Analysis of Total Variation Flow and Its Finite Element Approximations*

Xiaobing Feng
The University of Tennessee

Andreas Prohl
ETH Zürich

Abstract

We study the gradient flow for the total variation functional, which arises in image processing and geometric applications. We propose a variational inequality weak formulation for the gradient flow, and establish well-posedness of the problem by the energy method. The main idea of our approach is to exploit the relationship between the regularized gradient flow (characterized by a small positive parameter ε, see (1.7)) and the minimal surface flow [19] and the prescribed mean curvature flow [15]. Since our approach is constructive and variational, finite element methods can be naturally applied to approximate weak solutions of the limiting gradient flow problem. We propose a fully discrete finite element method and establish convergence to the regularized gradient flow problem as $h, k \to 0$, and to the total variation gradient flow problem as $h, k, \varepsilon \to 0$ in general cases. Provided that the regularized gradient flow problem possesses strong solutions, which is proved possible if the datum functions are regular enough, we establish practical a priori error estimates for the fully discrete finite element solution, in particular, by focusing on the dependence of the error bounds on the regularization parameter ε. Optimal order error bounds are derived for the numerical solution under the mesh relation $k = O(h^2)$. In particular, it is shown that all error bounds depend on $\frac{1}{\varepsilon}$ only in some lower polynomial order for small ε.

1 Introduction and Summary

One of the best known and most successful noise removal and image restoration model in image processing is the total variation (TV) model due to Rudin, Osher and Fatemi [22]. Let $u : \Omega \subset \mathbf{R}^2 \to \mathbf{R}$ denote the gray level of an image describing a real scene, and g be the observed image of the same scene, which usually is a degradation of u. The total variation model recovers the image u by minimizing the total variation functional

$$J(u) := \int_\Omega |\, \nabla u \,|\, dx \tag{1.1}$$

on $BV(\Omega)$, the space of functions of bounded variation (see Section 2 for the precise definition), subject to the constraint

$$Au + \eta = g\,. \tag{1.2}$$

*Submitted to Comm. Pure & Applied Math.

Here, A is a (known) linear operator representing the blur and η denotes an additive white Gaussian noise. For the sake of clarity of the presentation, in this paper we set $A = I$, the identity operator.

To avoid solving the constrained minimization problem, one strategy is to enforce the constraint weakly and reformulate the problem as an unconstrained minimization problem which minimizes the following penalized functional

$$J_\lambda(u) := \int_\Omega |\,\nabla u\,|\,dx + \frac{\lambda}{2}\int_\Omega |u-g|^2\,dx\,, \tag{1.3}$$

where $\lambda \geq 0$ is the penalization parameter which controls the trade-off between goodness of fit-to-the-data and variability in u.

A well-known method for solving the above minimization problem is the steepest descent method, which motivates to consider its gradient flow:

$$\frac{\partial u}{\partial t} = \operatorname{div}\Big(\frac{\nabla u}{|\,\nabla u\,|}\Big) - \lambda(u-g) \qquad \text{in } \Omega_T \equiv \Omega\times(0,T), \tag{1.4}$$

$$\frac{\partial u}{\partial n} = 0 \qquad \text{on } \partial\Omega_T \equiv \partial\Omega\times(0,T), \tag{1.5}$$

$$u(\cdot,0) = u_0(\cdot) \qquad \text{in } \Omega, \tag{1.6}$$

for a positive number T and an initial guess u_0. The above gradient flow will be referred to as *TV flow* in the rest of this paper. We remark that the above initial-boundary value problem with $\lambda = 0$ also arises in geometric measure theory for studying the evolution of a set with finite perimeter without distortion of the boundary [5].

Although the above TV flow has been addressed and approximated numerically by many authors (see [13, 9, 10] and references therein), its rigorous mathematical analysis has appeared in the literature very recently. The first such work was given by Andreu-Ballester-Caselles-Mazón in [3], in which (1.4)-(1.6) (for $\lambda = 0$, and $u_0 \in L^1(\Omega)$) was defined on the space $L^1\big((0,T);BV(\Omega)\big)$ as a variational inequality problem. Besides several other results, existence and uniqueness of weak solutions was proved by using Crandall-Liggett's semigroup generation theory [12]. On the other hand, numerical simulations were done by computing the solution of the following regularized problem

$$\frac{\partial u^\varepsilon}{\partial t} = \operatorname{div}\Big(\frac{\nabla u^\varepsilon}{\sqrt{|\,\nabla u^\varepsilon\,|^2+\varepsilon^2}}\Big) - \lambda(u^\varepsilon - g) \qquad \text{in } \Omega_T, \tag{1.7}$$

$$\frac{\partial u^\varepsilon}{\partial n} = 0 \qquad \text{on } \partial\Omega_T, \tag{1.8}$$

$$u^\varepsilon(\cdot,0) = u_0(\cdot) \qquad \text{in } \Omega, \tag{1.9}$$

for small $\varepsilon > 0$. Later, the results of [3] were extended to the TV flow with Dirichlet boundary conditions in [2], and further qualitative properties of the flow were addressed in [4].

It is easy to see that the equation (1.7) corresponds to the gradient flow for the energy functional

$$J_{\lambda,\varepsilon}(u) := \int_\Omega \sqrt{|\,\nabla u\,|^2+\varepsilon^2}\,dx + \frac{\lambda}{2}\int_\Omega |u-g|^2\,dx\,, \tag{1.10}$$

which is a (strictly) convex regularization to the total variation functional (1.1). In fact, this is the mostly used regularization technique to approximate and compute the minimizer of the total variation energy and its variants (cf. [9, 13, 8] and references therein).

The goal of this paper is to present an L^2-variational theory for the TV flow, based on the regularized gradient flow (1.7)-(1.9). In comparison with the approach of [3, 2], we give a simpler and more natural notion of weak solution for the TV flow, and establish well-posedness and regularities for the problem using *the energy method*, which is surprisingly "easy and short". Since this approach is based on analyzing the regularized gradient flow (1.7)-(1.9) and establishing the connection between regularized and limiting gradient flows, as a result, this paper also provides a qualitative analysis for the most widely used regularization technique for approximating and computing the solutions of the TV flow (and its stationary counterpart) through approximation of the regularized flow (and its stationary counterpart). In addition, our analytical results lay down the theoretical basis for analyzing convergence and error estimates for finite element and other numerical approximations of both gradient flows (and their stationary counterparts). The crux of our approach is to exploit the fact that equation (1.7) resembles the minimal surface flow [19] if $\lambda = 0$ and the prescribed mean curvature flow [15] if $\lambda \neq 0$. Both problems correspond to the case $\varepsilon = 1$. For other values of $\varepsilon > 0$, we introduce the scaling

$$\tau = \varepsilon\, t, \qquad y = \varepsilon\, x, \qquad T^\varepsilon = \varepsilon\, T, \qquad \Omega^\varepsilon = \varepsilon\, \Omega, \tag{1.11}$$

and define $v^\varepsilon(y,\tau) = u^\varepsilon(x,t)$, $g^\varepsilon(y) = g(x)$ and $v_0(y) = u_0(x)$. It is then easy to check that the function v^ε satisfies

$$\frac{\partial v^\varepsilon}{\partial \tau} = \operatorname{div}\Bigl(\frac{\nabla v^\varepsilon}{\sqrt{|\nabla v^\varepsilon|^2 + 1}}\Bigr) - \frac{\lambda}{\varepsilon}(v^\varepsilon - g^\varepsilon) \qquad \text{in } \Omega^\varepsilon_{T^\varepsilon} \equiv \Omega^\varepsilon \times (0, T^\varepsilon), \tag{1.12}$$

$$\frac{\partial v^\varepsilon}{\partial n_y} = 0 \qquad \text{on } \partial\Omega^\varepsilon_{T^\varepsilon} \equiv \partial\Omega^\varepsilon \times (0, T^\varepsilon), \tag{1.13}$$

$$v^\varepsilon(\cdot, 0) = v_0(\cdot) \qquad \text{in } \Omega^\varepsilon. \tag{1.14}$$

The above simple observation states that for each fixed $\varepsilon > 0$, the function $u^\varepsilon = v^\varepsilon$ evolves as a minimal surface flow if $\lambda = 0$, and a prescribed mean curvature flow if $\lambda \neq 0$ in the *scaled* coordinates (y, τ). We remark that since $\varepsilon \ll 1$, hence $\tau = \varepsilon\, t$ represents a *slow time.*

The minimal surface flow and the prescribed mean curvature flow on a fixed domain have been understood for both Dirichlet and Neumann boundary conditions. We refer to [19, 15] for detailed discussions. For the corresponding stationary problems, extensive research has been carried out in the past forty years, we refer to [16, 17] and the references therein for detailed expositions. In order to analyze the gradient flow (1.7)-(1.9) and its limiting flow (1.4)-(1.6) as $\varepsilon \to 0$, it is crucial for us to keep track of the dependence of the solution u^ε on the regularization parameter ε. Also, it is worth noting that both stationary and evolutionary surface of prescribed mean curvature problems do not have "regular" solutions unless the mean curvature of the boundary $\partial\Omega$ of the domain Ω is everywhere non-negative (cf. Chapter 16 of [16]), and solutions $u(t) \in BV(\Omega)$ for a.e. $t \geq 0$ are what one can only get in general (cf. [17, 15]). Knowledge of these facts is helpful for developing an appropriate analytical setting for the gradient flow (1.7)-(1.9) and particularly its limiting flow (1.4)-(1.6), as $\varepsilon \to 0$.

In the rest of this section, we shall summarize the main results of this paper. Our first theorem addresses existence and uniqueness for the gradient flow (1.7)-(1.9). The statement is in the spirit of [19, 15] for minimal surface and prescribed mean curvature flow, respectively. The main difference in this paper is that we deal with less regular data u_0 and g, and we also trace dependence of a priori estimates on the parameter ε.

Theorem 1.1 *Let $\Omega \subset \mathbf{R}^N$ $(N \geq 2)$ be a bounded open domain with Lipschitz boundary $\partial\Omega$. Suppose that $u_0, g \in L^2(\Omega)$. Then, there exists a unique function $u^\varepsilon \in L^1((0,T); BV(\Omega)) \cap C^0([0,T]; L^2(\Omega))$ such that*

$$u^\varepsilon(0) = u_0, \qquad u^\varepsilon_t \in L^2\bigl((0,T); H^{-1}(\Omega)\bigr), \tag{1.15}$$

and for any $s \in [0,T]$

$$\int_0^s \int_\Omega v_t(v-u^\varepsilon)\,dxdt + \int_0^s \big[J_{\lambda,\varepsilon}(v) - J_{\lambda,\varepsilon}(u^\varepsilon)\big]\,dt \geq \frac{1}{2}\big[\,\| v(s)-u^\varepsilon(s)\|_{L^2}^2 - \| v(0)-u_0\|_{L^2}^2\,\big] \tag{1.16}$$

$$\forall\, v \in L^1\big((0,T);BV(\Omega)\big)\cap L^2(\Omega_T) \text{ such that } v_t \in L^2(\Omega_T).$$

Moreover, suppose u_i^ε $(i=1,2)$ *are two functions which satisfy (1.16) with respective datum functions* $u_i^\varepsilon(0)$, g_i^ε $(i=1,2)$*. Then, there holds*

$$\| u_1^\varepsilon(s) - u_2^\varepsilon(s)\|_{L^2} \leq \| u_1^\varepsilon(0) - u_2^\varepsilon(0)\|_{L^2} + \sqrt{\lambda}\| g_1^\varepsilon - g_2^\varepsilon\|_{L^2} \qquad \forall s \in [0,T]. \tag{1.17}$$

Remark 1.1 *Borrowing the idea of [19, 15], we shall define a weak solution of the gradient flow (1.7)-(1.9) as a function* $u^\varepsilon \in L^1((0,T);BV(\Omega))\cap C^0([0,T];L^2(\Omega))$ *which satisfies (1.15)-(1.16). The motivation for such a definition is nicely explained in [19].*

Remark 1.2 *Following [19, 15], the results of Theorem 1.1 can be easily generalized to the cases of non-homogeneous Neumann and Dirichlet boundary conditions, under some appropriate assumptions on the boundary data. We particularly mention that in the case of the non-homogeneous Dirichlet boundary condition*

$$u^\varepsilon = \phi \qquad \text{on } \partial\Omega\times(0,T),$$

the only modification that needs to be done is to replace the energy functional $J_{\lambda,\varepsilon}(\cdot)$ *by the energy functional*

$$\Phi_{\lambda,\varepsilon}(u) := J_{\lambda,\varepsilon}(u) + \int_{\partial\Omega} |u-\phi|dx\,, \tag{1.18}$$

where the Dirichlet datum is enforced weakly (see [17, 19, 15] for more discussions). Then, all results of Theorem 1.1 can be extended to this case under some suitable assumptions on ϕ*, and particularly the analysis remains same.*

Our second theorem states some regularity results and a priori estimates (with emphasis on their dependence on ε) for the solution of the gradient flow (1.7)-(1.9).

Theorem 1.2 *Let* $\Omega \subset \mathbf{R}^N$ $(N\geq 2)$ *be a bounded open domain with Lipschitz boundary* $\partial\Omega$ *and* u^ε *is the function whose existence is given by Theorem 1.1 above.*

(i). If $u_0 \in BV(\Omega)$ *and* $g \in L^2(\Omega)$*, then*

$$u^\varepsilon \in L^\infty\big((0,T);BV(\Omega)\big), \qquad u_t^\varepsilon \in L^2(\Omega_T), \tag{1.19}$$

and for any $s \in [0,T]$

$$\int_0^s \int_\Omega u_t^\varepsilon(v-u^\varepsilon)\,dxdt + \int_0^s \big[J_{\lambda,\varepsilon}(v) - J_{\lambda,\varepsilon}(u^\varepsilon)\big] \geq 0 \tag{1.20}$$

$$\forall v \in L^1\big((0,T);BV(\Omega)\big)\cap L^2(\Omega_T).$$

(ii). If $\partial\Omega$ *satisfies an internal sphere condition (ISC) of radius* R *(cf. Definition 2.4 of [15]),* $u_0 \in BV(\Omega)\cap L^\infty(\Omega)$ *and* $g \in L^\infty(\Omega)$*, then*

$$\sup_{(x,t)\in\Omega_T} |u^\varepsilon(x,t)| \leq \sup_{x\in\Omega} |u_0(x)| + R\varepsilon + T\Big(\frac{N}{R} + \lambda\| g\|_{L^\infty(\Omega)}\Big). \tag{1.21}$$

(iii). If $u_0 \in H^1_{loc}(\Omega) \cap W^{1,1}(\Omega)$, $g \in L^2(\Omega) \cap H^1_{loc}(\Omega)$ and $\partial\Omega \in C^2$, then, in addition to (1.19)-(1.20), $u^\varepsilon \in L^\infty\big((0,T); W^{1,1}(\Omega)\big) \cap L^\infty\big((0,T); H^1_{loc}(\Omega)\big)$, and the equations (1.7)-(1.9) hold in Ω_T and on $\partial\Omega_T$, respectively in the distributional sense. Moreover, u^ε satisfies the following dissipative energy law:

$$\frac{d}{dt} J_{\lambda,\varepsilon}(u^\varepsilon) = -\| u^\varepsilon_t \|^2_{L^2} \qquad \text{for a.e. } t \in [0,T]. \tag{1.22}$$

(iv). If $u_0 \in C^2(\overline{\Omega})$, $g \in W^{1,\infty}(\Omega)$ and $\partial\Omega \in C^3$, then $u^\varepsilon \in W^{1,\infty}\big(\Omega_T\big) \cap L^2\big((0,T); H^2(\Omega)\big)$. Moreover, the following high order dissipative energy law is valid:

$$\frac{d}{dt}\| u^\varepsilon_t \|^2_{L^2} = -2\| \nabla u^\varepsilon_t \, (|\nabla u^\varepsilon|^2 + \varepsilon^2)^{-\frac{1}{4}} \|^2_{L^2} - 2\| \nabla u^\varepsilon \cdot \nabla u^\varepsilon_t \, (|\nabla u^\varepsilon|^2 + \varepsilon^2)^{-\frac{3}{4}} \|^2_{L^2} \\ -2\lambda \| u^\varepsilon_t \|^2_{L^2}, \qquad \text{for a.e. } t \in [0,T], \tag{1.23}$$

as well as the following estimates:

$$\| \nabla u^\varepsilon \|_{L^\infty(\Omega_T)} \leq \hat{C}_0(\varepsilon^{-1}), \tag{1.24}$$

$$\| u^\varepsilon \|_{L^2(H^2)} \leq \hat{C}_1(\varepsilon^{-1}), \tag{1.25}$$

$$\| u^\varepsilon_{tt} \|_{L^2(H^{-1})} \leq \hat{C}_2(\varepsilon^{-1}), \tag{1.26}$$

where $\hat{C}_j(\varepsilon^{-1})$ $(j = 0,1,2)$ are some positive, low order polynomial functions in ε^{-1}.

Remark 1.3 *(a). Remarks 1.1 and 1.2 remain valid for Theorem 1.2.*

(b). In the context of image processing, $\partial\Omega$ is usually piecewise smooth and the observed image $g \in L^\infty(\Omega)$, although the initial value is less restrictive. Hence, only weak solutions are expected for the gradient flow (1.7)-(1.9) in general.

(c). It is possible to obtain the regularity results of Theorem 1.2 under weaker assumptions (still too strong for image processing applications) on the function g and on the boundary $\partial\Omega$. However, no attempt is made to address this issue in the present paper.

Our third main theorem establishes existence and uniqueness of solutions for the TV flow; it also proves that the TV flow is indeed the limiting problem of the gradient flow (1.7)-(1.9) as the parameter $\varepsilon \to 0$. To the best of our knowledge, this latter fact has been assumed and used in the literature for numerical simulations without rigorous justification.

Theorem 1.3 *Let $\Omega \subset \mathbf{R}^N$ $(N \geq 2)$ be a bounded open domain with Lipschitz boundary $\partial\Omega$ and u_0, $g \in L^2(\Omega)$.*

(i). There exists a unique function $u \in L^1\big((0,T); BV(\Omega)\big) \cap C^0\big([0,T]; L^2(\Omega)\big)$ such that

$$u(0) = u_0, \qquad u_t \in L^2\big((0,T); H^{-1}(\Omega)\big), \tag{1.27}$$

and for any $s \in [0,T]$

$$\int_0^s \int_\Omega v_t (v - u)\, dxdt + \int_0^s \big[J_\lambda(v) - J_\lambda(u) \big]\, dt \\ \geq \frac{1}{2}\big[\, \| v(s) - u(s) \|^2_{L^2} - \| v(0) - u_0 \|^2_{L^2} \,\big] \\ \forall\, v \in L^1\big((0,T); BV(\Omega)\big) \cap L^2(\Omega_T) \text{ such that } v_t \in L^2(\Omega_T). \tag{1.28}$$

(ii). Suppose $u_i\,(i=1,2)$ are two functions which satisfy (1.28) with respect to given data $u_i(0)$, $g_i\,(i=1,2)$. Then

$$\|\, u_1(s)-u_2(s)\,\|_{L^2} \le \|\, u_1(0)-u_2(0)\,\|_{L^2} + \sqrt{\lambda}\|\, g_1-g_2\,\|_{L^2} \qquad \forall s\in[0,T]. \tag{1.29}$$

(iii). Let u^ε be the weak solution of the gradient flow (1.7)-(1.9) as stated in Theorem 1.1, then there holds

$$\lim_{\varepsilon\to 0} \|\, u^\varepsilon(t)-u(t)\,\|_{L^p(\Omega)} = 0 \quad \textit{for a.e. } t\in(0,T),\ \forall\, p\in[1,\frac{N}{N-1}), \tag{1.30}$$

$$u_t^\varepsilon \longrightarrow u_t \qquad \textit{weakly in } L^2\big((0,T);H^{-1}(\Omega)\big). \tag{1.31}$$

Remark 1.4 *For the same reason as stated in Remark 1.1, a weak solution of the TV flow (1.4)-(1.6) will be defined as a function $u\in L^1\big((0,T);BV(\Omega)\big)$ which satisfies (1.27)-(1.28). Clearly, this definition comes naturally in view of Theorem 1.1 and the convergence result (1.30).*

The following result is related to Theorem 1.3 like Theorem 1.2 to Theorem 1.1.

Theorem 1.4 *Let $\Omega\subset\mathbf{R}^N\,(N\ge 2)$ be a bounded open domain with Lipschitz boundary $\partial\Omega$. Let u be the function whose existence is given by Theorem 1.3 above.*

(i). If $u_0\in BV(\Omega)$ and $g\in L^2(\Omega)$, then

$$u\in L^\infty\big((0,T);BV(\Omega)\big), \qquad u_t\in L^2(\Omega_T), \tag{1.32}$$

and for any $s\in[0,T]$

$$\int_0^s\int_\Omega u_t(v-u)\,dxdt + \int_0^s \big[J_\lambda(v)-J_\lambda(u)\big]\,dt \ge 0 \qquad \forall\, v\in L^1\big((0,T);BV(\Omega)\big)\cap L^2(\Omega_T). \tag{1.33}$$

(ii). Let u^ε be the weak solution of the gradient flow (1.7)-(1.9) as stated in Theorem 1.2. Under the assumptions of (i) above, we have

$$\lim_{\varepsilon\to 0}\big(\,\|\, u^\varepsilon-u\,\|_{L^\infty((0,T);L^p(\Omega))} + \|\, u_t^\varepsilon-u_t\,\|_{L^2((0,T);H^{-1}(\Omega))}\,\big)=0, \tag{1.34}$$

$$u_t^\varepsilon \longrightarrow u_t \qquad \textit{weakly in } L^2\big((0,T);L^2(\Omega)\big), \tag{1.35}$$

for any $p\in[1,\frac{N}{N-1})$.

(iii). If $\partial\Omega$ satisfies an internal sphere condition (ISC) of radius R, $g\in L^\infty(\Omega)$ and $u_0\in L^\infty(\Omega)\cap BV(\Omega)$, then,

$$\sup_{(x,t)\in\Omega_T}|u(x,t)| \le \sup_{x\in\Omega}|u_0(x)| + T\Big(\frac{N}{R}+\lambda\|\, g\,\|_{L^\infty(\Omega)}\Big). \tag{1.36}$$

(iv). If $u_0\in H^1_{loc}(\Omega)\cap W^{1,1}(\Omega)$, $g\in L^2(\Omega)\cap H^1_{loc}(\Omega)$ and $\partial\Omega\in C^2$, then, in addition to the conclusion of (i), we also have $u\in L^\infty\big((0,T);W^{1,1}(\Omega)\big)\cap L^\infty\big((0,T);H^1_{loc}(\Omega)\big)$, and

$$\frac{\partial u}{\partial t} = \operatorname{div}\big(\mathrm{Sgn}(\nabla u)\big) - \lambda(u-g) \qquad \textit{in } \Omega_T, \tag{1.37}$$

$$\frac{\partial u}{\partial n} = 0 \qquad \textit{on } \partial\Omega_T, \tag{1.38}$$

in the distributional sense. Moreover, u satisfies the following dissipative energy law:

$$\frac{d}{dt}J_\lambda(u) = -\|\, u_t\,\|_{L^2}^2 \qquad \textit{for a.e. } t\in[0,T]. \tag{1.39}$$

Remark 1.5 *The remarks concerning non-homogeneous Dirichlet and Neumann boundary conditions given in Remarks 1.2 and 1.3 remain valid for the TV flow.*

Our last group of main results concerns quality of the finite element method for approximating the gradient flow (1.7)-(1.9) and the TV flow (1.4)-(1.6) in view of Theorems 1.3 and 1.4. To state the theorems, we need some preparations.

Let $\mathcal{T}_h$ be a quasi-uniform triangulation of Ω ($K \in \mathcal{T}_h$ are tetrahedrons when $N=3$) with mesh size $h \in (0,1)$. Let V^h denote the finite element space of continuous, piecewise linear functions associated with $\mathcal{T}_h$, that is,

$$V^h := \left\{ v_h \in C^0(\overline{\Omega});\ v_h|_K \in P_1(K),\ \forall K \in \mathcal{T}_h \right\}.$$

Let $\{t_m\}_{m=0}^M$ be an equidistant partition of $[0,T]$ of mesh size $k \in (0,1)$ and introduce the notation $d_t u^m := (u^m - u^{m-1})/k$. Then our fully discrete finite element discretization for the gradient flow (1.7)-(1.9) is defined as follows: find $U^m \in V^h$ for $m = 0,1,2,\cdots,M$ such that

$$\int_\Omega \Big[d_t U^m\, v_h + \frac{f'_\varepsilon(|\nabla U^m|)}{|\nabla U^m|} \nabla U^m \cdot \nabla v_h + \lambda (U^m - g)\, v_h \Big]\, dx = 0 \quad \forall\, v_h \in V^h\,, \tag{1.40}$$

with some starting value $U^0 \in V^h$ that approximates u_0. Here

$$f_\varepsilon(z) = \sqrt{z^2 + \varepsilon^2} \qquad \forall\, z \in \mathbf{R}. \tag{1.41}$$

Remark 1.6 *(a). Since* $f'_\varepsilon(z) = \frac{z}{\sqrt{z^2+\varepsilon^2}}$*, the second term in (1.40) is well-defined for all values of* $|\nabla U^m|$.

(b). Since $f_\varepsilon(z)$ *is strictly convex, it easy to check that (1.40) has a unique solution* $\{U^m\}$ *for a given starting value* U^0*. In fact, it is not hard to show that the finite element scheme (1.40) satisfies the following stability estimate:*

$$\| U_1^m - U_2^m \|_{L^2} \leq \| U_1^0 - U_2^0 \|_{L^2} + \sqrt{\lambda} \| g_1 - g_1 \|_{L^2} \qquad 0 \leq m \leq M, \tag{1.42}$$

where $\{U_i^m\}$, $(i = 1,2)$ *is the solution of (1.40) for initial data* U_i^0, $g_i\,(i = 1,2)$*, respectively. Clearly, (1.42) is the discrete counterpart of (1.17).*

(c). The fully discrete finite element scheme is based on a weak formulation of (1.7)-(1.9); for the purpose of error analysis this requires some regularity of the solution u^ε*, which in turn asks for some regularity of* u_0 *and* g*. One way to get around this technical issue is first to smooth the datum functions* u_0 *and* g*, denote the mollified functions by* $\hat{u}_0$ *and* $\hat{g}$*, respectively, then to work with the same differential equation with the new data* $\hat{u}_0$ *and* $\hat{g}$*. This approach is possible because of the stability estimate (1.17).*

On the other hand, if one prefers not using this smoothing approach when $u_0, g \in L^2(\Omega)$*, it is necessary to formulate a* space-time *finite element discretization which is based on the variational inequality weak formulation (1.16), which results in analyzing and solving a nonlinear algebraic inequality. We will report on the numerical analysis and simulation results for that discretization based on a slicing strategy later elsewhere.*

For the fully discrete finite element solution $\{U^m\}$, we define its constant and linear interpolations in t as follows:

$$\overline{U}^{\varepsilon,h,k}(\cdot,t) := U^{m-1} \qquad \forall\, t \in [t_{m-1}, t_m), \quad 1 \leq m \leq M, \tag{1.43}$$

$$\overline{\overline{U}}^{\varepsilon,h,k}(\cdot,t) := \frac{t - t_{m-1}}{k} U^m(x) + \frac{t_m - t}{k} U^{m-1}(x)\ \ \forall\, t \in [t_{m-1}, t_m], \quad 1 \leq m \leq M. \tag{1.44}$$

Clearly, $\overline{U}^{\varepsilon,h,k}$ is continuous in x but discontinuous in t. On the other hand, $\overline{\overline{U}}^{\varepsilon,h,k}$ is continuous in both x and t.

We are now ready to state our last three main theorems of the paper, which give an error analysis for the above fully discrete finite element solution.

Theorem 1.5 *Suppose that u_0, g and $\partial\Omega$ are sufficiently regular such that the solution u^ε of the gradient flow (1.7)-(1.9) belongs to $L^\infty\big((0,T);W^{1,1}(\Omega)\big)\cap L^\infty\big((0,T);H^1_{loc}(\Omega)\big)$ (cf. Theorem 1.2). Then, for each fixed $\varepsilon>0$, $\{U^m\}$ satisfies*

$$k\sum_{m=1}^{\ell}\big[\,\|\,d_tU^m\,\|_{L^2}^2+\frac{\lambda k}{2}\|\,d_t(U^m-g)\,\|_{L^2}^2\,\big]+J_{0,\varepsilon}(U^\ell)\le J_{0,\varepsilon}(U^0),\quad 1\le\ell\le M. \tag{1.45}$$

Moreover, under the following starting value constraint:

$$\lim_{h\to0}\|\,u_0-U^0\,\|_{L^2}=0,$$

there also hold

$$\lim_{h,k\to0}\|\,u^\varepsilon-\overline{U}^{\varepsilon,h,k}\,\|_{L^\infty((0,T);L^p(\Omega))}=0, \tag{1.46}$$

$$\lim_{h,k\to0}\|\,u^\varepsilon-\overline{\overline{U}}^{\varepsilon,h,k}\,\|_{L^\infty((0,T);L^p(\Omega))}=0, \tag{1.47}$$

uniformly in ε for any $p\in[1,\frac{N}{N-1})$.

An immediate consequence of (1.34) and (1.46)-(1.47) is the following convergence theorem.

Theorem 1.6 *Let u stand for the weak solution of the TV flow (1.4)-(1.6). Under the assumptions of Theorem 1.5 there hold*

$$\lim_{\varepsilon,h,k\to0}\|\,u-\overline{U}^{\varepsilon,h,k}\,\|_{L^\infty((0,T);L^p(\Omega))}=0, \tag{1.48}$$

$$\lim_{\varepsilon,h,k\to0}\|\,u-\overline{\overline{U}}^{\varepsilon,h,k}\,\|_{L^\infty((0,T);L^p(\Omega))}=0. \tag{1.49}$$

for any $p\in[1,\frac{N}{N-1})$.

Our last theorem gives the optimal order error estimates for the finite element solution $\{U^m\}$, provided that u^ε is regular enough (cf. Theorem 1.2).

Theorem 1.7 *Suppose that u_0, g and $\partial\Omega$ are sufficiently regular such that the solution u^ε of the gradient flow (1.7)-(1.9) belongs to $L^2((0,T);H^2(\Omega))\cap H^2((0,T);H^{-1}(\Omega))\cap W^{1,\infty}(\Omega_T)$ (cf. Theorem 1.2). Then, under the following starting value constraint:*

$$\|\,u_0-U^0\,\|_{L^2}\le C\,h^2, \tag{1.50}$$

and the parabolic mesh relation $k=O(h^2)$, the finite element solution $\{U^m\}$ also satisfies

$$\max_{0\le m\le M}\|\,u^\varepsilon(t_m)-U^m\,\|_{L^2}+\Big\{k\sum_{m=1}^{M}k\|\,d_t(u^\varepsilon(t_m)-U^m)\,\|_{L^2}^2\Big\}^{\frac12}\le\hat C_3(\varepsilon)\big(h^2+k\big), \tag{1.51}$$

$$\Big(k\sum_{m=1}^{M}\|\,\nabla(u^\varepsilon(t_m)-U^m)\,\|_{L^2}^2\Big)\le\hat C_4(\varepsilon)\big(h+k\big)\,. \tag{1.52}$$

Here, $\hat C_3$ and $\hat C_4$ are some positive constants which are (low order) polynomial functions of the constants $\hat C_j$ $(j=0,1,2)$ given in (1.24)-(1.26).

Remark 1.7 *The parabolic mesh relation* $k = O(h^2)$ *is attributed to the singular character of the TV-flow problem to verify* (1.52)*; exploiting monotonicity, this condition may be dropped if one confines to* (1.51)*.*

The rest of the paper is organized as follows: In Section 2, we introduce some notation and then provide proofs for Theorems 1.1 and 1.2. Section 3 devotes to proving Theorem 1.3 and Theorem 1.4. Finally, in Section 4, we first recall some approximation properties of the continuous piecewise linear finite element space, and then present proofs for Theorems 1.5-1.7.

2 Proofs of Theorems 1.1 and 1.2

Recall that a function $u \in L^1(\Omega)$ is called a function of *bounded variation* if all of its first order partial derivatives (in the distributional sense) are measures with finite total variations in Ω. Hence, the gradient of such a function u, still denoted by ∇u, is a bounded vector-valued measure, with the finite total variation

$$\int_\Omega |\nabla u|\,dx \equiv \|\nabla u\| := \sup\Big\{ \int_\Omega u \operatorname{div} \mathbf{v}\,dx\,;\ \mathbf{v} \in [C_0^1(\Omega)]^N,\ \|\mathbf{v}\|_{L^\infty} \le 1 \Big\}. \tag{2.1}$$

The space of functions of bounded variation is denoted by $BV(\Omega)$, it is a Banach space which is endowed with norm

$$\|u\|_{BV} := \|u\|_{L^1} + \|\nabla u\|. \tag{2.2}$$

We refer to [1, 17] for detailed discussions about the space $BV(\Omega)$, and we also refer to [20, 16, 17] for definitions of standard space notation which is used in this paper.

Proof of Theorem 1.1:

For each $\delta > 0$, we consider the regularized convex functional

$$J_{\lambda,\varepsilon,\delta}(v) := J_{\lambda,\varepsilon}(v) + \frac{\delta}{2}\|\nabla v\|_{L^2}^2\,, \tag{2.3}$$

whose gradient flow is the following parabolic problem:

$$\frac{\partial u^{\varepsilon,\delta}}{\partial t} = \delta\Delta u^{\varepsilon,\delta} + \operatorname{div}\Big(\frac{\nabla u^{\varepsilon,\delta}}{\sqrt{|\nabla u^{\varepsilon,\delta}|^2+\varepsilon^2}}\Big) - \lambda(u^{\varepsilon,\delta} - g) \qquad \text{in } \Omega_T, \tag{2.4}$$

$$\frac{\partial u^{\varepsilon,\delta}}{\partial n} = 0 \qquad \text{on } \partial\Omega_T, \tag{2.5}$$

$$u^{\varepsilon,\delta}(\cdot,0) = u_0(\cdot) \qquad \text{in } \Omega. \tag{2.6}$$

The general theory of monotone nonlinear equations (cf. [7, 20]) provides existence and uniqueness of solutions $u^{\varepsilon,\delta} \in L^\infty\big((0,T);L^2(\Omega)\big)\cap H^1\big((0,T);L^2(\Omega)\big)\cap L^2\big((0,T);H^1(\Omega)\big)$ for the above parabolic problem.

Next, we are going to derive some δ-independent *and* ε-independent a priori estimates for $u^{\varepsilon,\delta}$. We remark that ε is fixed throughout this section, and the ε-independent a priori estimates will be utilized in the next section for proving Theorems 1.3 and 1.4.

Test (2.4) by $u^{\varepsilon,\delta}$ we get

$$\frac{1}{2}\frac{d}{dt}\| u^{\varepsilon,\delta} \|_{L^2}^2 + \delta\| \nabla u^{\varepsilon,\delta} \|_{L^2}^2 + \int_\Omega \frac{| \nabla u^{\varepsilon,\delta} |^2}{\sqrt{| \nabla u^{\varepsilon,\delta} |^2+\varepsilon^2}}\,dx + \frac{\lambda}{2}\| u^{\varepsilon,\delta} \|_{L^2}^2 \le \frac{\lambda}{2}\| g \|_{L^2}^2\,,$$

which implies

$$\| u^{\varepsilon,\delta} \|_{L^\infty(L^2)} + \sqrt{\delta}\| \nabla u^{\varepsilon,\delta} \|_{L^2(L^2)} + \| u^{\varepsilon,\delta} \|_{L^1(W^{1,1})} \le C_0 \equiv \sqrt{\lambda T}\| g \|_{L^2} + \| u_0 \|_{L^2}\,. \tag{2.7}$$

Test (2.4) by any function $\phi\in H^1_0(\Omega)$ and use the uniform estimate (2.7) gives

$$\| u^{\varepsilon,\delta}_t \|_{L^2(H^{-1})} \le C_1, \tag{2.8}$$

where C_1 depends on C_0 linearly.

Then, we test (2.4) by $t\,u^{\varepsilon,\delta}_t$. After some manipulations of the nonlinear term we get

$$\begin{aligned}\frac{t}{2}\| u^{\varepsilon,\delta}_t \|_{L^2}^2 + \frac{\delta}{2}\frac{d}{dt}\big(t\,\| \nabla u^{\varepsilon,\delta} \|_{L^2}^2\big) + \frac{d}{dt}\big(t\,J_{\lambda,\varepsilon}(u^{\varepsilon,\delta})\big) + \frac{\lambda}{2}\frac{d}{dt}\big(t\,\| u^{\varepsilon,\delta} \|_{L^2}^2\big)&\\ \le \frac{\delta}{2}\| \nabla u^{\varepsilon,\delta} \|_{L^2}^2 + J_{\lambda,\varepsilon}(u^{\varepsilon,\delta}) + \frac{\lambda}{2}\| u^{\varepsilon,\delta} \|_{L^2}^2 + \frac{\lambda\,t}{2}\| g \|_{L^2}^2\,,&\end{aligned}$$

which together with (2.7) implies

$$\| \sqrt{t}\,u^{\varepsilon,\delta}_t \|_{L^2(L^2)} + \sqrt{\delta}\,\| \sqrt{t}\,\nabla u^{\varepsilon,\delta} \|_{L^\infty(L^2)} + \| \sqrt{t}\,u^{\varepsilon,\delta} \|_{L^\infty(W^{1,1})} \le C_2\,. \tag{2.9}$$

The constant C_2 depends on C_0 linearly.

From [23] we know that the uniform estimates (2.7)-(2.9) imply there exists a function $u^\varepsilon \in L^1\big((0,T);\,BV(\Omega)\big)\cap L^2(\Omega_T)$ and a subsequence of $\{u^{\varepsilon,\delta}\}$ (denoted by the same notation) such that as $\delta\to 0$

$$\begin{aligned}
&(2.10) && u^{\varepsilon,\delta} \longrightarrow u^\varepsilon && \text{weakly}\star \text{ in } L^\infty\big((0,T);L^2(\Omega)\big),\\
&&&&& \text{weakly in } L^2\big((0,T);L^2(\Omega)\big)\\
&&&&& \text{strongly in } L^1\big((0,T);L^p(\Omega)\big),\ 1\le p<\frac{N}{N-1},\\
&(2.11) && \sqrt{t}\,u^{\varepsilon,\delta} \longrightarrow \sqrt{t}\,u^\varepsilon && \text{strongly in } L^p(\Omega),\ 1\le p<\frac{N}{N-1},\ \text{ for a.e. } t\in[0,T],\\
&(2.12) && u^{\varepsilon,\delta}_t \longrightarrow u^\varepsilon_t && \text{weakly in } L^2\big((0,T);H^{-1}(\Omega)\big),\\
&&&&& \text{weakly in } L^2\big((t_0,T);L^2(\Omega)\big),\ \forall\, t_0\in(0,T),\\
&(2.13) && \sqrt{t}\,u^{\varepsilon,\delta}_t \longrightarrow \sqrt{t}\,u^\varepsilon_t && \text{weakly in } L^2\big((0,T);L^2(\Omega)\big)\,.
\end{aligned}$$

Here we have used the fact that $BV(\Omega)$ is compactly embedded in $L^p(\Omega)$ for $1\le p<\frac{N}{N-1}$ (cf. [1, 17]).

In order to show that u^ε satisfies (1.16), multiply (2.4) by $v-u^{\varepsilon,\delta}$, integrate in x over Ω and in t from 0 to $s(\le T)$, use the convexity of the function $\sqrt{z^2+\varepsilon^2}$, and drop the positive term $\delta\int_0^s\int_\Omega |\nabla u^{\varepsilon,\delta}|^2\,dxdt$ on the right-hand side to get

$$\begin{aligned}\int_0^s\int_\Omega v_t(v-u^{\varepsilon,\delta})\,dxdt + \int_0^s \big[J_{\lambda,\varepsilon}(v) - J_{\lambda,\varepsilon}(u^{\varepsilon,\delta})\big]\,dt + \delta\int_0^s\int_\Omega \nabla u^{\varepsilon,\delta}\cdot\nabla v\,dxdt&\\ \ge \frac{1}{2}\big[\,\| v(s)-u^{\varepsilon,\delta}(s) \|_{L^2}^2 - \| v(0)-u_0 \|_{L^2}^2\,\big]&\\ \forall\, v\in L^1\big((0,T);C^1(\Omega)\big)\cap L^2(\Omega_T) \text{ such that } v_t\in L^2(\Omega_T).&\end{aligned} \tag{2.14}$$

Since $J_{\lambda,\varepsilon}(v)$ is convex, it is lower semicontinuous in $BV(\Omega)$ with respect to convergence in $L^1(\Omega)$ (cf. [1, 17]). Hence, $(2.10)_3$, (2.11) and Fatou's Lemma imply

$$\liminf_{\delta\to 0} \int_0^s J_{\lambda,\varepsilon}(u^{\varepsilon,\delta})\,dt \geq \int_0^s J_{\lambda,\varepsilon}(u^{\varepsilon})\,dt \qquad \forall s\in[0,T]. \tag{2.15}$$

We also recall the fact that the L^2-norm $\|\, v\,\|_{L^2}^2$ is a lower semicontinuous functional with respect to weak convergence in $L^2(\Omega)$, that is,

$$\liminf_{\delta\to 0} \|\, u^{\varepsilon,\delta}(s)\,\|_{L^2}^2 \geq \|\, u^{\varepsilon}(s)\,\|_{L^2}^2 \qquad \forall s\in[0,T]. \tag{2.16}$$

Now, setting $\delta\to 0$ in (2.14) and using $(2.10)_2$, (2.15), (2.16) and the boundedness of $\sqrt{\delta}\|\,\nabla u^{\varepsilon,\delta}\,\|_{L^2}$ (cf. (2.7)), we show that (1.16) holds for all $v\in L^1\big((0,T);C^1(\Omega)\big)\cap L^2(\Omega_T)$ with $v_t\in L^2(\Omega_T)$. Since $C^1(\Omega)$ is dense in $BV(\Omega)$ (cf. [1, 17]), hence, a standard density argument shows that (1.16) also holds for all $v\in L^1\big((0,T);BV(\Omega)\big)\cap L^2(\Omega_T)$ with $v_t\in L^2(\Omega_T)$. So the existence part of the proof is complete.

It is trivial to see that the stability inequality (1.17) immediately implies the uniqueness, so it remains to show (1.17), this will be done by adapting a nice test function trick given in Section 2.5 of [19].

Let u_i^ε $(i=1,2)$ be two functions which satisfy (1.16) for given data $u_i^\varepsilon(0)$, $g_i^\varepsilon(i=1,2)$. Set

$$\hat{u} := \frac{u_1^\varepsilon+u_2^\varepsilon}{2}, \qquad \hat{u}(0) := \frac{u_1^\varepsilon(0)+u_2^\varepsilon(0)}{2}.$$

For any $\beta>0$, define w^β to be the solution of the following initial value problem:

$$\begin{aligned} \beta w_t^\beta + w^\beta &= \hat{u} \qquad\qquad \forall\, t\in(0,T),\\ w^\beta(0) &= \hat{u}(0). \end{aligned}$$

Now, take $v=w^\beta$ in each inequality (1.16) with u_i^ε in place of u^ε, and add the two resulting inequalities. We employ the definitions of $\hat{u}$, $\hat{u}(0)$ and w^β to get

$$\begin{aligned} -2\beta\int_0^s \|\, w_t^\beta\,\|_{L^2}^2\,dt &+ \int_0^s \big[2J_{0,\varepsilon}(w^\beta) - J_{0,\varepsilon}(u_1^\varepsilon) - J_{0,\varepsilon}(u_2^\varepsilon)\big]\,dt \\ &+\frac{\lambda}{2}\int_0^s \big[(w^\beta-g_1^\varepsilon)^2 + (w^\beta-g_2^\varepsilon)^2 - (u_1^\varepsilon-g_1^\varepsilon)^2 - (u_2^\varepsilon-g_2^\varepsilon)^2\big]\,dt \\ &\geq \frac{1}{2}\big[\,\|\, w^\beta(s)-u_1^\varepsilon(s)\,\|_{L^2}^2 + \|\, w^\beta(s)-u_2^\varepsilon(s)\,\|_{L^2}^2\,\big] - \frac{1}{4}\|\, u_1^\varepsilon(0)-u_2^\varepsilon(0)\,\|_{L^2}^2. \end{aligned} \tag{2.17}$$

From [19] (see pages 352-353 of [19]), we have

$$w^\beta \longrightarrow \hat{u} \qquad \text{strongly in } L^2(\Omega_T) \quad \text{as } \beta\to 0, \tag{2.18}$$

$$w^\beta(s) \longrightarrow \hat{u}(s) \quad \text{strongly in } L^2(\Omega), \quad \forall s\in(0,T) \quad \text{as } \beta\to 0. \tag{2.19}$$

Following the same proof we also can show that

$$w^\beta \longrightarrow \hat{u} \quad \text{strictly in } L^1\big((0,T);BV(\Omega)\big) \quad \text{as } \beta\to 0. \tag{2.20}$$

Taking $\beta\to 0$ in (2.17) we get

$$\begin{aligned} \int_0^s \big[2J_{0,\varepsilon}(\hat{u}) - J_{0,\varepsilon}(u_1^\varepsilon) - J_{0,\varepsilon}(u_2^\varepsilon)\big]\,dt &+ \frac{\lambda}{4}\|\, g_1^\varepsilon - g_2^\varepsilon\,\|_{L^2}^2 \\ &\geq \frac{1}{4}\|\, u_1^\varepsilon(s)-u_2^\varepsilon(s)\,\|_{L^2}^2 - \frac{1}{4}\|\, u_1^\varepsilon(0)-u_2^\varepsilon(0)\,\|_{L^2}^2. \end{aligned} \tag{2.21}$$

By the convexity of $J_{0,\varepsilon}(\cdot)$ we have

$$2J_{0,\varepsilon}\Big(\frac{u_1^\varepsilon+u_2^\varepsilon}{2}\Big) \le J_{0,\varepsilon}(u_1^\varepsilon)+J_{0,\varepsilon}(u_2^\varepsilon), \tag{2.22}$$

which and (2.21) imply (1.17).

Finally, it remains to show $u^\varepsilon \in C^0([0,T];L^2(\Omega))$ and $u^\varepsilon(0)=u_0$. Again, this will be done using the test function trick due to Lichnewsky and Temam [19].

For any $\beta>0$, define u_β^ε to be the solution of the following initial value problem:

$$\begin{aligned} \beta\,\frac{\partial u_\beta^\varepsilon}{\partial t}+u_\beta^\varepsilon &= u^\varepsilon \qquad \forall\, t\in(0,T),\\ u_\beta^\varepsilon(0) &= u_0\,. \end{aligned}$$

Now, take $v=u_\beta^\varepsilon$ in (1.16), by the boundedness of the function $f_\varepsilon'(z)$ (cf. (1.41)) and a direct calculation we obtain for any $s\in[0,T]$

$$\frac12\|\,u_\beta^\varepsilon(s)-u^\varepsilon(s)\,\|_{L^2}^2 \le \int_0^s\int_\Omega |u_\beta^\varepsilon-u^\varepsilon|\,dx\,dt+\frac{\lambda}{2}\int_0^s \|\,u_\beta^\varepsilon-u^\varepsilon\,\|_{L^2}\,\|\,u_\beta^\varepsilon+u^\varepsilon-2g\,\|_{L^2}^2\,dt\,. \tag{2.23}$$

It is not hard to check that the convergences given in (2.18)-(2.20) are still valid for the sequence $\{u_\beta^\varepsilon\}$. Hence, taking $\beta\to 0$ in (2.23) yields

$$\lim_{\beta\to 0}\|\,u_\beta^\varepsilon(s)-u^\varepsilon(s)\,\|_{L^2}=0 \qquad \text{uniformly with respect to } s\in[0,T]\,.$$

Since for each $\beta>0$, u_β^ε satisfies

$$u_\beta^\varepsilon\in C^0\big([0,T];L^2(\Omega)\big) \qquad\text{and}\qquad u_\beta^\varepsilon(0)=u_0\,,$$

so does u^ε. The proof is complete. □

An immediate consequence of (2.7)-(2.9) is the following corollary, which will play a crucial role in the proof of Theorem 1.3 in the next section.

Corollary 2.1 *The function u^ε constructed in the proof of Theorem 1.1 satisfies the following (uniform in ε) estimates:*

$$\begin{aligned} \|\,u^\varepsilon\,\|_{L^\infty(L^2)}+\|\,u^\varepsilon\,\|_{L^1(BV)} &\le C_0\equiv\sqrt{\lambda}\|\,g\,\|_{L^2}+\|\,u_0\,\|_{L^2}\,, && (2.24)\\ \|\,u_t^\varepsilon\,\|_{L^2(H^{-1})} &\le C_1, && (2.25)\\ \|\,\sqrt{t}\,u_t^\varepsilon\,\|_{L^2(L^2)}+\|\,\sqrt{t}\,u^\varepsilon\,\|_{L^\infty(BV)} &\le C_2\,, && (2.26) \end{aligned}$$

where C_1 and C_2 are positive constants depending on C_0 linearly.

In the remainder of this section, we will give a proof for Theorem 1.2.

Proof of Theorem 1.2:

The proof is based on studying (1.7)-(1.9) in the scaled coordinates defined in (1.11), and making use of the results of [19, 15] for the minimal surface flow and the prescribed mean curvature flow, whose proofs were carried out using the same approach as demonstrated in the proof of Theorem 1.1 above.

(i). As noted in Section 1, the initial-boundary value problem (1.12)-(1.14) corresponds to the prescribed mean curvature flow problem studied in [19, 15], and the "mean curvature" function is given by

$$H(y, v^\varepsilon) = \frac{\lambda}{\varepsilon}\big[\, v^\varepsilon - g^\varepsilon(y)\,\big]\,. \tag{2.27}$$

Since $H(x, v^\varepsilon)$ is linear in the second argument, hence, it is a monotone increasing function if $\lambda \geq 0$. Therefore, $H(x, v^\varepsilon)$ satisfies the assumption on the "mean curvature" function required in [15].

From Theorem 2.5 of [15] we know there exists a unique function $v^\varepsilon \in L^\infty\big((0,T^\varepsilon); BV(\Omega^\varepsilon)\big) \cap H^1\big((0,T^\varepsilon); L^2(\Omega^\varepsilon)\big)$ and for any $s \in [0, T^\varepsilon]$

$$\int_0^s \int_{\Omega^\varepsilon} v^\varepsilon_\tau (w - v^\varepsilon)\, dy d\tau + \int_0^s \big[I_{\lambda,\varepsilon}(w) - I_{\lambda,\varepsilon}(v^\varepsilon)\big]\, d\tau \geq 0 \qquad \forall\, w \in L^1\big((0,T^\varepsilon); BV(\Omega^\varepsilon)\big) \cap L^2(\Omega^\varepsilon_{T^\varepsilon}), \tag{2.28}$$

where

$$I_{\lambda,\varepsilon}(w) := \int_{\Omega^\varepsilon} \sqrt{|\nabla w|^2 + 1}\, dy + \frac{\lambda}{2\varepsilon} \int_{\Omega^\varepsilon} (w - g^\varepsilon)^2\, dy. \tag{2.29}$$

Now going back to the original (x,t) space-time domain using (1.11), it is easy to check that (2.28) exactly gives (1.20), and $v^\varepsilon \in L^\infty((0,T^\varepsilon); BV(\Omega^\varepsilon)) \cap H^1((0,T^\varepsilon); L^2(\Omega^\varepsilon))$ implies (1.19). This completes the proof of (i).

(ii). From Theorem 2.3 of [15] we have

$$\sup_{(y,\tau)\in\Omega^\varepsilon_{T^\varepsilon}} |v^\varepsilon(y,\tau)| \leq \sup_{y\in\Omega^\varepsilon} |v_0(y)| + R\varepsilon + T^\varepsilon \Big(\frac{N}{R\varepsilon} + \frac{\lambda}{\varepsilon} \|\, g^\varepsilon \,\|_{L^\infty(\Omega^\varepsilon)} \Big), \tag{2.30}$$

which is equivalent to (1.21) in (x,t) space-time domain. Hence, (ii) is proved.

(iii). It follows from Theorems 1.1 and 1.2 of [19] that $v^\varepsilon(y,\tau)$ also satisfies $v^\varepsilon(y,\tau) \in L^\infty\big((0,T^\varepsilon); W^{1,1}(\Omega^\varepsilon)\big) \cap L^\infty\big((0,T^\varepsilon); H^1_{\rm loc}(\Omega^\varepsilon)\big)$ and the equations (1.12)-(1.14) hold a.e. in $\Omega^\varepsilon_{T^\varepsilon}$ and on $\partial\Omega^\varepsilon_{T^\varepsilon}$, respectively. In (x,t) space-time domain, this means that $u^\varepsilon(x,t) \in L^\infty((0,T); W^{1,1}(\Omega)) \cap L^\infty((0,T); H^1_{\rm loc}(\Omega))$ and the equations (1.7)-(1.9) hold a.e. in Ω_T and on $\partial\Omega_T$, respectively.

Now, since (1.7) holds a.e. in Ω_T and $u^\varepsilon_t \in L^2(\Omega)$, testing (1.7) by u^ε_t immediately gives the desired energy law (1.22). Hence, the assertions of (iii) hold.

(iv). It turns out that for fixed $\varepsilon > 0$, the regularity $u^\varepsilon \in L^2\big((0,T); H^2(\Omega)\big)$ follows easily from Theorem 4.5 of [15], however, it is not easy to derive its a priori estimate which shows precise dependence on the parameter ε.

Applying Theorem 4.2 of [15] to the "prescribed mean curvature" flow (1.12)-(1.14), we have

$$\|\, \nabla v^\varepsilon \,\|_{L^\infty(\Omega^\varepsilon_{T^\varepsilon})} \leq \tilde{C}(T^\varepsilon, M_0, M_1), \tag{2.31}$$

where $\tilde{C}$ is some positive constant which is a polynomial function in each of its arguments, and $M_j\, (j = 0,1)$ are defined as (see (4.5) and (4.6) of [15])

$$M_0 := \sup_{y\in\Omega^\varepsilon} \Big\{ |v_0(y)| + |\nabla v_0(y)| \Big\}, \tag{2.32}$$

$$M_1 := \sup_{(y,\tau)\in\Omega^\varepsilon_{T^\varepsilon}} \Big\{ |v^\varepsilon(y,\tau)| + |v^\varepsilon_\tau(y,\tau)| + |H(y,v^\varepsilon)| + \Big|\frac{\partial}{\partial y} H(y, v^\varepsilon)\Big| \Big\}, \tag{2.33}$$

and $H(y, v^\varepsilon)$ is given by (2.27).

It is not hard to check that

$$M_0 = O\Big(\frac{1}{\varepsilon}\Big), \qquad \text{and} \qquad M_1 = O\Big(\frac{1}{\varepsilon^2}\Big).$$

under the assumptions on u_0 and g. Hence, the estimate (2.31) in (x,t) space-time domain becomes

$$\|\nabla u^\varepsilon\|_{L^\infty(\Omega_T)} \leq \varepsilon\, \tilde{C}(\varepsilon T, M_0, M_1), \tag{2.34}$$

which gives (1.24) with $\hat{C}_0 = \varepsilon\, \tilde{C}(\varepsilon T, M_0, M_1)$.

The proof of (1.25) is more technical, we skip the proof and refer the interested reader to Theorem 3.1 of [14] to get the idea of the proof.

Finally, to show (1.23) and (1.26), first differentiate (1.7) with respect to t to get

$$u^\varepsilon_{tt} = \operatorname{div}\Big(\frac{\nabla u^\varepsilon_t}{\sqrt{|\nabla u^\varepsilon|^2 + \varepsilon^2}}\Big) + \operatorname{div}\Big(\frac{(\nabla u^\varepsilon \cdot \nabla u^\varepsilon_t)\,\nabla u^\varepsilon}{\sqrt{(|\nabla u^\varepsilon|^2 + \varepsilon^2)^3}}\Big) - \lambda u^\varepsilon_t\,. \tag{2.35}$$

Then, test the above equation by u^ε_t gives the high order dissipative energy law (1.23). In turn, this high energy law immediately gives bounds for $\|u^\varepsilon_t\|_{L^\infty(L^2)}$ and $\|\nabla u^\varepsilon_t\|_{L^2(L^2)}$ in terms of $\|\nabla u^\varepsilon\|_{L^\infty(\Omega_T)}$ given by (2.34). As a result, the desired estimate (1.26) can be obtained easily by testing (2.35) by an arbitrary function $\varphi \in H^1_0(\Omega)$. The proof is complete. $\Box$

From the above proof, we have

Corollary 2.2 *The weak solution u^ε of the gradient flow (1.7)-(1.9) satisfies the following (uniform in ε) estimates:*

(i). If $u_0 \in BV(\Omega)$ and $g \in L^2(\Omega)$, then

$$\|u^\varepsilon\|_{L^\infty(BV)} + \|u^\varepsilon_t\|_{L^2(L^2)} \leq C_3. \tag{2.36}$$

(ii). If $\partial\Omega$ satisfies an internal sphere condition (ISC) of radius R, $u_0 \in BV(\Omega) \cap L^\infty(\Omega)$ and $g \in L^\infty(\Omega)$, then

$$\|u^\varepsilon\|_{L^\infty(L^\infty)} \leq C_4. \tag{2.37}$$

(iii). If $u_0 \in H^1_{loc}(\Omega) \cap W^{1,1}(\Omega)$, $g \in L^2(\Omega) \cap H^1_{loc}(\Omega)$ and $\partial\Omega \in C^2$, then

$$\|u^\varepsilon\|_{L^\infty(W^{1,1})} + \|u^\varepsilon\|_{L^\infty(H^1_{loc})} \leq C_5\,. \tag{2.38}$$

Here, $C_j\ (j = 3, 4, 5)$ are some positive constants independent of ε.

3 Proofs of Theorems 1.3 and 1.4

In this section, we present proofs for Theorems 1.3 and 1.4.

Proof of Theorem 1.3:

The idea of the proof is straightforward: first extract convergent subsequence using the ε-independent estimates (2.24)-(2.26), and then pass to the limit in (1.16). The only subtle point that needs to be addressed is convergence of the *whole* sequence, which is done with the help of the uniqueness of the limiting problem (1.27)-(1.28).

Mentioning about the uniqueness of the problem (1.27)-(1.28), it suffices to show the stability inequality (1.29). Since the proof of (1.29) is exactly same as that of (1.17), which was given in the previous section using a test function trick due to Lichnewsky and Temam [19], hence, we omit the proof. On the other hand, we remark that now $J(\cdot)$ replaces $J_{0,\varepsilon}(\cdot)$, and convexity of the total variation functional $J(\cdot)$ ensures that the same inequality as (2.22) holds for it. Hence, the proof goes through.

Since $BV(\Omega)$ is compactly embedded in $L^p(\Omega)$ for $1 \leq p < \frac{N}{N-1}$ (cf. [1, 17]), then it follows from (2.24)-(2.26) that there exists $u \in L^1\big((0,T); BV(\Omega)\big) \cap C^0\big((0,T); L^2(\Omega)\big)$ with $u_t \in L^2\big((0,T); H^{-1}(\Omega)\big)$, and a subsequence of $\{u^\varepsilon\}$ (denoted by the same notation) such that as $\varepsilon \to 0$

$$\begin{aligned}
&(3.1) && u^\varepsilon \longrightarrow u && \text{weakly}\star \text{ in } L^\infty\big((0,T); L^2(\Omega)\big),\\
& && && \text{weakly in } L^2((0,T); L^2(\Omega))\\
& && && \text{strongly in } L^1\big((0,T); L^p(\Omega)\big),\ 1 \leq p < \frac{N}{N-1},\\
&(3.2) && \sqrt{t}\, u^\varepsilon \longrightarrow \sqrt{t}\, u && \text{strongly in } L^p(\Omega),\ 1 \leq p < \frac{N}{N-1},\ \text{ for a.e. } t \in [0,T],\\
&(3.3) && u_t^\varepsilon \longrightarrow u_t && \text{weakly in } L^2\big((0,T); H^{-1}(\Omega)\big),\\
& && && \text{weakly in } L^2\big((t_0,T); L^2(\Omega)\big),\ \forall t_0 \in (0,T),\\
&(3.4) && \sqrt{t}\, u_t^\varepsilon \longrightarrow \sqrt{t}\, u_t && \text{weakly in } L^2((0,T); L^2(\Omega))\,.
\end{aligned}$$

Clearly, (1.30) and (1.31) are contained in (3.1)-(3.4). To show (1.28), we write

$$J_{\lambda,\varepsilon}(u^\varepsilon) = J_\lambda(u^\varepsilon) + \int_\Omega \frac{\varepsilon^2}{\sqrt{|\nabla u^\varepsilon|^2 + \varepsilon^2} + |\nabla u^\varepsilon|}\, dx. \tag{3.5}$$

Since the integrand of the second term on the right-hand side is bounded by ε, by Lebesgue Dominated Convergence Theorem we know that the integral converges to zero. Hence,

$$\liminf_{\varepsilon \to 0} J_{\lambda,\varepsilon}(u^\varepsilon) \geq \liminf_{\varepsilon \to 0} J_\lambda(u^\varepsilon) \geq J_\lambda(u). \tag{3.6}$$

We note that the above argumentation clearly applies to any convergent subsequence in $W^{1,1}(\Omega)$. Correspondingly, this argumentation applies to any convergent subsequence in $BV(\Omega)$ by a density argument. Now, (1.28) follows from taking $\varepsilon \to 0$ in (1.16) (with subsequence in the place of u^ε) and applying (3.1)-(3.4) and (3.6).

Finally, since the solution of the problem (1.27)-(1.28) is unique (see the beginning of this proof), then, every convergent subsequence of $\{u^\varepsilon\}$ converges to the same limit, which is the solution of problem (1.27)-(1.28). Hence, the *whole* sequence $\{u^\varepsilon\}$ converges to the same limit, too. The proof is complete. □

Next, we give a proof for Theorem 1.4.

Proof of Theorem 1.4:

The proof follows along the same lines as that of Theorem 1.3; the only difference is that now we work with a function set $\{u^\varepsilon\}$ whose members are bounded uniformly in ε in stronger norms due to the stronger assumptions on the datum functions u_0 and g, as well as the boundary $\partial\Omega$.

These new estimates are given by (2.36)-(2.38). Because of the similarities, we only highlight the main steps of the proof.

(i). From the uniform estimate (2.36), we know that the limit u of any convergent subsequence of $\{u^\varepsilon\}$ now resides in $L^\infty((0,T);BV(\Omega))\cap H^1((0,T);L^2(\Omega))$. Hence, (1.32) holds, and u still satisfies (1.28). To verify (1.33), integrating by parts in the first term on the left-hand side of (1.28) we get

$$\int_0^s\int_\Omega v_t(v-u)\,dx\,dt=\int_0^s\Big[\int_\Omega u_t(v-u)\,dx+\frac{1}{2}\frac{d}{dt}\|\,v-u\,\|_{L^2}^2\Big]\,dt. \tag{3.7}$$

Substituting (3.7) into (1.28) gives (1.33).

(ii). The existence of such a convergent subsequence of $\{u^\varepsilon\}$ is immediate from the uniform estimate (2.36) and the compact embedding $L^2(\Omega)\hookrightarrow H^{-1}(\Omega)$. Repeating the same argumentation as in the proof of Theorem 1.3 to show (1.34) and (1.35) hold for the whole sequence $\{u^\varepsilon\}$.

(iii). (1.36) follows directly from the uniform estimates (2.37) and (1.21).

(iv). By (2.38) and a compactness argument we conclude that $u\in L^\infty((0,T);W^{1,1}(\Omega))\cap L^\infty((0,T);H^1_{\rm loc}(\Omega))$. To show (1.37), for all $w\in C_0^\infty(\Omega_T)$ and $\eta>0$, setting $v=u\pm\eta w$ in (1.33) we get

$$\pm\int_0^s\int_\Omega v_t w\,dx\,dt+\int_0^s\frac{J_\lambda(u\pm\eta w)-J_\lambda(u)}{\eta}\,dt\geq 0. \tag{3.8}$$

Using the definition of $J_\lambda(\cdot)$ in (1.3) we can rewrite the second term in (3.8) as

$$\int_0^s\frac{J_\lambda(u\pm\eta w)-J_\lambda(u)}{\eta}\,dt=\int_0^s\int_\Omega\frac{|\nabla u\pm\eta\nabla w|-|\nabla u|}{\eta}\,dx\,dt\pm\lambda\int_0^s\int_\Omega(u-g)w\,dx\,dt.$$

Since (cf. [7])

$$\lim_{\eta\to 0}\int_0^s\int_\Omega\frac{|\nabla u\pm\eta\nabla w|-|\nabla u|}{\eta}\,dx\,dt=\pm\int_0^s\int_\Omega{\rm Sgn}(\nabla u)\cdot\nabla w\,dx\,dt, \tag{3.9}$$

by setting $\eta\to 0$ in (3.8) and using (3.9) we obtain

$$\pm\int_0^s\int_\Omega v_t w\,dx\,dt\pm\int_0^s\int_\Omega{\rm Sgn}(\nabla u)\cdot\nabla w\,dx\,dt\pm\lambda\int_0^s\int_\Omega(u-g)w\,dx\,dt\geq 0,$$

and equivalently,

$$\int_0^s\int_\Omega v_t w\,dx\,dt+\int_0^s\int_\Omega{\rm Sgn}(\nabla u)\cdot\nabla w\,dx\,dt+\lambda\int_0^s\int_\Omega(u-g)w\,dx\,dt=0, \tag{3.10}$$

for any $\forall w\in C_0^\infty(\Omega_T)$, which gives (1.37).

(1.38) can be shown by the same argumentation as above, the only difference is to set $v=u\pm\eta w$, for all $w\in C^\infty(\overline{\Omega}_T)$ in (1.33). So we omit the proof.

It remains to show (1.39), which comes immediately from testing (1.37) by $u_t\in L^2(\Omega_T)$ and the standard density argumentation. The proof is complete. $\Box$.

4 Proofs of Theorems 1.5-1.7

The goal of this last section is to give proofs for Theorems 1.5 through 1.7. To this end, we need some preparations.

Define the L^2 projection operator $Q_h : L^2(\Omega) \to V^h$

$$(v - Q_h v, w_h) = 0 \quad \forall\, w_h \in V^h, \tag{4.1}$$

and the (nonlinear) elliptic projection operator $P_h : H^1(\Omega) \to V^h \cap H^1_0(\Omega)$

$$\Big(\frac{f_\varepsilon'(|\nabla P_h v|)}{|\nabla P_h v|}\nabla P_h v, \nabla w_h\Big) = \Big(\frac{f_\varepsilon'(|\nabla v|)}{|\nabla v|}\nabla v, \nabla w_h\Big) \qquad \forall w_h \in V^h, \tag{4.2}$$

where $(\cdot,\cdot)$ denotes the standard inner product of $L^2(\Omega)$.

It is well-known that Q_h has the following approximation properties [11, 6]:

$$\|\, v - Q_h v \,\|_{L^2} + h\|\, \nabla(v - Q_h v) \,\|_{L^2} \le Ch\|v\|_{H^1} \quad \forall\, v \in H^1(\Omega), \tag{4.3}$$

$$\|\, v - Q_h v \,\|_{L^2} \le Ch^2\|v\|_{H^2} \quad \forall\, v \in H^2(\Omega). \tag{4.4}$$

For the elliptic projection P_h, the following approximation properties can be obtained using the ideas of [18, 21].

Lemma 4.1 *Let $u^\varepsilon \in W^{1,\infty}(\Omega_T) \cap L^2\big((0,T); H^2(\Omega)\big)$ be the solution of (1.7)-(1.9). There holds for $h \in (0,1)$*

$$\|\, \nabla P_h u^\varepsilon \,\|_{L^\infty(L^\infty)} \le \hat{C}_6(\varepsilon)\,, \tag{4.5}$$

$$\|\, u^\varepsilon - P_h u^\varepsilon \,\|_{L^2(L^2)} + h\,\|\, \nabla(u^\varepsilon - P_h u^\varepsilon) \,\|_{L^2(L^2)} \le \hat{C}_7(\varepsilon)\, h^2, \tag{4.6}$$

where

$$\hat{C}_6(\varepsilon) := T\,\|\, \nabla u^\varepsilon \,\|_{L^\infty(L^\infty)} + \|\, u^\varepsilon \,\|_{L^2(H^2)}, \qquad \hat{C}_7(\varepsilon) := \|\, \nabla u^\varepsilon \,\|_{L^\infty(L^\infty)}\,\|\, u^\varepsilon \,\|_{L^2(H^2)}.$$

We are ready to give a proof for Theorem 1.5.

Proof of Theorem 1.5:

To show (1.45), testing (1.40) with $d_t U^m$ yields

$$\begin{aligned}\|\, d_t U^m \,\|_{L^2}^2 + \frac{\lambda}{2} d_t \|\, U^m - g \,\|_{L^2}^2 + \frac{\lambda k}{2}\|\, d_t(U^m - g) \,\|_{L^2}^2 \\ + \frac{1}{2}\int_\Omega \frac{f_\varepsilon'(|\nabla U^m|)}{|\nabla U^m|}\Big(d_t|\nabla U^m|^2 + k|\nabla d_t U^m|^2\Big)\, dx = 0.\end{aligned} \tag{4.7}$$

Rewrite the fourth term on the left-hand side as

$$\begin{aligned}\frac{1}{2}\int_\Omega \frac{f_\varepsilon'(|\nabla U^m|)}{|\nabla U^m|} d_t|\nabla U^m|^2\, dx = \frac{1}{k}\int_\Omega f_\varepsilon'(|\nabla U^m|)\,\big(|\nabla U^m| - |\nabla U^{m-1}|\big)\, dx \\ - \frac{1}{2k}\int_\Omega \frac{f_\varepsilon'(|\nabla U^m|)}{|\nabla U^m|}\big(|\nabla U^m| - |\nabla U^{m-1}|\big)^2\, dx.\end{aligned} \tag{4.8}$$

By convexity of $f_\varepsilon(z)$, the first term on the right-hand side of (4.8) is bounded by

$$\frac{1}{k}\int_\Omega f_\varepsilon'(|\nabla U^m|)\,\big(|\nabla U^m| - |\nabla U^{m-1}|\big)\, dx \ge d_t\Big[\int_\Omega f_\varepsilon(|\nabla U^m|)\, dx\Big]. \tag{4.9}$$

Using the formula $(|\mathbf{a}|-|\mathbf{b}|)^2 \le |\mathbf{a}-\mathbf{b}|^2$ we have

$$-\frac{1}{2k}\int_\Omega \frac{f'_\varepsilon(|\nabla U^m|)}{|\nabla U^m|}\bigl(|\nabla U^m|-|\nabla U^{m-1}|\bigr)^2\,dx \ge -\frac{1}{2}\int_\Omega \frac{f'_\varepsilon(|\nabla U^m|)}{|\nabla U^m|}\,k|\nabla d_t U^m|^2\,dx. \tag{4.10}$$

(1.45) now follows from substituting (4.8)-(4.10) into (4.7) and applying the summation operator $k\sum_{m=1}^M$ to the resulting inequality.

To show (1.46)-(1.47), we first notice that (1.45) implies the following (uniform in both h,k and ε) estimates:

$$\| \overline{\overline{U}}_t^{\varepsilon,h,k} \|_{L^2(L^2)} = \Bigl(k\sum_{m=1}^M \| d_t U^m \|_{L^2}^2\Bigr)^{\frac{1}{2}} \le C, \tag{4.11}$$

$$\| \overline{\overline{U}}^{\varepsilon,h,k} \|_{L^\infty(L^2)} \le \| \overline{U}^{\varepsilon,h,k} \|_{L^\infty(L^2)} = \max_{0\le m\le M} \| U^m \|_{L^2} \le C, \tag{4.12}$$

$$\| \nabla\overline{\overline{U}}^{\varepsilon,h,k} \|_{L^\infty(L^1)} \le \| \nabla\overline{U}^{\varepsilon,h,k} \|_{L^\infty(L^1)} = \max_{0\le m\le M} \| \nabla U^m \|_{L^1} \le C\,, \tag{4.13}$$

$$\sum_{i=1}^M \| U^m - U^{m-1} \|_{L^2}^2 \le C \qquad \text{if } \lambda \ne 0\,. \tag{4.14}$$

Then there exists a convergent subsequence of $\{\overline{\overline{U}}^{\varepsilon,h,k}\}$ (denoted by the same notation) and a function $\hat{u}^\varepsilon \in L^\infty\bigl((0,T);BV(\Omega)\bigr)\cap H^1\bigl((0,T);L^2(\Omega)\bigr)$ such that as $h,k\to 0$

$$\begin{aligned} \overline{\overline{U}}^{\varepsilon,h,k} &\longrightarrow \hat{u}^\varepsilon \quad && \text{weakly}\star \text{ in } L^\infty\bigl((0,T);L^2(\Omega)\bigr), \\ & && \text{weakly in } L^2\bigl((0,T);L^2(\Omega)\bigr) \\ & && \text{strongly in } L^p(\Omega),\ 1\le p<\frac{N}{N-1},\ \text{ for a.e. } t\in[0,T], \end{aligned} \tag{4.15}$$

$$\overline{\overline{U}}_t^{\varepsilon,h,k} \longrightarrow \hat{u}_t^\varepsilon \quad \text{weakly in } L^2\bigl((0,T);L^2(\Omega)\bigr), \tag{4.16}$$

Here we have used the fact that $BV(\Omega)$ is compactly embedded in $L^p(\Omega)$ for $1\le p<\frac{N}{N-1}$ (cf. [1, 17]). Notice that the assumption on U^0 implies that $\hat{u}^\varepsilon(0)=u_0$.

Next, we like to pass to the limit in (1.40) and show that $\hat{u}^\varepsilon$ is indeed a solution of the differential problem (1.7)-(1.9) with the initial value u_0. Clearly, it suffices to show that $\hat{u}^\varepsilon$ satisfies (1.20). To the end, we rewrite (1.40) as the following equivalent variational inequality formulation:

$$\Bigl(\overline{\overline{U}}_t^{\varepsilon,h,k},(t^r\,v_h-\overline{\overline{U}}^{\varepsilon,h,k})\Bigr)+J_{\lambda,\varepsilon}(t^r\,v_h)-J_{\lambda,\varepsilon}\bigl(\overline{\overline{U}}^{\varepsilon,h,k}\bigr)\ge 0 \quad \forall\, v_h\in V^h,\ r\ge 0.$$

Setting $h,k\to 0$, using (4.15)-(4.16), the lower semicontinuity of $J_{\lambda,\varepsilon}(\cdot)$ in $BV(\Omega)$ with respect to convergence in L^1 (cf. [1]), and Fatou's lemma we get for $s\in[0,T]$

$$\int_0^s \Bigl(\hat{u}_t^\varepsilon,(t^r\,v_h-\hat{u}^\varepsilon)\Bigr)\,dt+\int_0^s\Bigl[J_{\lambda,\varepsilon}(t^r\,v_h)-J_{\lambda,\varepsilon}\bigl(\hat{u}^\varepsilon\bigr)\Bigr]\,dt\ge 0 \quad \forall\, v_h\in V^h. \tag{4.17}$$

To show (4.17) holds for all $v\in L^1\bigl((0,T);BV(\Omega)\bigr)\cap L^2(\Omega_T)$, we appeal to the fact that the polynomials $\varphi:\ [0,T]\to V^h$

$$\varphi(t) := a_0^h + a_1^h\,t+\cdots+a_r^h\,t^r$$

with $a_j^h\in V^h$ for all $j=0,1,2,\cdots,r$, $r\in\mathbf{N}$ are dense in $L^1\bigl((0,T);BV(\Omega)\bigr)\cap L^2(\Omega_T)$.

Hence, $\hat{u}^\varepsilon$ satisfies (1.20) with $\hat{u}^\varepsilon(0) = u_0$. By uniqueness of solutions for (1.19)-(1.20), we conclude that $\hat{u}^\varepsilon = u^\varepsilon$ and the *whole* sequence $\{\overline{\overline{U}}^{\varepsilon,h,k}\}$ converges to u^ε. This completes the proof of (1.47).

Finally, it remains to show (1.46). From (4.11) we obtain

$$
\begin{aligned}
\| \overline{\overline{U}}^{\varepsilon,h,k} - \overline{U}^{\varepsilon,h,k} \|^2_{L^2(L^2)} &= \int_0^T \| \overline{\overline{U}}^{\varepsilon,h,k} - \overline{U}^{\varepsilon,h,k} \|^2_{L^2} \, dt \\
&= \sum_{m=1}^M \| U^m - U^{m-1} \|^2_{L^2} \int_{t_{m-1}}^{t_m} \Big(\frac{t - t_{m-1}}{k}\Big)^2 \, dt \\
&= \frac{k^3}{3} \sum_{m=1}^M \| d_t U^m \|^2_{L^2} \\
&\leq C\, k^2.
\end{aligned}
\tag{4.18}
$$

This and (1.47) immediately give (1.46). The proof is complete. □

Remark 4.1 *(a). Since the proof of (1.45) only relies on the structure of the function $f_\varepsilon(z)$, it is easy to see that (1.45) holds for* any gradient flow *whose energy density function $f(z)$ satisfies (i) $f(z)$ is convex, (ii) $f'(z)$ is bounded.*

(b). The convergence in (1.46)-(1.47) also hold for the case $p = 2$. This can be proved by analyzing an equation satisfied by the global error $E := u^\varepsilon - \overline{U}^{\varepsilon,h,k}$ (see Proof of Theorem 1.7).

Proof of Theorem 1.6:

Since the limits in (1.46)-(1.47) hold *uniformly* in ε, the assertions (1.51) and (1.52) follows immediately from (1.34), (1.46)-(1.47) and the triangle inequality. The proof is complete. □

Proof of Theorem 1.7:

To show (1.51) and (1.52), we decompose the global error $E^m := U^m - u^\varepsilon(t_m)$ into $E^m = \Phi^m + \Theta^m$, where

$$\Phi^m := U^m - P_h u^\varepsilon(t_m), \qquad \Theta^m := P_h u^\varepsilon(t_m) - u^\varepsilon(t_m).$$

Test (1.7) with any $v_h \in V^h$ and subtract it from (1.40) gives the following error equation:

$$
\begin{aligned}
\big(d_t E^m, v_h\big) + \Big(\frac{f'_\varepsilon(|\nabla U^m|)}{|\nabla U^m|}\nabla U^m - \frac{f'_\varepsilon(|\nabla u^\varepsilon(t^m)|)}{|\nabla u^\varepsilon(t^m)|}\nabla u^\varepsilon(t_m), \nabla v_h\Big)& \\
+\lambda\big(E^m, v_h\big) = \big(\mathcal{R}_m, v_h\big)&,
\end{aligned}
\tag{4.19}
$$

where

$$\mathcal{R}_m := u^\varepsilon_t(t_m) - d_t u^\varepsilon(t_m). \tag{4.20}$$

Now, taking $v_h = \Phi^m$ in (4.19) and using (4.2) gives

$$
\begin{aligned}
&\frac{1}{2} d_t \| \Phi^m \|^2_{L^2} + \frac{k}{2} \| d_t \Phi^m \|^2_{L^2} + \lambda \| \Phi^m \|^2_{L^2} \\
&\qquad + \Big(\frac{f'_\varepsilon(|\nabla U^m|)}{|\nabla U^m|}\nabla U^m - \frac{f'_\varepsilon(|\nabla P_h u^\varepsilon(t^m)|)}{|\nabla P_h u^\varepsilon(t^m)|}\nabla P_h u^\varepsilon(t_m), \nabla \Phi^m\Big) \\
&\qquad = \big(\mathcal{R}_m, \Phi^m\big) - \big(d_t \Theta^m, \Phi^m\big) - \lambda\big(\Theta^m, \Phi^m\big).
\end{aligned}
\tag{4.21}
$$

The convexity of the function $f_\varepsilon(z)$ immediately implies that the fourth term on the left-hand side of (4.21) is non-negative. However, this only leads to some estimate in $L^\infty\big((0,T);L^2(\Omega)\big)$, *provided that* $u^\varepsilon_{tt}\in L^2(\Omega_T)$, which is not assumed in the statement of the theorem. It turns out that in order to get error estimates not only in $L^\infty\big((0,T);L^2(\Omega)\big)$ but also in $L^2\big((0,T);H^1(\Omega)\big)$, especially under the weaker assumption $u^\varepsilon_{tt}\in L^2\big((0,T);H^{-1}(\Omega)\big)$, we need to extract a positive contribution from that term.

$$\begin{aligned}(4.22)\qquad \Big(\frac{f'_\varepsilon(|\nabla U^m|)}{|\nabla U^m|}\nabla U^m-\frac{f'_\varepsilon(|\nabla P_h u^\varepsilon(t^m)|)}{|\nabla P_h u^\varepsilon(t^m)|}\nabla P_h u^\varepsilon(t_m),\nabla\Phi^m\Big)&=\int_\Omega\frac{|\nabla\Phi^m|^2}{\sqrt{|\nabla U^m|^2+\varepsilon^2}}\,dx+\\ +\int_\Omega\Big[\frac{1}{\sqrt{|\nabla U^m|^2+\varepsilon^2}}-\frac{1}{\sqrt{|\nabla P_h u^\varepsilon(t_m)|^2+\varepsilon^2}}\Big]\,\nabla P_h u^\varepsilon(t_m)\cdot\nabla\Phi^m\,dx\,.\end{aligned}$$

Define

$$\gamma:=\max_{(x,t)\in\overline{\Omega}_T}\Big[\frac{|\nabla P_h u^\varepsilon(t_m)|}{\sqrt{|\nabla P_h u^\varepsilon(t_m)|^2+\varepsilon^2}}\Big].$$

We remark that $\gamma<1$ in view of the estimate (4.5). Then, using the elementary inequality $|\,|\mathbf{a}|-|\mathbf{b}|\,|\le|\mathbf{a}-\mathbf{b}|$, the second term on the right-hand side of (4.22) can be bounded as follows

$$\begin{aligned}(4.23)\qquad &\int_\Omega\Big[\frac{1}{\sqrt{|\nabla U^m|^2+\varepsilon^2}}-\frac{1}{\sqrt{|\nabla P_h u^\varepsilon(t_m)|^2+\varepsilon^2}}\Big]\,\nabla P_h u^\varepsilon(t_m)\cdot\nabla\Phi^m\,dx\\ &=\int_\Omega\frac{\nabla P_h u^\varepsilon(t_m)\cdot\nabla\Phi^m}{\sqrt{|\nabla P_h u^\varepsilon(t_m)|^2+\varepsilon^2}\,\sqrt{|\nabla U^m|^2+\varepsilon^2}}\,\frac{|\nabla P_h u^\varepsilon(t_m)|^2-|\nabla U^m|^2}{\sqrt{|\nabla P_h u^\varepsilon(t_m)|^2+\varepsilon^2}+\sqrt{|\nabla U^m|^2+\varepsilon^2}}\,dx\\ &<\gamma\int_\Omega\frac{|\nabla\Phi^m|^2}{\sqrt{|\nabla U^m|^2+\varepsilon^2}}\,dx.\end{aligned}$$

Substituting (4.22)-(4.23) into (4.21) and applying the operator $k\sum_{m=1}^{\ell}$ $(\ell\le M)$ to the inequality we get

$$\begin{aligned}(4.24)\qquad &\|\,\Phi^\ell\,\|_{L^2}^2+k\sum_{m=1}^{\ell}\Big[k\|\,d_t\Phi^m\,\|_{L^2}^2+\lambda\|\,\Phi^m\,\|_{L^2}^2+(1-\gamma)\int_\Omega\frac{|\nabla\Phi^m|^2}{\sqrt{|\nabla U^m|^2+\varepsilon^2}}\,dx\Big]\\ &\le\|\,\Phi^0\,\|_{L^2}^2+k\sum_{m=1}^{\ell}\Big[\frac{1}{\eta}\big(\|\,\mathcal{R}_m\,\|_{H^{-1}}^2+\|\,d_t\Theta^m\,\|_{H^{-1}}^2\big)+\lambda\|\,\Theta^m\,\|_{L^2}^2+\eta\|\,\nabla\Phi^m\,\|_{L^2}^2\Big],\end{aligned}$$

where $\eta>$ is some positive constant to be chosen later.

Since $\mathcal{R}_m$ can be written as

$$\mathcal{R}_m=\frac{1}{k}\int_{t_m}^{t_{m+1}}(s-t_m)\,u_{tt}(s)\,\mathrm{d}s\,,$$

from (4.6) we get

$$\begin{aligned}(4.25)\qquad \frac{k}{\eta}\sum_{m=0}^{M}\|\,\mathcal{R}_m\,\|_{H^{-1}}^2\quad&\le\quad\frac{k}{\eta}\sum_{m=0}^{M}\Big[\int_{t_m}^{t_{m+1}}(s-t_m)^2\,\mathrm{d}s\Big]\Big[\int_{t_m}^{t_{m+1}}\|\,u_{tt}(s)\,\|_{H^{-1}}^2\,\mathrm{d}s\Big]\\ &\le\quad\frac{\hat{C}_7^2}{\eta}\,k^2\,.\end{aligned}$$

It follows from (1.50) and (4.5) that

$$\| \Phi^0 \|_{L^2}^2 \leq \left(C + \hat{C}_7^2\right) h^4, \tag{4.26}$$

$$k \sum_{m=1}^{M} \Big[\frac{1}{\eta}\| d_t\Theta^m \|_{H^{-1}}^2 + \lambda\| \Theta^m \|_{L^2}^2\Big] \leq \left(\lambda + \frac{1}{\eta}\right)\hat{C}_7^2\, h^4. \tag{4.27}$$

Substituting (4.25)-(4.27) into (4.24) gives

$$\begin{aligned} \| \Phi^\ell \|_{L^2}^2 + k \sum_{m=1}^{\ell} \Big[k\| d_t\Phi^m \|_{L^2}^2 + \lambda\| \Phi^m \|_{L^2}^2 + (1-\gamma) \int_\Omega \frac{|\nabla \Phi^m|^2}{\sqrt{|\nabla U^m|^2 + \varepsilon^2}}\, dx\Big] \\ \leq \eta\, k \sum_{m=1}^{\ell} \| \nabla\Phi^m \|_{L^2}^2 + \frac{\hat{C}_7^2}{\eta}\, k^2 + \Big[\, C + \Big(1 + \lambda + \frac{1}{\eta}\Big)\hat{C}_7^2\, \Big]\, h^4. \end{aligned} \tag{4.28}$$

The last step of the proof involves a fixed point argumentation. Suppose that

$$\max_{0\leq m\leq M} \| \nabla U^m \|_{L^\infty(\Omega)} \leq C^* := 2\,\hat{C}_6. \tag{4.29}$$

Then we can choose

$$\eta = \frac{1-\gamma}{4\,C^*}$$

so that the first term on the right-hand side of (4.29) can be absorbed by the last term on the left-hand side. Applying Grownwall's inequality immediately gives

$$\max_{0\leq m\leq M} \| \Phi^m \|_{L^2}^2 + k \sum_{m=1}^{M} k\| d_t\Phi^m \|_{L^2}^2 \leq \frac{\hat{C}_7^2}{\eta}\, k^2 + \Big[\, C + \Big(1 + \lambda + \frac{1}{\eta}\Big)\hat{C}_7^2\, \Big]\, h^4, \tag{4.30}$$

$$\frac{(1-\gamma)k}{4\,C^*} \sum_{m=1}^{M} \| \nabla\Phi^m \|_{L^2}^2 \leq \frac{\hat{C}_7^2}{\eta}\, k^2 + \Big[\, C + \Big(1 + \lambda + \frac{1}{\eta}\Big)\hat{C}_7^2\, \Big]\, h^4. \tag{4.31}$$

Now, applying the triangle inequality on $E^m = \Phi^m + \Theta^m$, using the inequalities (4.6), (4.30) and (4.31) then leads to the desired estimates (1.51) and (1.52), *provided that* we can justify the induction assumption (4.29). This can be done easily as follows.

Let $\Phi(x,t)$ denote linear interpolation of $\{\Phi^m\}$ in t; hence, $\Phi(x,t)$ is continuous, piecewise linear in both x and t, and the inequality (4.31) implies

$$\frac{(1-\gamma)}{4\,C^*} \| \nabla\Phi \|_{L^2(\Omega_T)}^2 = O(h^4 + k^2).$$

By the inverse inequality bounding $L^\infty(\Omega_T)$ norm in terms of $L^2(\Omega_T)$ norm [6, 11], we get

$$\| \nabla\Phi \|_{L^\infty(\Omega_T)}^2 \leq C\,(hk)^{-1} \| \nabla\Phi \|_{L^2(\Omega_T)}^2 = O(h^3\, k^{-1} + h^{-1}\, k). \tag{4.32}$$

Clearly, for sufficient small h, the right-hand side of (4.32) is bounded by $\hat{C}_6^2$, under the parabolic mesh relation $k = O(h^2)$. Hence,

$$\| \nabla U^m \|_{L^\infty(\Omega_T)} \leq \| \nabla\Phi^m \|_{L^\infty(\Omega_T)} + \| \nabla P_h u^\varepsilon(t_m) \|_{L^\infty(\Omega_T)} \leq 2\,\hat{C}_6, \qquad 0 \leq m \leq M.$$

The proof is complete. □

Remark 4.2 *(a). In view of (4.3) and (4.6), the starting value $U^0 = Q_h u_0$, the L^2 projection of u_0 satisfies the condition (1.50).*

(b). (4.31) shows that $\nabla(U^m - P_h u^\varepsilon(t_m))$ exhibits a superconvergence property in h.

(c). Inequality (4.24) was obtained by interpreting the last term in (4.21) as a scalar product in L^2; alternatively, interpreting this term as dual pairing on $H^{-1} \times H_0^1$ then leads to a bound which involves a weaker norm of u^ε in (4.27).

Acknowledgment: The first author would like to thank his colleague Professor Bo Guan for bringing the reference [19] to his attention. Both authors would like to thank the Mathematisches Forschungsinstitut Oberwolfach for the kind hospitality and opportunity of its "Research in Pair" program.

References

[1] L. Ambrosio, N. Fusco, and D. Pallara. *Functions of bounded variation and free discontinuity problems.* The Clarendon Press Oxford University Press, New York, 2000.

[2] F. Andreu, C. Ballester, V. Caselles, and J. M. Mazón. The Dirichlet problem for the total variation flow. *J. Funct. Anal.*, 180(2):347–403, 2001.

[3] F. Andreu, C. Ballester, V. Caselles, and J. M. Mazón. Minimizing total variation flow. *Differential Integral Equations*, 14(3):321–360, 2001.

[4] F. Andreu, V. Caselles, J. I. Díaz, and J. M. Mazón. Some qualitative properties for the total variation flow. *J. Funct. Anal.*, 188(2):516–547, 2002.

[5] G. Bellettini and V. Caselles. The total variation flow in $\mathbf{R}^N$. J. Differential Equations (accepted), 2001.

[6] S. C. Brenner and L. R. Scott. *The mathematical theory of finite element methods.* Springer-Verlag, New York, second edition, 2002.

[7] H. Brézis. *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert.* North-Holland Publishing Co., Amsterdam, 1973. North-Holland Mathematics Studies, No. 5. Notas de Matemática (50).

[8] E. Casas, K. Kunisch, and C. Pola. Regularization by functions of bounded variation and applications to image enhancement. *Appl. Math. Optim.*, 40(2):229–257, 1999.

[9] T. Chan and J. Shen. On the role of the BV image model in image restoration. Tech. Report CAM 02-14, Department of Mathmatics, UCLA, 2002.

[10] T. F. Chan, G. H. Golub, and P. Mulet. A nonlinear primal-dual method for total variation-based image restoration. *SIAM J. Sci. Comput.*, 20(6):1964–1977 (electronic), 1999.

[11] P. G. Ciarlet. *The finite element method for elliptic problems.* North-Holland Publishing Co., Amsterdam, 1978. Studies in Mathematics and its Applications, Vol. 4.

[12] M. G. Crandall and T. M. Liggett. Generation of semi-groups of nonlinear transformations on general Banach spaces. *Amer. J. Math.*, 93:265–298, 1971.

[13] D. C. Dobson and C. R. Vogel. Convergence of an iterative method for total variation denoising. *SIAM J. Numer. Anal.*, 34(5):1779–1791, 1997.

[14] C. Gerhardt. Boundary value problems for surfaces of prescribed mean curvature. *J. Math. Pures Appl. (9)*, 58(1):75–109, 1979.

[15] C. Gerhardt. Evolutionary surfaces of prescribed mean curvature. *J. Differential Equations*, 36(1):139–172, 1980.

[16] D. Gilbarg and N. S. Trudinger. *Elliptic partial differential equations of second order.* Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition.

[17] E. Giusti. *Minimal surfaces and functions of bounded variation.* Birkhäuser Verlag, Basel, 1984.

[18] C. Johnson and V. Thomée. Error estimates for a finite element approximation of a minimal surface. *Math. Comp.*, 29:343–349, 1975.

[19] A. Lichnewsky and R. Temam. Pseudosolutions of the time-dependent minimal surface problem. *J. Differential Equations*, 30(3):340–364, 1978.

[20] J.-L. Lions. *Quelques méthodes de résolution des problèmes aux limites non linéaires.* Dunod, 1969.

[21] R. Rannacher. Some asymptotic error estimates for finite element approximation of minimal surfaces. *R.A.I.R.O.*, 11:181–196, 1977.

[22] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D.*, 60:259–268, 1992.

[23] J. Simon. Compact sets in the space $L^p(0,T;B)$. *Ann. Mat. Pura Appl. (4)*, 146:65–96, 1987.

Department of Mathematics
The University of Tennessee
Knoxville, TN 37996, U.S.A.
email: xfeng@math.utk.edu

Seminar of Applied Mathematics
ETH Zürich
8092 Zürich, Switzerland
email: andreas.prohl@math.ethz.ch