - r^n = \frac{1}{2\sqrt{E_1^n}} \int_\Omega g(u^n)(u^{n+1} - u^n) d\boldsymbol{x}. \tag{1.7c}$$

We note that the convergence and error estimates derived for the above scheme can be extended to second-order SAV schemes with a similar procedure.

The above SAV scheme leads to a linear equation of the form

$$Au^{n+1} + (u^{n+1}, b_1)b_2 = g,$$

where $A = I + \Delta t \mathcal{G}\Delta$. One can first solve two linear equations with constant coefficients to obtain $A^{-1}b_2$ and $A^{-1}g$. Then, from

$$u^{n+1} + (u^{n+1}, b_1)A^{-1}b_2 = A^{-1}g,$$

one can compute (u^{n+1}, b_1) by taking the inner product with b_1, we then obtain u^{n+1} (see [17, 18] for details). Hence the scheme is easy to implement and very efficient.

Taking the inner product of the first two equations with μ^{n+1} and $(u^{n+1}-u^n)/\Delta t$, respectively, and multiplying the third equation by $2r^{n+1}/\Delta t$, we derive that the above SAV scheme satisfies the following discrete energy law:

$$\begin{aligned}
&\int_\Omega \left(\frac{\lambda}{2}(u^{n+1})^2 + \frac{1}{2}|\nabla u^{n+1}|^2\right) d\boldsymbol{x} + (r^{n+1})^2 - \int_\Omega \left(\frac{\lambda}{2}(u^n)^2 + \frac{1}{2}|\nabla u^n|^2\right) d\boldsymbol{x} - (r^n)^2 \\
&\qquad + \int_\Omega \left(\frac{\lambda}{2}(u^{n+1} - u^n)^2 + \frac{1}{2}|\nabla(u^{n+1} - u^n)|^2\right) d\boldsymbol{x} + (r^{n+1} - r^n)^2 \\
&\qquad = \Delta t \int_\Omega \mu^{n+1} \mathcal{G}\mu^{n+1} d\boldsymbol{x} \le 0.
\end{aligned} \tag{1.8}$$

Hence, the SAV is unconditionally energy stable with the modified energy

$$\tilde{E}^n = \int_\Omega \left(\frac{\lambda}{2}(u^n)^2 + \frac{1}{2}|\nabla u^n|^2\right) d\boldsymbol{x} + (r^n)^2. \tag{1.9}$$

However, energy stability alone is not sufficient for convergence which typically needs bounds in higher norms, but it plays an important role in deriving needed estimates. In particular, we shall start from the energy stability (1.8) to derive H^2 estimates for the numerical solution u^n, which imply an L^∞ bound for u^n. This is an essential ingredient to allow us to pass to the limit ($\Delta t \to 0$) and show that u^n converges to the exact solution u in suitable norms.

The rest of the paper is organized as follows. In section 2, we derive H^2 bounds for both PDEs and the corresponding SAV schemes. The convergence of SAV is proved in section 3, followed by an error estimate in section 4. In section 5, we discuss extensions to second-order SAV schemes and to several other gradient flows that cannot be cast in the general framework used in previous sections. Some concluding remarks are given in the last section.

Below is some notation to be used throughout the paper. We denote the spaces $L^p(\Omega)$ by L^p in short. The Sobolev spaces H^s with the noninteger order s will also be used. The space $L^p(0,T;V)$ represents the L^p space on the interval $(0,T)$ with values in the function space V. The dual space of V is denoted by V'. We use $\|\cdot\|_V$ to denote the norm in the space V, and the L^2 norm without subscript. We denote by $(\cdot,\cdot)$ and $\|\cdot\|$ the inner product and the norm in L^2, and by C any constant depending only on Ω, u_0 and the lower bound of E_1.

2. H^2 bounds.

We assume that $F \in C^3(\mathbb{R})$. We first recall the existence, uniqueness, and regularity results about L^2 and H^{-1} gradient flows (cf., for instance, [21]). In some cases, the following assumption is needed to ensure the uniqueness: there exists a constant $c_1 > 0$ such that

$$F''(s) = g'(s) \geq -c_1. \tag{2.1}$$

PROPOSITION 2.1. *Let $\mathcal{G} = -I$ (i.e., the L^2 gradient flow). If (2.1) holds, $u^0 \in L^2$ and there exists $p_0 > 0$ such that*

$$sg(s) \geq b|s|^{p_0} - c, \tag{2.2}$$

where $b > 0$ and c are constants. Then there exists a unique solution u for (1.3) such that

$$u \in L^2(0,T;H^1) \cap L^{p_0}(0,T;L^{p_0}) \cap C([0,+\infty);L^2) \quad \forall T > 0. \tag{2.3}$$

Furthermore if $u^0 \in H^1$, we have

$$u \in C([0,T];H^1) \cap L^2(0,T;H^2) \quad \forall T > 0. \tag{2.4}$$

PROPOSITION 2.2. *Let $\mathcal{G} = -\Delta$ (i.e., the H^{-1} gradient flow).*

(i) *If $u^0 \in L^2$, (2.1) holds, and there exists $p_0 > 0$ such that (2.2) holds. Then there exists a unique solution u for (1.3) such that*

$$u \in L^2(0,T;H^2) \cap L^{p_0}(0,T;L^{p_0}) \cap C([0,T];L^2) \quad \forall T > 0. \tag{2.5}$$

(ii) *If $u^0 \in H^2$, and*

$$|g'(x)| < C(|x|^p + 1), \quad p > 0 \text{ arbitrary if } n = 1,2; \quad 0 < p < 4 \text{ if } n = 3; \tag{2.6}$$
$$|g''(x)| < C(|x|^{p'} + 1), \quad p' > 0 \text{ arbitrary if } n = 1,2; \quad 0 < p' < 3 \text{ if } n = 3, \tag{2.7}$$

then, for any $T > 0$, there exists a unique solution u for (1.3) in the space

$$C([0,T];H^2) \cap L^2(0,T;H^4). \tag{2.8}$$

2.1. A technical lemma. A common strategy to establish the convergence of a time discretization numerical scheme is to derive bounds in norms similar to those of the PDE system. For fully implicit or nonlinearly implicit schemes, it is often possible to derive such bounds following a similar procedure for the PDE system. However, for semi-implicit or linearly implicit schemes such as the SAV schemes, this procedure cannot be followed since the nonlinear terms are treated explicitly. This is the main reason why a Lipschitz condition on F' is assumed in many works for semi-implicit or linearly implicit schemes so that necessary bounds can be derived.

Below, we shall derive an H^2 bound for u^n without assuming the Lipschitz condition, using the unconditionally energy stability (1.8) which implies, in particular, that there exists a constant M depending only on Ω and u_0 such that

$$\|u^n\|_{H^1} + |r^n| \leq M \quad \forall n \quad \text{if } u^0 \in H^1(\Omega). \tag{2.9}$$

We start with a technical lemma which will help us to derive H^2 bounds for the solution of the SAV scheme. The second part of the lemma is given in [21]. For the reader's convenience, we still write down the proof.

LEMMA 2.3. *Assume that* $\|u\|_{H^1} \leq M$.

1. *Assume that* (2.6) *holds. Then, for any* $u \in H^3$, *there exist* $0 \leq \sigma < 1$ *and a constant* $C(M)$ *such that the following inequality holds:*

$$\|\nabla g(u)\|^2 \leq C(M)(1 + \|\nabla \Delta u\|^{2\sigma}). \tag{2.10}$$

2. *Assume that* (2.6) *and* (2.7) *hold. Then, for any* $u \in H^4$, *there exist* $0 \leq \sigma < 1$ *and a constant* $C(M)$ *such that the following inequality holds:*

$$\|\Delta g(u)\|^2 \leq C(M)(1 + \|\Delta^2 u\|^{2\sigma}). \tag{2.11}$$

Proof. Since

$$\nabla g(u) = g'(u)\nabla u,$$

then, by $\|u\|_{H^1} \leq M$ and (2.6), we have

$$\|\nabla g(u)\| \leq \|g'(u)\|_{L^\infty} \|\nabla u\| \leq C(M)(1 + \|u\|^p_{L^\infty}). \tag{2.12}$$

Denote $m(u) = \frac{1}{|\Omega|}\int_\Omega u d\boldsymbol{x}$, which is bounded by

$$|m(u)|^2 \leq \frac{1}{|\Omega|}\|u\|^2 \leq \frac{M^2}{|\Omega|}. \tag{2.13}$$

By Sobolev embedding theorems, $H^1 \subseteq L^\infty$ when $n = 1$, and $H^{1+2\delta} \subseteq L^\infty \, \forall \delta > 0$ when $n = 2$. Together with the interpolation inequality about the spaces H^s (see, for example, Chapter II, section 2.1 in [21]), we deduce that

$$\begin{aligned}
&\|u - m(u)\|_{L^\infty} \leq C\|\nabla u\| \leq C(M), \quad n = 1,\\
&\|u - m(u)\|_{L^\infty} \leq C\|u - m(u)\|_{H^{1+2\delta}} \leq C\|\nabla u\|^{1-\delta}\|\nabla \Delta u\|^\delta\\
&\qquad \leq C(M)\|\nabla\Delta u\|^\delta \quad \forall \delta > 0, \quad n = 2.
\end{aligned}$$

For $n = 3$, we use Agmon's inequality (Chapter II, section 1.4 in [21]) and interpolation inequality to derive that

$$\|u - m(u)\|_{L^\infty} \leq C\|\nabla u\|^{1/2}\|\Delta u\|^{1/2} \leq C\|\nabla u\|^{3/4}\|\nabla \Delta u\|^{1/4} \leq C(M)\|\nabla \Delta u\|^{1/4}.$$

Combining the above results with (2.12), we obtain (2.10).

From

$$\Delta g(u) = g'(u)\Delta u + g''(u)|\nabla u|^2,$$

we derive

$$\begin{aligned}\|\Delta g(u)\| \leq& \|g'(u)\|_{L^\infty}\|\Delta u\| + \|g''(u)\|_{L^\infty}\|\nabla u\|_{L^4}^2\\ \leq& C\Big[(1+\|u\|_{L^\infty}^{p})\|\Delta u\| + (1+\|u\|_{L^\infty}^{p'})\|\nabla u\|_{L^4}^2\Big].\end{aligned}\tag{2.14}$$

By interpolation inequality, we have

$$\|\Delta u\| \leq C\|\nabla u\|^{2/3}\|\Delta^2 u\|^{1/3} \leq C(M)\|\Delta^2 u\|^{1/3}.$$

Together with Sobolev embedding, we obtain

$$\|\nabla u\|_{L^4} \leq C\|\nabla u\|_{H^{n/4}} \leq C\|\nabla u\|^{1-n/12}\|\Delta^2 u\|^{n/12} \leq C(M)\|\Delta^2 u\|^{n/12}.$$

Similar to the first part of the proof, we deduce that

$$\begin{aligned}&\|u - m(u)\|_{L^\infty} \leq C(M), \quad n = 1,\\ &\|u - m(u)\|_{L^\infty} \leq C(M)\|\Delta^2 u\|^{\delta} \quad \forall \delta > 0, \quad n = 2.\end{aligned}$$

For $n = 3$, we have

$$\|u - m(u)\|_{L^\infty} \leq C\|\nabla u\|^{5/6}\|\Delta^2 u\|^{1/6} \leq C(M)\|\Delta^2 u\|^{1/6}.$$

Combining the above inequalities with (2.14), we obtain (2.11). ∎

2.2. H^{-1} gradient flow. We give an H^2 bound for the SAV scheme similar to that for the PDE (cf. [21, Chapter III, section 4.2.3]).

LEMMA 2.4. *For the H^{-1} gradient flow, assume both* (2.6) *and* (2.7) *hold, and $u^0 \in H^4$. Let M be given in* (2.9). *Then for all $n \leq T/\Delta t$, we have*

$$\|\Delta u^n\|^2 + \frac{\Delta t}{2}\sum_{k=0}^{n}\|\Delta^2 u^k\|^2 \leq C(M)(T+1) + \|\Delta u^0\|^2 + \Delta t\|\Delta^2 u^0\|^2.\tag{2.15}$$

Proof. We observe from (1.7a), (1.7b), and the regularity of elliptical equations that $u^n \in H^4$ for any n. We multiply (1.7a) with $\Delta^2 u^{n+1}$ and combine (1.7b) with (1.7a). Note that $|r^n| \leq M$ and $E_1^n \geq C_0$. So we have

$$\begin{aligned}&\frac{1}{2\Delta t}(\|\Delta u^{n+1}\|^2 - \|\Delta u^n\|^2 + \|\Delta(u^{n+1}-u^n)\|^2) + \|\Delta^2 u^{n+1}\|^2 + \lambda\|\nabla\Delta u^{n+1}\|^2\\ &\quad = -\frac{r^{n+1}}{\sqrt{E_1^n}}(\Delta g(u^n), \Delta^2 u^{n+1})\\ &\quad \leq C(M)\|\Delta g(u^n)\|^2 + \frac{1}{2}\|\Delta^2 u^{n+1}\|^2.\end{aligned}\tag{2.16}$$

By Lemma 2.3, for any $\epsilon > 0$, there exists a constant $C(\epsilon, M)$ depending on ϵ, such that the following inequality holds:

$$\|\Delta g(u^n)\|^2 \leq C(M)(1+\|\Delta^2 u^n\|^{2\sigma}) \leq \epsilon\|\Delta^2 u^n\|^2 + C(\epsilon, M).\tag{2.17}$$

We choose $\epsilon = 1/4$ to arrive at

$$\|\Delta u^{n+1}\|^2 - \|\Delta u^n\|^2 + \|\Delta(u^{n+1} - u^n)\|^2 + \Delta t\|\Delta^2 u^{n+1}\|^2 - \frac{\Delta t}{2}\|\Delta^2 u^n\|^2 \le C(M). \tag{2.18}$$

We conclude the proof by taking the sum from 0 to $n-1$. ∎

Remark 2.5. We present the following points.

- The energy stability (1.8) is crucial for the H^2 bound, because the constant in Lemma 2.3 depends on the H^1 bound (2.9).
- To obtain the H^2 bound for the PDE, we only need $u^0 \in H^2$. But for the SAV scheme, we need to assume higher regularity of u^0 because we cannot cancel the $O(\Delta t)$ term on the right-hand side of (2.15).

2.3. L^2 gradient flow. We first derive a regularity result for the L^2 gradient flow using the H^2 bound based on Lemma 2.3. To this end, we need to assume (2.6) which is slightly stronger than condition (2.2) in Proposition 2.1.

THEOREM 2.6. *Assume $u^0 \in H^2$ and* (2.6) *holds. Then for any $T > 0$, the problem* (1.3) *with $\mathcal{G} = -I$ has a unique solution in the space*

$$C([0,T]; H^2) \cap L^2(0,T; H^3). \tag{2.19}$$

Proof. We use the Galerkin method. Denote by $\{w_j\}$ the orthonormal basis in $L^2(\Omega)$ consisting of the eigenfunctions of $-\Delta$, i.e.,

$$-\Delta w_j = \lambda_j w_j.$$

Consider the approximate solution constructed by

$$u_m(\cdot, t) = \sum_{j=1}^{m} g_{jm}(t) w_j, \tag{2.20}$$

which satisfies

$$(u_m'(\cdot,t), w_j) + (\nabla u_m, \nabla w_j) + \lambda(u_m, w_j) + (g(u_m), w_j) = 0, \tag{2.21}$$

with $u_m(0)$ given by the projection of u^0 in $L^2(\Omega)$ on the space spanned by $\{w_j\}$.

Then, by multiplying (2.21) by $g_{jm}'(t)$ and summing up for $j = 1, \dots, m$, we obtain

$$\frac{d}{dt}\left[\frac{1}{2}\Big(\|\nabla u_m\|^2 + \lambda\|u_m\|^2\Big) + E_1(u_m)\right] = -\|u_m'(\cdot,t)\|^2 \le 0. \tag{2.22}$$

Thus, together with $u^0 \in H^2 \subseteq L^\infty$, $E_1 > -C_0$, and $F \in C^3(\mathbb{R})$, we deduce that

$$\|u_m\|_{H^1} \le CE[u_m(t)] + C \le CE[u_m(0)] + C \le M,$$

where M depends on $\|u^0\|_{H^2}$.

Next, we multiply (2.21) with $\lambda_j^2 g_{jm}(t)$ and sum up for $j = 1, \dots, m$ to obtain

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\|\Delta u_m\|^2 + \|\nabla\Delta u_m\|^2 + \lambda\|\Delta u_m\|^2 &= -(\nabla g(u_m), \nabla\Delta u_m)\\ &\le \frac{1}{2}\|\nabla g(u_m)\|^2 + \frac{1}{2}\|\nabla\Delta u_m\|^2.\end{aligned} \tag{2.23}$$

Using Lemma 2.3, we obtain

$$\|\nabla g(u_m)\|^2 \leq \frac{1}{2}\|\nabla\Delta u_m\|^2 + C(M).$$

Thus, we arrive at

$$\frac{d}{dt}\|\Delta u_m\|^2 + \frac{1}{2}\|\nabla\Delta u_m\|^2 \leq C(M). \tag{2.24}$$

So we know that u_m is bounded independently of m in $L^\infty(0,T;H^2) \cap L^2(0,T;H^3)$.

Next, we can select a subsequence, still denoted by u_m, such that

$$u_m \to u \text{ in } L^2(0,T;H^3) \text{ weakly, in } L^\infty(0,T;H^2) \text{ weak-star.}$$

By the Aubin–Lions lemma (see, for example, Chapter 3, section 2 in [20]), $u_m \to u$ strongly in $L^2(0,T;H^1)$. We know that $\|u_m\|_{L^\infty((0,T)\times\Omega)} \leq C\|u_m\|_{L^\infty(0,T;H^2)} \leq C$. Since $g(s) \in C^2(\mathbb{R})$, we can find a constant L such that $|g'(\xi u_m + (1-\xi)u)| \leq L\, \forall 0 \leq \xi \leq 1$. Let $\psi(t)$ be any function in $C^1([0,T])$ with $\psi(T) = 0$. We have

$$\begin{aligned}&\int_0^T (g(u_m) - g(u), v\psi(t))dt \\ &\quad \leq \|g(u_m) - g(u)\|_{L^2(0,T;L^2)}\|v\psi(t)\|_{L^2(0,T;L^2)} \\ &\quad \leq L\|u_m - u\|_{L^2(0,T;L^2)}\|v\psi(t)\|_{L^2(0,T;L^2)}.\end{aligned}$$

So we take the limit $m \to +\infty$ in (2.21), obtaining

$$\frac{d}{dt}(u,v) + (\nabla u, \nabla v) + \lambda(u,v) + (g(u),v) = 0 \;\; \forall v \in H^1, \tag{2.25}$$

in the distribution sense in $(0,T)$. We also know from the above equality that $u' = \Delta u - g(u) \in L^2(0,T;H^1) \subseteq L^2(0,T;(H^3)')$ by $u \in L^2(0,T;H^3)$ and (2.10). Thus, the continuity of $u(t)$ about t comes from a standard result (Lemma 3.2, Chapter II in [21]). Then, it is easy to check that $u(0) = u^0$.

For the uniqueness, suppose u and v are two solutions that lie in the space (2.19). Denote $w = u - v$. We take the inner product about w and

$$w_t - \Delta w = g(v) - g(u) = -g'(\xi u + (1-\xi)v)w, \quad 0 \leq \xi \leq 1. \tag{2.26}$$

Since $u, v \in C([0,T];H^2)$, we can find $L > 0$ such that $g'(\xi u + (1-\xi)v) \geq -L$ for $t \in [0,T]$. So we have

$$\frac{d}{dt}\|w\|^2 + \|\nabla w\|^2 \leq L\|w\|^2, \quad t \in [0,T], \tag{2.27}$$

yielding

$$\|w(t)\|^2 \leq \exp(Lt)\|w(0)\|^2, \quad t \in [0,T]. \tag{2.28}$$

Therefore, if $w(0) = 0$, then for any $t \in [0,T]$ we have $w(t) = 0$. $\square$

Next, we derive an analogous H^2 bound for the SAV scheme in the case of L^2 gradient flow.

Lemma 2.7. *Assume* (2.6) *holds and the initial value* $u^0 \in H^3$. *Let* M *be given in* (2.9). *Then for all* $n \leq T/\Delta t$, *the solution of* (1.7) *with* $\mathcal{G} = -I$ *satisfies*

$$\|\Delta u^n\|^2 + \frac{\Delta t}{2}\sum_{k=0}^{n}\|\nabla\Delta u^k\|^2 \leq C(M)(T+1) + \|\Delta u^0\|^2 + \Delta t\|\nabla\Delta u^0\|^2. \tag{2.29}$$

The proof is essentially the same as that of Lemma 2.4, so we leave it as an exercise for the interested reader.

3. Convergence. In order to describe convergence, we define several functions in time based on the numerical solution of (1.7):

1. $u_1(t) = u^{n+1}$ for $n\Delta t \le t < (n+1)\Delta t$.
2. $u_2(t) = u^n$ for $n\Delta t \le t < (n+1)\Delta t$.
3. $u_3(t)$ is piecewise linear with $u_3(n\Delta t) = u^n$.
4. $\mu_1(t)$, $\mu_2(t)$, $\mu_3(t)$, $r_1(t)$, $r_2(t)$, $r_3(t)$ are similarly defined.

3.1. H^{-1} gradient flow.

THEOREM 3.1. *Assume $u^0 \in H^4$ and (2.6), (2.7) hold. When $\Delta t \to 0$, we have $u_i \to u$ strongly in $L^2(0,T;H^{4-\epsilon})$, weakly in $L^2(0,T;H^4)$, weak-star in $L^\infty(0,T;H^2)$; $r_i \to r = \sqrt{E_1}$ weak-star in $L^\infty(0,T)$; and $\mu_1 \to \mu$ weakly in $L^2(0,T;H^1)$.*

Proof. Let $\psi(t) \in C^\infty([0,T])$ with $\psi(T) = 0$. Multiplying (1.7a) by $v\psi(t)$ and integrating in space-time, and multiplying (1.7c) by $\psi(t)$ and integrating in time, we find, after integration by parts, that (u_i, r_i) satisfy the following equations:

$$\begin{aligned}
&-(u^0, v\psi(0)) + \int_0^T -\psi'(t)(u_3, v) + \psi(t)\Bigg[(\Delta u_1, \Delta v) + \lambda(\nabla u_1, \nabla v) \\
&\qquad + \frac{r_1}{\sqrt{E_1[u_2]}}(\nabla g(u_2), \nabla v)\Bigg] dt = 0 \quad \forall v \in H^2, \qquad (3.1)
\end{aligned}$$

$$-r^0\psi(0) + \int_0^T -\psi'(t) r_3 - \psi(t)\frac{1}{2\sqrt{E_1[u_2]}}\left(g(u_2), \frac{\partial u_3}{\partial t}\right) dt = 0. \tag{3.2}$$

Let us denote $X_0 = H^2$, $X_1 = H^1$, and $X_2 = H^4$. The following can be derived from (1.8) and Lemma 2.4:
(A1) u_i are bounded in $L^\infty(0,T;X_0)$ and $L^2(0,T;X_2)$.
(A2) r_i are bounded in $L^\infty(0,T)$.
(A3) $\partial u_3/\partial t$ is bounded in $L^2(0,T;X_1')$, since

$$\frac{\partial u_3}{\partial t} = \frac{u^{n+1} - u^n}{\Delta t} = \Delta\mu^{n+1} = \Delta\mu_1(t), \quad n\Delta t \le t < (n+1)\Delta t. \tag{3.3}$$

Here X_1' denotes the dual space of X_1.
(A4) $\|u_i - u_j\|_{L^2(0,T;L^2)}$, $\|r_i - r_j\|_{L^2(0,T)} \le C\sqrt{\Delta t}$.
Hence, there exist U_i and R_i and a subsequence of $\{\Delta t_k\}$, such that when $k \to \infty$,
(B1) $u_i \to U_i$ weak-star in $L^\infty(0,T;X_0)$, weakly in $L^2(0,T;X_2)$;
(B2) $r_i \to R_i$ weak-star in $L^\infty(0,T)$;
(B3) $\partial u_3/\partial t \to \partial U_3/\partial t$ weakly in $L^2(0,T;X_1')$.
Using the Aubin–Lions lemma [20], we derive that $u_i \to U_i$ strongly in $L^2(0,T;H^{4-\epsilon})$ for any $\epsilon > 0$. Furthermore, we can conclude from (A4) that $U_1 = U_2 = U_3 = u_*$ and $R_1 = R_2 = R_3 = r_*$.

In order to take the limit $k \to \infty$ in (3.1) and (3.2), we need to show first that as $k \to \infty$,
(C1) $E_1[u_2] \to E_1[u_*]$ weak-star in $L^\infty(0,T)$;
(C2) $g(u_2) \to g(u_*)$ strongly in $L^2(0,T;X_1)$.
Indeed, since H^2 is embedded in L^∞, and thanks to Lemma 2.4, u_i are uniformly bounded in $L^\infty(0,T;H^2)$, we have $\|u_i\|_{L^\infty((0,T)\times\Omega)} < C$. Thus we can find a constant L such that $|g(u_i)|$, $|g'(u_i)|$, $|g''(u_i)| \le L$. Therefore, (C1) and (C2) follow from the following inequalities and the strong convergence of u_2 in $L^2(0,T;H^{4-\epsilon})$:

$$|E_1[u_2] - E_1[u_*]| \le L\|u_2 - u_*\|_{L^1},$$

$$
\begin{aligned}
\|\nabla g(u_2)-\nabla g(u_*)\| \leq& \|(g'(u_2)-g'(u_*))\nabla u\| + \|g'(u_2)\nabla(u_2-u_*)\| \\
\leq& L\|u_2-u_*\|_{L^\infty}\|u_*\|_{H^1} + L\|u_2-u_*\|_{H^1} \\
\leq& CL\|u_2-u_*\|_{H^2}\|u_*\|_{H^1} + L\|u_2-u_*\|_{H^1}.
\end{aligned}
$$

With (C1) and (C2), we can conclude that (u_*, r_*) is a solution of (1.6). We shall only show the convergence of the last term in (3.2) since other terms can be treated similarly. Indeed, by noting $E_1[u_2], E_1[u_*] > C_0$, we have

$$
\begin{aligned}
&\left|\int_0^T \psi(t)\frac{1}{2\sqrt{E_1[u_2]}}\left(g(u_2), \frac{\partial u_3}{\partial t}\right) - \psi(t)\frac{1}{2\sqrt{E_1[u_*]}}\left(g(u_*), \frac{\partial u_*}{\partial t}\right)dt\right| \\
\leq& \int_0^T \left|\psi(t)\frac{1}{2\sqrt{E_1[u_2]}}\right|\left|\left(g(u_2)-g(u_*), \frac{\partial u_3}{\partial t}\right)\right|dt \\
&+ \int_0^T \left|\psi(t)\frac{1}{2\sqrt{E_1[u_2]}}\left(g(u_*), \frac{\partial (u_3-u_*)}{\partial t}\right)\right|dt \\
&+ \int_0^T |\psi(t)|\left|\frac{1}{2\sqrt{E_1[u_2]}} - \frac{1}{2\sqrt{E_1[u_*]}}\right|\left|\left(g(u_*), \frac{\partial u_*}{\partial t}\right)\right|dt \\
\leq& C\|g(u_2)-g(u_*)\|_{L^2(0,T;X_1)}\left\|\frac{\partial u_3}{\partial t}\right\|_{L^2(0,T;X_1')} \\
&+ C\int_0^T \left|\left(g(u_*), \frac{\partial(u_3-u_*)}{\partial t}\right)\right|dt \\
&+ C\int_0^T |E_1[u_2]-E_1[u_*]|\left|\left(g(u_*), \frac{\partial u_*}{\partial t}\right)\right|dt. \qquad \square
\end{aligned}
$$

The right-hand side goes to zero by (B3), (C1), and (C2). Thus, we can conclude from the uniqueness that for all sequences, $(u_i, r_i) \to (u_*, r_*)$.

3.2. L^2 gradient flow. For the L^2 gradient flow, the weak form satisfied by the solution of SAV scheme is

$$
\begin{aligned}
&-(u^0, v\psi(0)) + \int_0^T -\psi'(t)(u_3, v) + \psi(t)\Bigg[(\nabla u_1, \nabla v) + \lambda(u_1, v) \\
&+ \frac{r_1}{\sqrt{E_1[u_2]}}(g(u_2), v)\Bigg]dt = 0 \quad \forall v \in H^1, \qquad (3.4)
\end{aligned}
$$

$$
-r^0\psi(0) + \int_0^T -\psi'(t)r_3 - \psi(t)\frac{1}{2\sqrt{E_1[u_2]}}\left(g(u_2), \frac{\partial u_3}{\partial t}\right)dt = 0. \qquad (3.5)
$$

We first give a result analogous to Theorem 3.1.

THEOREM 3.2. *Assume $u^0 \in H^3$ and (2.6) holds. When $\Delta t \to 0$, we have $u_i \to u$ strongly in $L^2(0,T;H^{3-\epsilon})\,\forall\epsilon > 0$, weakly in $L^2(0,T;H^3)$, weak-star in $L^\infty(0,T;H^2)$; $r_i \to r = \sqrt{E_1}$ weak-star in $L^\infty(0,T)$; and $\mu_1 \to \mu$ weakly in $L^2(0,T;L^2)$.*

Proof. The proof is almost the same as that of Theorem 3.1. The difference is that we need to let $X_1 = L^2$ in (A3), (B3), and (C2), and let $X_2 = H^3$ in (A1) and (B1). Thanks to Lemma 2.7, we can find a constant L such that $|g(u_i)|, |g'(u_i)| \leq L$.

Therefore, (C1) and (C2) follow from the following estimates:

$$|E_1[u_2] - E_1[u_*]| \le \int |g(\xi u_2 + (1-\xi)u_*)||u_2 - u_*| d\boldsymbol{x} \le L\|u_2 - u_*\|_{L^1}, \quad 0 \le \xi \le 1,$$
$$\|g(u_2) - g(u_*)\| = \|g'(\xi u_2 + (1-\xi)u_*)(u_2 - u_*)\| \le L\|u_2 - u_*\|.$$

Thanks to the existence and uniqueness of the PDE (Theorem 2.6), we can then prove the desired convergence by passing to the limit as in the proof of Theorem 3.1. ∎

Next we give a result with less regular u^0.

THEOREM 3.3. *Assume $u^0 \in H^1$ and (2.1), (2.6) hold. When $\Delta t \to 0$, we have $u_i \to u$ strongly in $L^2(0,T;H^{1-\epsilon})$, weak-star in $L^\infty(0,T;H^1)$; $r_i \to r = \sqrt{E_1}$ weak-star in $L^\infty(0,T)$; and $\mu_1 \to \mu$ weakly in $L^2(0,T;L^2)$.*

Proof. We follow the same procedure as above. In this case, we need to let $X_1 = L^2$, and we only have $X_0 = X_2 = H^1$ in (A1) and (B1). To pass to the limit, we also need (C1) and (C2). Let p be given in (2.6). We set $q - 1 = p/2 > 0$, which satisfies $q < 3$ for $n = 3$. Then, we can choose $\epsilon > 0$ such that we have the embedding $H^{1-\epsilon} \subseteq L^{2q}$. Using the Aubin–Lions lemma, we know that $u_i \to u_*$ strongly in $L^2(0,T;H^{1-\epsilon})$. Let $q^* = q/(q-1)$. Using Hölder's inequality and $|g'(s)| \le C(|s|^p + 1) = C(|s|^{2q-2} + 1)$, we obtain

$$\begin{aligned}
\|g(u_2) - g(u_*)\|_{L^2} &\le \|u_2 - u_*\|_{L^{2q}} \left\|g'(\xi u_2 + (1-\xi)u_*)\right\|_{L^{2q^*}} \\
&\le C\|u_2 - u_*\|_{L^{2q}} \Big(\|u_2\|_{L^{2q}} + \|u_*\|_{L^{2q}} + 1\Big)^{q-1} \\
&\le C\|u_2 - u_*\|_{H^{1-\epsilon}} \Big(\|u_2\|_{H^1} + \|u_*\|_{H^1} + 1\Big)^{q-1}, \\
|E_1[u_2] - E_1[u_*]| &\le C(|u_*|^{2q-2} + |u_2|^{2q-2} + 1, |u_2 - u_*|) \\
&\le C\|u_2 - u_*\|_{L^{2q}} \Big(\|u_2\|_{L^{2q}} + \|u_*\|_{L^{2q}} + 1\Big)^{q-1}.
\end{aligned}$$

Note that u_2, u_* are bounded in $L^\infty(0,T;H^1)$, and (C1) and (C2) follow from these estimates. The proof is complete by noting the uniqueness of the exact solution. ∎

4. Error estimate. In the last section, we have established convergence results with minimum assumptions. In this section, we shall derive error estimates with further smoothness requirements of the exact solution. Denote $e^n = u^n - u(t^n)$, $s^n = r^n - r(t^n)$, and $w^{n+1} = \mu^{n+1} - \mu(t^{n+1})$.

4.1. H^{-1} gradient flow.

THEOREM 4.1. *For the H^{-1} gradient flow, assume that $u^0 \in H^4$ and (2.6), (2.7) hold. In addition, we assume that*

$$u \in L^\infty(0,T;W^{1,\infty}), \quad u_t \in L^\infty(0,T;H^{-1}) \cap L^2(0,T;H^1), \quad u_{tt} \in L^2(0,T;H^{-1}). \tag{4.1}$$

Then for all $n \le T/\Delta t$, we have

$$\begin{aligned}
&\frac{1}{2}\|\nabla e^n\|^2 + \frac{\lambda}{2}\|e^n\|^2 + (s^n)^2 \\
&\le C\exp\Big((1 - C\Delta t)^{-1}t^n\Big)\Delta t^2 \int_0^{t^n} (\|u_{tt}(s)\|^2_{H^{-1}} + \|u_t(s)\|^2_{H^1})ds.
\end{aligned} \tag{4.2}$$

The constant C is dependent on T, u^0, Ω, $\|u\|_{L^\infty(0,T;W^{1,\infty})}$, and $\|u_t\|_{L^\infty(0,T;H^{-1})}$.

Proof. We know from Proposition 2.2 and Lemma 2.4 that

$$\|u(t)\|_{H^2},\ \|u^n\|_{H^2} \le C, \tag{4.3}$$

where C is dependent on u^0, Ω, and T. Note that $H^2 \subseteq L^\infty$. Therefore, we can find a constant C such that

$$|g(u)|,\ |g'(u)|,\ |g''(u)|,\ |g(u^n)|,\ |g'(u^n)|,\ |g''(u^n)| \le C. \tag{4.4}$$

By direct calculation,

$$r_{tt} = -\frac{1}{4\sqrt{E_1[u]^3}}\left(\int_\Omega g(u)u_t d\boldsymbol{x}\right)^2 + \frac{1}{2\sqrt{E_1[u]}}\int_\Omega (g'(u)u_t^2 + g(u)u_{tt})d\boldsymbol{x}. \tag{4.5}$$

Together with (4.3), (4.4), and (4.1), we deduce that

$$\begin{aligned}\int_0^T |r_{tt}|^2 dt &\le C\int_0^T (\|u_t\|_{L^4}^2 + \|g'(u)\nabla u\|_{L^\infty}^2\|u_{tt}\|_{H^{-1}}^2)dt\\ &\le C\|\nabla u\|^2_{L^\infty((0,T)\times\Omega)}\int_0^T(\|u_t\|_{H^1}^2 + \|u_{tt}\|_{H^{-1}}^2)dt.\end{aligned} \tag{4.6}$$

The equations for the errors are written as

$$e^{n+1} - e^n = \Delta t\Delta w^{n+1} + T_1^n, \tag{4.7}$$

$$\begin{aligned}w^{n+1} =& -\Delta e^{n+1} + \lambda e^{n+1} + \frac{s^{n+1}}{\sqrt{E_1^n}}g(u^n)\\ &+ r(t^{n+1})\left(\frac{g(u^n)}{\sqrt{E_1^n}} - \frac{g(u(t^n))}{\sqrt{E_1(t^n)}}\right) + T_2^n,\end{aligned} \tag{4.8}$$

$$\begin{aligned}s^{n+1} - s^n =& \int_\Omega \frac{g(u^n)}{2\sqrt{E_1^n}}(e^{n+1}-e^n)\\ &+ \frac{1}{2}\left(\frac{g(u^n)}{\sqrt{E_1^n}} - \frac{g(u(t^n))}{\sqrt{E_1(t^n)}}\right)(u(t^{n+1}) - u(t^n))d\boldsymbol{x}\\ &- v_1^n + v_2^n.\end{aligned} \tag{4.9}$$

The truncation errors are given by

$$T_1^n = u(t^{n+1}) - u(t^n) - \Delta t u_t(t^{n+1}) = \int_{t^n}^{t^{n+1}} (t^n - s)u_{tt}(s)ds, \tag{4.10}$$

$$T_2^n = r(t^{n+1})\left(\frac{g(u(t^n))}{\sqrt{E_1(t^n)}} - \frac{g(u(t^{n+1}))}{\sqrt{E_1(t^{n+1})}}\right), \tag{4.11}$$

$$v_1^n = r(t^{n+1}) - r(t^n) - \Delta t\, r_t(t^n) = \int_{t^n}^{t^{n+1}} (t^{n+1}-s)r_{tt}(s)ds, \tag{4.12}$$

$$v_2^n = \frac{1}{2}\left(\frac{g(u(t^n))}{\sqrt{E_1(t^n)}}, \int_{t^n}^{t^{n+1}}(t^{n+1}-s)u_{tt}(s)ds\right). \tag{4.13}$$

Multiplying (4.7) with w^{n+1}, (4.8) with $e^{n+1}-e^n$, and (4.9) with $2s^{n+1}$, then summing up three equalities, we get

$$\frac{\lambda}{2}(\|e^{n+1}\|^2 - \|e^n\|^2) + \frac{1}{2}(\|\nabla e^{n+1}\|^2 - \|\nabla e^n\|^2) + (s^{n+1})^2 - (s^n)^2$$

$$
\begin{aligned}
&+\frac{\lambda}{2}(\|e^{n+1}-e^n\|^2)+\frac{1}{2}(\|\nabla e^{n+1}-\nabla e^n\|^2)+(s^{n+1}-s^n)^2+\Delta t\|\nabla w^{n+1}\|^2\\
=&-r(t^{n+1})\left(e^{n+1}-e^n,\frac{g(u^n)}{\sqrt{E_1^n}}-\frac{g(u(t^n))}{\sqrt{E_1(t^n)}}\right)\\
&+s^{n+1}\left(u(t^{n+1})-u(t^n),\frac{g(u^n)}{\sqrt{E_1^n}}-\frac{g(u(t^n))}{\sqrt{E_1(t^n)}}\right)\\
&-(e^{n+1}-e^n,T_2^n)+2s^{n+1}(-v_1^n+v_2^n)+(w^{n+1},T_1^n). \qquad (4.14)
\end{aligned}
$$

Note that $|r(t)|<C$. We have the following estimates:

$$
\begin{aligned}
&r(t^{n+1})\left(e^{n+1}-e^n,\frac{g(u^n)}{\sqrt{E_1^n}}-\frac{g(u(t^n))}{\sqrt{E_1(t^n)}}\right)\\
&=r(t^{n+1})\Delta t\left(\Delta w^{n+1}+\frac{T_1^n}{\Delta t},\frac{g(u^n)}{\sqrt{E_1^n}}-\frac{g(u(t^n))}{\sqrt{E_1(t^n)}}\right)\\
&\le\frac{\Delta t}{2}\|\nabla w^{n+1}\|^2+C\Delta t\left\|\frac{\nabla g(u^n)}{\sqrt{E_1^n}}-\frac{\nabla g(u(t^n))}{\sqrt{E_1(t^n)}}\right\|^2+\frac{C}{\Delta t}\|(-\Delta)^{-1/2}T_1^n\|^2,\\
&s^{n+1}\left(u(t^{n+1})-u(t^n),\frac{g(u^n)}{\sqrt{E_1^n}}-\frac{g(u(t^n))}{\sqrt{E_1(t^n)}}\right)\\
&\le C\Delta t\|u_t\|_{L^\infty(0,T;H^{-1})}\left((s^{n+1})^2+\left\|\frac{\nabla g(u^n)}{\sqrt{E_1^n}}-\frac{\nabla g(u(t^n))}{\sqrt{E_1(t^n)}}\right\|^2\right). \qquad (4.15)
\end{aligned}
$$

In the above, we define $(-\Delta)^{-1/2}$ by the power of $-\Delta$ by spectral theory of self-adjoint operators, with noticing that $\int_\Omega T_1^n\,d\boldsymbol{x}=0$ because $\int_\Omega u(t)\,d\boldsymbol{x}$ is a constant for the H^{-1} gradient flow.

Now we estimate

$$
\begin{aligned}
&\frac{\nabla g(u^n)}{\sqrt{E_1^n}}-\frac{\nabla g(u(t^n))}{\sqrt{E_1(t^n)}}\\
&=\nabla g(u(t^n))\frac{E_1(t^n)-E_1^n}{\sqrt{E_1(t^n)E_1^n}\cdot(\sqrt{E_1(t^n)}+\sqrt{E_1^n})}+\frac{\nabla g(u^n)-\nabla g(u(t^n))}{\sqrt{E_1^n}}=A_1+A_2. \qquad (4.16)
\end{aligned}
$$

Note that we have (4.3) and (4.4). The first term is bounded by

$$
\|A_1\|\le C\|\nabla g(u(t^n))\|\|e^n\|\le C\|\nabla u\|\|e^n\|\le C\|e^n\|. \qquad (4.17)
$$

For the second term, we have

$$
\begin{aligned}
\|A_2\|\le&C\|\nabla g(u^n)-\nabla g(u(t^n))\|\\
\le&C\|(g'(u^n)-g'(u(t^n)))\nabla u(t^n)\|+C\|g'(u^n)\nabla e^n\|\\
\le&C(\|\nabla u(t^n)e^n\|+\|\nabla e^n\|).
\end{aligned}
$$

Then, by Hölder's inequality and Sobolev embedding, we deduce that

$$
\|A_2\|\le C(\|\nabla u(t^n)\|_{L^3}\|e^n\|_{L^6}+\|\nabla e^n\|)
$$

$$
\begin{aligned}
&\leq C(\|u(t^n)\|_{H^2}\|e^n\|_{H^1} + \|\nabla e^n\|) \\
&\leq C(\|e^n\| + \|\nabla e^n\|). \qquad (4.18)
\end{aligned}
$$

Therefore,

$$
\left\| \frac{\nabla g(u^n)}{\sqrt{E_1^n}} - \frac{\nabla g(u(t^n))}{\sqrt{E_1(t^n)}} \right\|^2 \leq C(\|\nabla e^n\|^2 + \|e^n\|^2). \qquad (4.19)
$$

For the truncation errors, we have the following estimates:

$$
\begin{aligned}
\|(-\Delta)^{-1/2}T_1^n\|^2 \leq & C\Delta t^3 \int_{t^n}^{t^{n+1}} \|(-\Delta)^{-1/2}u_{tt}(s)\|^2\, ds, \\
\|\nabla T_2^n\|^2 \leq & C\Big(\|\nabla\big(u(t^n) - u(t^{n+1})\big)\|^2 + \|u(t^n) - u(t^{n+1})\|^2\Big) \\
\leq & C\Delta t \int_{t^n}^{t^{n+1}} \|u_t(s)\|_{H^1}^2\, ds, \\
|v_1^n|^2 \leq & C\Delta t^3 \int_{t^n}^{t^{n+1}} |r_{tt}(s)|^2\, ds, \\
|v_2^n|^2 \leq & C\Delta t^3 \|\nabla g(u(t^n))\|^2 \int_{t^n}^{t^{n+1}} \|(-\Delta)^{-1/2}u_{tt}(s)\|^2\, ds \\
\leq & C\Delta t^3 \|\nabla u(t^n)\|^2 \int_{t^n}^{t^{n+1}} \|(-\Delta)^{-1/2}u_{tt}(s)\|^2\, ds \\
\leq & C\Delta t^3 \int_{t^n}^{t^{n+1}} \|(-\Delta)^{-1/2}u_{tt}(s)\|^2\, ds,
\end{aligned}
$$

where we utilized (4.19) for T_2^n, and (4.3) and (4.4) for v_2^n. Therefore,

$$
\begin{aligned}
2s^{n+1}(-v_1^n + v_2^n) \leq & \Delta t (s^{n+1})^2 + \frac{2}{\Delta t}\big((v_1^n)^2 + (v_2^n)^2\big) \\
\leq & \Delta t (s^{n+1})^2 + C\Delta t^2 \int_{t^n}^{t^{n+1}} |r_{tt}(s)|^2 + \|(-\Delta)^{-1/2}u_{tt}(s)\|^2 ds, \\
(w^{n+1}, T_1^n) \leq & \frac{\Delta t}{4}\|\nabla w^{n+1}\|^2 + \frac{1}{\Delta t}\|(-\Delta)^{-1/2}T_1^n\|^2 \\
\leq & \frac{\Delta t}{4}\|\nabla w^{n+1}\|^2 + C\Delta t^2 \int_{t^n}^{t^{n+1}} \|(-\Delta)^{-1/2}u_{tt}(s)\|^2 ds, \\
-(e^{n+1} - e^n, T_2^n) = & -\Delta t \bigg(\Delta w^{n+1} + \frac{T_1^n}{\Delta t}, T_2^n\bigg) \\
\leq & \frac{\Delta t}{4}\|\nabla w^{n+1}\|^2 + C\Delta t \|\nabla T_2^n\|^2 + \frac{C}{\Delta t}\|(-\Delta)^{-1/2}T_1^n\|^2 \\
\leq & \frac{\Delta t}{4}\|\nabla w^{n+1}\|^2 + C\Delta t^2 \int_{t^n}^{t^{n+1}} \|u_t(s)\|_{H^1}^2 + \|(-\Delta)^{-1/2}u_{tt}(s)\|^2\, ds. \qquad (4.20)
\end{aligned}
$$

Combining (4.14), (4.15), and (4.20), we obtain

$$
\begin{aligned}
&\frac{\lambda}{2}(\|e^{n+1}\|^2 - \|e^n\|^2) + \frac{1}{2}(\|\nabla e^{n+1}\|^2 - \|\nabla e^n\|^2) + (s^{n+1})^2 - (s^n)^2 \\
&\leq C\Delta t(\|e^{n+1}\|^2 + \|e^n\|^2 + \|\nabla e^{n+1}\|^2 + \|\nabla e^n\|^2 + (s^{n+1})^2 + (s^n)^2)
\end{aligned}
$$

$$+ C\Delta t^2 \int_{t^n}^{t^{n+1}} (|r_{tt}(s)|^2 + \|(-\Delta)^{-1/2}u_{tt}(s)\|^2 + \|u_t(s)\|^2_{H^1})ds. \tag{4.21}$$

By noting (4.6), we can conclude the proof by applying the discrete Gronwall's inequality (see, for example, [16, p. 15]) to the above. The constant in (4.2) also depends on $\|\nabla u\|_{L^\infty((0,T)\times\Omega)}$ and $\|u_t\|_{L^\infty(0,T;H^{-1})}$ as they appear in (4.6) and (4.15), respectively. ∎

4.2. L^2 gradient flow. For L^2 gradient flow, we shall only state the error estimates below, as their proofs are essentially the same as for the H^{-1} gradient flow.

THEOREM 4.2. *For the L^2 gradient flow, assume $u^0 \in H^3$ and (2.6) holds. In addition, we assume that*

$$u_t \in L^\infty(0,T;L^2) \cap L^2(0,T;L^4), \quad u_{tt} \in L^2(0,T;L^2). \tag{4.22}$$

Then for all $n \leq T/\Delta t$, we have

$$\begin{aligned} &\frac{1}{2}\|\nabla e^n\|^2 + \frac{\lambda}{2}\|e^n\|^2 + (s^n)^2 \\ &\leq C \exp\Big((1-C\Delta t)^{-1}t^n\Big)\Delta t^2 \int_0^{t^n} (\|u_{tt}(s)\|^2 + \|u_t(s)\|^2_{L^4})ds. \end{aligned} \tag{4.23}$$

The constant C is dependent on T, u^0, Ω, and $\|u_t\|_{L^\infty(0,T;L^2)}$.

5. Miscellaneous extensions. In this section, we discuss some miscellaneous extensions. First, we discuss how the convergence and error analysis can be extended to second-order SAV schemes. Then, we consider several other gradient flows, which are not in the form of (1.3), but can still be dealt with similarly as above.

5.1. Second-order schemes. Since the second order BDF2 and Crank–Nicolson schemes [17, 18] also enjoy the unconditional energy stability similar to (1.8), we can derive results similar to Lemmas 2.3 and 2.4. Therefore, we can also establish error estimates for the second-order schemes using a similar procedure, but with stronger regularity assumptions than (4.1). We state below the result for the Crank–Nicolson SAV scheme to the H^{-1} gradient flow, given by

$$\frac{u^{n+1}-u^n}{\Delta t} = \Delta \mu^{n+1/2}, \tag{5.1a}$$

$$\mu^{n+1/2} = -\Delta u^{n+1/2} + \lambda u^{n+1/2} + \frac{r^{n+1/2}}{\sqrt{E_1[\bar{u}]}} g(\bar{u}), \tag{5.1b}$$

$$r^{n+1} - r^n = \frac{1}{2\sqrt{E_1[\bar{u}]}} \int_\Omega g(\bar{u})(u^{n+1}-u^n)d\boldsymbol{x}, \tag{5.1c}$$

with

$$u^{n+1/2} = \frac{1}{2}(u^{n+1}+u^n), \quad r^{n+1/2} = \frac{1}{2}(r^{n+1}+r^n), \quad \bar{u} = \frac{1}{2}(3u^n - u^{n-1}). \tag{5.2}$$

THEOREM 5.1. *For the H^{-1} gradient flow, assume that $u^0 \in H^4$ and (2.6), (2.7) hold. In addition, we assume that*

$$u \in L^\infty(0,T;W^{1,\infty}), \quad u_t \in L^\infty(0,T;H^{-1}) \cap L^2(0,T;H^1),$$

$$u_{tt} \in L^2(0,T;H^3), \quad u_{ttt} \in L^2(0,T;H^{-1}). \tag{5.3}$$

Then, for the Crank–Nicolson SAV scheme (5.1) *with* $n \le T/\Delta t$, *we have*

$$\frac{1}{2}\|\nabla e^n\|^2 + \frac{\lambda}{2}\|e^n\|^2 + (s^n)^2 \le C\exp\Big((1-C\Delta t)^{-1}t^n\Big)\Delta t^4$$
$$\int_0^{t^n} (\|u_t(s)\|^2_{H^1} + \|u_{ttt}(s)\|^2_{H^{-1}} + \|u_{tt}(s)\|^2_{H^3})ds. \tag{5.4}$$

The constant C *is dependent on* T, u^0, Ω, $\|u\|_{L^\infty(0,T;W^{1,\infty})}$, *and* $\|u_t\|_{L^\infty(0,T;H^{-1})}$.

Similar results can be derived for the Crank–Nicolson SAV scheme to the L^2 gradient flow.

5.2. Gradient flows about several functions. Many physical systems are described by several functions, such as multiphase flows (see [3] and the references therein). We consider the following energy functional:

$$E(\phi_1,\dots,\phi_l) = \sum_{i=1}^{k}\int_\Omega \frac{1}{2}|\nabla\phi_i|^2 d\boldsymbol{x} + E_1(\phi_1,\dots,\phi_l). \tag{5.5}$$

Denote $\Phi = (\phi_1,\dots,\phi_l)^t$. We assume that $E_1(\Phi) = \int_\Omega F(\Phi)d\boldsymbol{x} > -c_0$. Let $U_i = \partial F/\partial\phi_i$. Then the H^{-1} gradient flow is given by

$$\frac{\partial\phi_i}{\partial t} = \Delta\mu_i = \Delta(-\Delta\phi_i + U_i). \tag{5.6}$$

If the nonlinear term F satisfies conditions similar to (2.6) and (2.7) about ϕ_k, i.e.,

$$\left|\frac{\partial U_i}{\partial\phi_j}\right| \le C\bigg(\sum_k |\phi_k|^p + 1\bigg), \quad p>0 \text{ arbitrary if } n=1,2; \quad 0<p<4 \text{ if } n=3; \tag{5.7}$$

$$\left|\frac{\partial^2 U_i}{\partial\phi_j\phi_{j'}}\right| \le C\bigg(\sum_k |\phi_k|^{p'} + 1\bigg), \quad p'>0 \text{ arbitrary if } n=1,2; \quad 0<p'<3 \text{ if } n=3, \tag{5.8}$$

then we can repeat the same procedure to obtain the convergence and error estimate.

5.3. Phase field crystal. In models that describe modulated structures, higher-order linear operators will take the place of the Laplacian. A typical example is the phase field crystal equation [8],

$$\frac{\partial u}{\partial t} = \Delta\mu = \Delta((\Delta+1)^2 u + \lambda u + g(u)), \tag{5.9}$$

where we require $\lambda > 0$. Another example is the Lifshitz–Petrich model [15] for quasi-crystals, where $(\Delta+1)^2$ is substituted with $(\Delta+1)^2(\Delta+q^2)^2$. In these cases, the energy dissipation itself gives higher regularity. Take (5.9), for example. If $u^0 \in H^2$, then the energy dissipation indicates that $u \in L^\infty(0,T;H^2)$, both for the exact and SAV solutions. Therefore, we can follow the same procedure to obtain the convergence like in Theorem 3.1, without any further assumptions about u^0 and g. For the error estimate, we need only assume (4.1).

5.4. *Q*-tensor theory. Q-tensor theory [6] is a widely used model describing nematic phases of rod-like liquid crystals. We consider the L^2 gradient flow of $E[Q(\boldsymbol{x})] = E_b + E_e$, where $Q \in \mathbb{R}^{3\times 3}$ is a symmetric traceless second-order tensor, and

$$E_b = \int_\Omega f_b(Q)d\boldsymbol{x} = \int_\Omega \left[\frac{a}{2}\mathrm{tr}(Q^2) - \frac{b}{3}\mathrm{tr}(Q^3) + \frac{c}{4}(\mathrm{tr}(Q^2))^2\right]d\boldsymbol{x}, \tag{5.10}$$

$$E_e = \int_\Omega \left[\frac{L_1}{2}|\nabla Q|^2 + \frac{L_2}{2}\partial_i Q_{ik}\partial_j Q_{jk} + \frac{L_3}{2}\partial_i Q_{jk}\partial_j Q_{ik}\right]d\boldsymbol{x}. \tag{5.11}$$

To ensure the lower-boundedness, it requires $c, L_1, L_1 + L_2 + L_3 > 0$ so that we have $E_b,\ E_e \geq 0$.

We note that the nonlinear terms are fourth-order polynomials. This will be sufficient for us to derive the estimate like (2.10). Then we can derive similar H^2 estimates like Theorem 2.6 and Lemma 2.7. The convergence and error analysis will follow from these estimates.

5.5. Molecular beam epitaxy (MBE) equation. The MBE equation (see, for example, [14]) describes the evolution of the height of a thin film. We consider the L^2 gradient flow of the energy,

$$E[u] = \int_\Omega \left[\frac{1}{4}(1 - |\nabla u|^2)^2 + \frac{\eta^2}{2}|\Delta u|^2\right]d\boldsymbol{x}. \tag{5.12}$$

The nonlinear term contains ∇u. To deal with this term, we can utilize the estimate (2.10) with u replaced by ∇u. Then we can derive an H^3 estimate, which is sufficient for the convergence and error analysis.

6. Concluding remarks. We carried out convergence and error analysis of the SAV schemes for L^2 and H^{-1} gradient flows with a typical form of free energy. Using the unconditional energy stability of the SAV schemes, we first derive H^2 estimates, which enabled us to prove convergence results under very mild conditions. We then derived error estimates by assuming more regularity on the exact solution. Note that these results are derived for a large class of free energies, in particular, without assuming the Lipschitz condition (1.1), which are usually required for the stability and error analysis of semi-implicit schemes.

We have also indicated that the convergence and error analysis presented in this paper can be extended to SAV schemes for several other gradient flows which cannot be cast in the general form considered in this paper.

REFERENCES

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

[2] A. Baskaran, J. S. Lowengrub, C. Wang, and S. M. Wise, *Convergence analysis of a second order convex splitting scheme for the modified phase field crystal equation*, SIAM J. Numer. Anal., 51 (2013), pp. 2851–2873, https://doi.org/10.1137/120880677.

[3] F. Boyer and S. Minjeaud, *Hierarchy of consistent n-component Cahn–Hilliard systems*, Math. Models Methods Appl. Sci., 24 (2014), pp. 2885–2928.

[4] 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.

[5] N. Condette, C. Melcher, and E. Süli, *Spectral approximation of pattern-forming nonlinear evolution equations with double-well potentials of quadratic growth*, Math. Comp., 80 (2011), pp. 205–223.

[6] P. De Gennes, *Short range order effects in the isotropic phase of nematics and cholesterics*, Mol. Cryst. Liq. Cryst., 12 (1971), pp. 193–214.

[7] Q. Du and R. A. Nicolaides, *Numerical analysis of a continuum model of phase transition*, SIAM J. Numer. Anal., 28 (1991), pp. 1310–1322, https://doi.org/10.1137/0728069.

[8] K. Elder and M. Grant, *Modeling elastic and plastic deformations in nonequilibrium processing using phase field crystals*, Phys. Rev. E, 70 (2004), 051605.

[9] C. M. Elliott and A. M. Stuart, *The global dynamics of discrete semilinear parabolic equations*, SIAM J. Numer. Anal., 30 (1993), pp. 1622–1663, https://doi.org/10.1137/0730084.

[10] 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.

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

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

[13] D. Kessler, R. H. Nochetto, and A. Schmidt, *A posteriori error control for the Allen–Cahn problem: Circumventing Gronwall's inequality*, M2AN Math. Model. Numer. Anal., 38 (2004), pp. 129–142.

[14] B. Li and J.-G. Liu, *Thin film epitaxy with or without slope selection*, European J. Appl. Math., 14 (2003), pp. 713–743.

[15] R. Lifshitz and D. M. Petrich, *Theoretical model for faraday waves with multiple-frequency forcing*, Phys. Rev. Lett., 79 (1997), 1261.

[16] A. Quarteroni and A. Valli, *Numerical Approximation of Partial Differential Equations*, Springer-Verlag, Berlin, 2008.

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

[18] J. Shen, J. Xu, and J. Yang, *A new class of efficient and robust energy stable schemes for gradient flows*, SIAM Rev., to appear; available online at https://arxiv.org/abs/1710.01331.

[19] J. Shen and X. Yang, *Energy stable schemes for Cahn–Hilliard phase-field model of two phase incompressible flows equations*, Chin. Ann. Math. Ser. B, 31 (2010), pp. 743–758.

[20] R. Temam, *Navier-Stokes Equations*, North-Holland, Amsterdam, 1984.

[21] R. Temam, *Infinite-Dimensional Dynamical Systems in Mechanics and Physics*, 2nd ed., Appl. Math. Sci. 68, Springer-Verlag, New York, 1997.

[22] 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.

[23] X. Ye, *The Legendre collocation method for the Cahn–Hilliard equation*, J. Comput. Appl. Math., 150 (2003), pp. 87–108.