

RESEARCH

Open Access



A class of singular diffusion equations based on the convex–nonconvex variation model for noise removal

Gang Dong^{1*} and Boying Wu²

*Correspondence:

mathdg0206@163.com

¹Department of Applied Mathematics, Harbin University of Science and Technology, GA, 150001 Harbin, China
Full list of author information is available at the end of the article

Abstract

This paper focuses on the problem of noise removal. First, we propose a new convex–nonconvex variation model for noise removal and consider the nonexistence of solutions of the variation model. Based on the new variation method, we propose a class of singular diffusion equations and prove the of solutions and comparison rule for the new equations. Finally, experimental results illustrate the effectiveness of the model in noise reduction.

Keywords: Convex–nonconvex variation; Forward–backward diffusion equation; Noise removal

1 Introduction and motivation

Image denoising is used to recover/decompose a true image from an observed noisy image. Specifically, let $f : \Omega \rightarrow \mathbb{R}$ be a given image defined on the domain $\Omega \subset \mathbb{R}^N$. Image denoising is used to decompose f into two functions u and n with $f = u + n$, where u contains the most meaningful signals depicted by f and n represents the noise. In the ideal case, the noise part n has no signal information. The task of removing noise can be accomplished in traditional ways such as employing linear filters, which, though very simple to implement, may cause the restored image to be blurred at the edges. Various adaptive filters for noise removal have been proposed. Among these the variational method is one of the most extensively used techniques. In genal, nonlinear PDEs associated to the variational method are used as anisotropic diffusion filters because they apply different strengths of diffusivity to different locations in the image. These variational methods can be classified into the following two cases.

1.1 Convex variational model and forward diffusion equation

A classical variational model for image denoising was proposed by Rudin, Osher, and Fatemi [1]. In [1], for a given noisy image $f \in L^2(\Omega)$, the image denoising problem is equivalent to the following minimization problem (the ROF model):

$$E(u) = \int_{\Omega} |\nabla u| + \frac{\lambda}{2} \int_{\Omega} (u - f)^2 dx,$$

© The Author(s) 2021. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

where $\lambda > 0$ is a tuning parameter. In [2], Vese proposed the following general framework of variational model for image denoising:

$$E(u) = \int_{\Omega} \phi(|\nabla u|) dx + \frac{\lambda}{2} \int_{\Omega} (u - f)^2 dx.$$

The author discussed the minimizing problem $\min_{u \in BV(\Omega)} E(u)$, when $\phi(s)$ is a strictly convex function. In order to use the direct method for the calculus of variations, the convexity of the function $\phi(s)$ is always assumed. The BV norm, i.e., the total variation, is well suited for $\phi(|\nabla u|)$. And the total variation has also been widely used in other tasks of image processing, since it can help prevent the noise from staying in the denoised image u because the noise part yields a large total variation of u .

The ROF model yields very satisfactory results for removing image noise while preserving edges and contours of objects. However, it also possesses some unfavorable properties under some circumstances, such as the loss of image contrast, the smearing of corners, and the staircase effect. For instance, in [3], Meyer showed that the ROF model cannot preserve image contrast (cf. Theorem 3, p. 32) and cannot keep corners (cf. Proposition 6, p. 39). A study of the loss of image contrast can also be found in [4]. And in [5], Bellettini, Caselles, and Novaga pointed out what kind of shapes can be preserved by the ROF model, which indicates that the ROF model will smear object corners. A full discussion of these undesirable properties of the ROF model can also be found in [6].

To remedy these unfavorable properties of the ROF model, new models or techniques have been proposed [7–17]. Chan and Strong [7] proposed an adaptive total variation based on a control factor. Chambolle and Lions [8] proposed to minimize a combination of total variation and the integral of the squared norm of the gradient. Yunmei et al. [9] observed that this model is successful in restoring images where homogeneous regions are separated by distinct edges, but may become sensitive to the thresholding parameter, in the event of nonuniform image intensities or heavy degradation. And Yunmei et al. [9] proposed a variable-exponent approach adaptive model which exploits the benefits of Gaussian smoothing and the strength of TV regularization. On the other hand, in [10, 11], the authors introduced new variational models based on high-order derivatives of the denoised image u . In addition to the basic requirements of images denoising, such as edge preservation and noise removal, these new models effectively ameliorate the staircase effect.

It is worth mentioning that the diffusion equations associated to these methods are the forward anisotropic diffusion equations, which smooth homogeneous regions while preserving edges. However, these diffusion equations cannot enhance the image, for example, by preserving corners, smoothing parts of objects, as well as image greyscale intensity contrasts.

1.2 Nonconvex variational model and backward diffusion equation

Most of those existing algorithms are based on a convex potential. For a convex potential, ϕ needs to increase near-linearly at least; but for better edge-preservation, ϕ needs to increase less-linearly, then ϕ becomes nonconvex, and such a form of ϕ has been suggested in [18–23]. It is interesting that Vese proposes several variational models for image denoising, when $\phi(s)$ is nonconvex and then implements numerical simulations for this case in [2]. Unfortunately, there is not necessarily a unique solution for the variational model

with nonconvex potential, and Chipot et al. [22] proved that there is no minimizer in any reasonable space if f is not a constant. And they introduced the following energy:

$$E_\varepsilon(u) = \lambda \int_{\Omega} \frac{|\nabla u|^2}{1 + |\nabla u|^2} dx + \varepsilon \int_{\Omega} |\nabla u|^2 dx + \int_{\Omega} (u - f)^2 dx.$$

They proved that $E_\varepsilon(u)$ is convex for $\varepsilon \geq \lambda/4$ and nonconvex for $\varepsilon < \lambda/4$; for $\varepsilon < \lambda/4$, $E_\varepsilon(u)$ has quadratic growth at infinity, and then they use convexification tools to obtain the existence of a minimizer for $E_\varepsilon(u)$ in the one-dimensional case. For dimensions greater than one, the problem is quite open. The behavior of the minimizing sequence is also a challenging problem, which is closely related to Perona and Malik anisotropic diffusion [23] whose associated potential is nonconvex also. In spite of the lack of a rigorous mathematical theory for the continuous minimization problem with nonconvex potential, its associated discrete version can be solved numerically, for example, with the gradient descent algorithm [23], the simulated annealing algorithm [24], the half-quadratic algorithms [18, 20, 21, 25–29], and so on. The nonconvex potential always leads to the backward diffusion equation or the forward–backward diffusion equation, which can sharpen the edges, corners, as well as the singular features.

In this paper, we intend to propose a new convex–nonconvex variational model for image denoising. In addition to removing noise and keeping edges and contours of objects, the new model aims at preserving corners, smoothing parts of objects, as well as image greyscale intensity contrasts. As corners and edges differ from ordinary points or contours in their singularities, a natural idea is to incorporate the related geometric quantities into the process of denoising. Our idea can be described as follows: First, inspired by [22], instead of $\varepsilon \int_{\Omega} |\nabla u|^2 dx$, we consider the linear growth functional $\varepsilon \int_{\Omega} \psi(|\nabla u|) dx$, which can also preserve edges and corners. The new variational model is a combination of the convex and nonconvex variational models. Second, based on the new idea, we propose a new variational framework for image denoising, which is under some basic hypotheses and may not satisfy the convexity condition. Third, we propose and analyze a class of singular diffusion equations associated with the new variational model. To efficiently solve the singular diffusion equations, one might employ some fast methods, such as AOS [30], etc. In this paper, we also use the standard time marching scheme and PM scheme [1].

In fact, the anisotropic diffusion equation has been widely used in the modeling of image processing during the last two decades. In the famous work [23], Perona and Malik proposed a framework to deal with the denoising problem based on the diffusion equation. To make the images more pleasing to the eye, it would be useful to reduce staircasing effects. Many models reducing this effect have been proposed in the literature. In [31, 32], Charbonnier and Weickert developed and studied the forward diffusion equation by proposing the different diffusivity. In [33], Catt et al. proposed the regularization of the Perona and Malik model to obtain a smoother image. In [34], Keeling et al. proposed the nonlinear anisotropic diffusion filtering for multiscale edge enhancement. In [35] Gilboa et al. proposed forward–backward diffusion processes for adaptive image enhancement and denoising. In [36], Smolka proposed combined forward and backward anisotropic diffusion filtering of color images. These forward–backward diffusions are related to nonconvex potentials. In all these works, the nonlinear anisotropic diffusion equation was considered, while in our present work, based on the new nonconvex variational model, we consider the singular forward–backward diffusion equation for denoising, which can obtain the

singular solution to preserve the singular parts of the image, such as edges, corners, and so on.

The rest of this paper is organized as follows. In Sect. 2, our convex–nonconvex variational model is introduced in detail. We discuss the ill-posed problem. We prove the existence of Young measure solutions in Sect. 3. Of course, we are very interested in the investigation of the properties of Young measure solutions in Sect. 4. Numerical implementation is then developed in Sect. 5. We list numerical experiments for synthetic image denoising as well as real world image denoising, and compare our results with those obtained by the ROF model. A conclusion is subsequently given in Sect. 6.

2 Convex–nonconvex variational framework for denoising model

The following new variational model is proposed:

$$E(u) = \mu_1 \int_{\Omega} \psi_C(|\nabla u|) dx + \mu_2 \int_{\Omega} \psi_{NC}(|\nabla u|) dx + \lambda \int_{\Omega} (u - f)^2 dx, \quad (1)$$

where $\mu_1 > 0$, $\mu_2 > 0$. The functions $\psi_C(s)$ and $\psi_{NC}(s)$ are convex and nonconvex, respectively. For the image processing, $\psi_C(s)$ satisfies the following assumptions [2]:

- ψ_C is a strictly convex, nondecreasing function from \mathbb{R}^+ to \mathbb{R}^+ , with $\psi_C(0) = 0$ (without a loss of generality);
- $\lim_{s \rightarrow +\infty} \psi_C(s) = +\infty$;
- There exist two constants $c > 0$ and $b \geq 0$ such that

$$cs - b \leq \psi_C(s) \leq cs + b, \quad \forall s \geq 0,$$

and $\psi_{NC}(s)$ satisfies the following assumptions [18]:

- ψ_{NC} is nonconvex;
- $\psi_{NC} \approx cs^2$ as $s \rightarrow 0^+$;
- $\lim_{s \rightarrow +\infty} \psi_{NC}(s) \approx \gamma > 0$.

Compared with the conditions about ϕ_C and ϕ_{NC} in [2] and [18], the hypotheses on ψ_{NC} and ψ_C in this paper are as follows:

(H1) $\psi_C \in C^1(\mathbb{R}^N)$. There exist two constants $0 < \lambda \leq \Lambda$ such that

$$(\lambda|X| - 1)^+ \leq \psi_C(|X|) \leq \Lambda|X| + 1, \quad \forall X \in \mathbb{R}^N;$$

(H2) $Z(X) = \nabla \psi_C(X)$ and $|Z(X)| \leq \Lambda$;

(H3) Moreover, we assume that there exist a sequence $\{\varphi_p\}_{1 < p < 2} \subset C^1(\mathbb{R}^N)$ and $C_0 > 0$ such that $\{Z_p = \nabla \varphi_p\}_{1 < p < 2}$ locally and uniformly converges to Z in \mathbb{R}^N . For all $p \in (1, 2)$, φ_p and Z_p satisfy the structure conditions

$$(\lambda|X|^p - 1)^+ \leq \varphi_p(X) \leq \Lambda|X|^p + 1, \quad \forall X \in \mathbb{R}^N,$$

and

$$|Z_p(X)| \leq \Lambda|X|^{p-1}, \quad \forall X \in \mathbb{R}^N;$$

(H4) $\psi_{NC} \in C^1(\mathbb{R}^N)$ and ψ_{NC} is a nonconvex function;

(H5) $\lim_{s \rightarrow +\infty} \frac{\psi_{NC}(s)}{s} = 0$.

The new variational model is a combination of the convex and nonconvex variational models. Hence, the new model is demonstrated to be capable of achieving a good trade-off between noise removal and edge preservation which the convex and nonconvex variational models are respectively good at. This is not a simple combination: ψ_C controls the growth of the new functional and the regularization of the solution; ψ_{NC} can't only influence the growth of the new functional, but also preserve the singular parts of the image, such as corners, image contrast, edges, and so on, and furthermore, ψ_{NC} controls the convexity of the functional.

In order to use the Young measure theory in [37–40], we have to assume Hypothesis (H3) on ψ_C . However, Hypothesis (H3) is easy to be satisfied, for example, if $\psi_C = \sqrt{1 + |s|^2}$. On the other hand, the new hypotheses are different from the assumptions in [2, 18], since the new hypotheses do not restrict the convexity of the functionals. Hence, based on the new framework, we can propose many interesting models. Assuming the hypotheses above, it is difficult to confirm the convexity of the new variational model, which may yield an ill-posed problem. The existence may indeed not be straightforward. In the next section, the existence of solutions of the singular diffusion equations based on the convex–nonconvex variation is considered. In [41, 42], Guidotti proposed two types of backward–forward regularization of the Perona–Malik equation. The two models are contained in the new framework when $\psi_{NC} = \ln(1 + |s|^2)$, $\psi_C = |s|^2$, and $\psi_C = |s|^{p-2}s$, respectively.

2.1 Some special examples

Nevertheless, the potentials satisfy Hypotheses (H1)–(H3), if

$$\begin{aligned}\psi_C(s) &= \sqrt{1 + s^2}, \\ \psi_C(s) &= \sqrt{\beta + s^2}, \quad 0 < \beta < 1, \\ \psi_C(s) &= \begin{cases} \frac{1}{2}s^2, & |s| < 1, \\ s - \frac{1}{2}, & |s| \geq 1, \end{cases} \\ \psi_C(s) &= s - \frac{1}{K} \ln(1 + Ks), \quad K > 1, \\ \psi_C(s) &= \begin{cases} \sqrt{1 + s^2} - \sqrt{1 - \beta^2}, & |s| < \frac{\beta}{\sqrt{1 - \beta^2}}, \\ \beta|s|, & |s| \geq \frac{\beta}{\sqrt{1 - \beta^2}}, \end{cases} \quad 0 < \beta < 1,\end{aligned}$$

and so on. On the other hand, the potentials satisfy Hypotheses (H4)–(H5), like in [2, 18–23], and there are a lot of nonconvex functionals, such as

$$\begin{aligned}\psi_{NC}(s) &= \frac{s^2}{1 + s^2}, \\ \psi_{NC}(s) &= |s|^\alpha, \quad 0 < \alpha < 1, \\ \psi_{NC}(s) &= \sqrt{1 + |s|^{2\alpha}}, \quad 0 < \alpha < 1, \\ \psi_{NC}(s) &= \ln(1 + |s|^2),\end{aligned}$$

and so on. Moreover,

$$E(u) = \int_{\Omega} \mu_1 \sqrt{1 + |\nabla u|^2} dx + \int_{\Omega} \mu_2 \psi_{\text{NC}}(|\nabla u|) dx + \lambda \int_{\Omega} (u - f)^2 dx,$$

where ψ_{NC} is any nonconvex functional which satisfies Hypotheses (H4)–(H5).

In this paper, the following model is considered:

$$E(u) = \int_{\Omega} \mu_1 \sqrt{1 + |\nabla u|^2} dx + \int_{\Omega} \mu_2 \ln(1 + |\nabla u|^2) dx + \lambda \int_{\Omega} (u - f)^2 dx, \quad (2)$$

which can be rewritten as

$$E(u) = E_C(u) + E_{\text{NC}}(u),$$

where

$$E_{\text{NC}}(u) = \mu_2 \int_{\Omega} \ln(1 + |\nabla u|^2) dx + \frac{\lambda}{2} \int_{\Omega} (u - f)^2 dx$$

and

$$E_C(u) = \mu_1 \int_{\Omega} \sqrt{1 + |\nabla u|^2} dx + \frac{\lambda}{2} \int_{\Omega} (u - f)^2 dx.$$

Following the proof given by Chipot et al. [22], we have

Theorem 1 *If $f(x)$ is not a constant and $f \in L^\infty(\Omega)$, then the function $E_{\text{NC}}(u)$ has no minimizer in $W^{1,2}(\Omega)$ and $\inf_{u \in W^{1,2}(\Omega)} E_{\text{NC}}(u) = 0$.*

Proof The theorem in the one-dimensional case $\Omega = (a, b)$ is proved for the clarity sake, and the same proof goes for $N \geq 2$. It's clear that

$$E_{\text{NC}}(u) \leq \mu_2 \int_{\Omega} |\nabla u|^{0.8} dx + \frac{\lambda}{2} \int_{\Omega} (u - f)^2 dx.$$

Let

$$E_\alpha(u) = \mu_2 \int_{\Omega} |\nabla u|^\alpha dx + \frac{\lambda}{2} \int_{\Omega} (u - f)^2 dx,$$

and then we will prove that the theorem is true for $E_\alpha(u)$ with $0 < \alpha < 1$.

By density, we always may find a sequence of step functions \tilde{u}_n such that

$$|\tilde{u}_n| \leq |f|_{L^\infty}, \quad \lim_{n \rightarrow +\infty} \|\tilde{u}_n - f\|_{L^2(\Omega)} = 0.$$

In fact, we can find a partition $a = x_0 < x_1 < \dots < x_n = b$ such that \tilde{u}_n is the constant $\tilde{u}_{n,i}$ on each interval (x_i, x_{i+1}) , $h_n = \max_i (x_i - x_{i-1}) < 1$ with $\lim_{n \rightarrow +\infty} h_n = 0$. Let us set $\sigma_i = x_i - x_{i-1}$. Next, we define a sequence of continuous functions u_n by

$$u_n(x) = \begin{cases} \tilde{u}_n & \text{if } x \in [x_{i-1}, x_i - \sigma_i^{2/(1-\alpha)}], \\ \frac{\tilde{u}_{n,i+1} - \tilde{u}_{n,i}}{\sigma_i^{2/(1-\alpha)}} (x - x_i) + \tilde{u}_{n,i+1} & \text{if } x \in [x_i - \sigma_i^{2/(1-\alpha)}, x_i]. \end{cases}$$

Note that

$$\begin{aligned} |\tilde{u}_n - u_n|_{L^2(\Omega)}^2 &= \sum_{i=1}^n \int_{x_i - \sigma_i^{2/(1-\alpha)}}^{x_i} (\tilde{u}_{n,i+1} - \tilde{u}_{n,i})^2 \left(\frac{x - x_i}{\sigma_i^{2/(1-\alpha)}} + 1 \right)^2 dx \\ &\leq \frac{4}{3} |f|_{L^\infty(\Omega)}^2 \sum_{i=1}^n \sigma_i^{2/(1-\alpha)} \\ &\leq \frac{4}{3} |f|_{L^\infty(\Omega)}^2 h_n^{(1+\alpha)/(1-\alpha)} \sum_{i=1}^n \sigma_i \\ &= \frac{4}{3} |f|_{L^\infty(\Omega)}^2 (b-a) h_n^{(1+\alpha)/(1-\alpha)}, \end{aligned}$$

and therefore,

$$\lim_{n \rightarrow +\infty} |\tilde{u}_n - u_n|_{L^2(\Omega)} = 0.$$

Since

$$|u_n - f|_{L^2(\Omega)} \leq |\tilde{u}_n - u_n|_{L^2(\Omega)} + |\tilde{u}_n - f|_{L^2(\Omega)},$$

taking the limit on both sides yields

$$\lim_{n \rightarrow +\infty} |u_n - f|_{L^2(\Omega)} = 0.$$

Moreover,

$$\begin{aligned} \frac{1}{\alpha} \int_a^b |\nabla u|^\alpha dx &= \frac{1}{\alpha} \sum_{i=1}^n \int_{x_i - \sigma_i^{2/(1-\alpha)}}^{x_i} \frac{(\tilde{u}_{n,i+1} - \tilde{u}_{n,i})^\alpha}{\sigma_i^{2\alpha/(1-\alpha)}} dx \\ &\leq \frac{1}{\alpha} 2^\alpha |f|_{L^\infty(\Omega)}^\alpha \sum_{i=1}^n \sigma_i^2 \leq \frac{1}{\alpha} 2^\alpha |f|_{L^\infty(\Omega)}^\alpha h_n \sum_{i=1}^n \sigma_i = \frac{1}{\alpha} 2^\alpha |f|_{L^\infty(\Omega)}^\alpha h_n (b-a). \end{aligned}$$

Thus

$$\lim_{n \rightarrow +\infty} \frac{1}{\alpha} \int_a^b |\nabla u|^\alpha dx = 0,$$

and finally,

$$0 \leq \inf_{u \in W^{1,2}(\Omega)} E_\alpha(u) \leq \lim_{n \rightarrow +\infty} E_\alpha(u_n) = 0,$$

i.e.,

$$\inf_{u \in W^{1,2}(\Omega)} E_\alpha(u) = 0.$$

Now, if there exists a minimizer $u \in W^{1,2}(\Omega)$, then necessarily $E_\alpha(u) = 0$, which implies

$$\int_a^b |u - f|^2 dx = 0 \quad \Leftrightarrow \quad u = f \quad \text{a.e.},$$

$$\frac{1}{\alpha} \int_a^b |\nabla u|^2 dx = 0 \quad \Leftrightarrow \quad u' = 0 \quad \text{a.e.}$$

The first equality is possible only if $f \in W^{1,2}(\Omega)$, and in this case the second equality implies $f' = 0$, which is possible only if f is a constant. Therefore, excluding this trivial case, $E_{\text{NC}}(u)$ has no minimizer in $W^{1,2}(\Omega)$. \square

Remark 1 As we know, if the region Ω is bounded,

$$W^{1,2}(\Omega) \subset W^{1,1}(\Omega) \subset \text{BV}(\Omega).$$

Then

$$\inf_{u \in W^{1,2}(\Omega)} E_{\text{NC}}(u) \geq \inf_{u \in W^{1,1}(\Omega)} E_{\text{NC}}(u) \geq \inf_{u \in \text{BV}(\Omega)} E_{\text{NC}}(u).$$

Note that $E_{\text{NC}}(u) \geq 0$, and therefore

$$\inf_{u \in \text{BV}(\Omega)} E_{\text{NC}}(u) = 0.$$

However, we cannot obtain any information about the minimizer of $E_{\text{NC}}(u)$ in $\text{BV}(\Omega)$.

3 A class of singular diffusion equations for denoising model

3.1 Singular diffusion equations based on the convex–nonconvex variation

Based on the new variational model, the following diffusion equation is proposed:

$$\frac{\partial u}{\partial t} = \mu_1 \operatorname{div} \left(\frac{\nabla u}{\sqrt{1 + |\nabla u|^2}} \right) + \mu_2 \operatorname{div} \left(\frac{\nabla u}{1 + |\nabla u|^2} \right) - \lambda(u - f), \quad (3)$$

$$u(x, 0) = f, \quad x \in \Omega, \quad (4)$$

$$\frac{\partial u}{\partial \vec{n}} = 0, \quad (x, t) \in \partial\Omega \times (0, T). \quad (5)$$

For this special equation, what we obtain in this paper will reveal another aspect for the existence, namely the existence of a discontinuous solution. Note that the equation is strongly degenerate at the discontinuous points of such a solution. On the other hand, the new equation can be considered as a perturbation of Perona–Malik model [23]. Such a perturbation is not the usual viscous one, for example, Δu or $\Delta^2 u$, which has standard regularity effects. The perturbation has no hazard for the equation to permit the existence of discontinuous solutions, which has particular meaning: with the new perturbation, the new model is still an anisotropic diffusion equation. That is to say, inside the regions where the magnitude of the gradient of u is weak, the new equation acts as Gaussian smoothing, resulting in isotropic smoothing; near the region's boundaries where the magnitude of the gradient is large, the regularization is “stopped” and the edges are preserved.

Let

$$Z(X) = \nabla \varphi(X) = \mu_1 \frac{X}{\sqrt{1 + |X|^2}} + \mu_2 \frac{X}{1 + |X|^2}$$

and

$$\varphi(X) = \mu_1 \sqrt{1 + |X|^2} + \frac{\mu_2}{2} \ln(1 + |X|^2),$$

for $X \in \mathbb{R}^N$. Therefore, the new diffusion equation can be rewritten as

$$\frac{\partial u}{\partial t} = \operatorname{div}(Z(\nabla u)) - \lambda(u - f), \quad (x, t) \in \Omega \times (0, T),$$

$$u(x, 0) = f, \quad x \in \Omega,$$

$$\frac{\partial u}{\partial \vec{n}} = 0, \quad (x, t) \in \partial\Omega \times (0, T).$$

Let φ^{**} denote the convexification of φ , namely,

$$\varphi^{**}(X) = \sup\{l(X) : l \leq \varphi, l \text{ is convex}\},$$

and

$$Z(X) = \nabla \varphi(X), \quad Z^{**}(X) = \nabla \varphi^{**}(X), \quad X \in \mathbb{R}^N.$$

Since $\varphi \in C^1(\mathbb{R}^N)$, $\varphi^{**} \in C^1(\mathbb{R}^N)$ is convex.

Definition 1 A Young measure solution to problem (3)–(5) is a function

$$u \in L^\infty((0, T); \operatorname{BV}(\Omega)) \cap L^\infty(Q_T), \quad \frac{\partial u}{\partial t} \in L^2(Q_T)$$

and there exists a $W^{1,1}(Q_T)$ -gradient Young measure $\nu = (\nu_{x,t})_{(x,t) \in Q_T}$ on \mathbb{R}^N such that

$$\iint_{Q_T} (\langle \nu, Z \rangle \cdot \nabla \zeta + u_t \zeta + \lambda(u - f_p) \zeta) dx dt = 0,$$

$$\forall \zeta \in C^\infty(\overline{Q_T}) \text{ with } \zeta(x, 0) = \zeta(x, T) = 0, \quad (6)$$

$$\nabla u(x, t) = \langle \nu_{x,t}, \operatorname{id} \rangle \quad \text{a.e. } (x, t) \in Q_T, \quad (7)$$

$$\langle \nu_{x,t}, Z \cdot \operatorname{id} \rangle \leq \langle \nu_{x,t}, Z \rangle \cdot \langle \nu_{x,t}, \operatorname{id} \rangle \quad \text{a.e. } (x, t) \in Q_T, \quad (8)$$

$$\operatorname{supp} \nu_{x,t} \subset \{X \in \mathbb{R}^N : \varphi(X) = \varphi^{**}(X)\} \quad \text{a.e. } (x, t) \in Q_T, \quad (9)$$

where id is the identity mapping in \mathbb{R}^N ,

$$\langle \nu, Z \rangle = \int_{\mathbb{R}^N} Z(X) d\nu(X), \quad \langle \nu, \operatorname{id} \rangle = \int_{\mathbb{R}^N} X d\nu(X), \quad \langle \nu, Z \cdot \operatorname{id} \rangle = \int_{\mathbb{R}^N} Z(X) d\nu(X),$$

and

$$u(x, 0) = f(x), \quad x \in \Omega, \quad (10)$$

in the sense of trace.

Theorem 2 Let $f \in \text{BV}(\Omega) \cap L^\infty(\Omega)$. Then problem (3)–(5) admits at least one Young measure solution.

3.2 Preliminaries

We use $C_0(\mathbb{R}^d)$ to denote the closure of the set of continuous functions on \mathbb{R}^d with compact supports. The dual of $C_0(\mathbb{R}^d)$ can be identified with the space $\mathcal{M}(\mathbb{R}^d)$ of signed Radon measures with finite mass via the pairing

$$\langle \nu, f \rangle = \int_{\mathbb{R}^d} f d\nu_x.$$

Let $D \subset \mathbb{R}^n$ be a measurable set of finite measure. A map $\nu : D \rightarrow \mathcal{M}(\mathbb{R}^d)$ is called weakly* measurable if the functions $x \mapsto \int_{\mathbb{R}^d} f d\nu_x$ are measurable for all $f \in C_0(\mathbb{R}^d)$, where $\nu_x = \mu(x)$.

For $p \geq 1$, define

$$\mathcal{E}_0^p(\mathbb{R}^d) = \left\{ \varphi \in C(\mathbb{R}^d) : \lim_{|X| \rightarrow +\infty} \frac{|\varphi(X)|}{1 + |X|^p} \text{ exists} \right\}.$$

As noted in [11], the space $\mathcal{E}_0^p(\mathbb{R}^d)$ is a separable Banach space with the norm

$$\|\psi\|_{\mathcal{E}_0^p} = \sup_{X \in \mathbb{R}^d} \frac{|\psi(X)|}{1 + |X|^p}.$$

We define

$$\mathcal{E}^p(\mathbb{R}^d) = \left\{ \varphi \in C(\mathbb{R}^d) : \sup_{|X| \in \mathbb{R}^d} \frac{|\varphi(X)|}{1 + |X|^p} < +\infty \right\},$$

which is an inseparable space with the above norm.

Definition 2 Let $p \geq 1$. A Young measure $\nu = (\nu_x)_{x \in D}$ on \mathbb{R}^d is called a $W^{1,p}$ -gradient Young measure if

- (i) $x \in D \mapsto \int_{\mathbb{R}^d} f d\nu_x \in \mathbb{R}$ is a Lebesgue measurable function for all f bounded and continuous on \mathbb{R}^d ;
- (ii) There is a sequence of functions $\{u^k\}_{k=1}^\infty \subset W^{1,p}(D)$ for which the representation formula

$$\lim_{k \rightarrow \infty} \int_E \psi(\nabla u^k(x)) dx = \int_E \langle \nu_x, \psi \rangle dx \quad (11)$$

holds for all measurable $E \subset D$ and all $\psi \in \mathcal{E}_0^p(\mathbb{R}^d)$, where $\langle \nu_x, \psi \rangle = \int_{\mathbb{R}^d} \psi d\nu_x$.

We also call ν the $W^{1,p}(D)$ -gradient Young measure generated by $\{\nabla u^k\}_{k=1}^\infty$ and $\{\nabla u^k\}_{k=1}^\infty$ the $W^{1,p}(D)$ -gradient generating sequence of ν . In addition, the representation formula (11) also holds for $\psi \in \mathcal{E}^p(\mathbb{R}^d)$. By the fundamental theorem for Young measure, we see that

$$\|\nu(x)\|_{\mathcal{M}(\mathbb{R}^d)} = 1 \quad \text{a.e. } x \in D.$$

Definition 3 Let $\{z^k\}_{k=1}^\infty \subset L^1(D)$ and $z \in L^1(D)$. We say that $\{z^k\}_{k=1}^\infty$ converges to z in the biting sense if there is a decreasing sequence of subsets $E_{j+1} \subset E_j$ of D with $\lim_{j \rightarrow \infty} \text{meas}(E_j) = 0$ such that $\{z^k\}_{k=1}^\infty$ converges weakly to z in $L^1(D \setminus E_j)$ for all j .

Definition 4 Let $p \geq 1$. A Young measure $\nu = (\nu_x)_{x \in D}$ on \mathbb{R}^d is called a $W^{1,p}(D)$ -bitting Young measure if there is a sequence $\{z^k\}_{k=1}^\infty \subset L^p(D)$ and $z \in L^1(D)$ such that $\{|z^k|^p\}_{k=1}^\infty$ converges to z and $\{\psi(z^k(x))\}_{k=1}^\infty$ converges to $\langle \nu_x, \psi \rangle$ in the biting sense for all $\psi \in \mathcal{E}_0^p(\mathbb{R}^d)$ (or $\mathcal{E}^p(\mathbb{R}^d)$).

We also call ν the $W^{1,p}(D)$ -bitting Young measure generated by $\{z^k\}_{k=1}^\infty$ and $\{z^k\}_{k=1}^\infty$ the $W^{1,p}(D)$ -bitting generating sequence of ν . By the fundamental theorem for Young measure, we see that

$$\|\nu(x)\|_{\mathcal{M}(\mathbb{R}^d)} = 1 \quad \text{a.e. } x \in D.$$

Kinderlehrer and Pedregal [37] showed a property which characterizes $W^{1,p}$ -gradient Young measures as described in the following lemma.

Lemma 1 Let $\nu = (\nu_x)_{x \in D}$ be a Young measure on \mathbb{R}^d . Then $\nu = (\nu_x)_{x \in D}$ is a $W^{1,p}(D)$ -gradient Young measure if and only if

- (i) There exists $u \in W^{1,p}(D)$ such that

$$\nabla u(x) = \int_{\mathbb{R}^d} A \, d\nu_x(x) \quad \text{a.e. } x \in D;$$

- (ii) Jensen's inequality

$$\psi(\nabla u(x)) \leq \int_{\mathbb{R}^d} \psi(A) \, d\nu_x(A)$$

holds for all $\psi \in \mathcal{E}^p(\mathbb{R}^d)$ continuous, quasiconvex, and bounded below;

- (iii) The function

$$x \mapsto \int_{\mathbb{R}^d} |A|^p \, d\nu_x(A)$$

is in $L^1(D)$.

We give the following two lemmas. The proofs can be found in [38, 39].

Lemma 2 Suppose $f \in \mathcal{E}^p(\mathbb{R}^d)$, for some $p \geq 1$, is quasiconvex and bounded below and let $\{u^k\}_{k=1}^\infty$ converge weakly to u in $W^{1,p}(D)$. Then

- (i) For all measurable $E \subset D$,

$$\int_E f(\nabla u) \, dx \leq \liminf_{k \rightarrow \infty} \int_E f(\nabla u^k) \, dx;$$

- (ii) If

$$\lim_{k \rightarrow \infty} \int_D f(\nabla u^k) \, dx = \int_D f(\nabla u) \, dx,$$

then $\{f(\nabla u^k)\}_{k=1}^\infty$ are weakly sequentially precompact in $L^1(D)$ and the sequence converges weakly to $f(\nabla u)$.

Lemma 3 Let f and $\{u^k\}_{k=1}^\infty$ be as in Lemma 2 (ii) and assume in addition that

$$(c|X|^p - 1)^+ \leq f(X) \leq C|X|^p + 1$$

for $0 < c \leq C$. Let ν be generated by the gradients $\{\nabla u\}_{k=1}^\infty$. Then ν is a $W^{1,p}(D)$ -gradient Young measure.

We now state and prove a result for the sequences of gradient-generated Young measures [40].

Lemma 4 Let $1 \leq p < 2$. Suppose that $\{\nu^\alpha = (\nu_x^\alpha)_{x \in D}\}_{\alpha > 0}$ is a family of $W^{1,p}(D)$ -gradient Young measures and each is generated by $\{\nabla u^{\alpha,m}\}_{m=1}^\infty$, where $u^{\alpha,m}$ is in $W^{1,p}(D)$ uniformly bounded in α and m . Then there exist a subsequence of $\{\nu^\alpha\}_{\alpha > 0}$, denoted by $\{\nu^{\alpha_i}\}_{i=1}^\infty$, and a $W^{1,p}(D)$ -gradient Young measure ν such that

- (i) $\{\nu^{\alpha_i}\}_{i=1}^\infty$ converges weakly* to ν in $L^\infty(D; \mathcal{M}(\mathbb{R}^d))$, namely, $\{\langle \nu^{\alpha_i}, \psi \rangle\}_{i=1}^\infty$ converges weakly* to $\langle \nu, \psi \rangle$ in $L^\infty(D)$ for all $\psi \in C_0(\mathbb{R}^d)$;
- (ii) For $1 \leq q < p$, $\{\nu^{\alpha_i}\}_{i=1}^\infty$ converges weakly to ν in $L^1(D; (\mathcal{E}_0^q(\mathbb{R}^d))')$, namely, $\{\langle \nu^{\alpha_i}, \psi \rangle\}_{i=1}^\infty$ converges weakly to $\langle \nu, \psi \rangle$ in $L^\infty(D)$ for all $\psi \in \mathcal{E}_0^q(\mathbb{R}^d)$;
- (iii) $\{\nu^{\alpha_i}\}_{i=1}^\infty$ converges to ν in $L^1(D; (\mathcal{E}_0^q(\mathbb{R}^d))')$ in the biting sense, namely, $\psi \in \mathcal{E}_0^p(\mathbb{R}^d)$, $\{\langle \nu^{\alpha_i}, \psi \rangle\}_{i=1}^\infty$ converges to $\langle \nu, \psi \rangle$ in the biting sense.

3.3 Existence of solution to the approximation problem

Since equation (3) is degenerate, singular, and forward–backward, some necessary approximations are required for discussing the existence of solutions. Our approximations will be divided into two steps. For this purpose, we need to approximate the initial datum f . By the density properties of BV functions in [7], there exists some subsequence $\{f_p\} \subset C_0^\infty(\Omega)$ such that $\|f_p\|_{L^\infty(\Omega)}$ and $\|\nabla f_p\|_{L^1(\Omega)}$ are uniformly bounded in p , and $\{f_p\}_{0 < p < 1}$ converges to f in $L^1(\Omega)$.

As the first step, we consider the following evolution problem:

$$\frac{\partial u}{\partial t} = \operatorname{div}(Z_p(\nabla u)) - \lambda(u - f_p), \quad (x, t) \in \Omega \times (0, T), \quad (12)$$

$$u(x, 0) = f_p, \quad x \in \Omega, \quad (13)$$

$$\frac{\partial u}{\partial \vec{n}} = 0, \quad (x, t) \in \partial\Omega \times (0, T), \quad (14)$$

where

$$Z_p(X) = \nabla \varphi_p(X) = \mu_1 \frac{|X|^{2(p-1)}X}{\sqrt{1 + \frac{1}{1+\delta}|X|^{2p}}} + \mu_2 \frac{X}{1 + |X|^2}$$

and

$$\varphi_p(X) = \mu_1 \sqrt{1 + \frac{1}{p}|X|^{2p}} + \frac{\mu_2}{2} \ln(1 + |X|^2).$$

Let φ_p^{**} denote the convexification of φ_p , namely,

$$\varphi_p^{**}(X) = \sup \{f(X) : f \leq \varphi_p, f \text{ is convex}\}$$

and

$$Z_p^{**} = \nabla \varphi_p^{**}.$$

Since $\varphi_p \in C^1(\mathbb{R}^N)$, $\varphi_p^{**} \in C^1(\mathbb{R}^N)$ is convex. In addition,

$$Z_p(X) \cdot \text{id}(X) \geq Z(X) \cdot \text{id}(X), \quad \forall X \in \{X \in \mathbb{R}^N : |X| \geq C_0\},$$

$$\{X \in \mathbb{R}^N : \varphi_p(X) = \varphi_p^{**}(X)\} \subset \{X \in \mathbb{R}^N : \varphi(X) = \varphi^{**}(X)\}.$$

Definition 5 A Young measure solution to problem (12)–(14) is a function

$$u \in L^\infty((0, T); W^{1,p}(\Omega)) \cap L^\infty(Q_T), \quad \frac{\partial u}{\partial t} \in L^2(Q_T)$$

and there exists a $W^{1,p}(Q_T)$ -gradient Young measure $\nu = (\nu_{x,t})_{(x,t) \in Q_T}$ on \mathbb{R}^N such that

$$\iint_{Q_T} (\langle \nu, Z_p \rangle \cdot \nabla \zeta + u_t \zeta + \lambda(u - f_p)\zeta) dx dt = 0,$$

$$\forall \zeta \in C^\infty(\overline{Q_T}) \text{ with } \zeta(x, 0) = \zeta(x, T) = 0, \quad (15)$$

$$\nabla u(x, t) = \langle \nu_{x,t}, \text{id} \rangle \quad \text{a.e. } (x, t) \in Q_T, \quad (16)$$

$$\langle \nu_{x,t}, Z_p \cdot \text{id} \rangle = \langle \nu_{x,t}, Z_p \rangle \cdot \langle \nu_{x,t}, \text{id} \rangle \quad \text{a.e. } (x, t) \in Q_T, \quad (17)$$

$$\text{supp } \nu_{x,t} \subset \{X \in \mathbb{R}^N : \varphi_p(X) = \varphi_p^{**}(X)\} \quad \text{a.e. } (x, t) \in Q_T, \quad (18)$$

and

$$u(x, 0) = f_p(x), \quad x \in \Omega, \quad (19)$$

in the sense of trace.

Theorem 3 Let $f_p \in W^{1,p}(\Omega) \cap L^\infty(\Omega)$. Then problem (12)–(14) admits at least one Young measure solution.

The following existence proof follows by the ideas due to Kinderlehrer and Pedregal [39], Demoulini [38], and Yin and Wang [40].

In order to obtain the theorem above, the following functionals defined on $W^{1,p}(\Omega)$ are considered:

$$\mathcal{F}_h(v; u^{h,j-1}) = \int_\Omega \varphi_p(\nabla v) dx + \frac{1}{2h} \int_\Omega (v - u^{h,j-1})^2 dx + \frac{\lambda}{2} \int_\Omega |v - f_p|^2 dx, \quad v \in W^{1,p}(\Omega),$$

and

$$\mathcal{F}_h^{**}(v; u^{h,j-1}) = \int_\Omega \varphi_p^{**}(\nabla v) dx + \frac{1}{2h} \int_\Omega (v - u^{h,j-1})^2 dx + \frac{\lambda}{2} \int_\Omega |v - f_p|^2 dx, \quad v \in W^{1,p}(\Omega),$$

where $0 < h < 1$, $u^{h,0} = f_p$, j is an integer and $1 \leq j \leq T/h + 1$.

Lemma 5 *There exists $u^{h,j} \in W^{1,p}(\Omega) \cap L^\infty(\Omega)$ such that $u^{h,j}$ is a minimum of $\mathcal{F}_h^{**}(v; u^{h,j-1})$ and*

$$\inf\{\mathcal{F}_h(v; u^{h,j-1}); v \in W^{1,p}(\Omega)\} = \min\{\mathcal{F}_h^{**}(v; u^{h,j-1}); v \in W^{1,p}(\Omega)\} = \mathcal{F}_h^{**}(u^{h,j}; u^{h,j-1}).$$

Moreover,

$$\|u^{h,j}\|_{W^{1,p}(\Omega)} + \|u^{h,j}\|_{L^\infty(\Omega)} \leq M_1$$

and

$$\mathcal{F}_h^{**}(u^{h,j}; u^{h,j-1}) \leq \Lambda |\nabla f_p|^p + 1,$$

where Λ only depends of μ_1 and μ_2 .

Proof By the relaxation theorem (cf. [43]), we get that

$$\begin{aligned} \mathcal{F}_h^{**}(v; u^{h,j-1}) &\leq \mathcal{F}_h(v; u^{h,j-1}), \quad \forall v \in W^{1,p}(\Omega), \\ \inf\{\mathcal{F}_h(v; u^{h,j-1}); v \in W^{1,p}(\Omega)\} &= \inf\{\mathcal{F}_h^{**}(v; u^{h,j-1}); v \in W^{1,p}(\Omega)\}. \end{aligned}$$

Let $\{u^{h,j,k}\}_{k=1}^\infty \subset W^{1,1+\delta}(\Omega)$ be a minimizing sequence of \mathcal{F}_h and \mathcal{F}_h^{**} . Then

$$\lim_{k \rightarrow \infty} \int_{\Omega} \varphi_p^{**}(\nabla u^{h,j,k}) dx = \lim_{k \rightarrow \infty} \int_{\Omega} \varphi_p(\nabla u^{h,j,k}) dx \quad (20)$$

and, for k sufficiently large,

$$\begin{aligned} \|u^{h,j,k}\|_{L^\infty(\Omega)} &\leq \|u^{h,j-1}\|_{L^\infty(\Omega)} + 2^{-j}, \\ \int_{\Omega} \varphi_p^{**}(\nabla u^{h,j,k}) dx + \frac{\lambda}{2} \int_{\Omega} (u^{h,j,k} - f^\delta)^2 dx \\ &\leq \mathcal{F}_h^{**}(u^{h,j,k}; u^{h,j-1}) \leq \mathcal{F}_h^{**}(u^{h,j-1}; u^{h,j-1}) + 2^{-j} \\ &= \int_{\Omega} \varphi_p^{**}(\nabla u^{h,j-1}) dx + \frac{\lambda}{2} \int_{\Omega} (u^{h,j-1} - f^\delta)^2 dx + 2^{-j}. \end{aligned}$$

From the growth condition, we see that $\{u^{h,j,k}\}_{k=1}^\infty$ is bounded in $W^{1,p}(\Omega) \cap L^\infty(\Omega)$, and therefore

$$\|u^{h,j,k}\|_{W^{1,p}(\Omega)} + \|u^{h,j,k}\|_{L^\infty(\Omega)} \leq M_1,$$

where M_1 is a constant independent of p , h , j , and k . Hence there exist $u^{h,j} \in W^{1,p}(\Omega) \cap L^\infty(\Omega)$ and a subsequence of $\{u^{h,j,k}\}_{k=1}^\infty$, denoted the same, such that

$$\begin{aligned} u^{h,j,k} &\rightharpoonup u^{h,j}, \quad \text{weakly in } W^{1,p}(\Omega), \\ u^{h,j,k} &\rightarrow u^{h,j}, \quad \text{strongly in } L^p(\Omega), \\ \mathcal{F}_h^{**}(u^{h,j}; u^{h,j-1}) &= \inf\{\mathcal{F}_h^{**}(v; u^{h,j-1}); v \in W^{1,p}(\Omega)\}, \end{aligned}$$

$$\|u^{h,j}\|_{W^{1,p}(\Omega)} + \|u^{h,j}\|_{L^\infty(\Omega)} \leq M_1,$$

and $\{u^{h,j,k}\}_{k=1}^\infty \subset L^\infty(\Omega)$ yields

$$\lim_{k \rightarrow \infty} \int_{\Omega} (u^{h,j,k} - u^{h,j-1})^2 dx = \int_{\Omega} (u^{h,j} - u^{h,j-1})^2 dx \quad (21)$$

and

$$\lim_{k \rightarrow \infty} \int_{\Omega} (u^{h,j,k} - f_p)^2 dx = \int_{\Omega} (u^{h,j} - f_p)^2 dx. \quad (22)$$

Thus $u^{h,j}$ is a minimum of $\mathcal{F}_h^{**}(v; u^{h,j-1})$ and

$$\begin{aligned} & \mathcal{F}_h^{**}(u^{h,j}; u^{h,j}) - \mathcal{F}_h^{**}(u^{h,j-1}; u^{h,j-1}) \\ &= \int_{\Omega} (\varphi_p^{**}(\nabla u^{h,j}) - \varphi_p^{**}(\nabla u^{h,j-1})) dx + \frac{\lambda}{2} \int_{\Omega} ((u^{h,j} - f_p)^2 - (u^{h,j-1} - f_p)^2) dx \\ &\leq \int_{\Omega} Z_p^{**}(\nabla u^{h,j}) \cdot (\nabla u^{h,j} - \nabla u^{h,j-1}) dx + \frac{\lambda}{2} \int_{\Omega} (u^{h,j} + u^{h,j-1} - 2f_p)(u^{h,j} - u^{h,j-1}) dx \\ &\leq - \int_{\Omega} \operatorname{div}(Z_p^{**}(\nabla u^{h,j}))(u^{h,j} - u^{h,j-1}) dx + \frac{\lambda}{2} \int_{\Omega} (u^{h,j} + u^{h,j-1} - 2f_p)(u^{h,j} - u^{h,j-1}) dx. \end{aligned}$$

Note that

$$\frac{u^{h,j} - u^{h,j-1}}{h} = \operatorname{div}(Z_p^{**}(\nabla u^{h,j})) - \lambda(u^{h,j} - f_p),$$

and then

$$\mathcal{F}_h^{**}(u^{h,j}; u^{h,j}) - \mathcal{F}_h^{**}(u^{h,j-1}; u^{h,j-1}) \leq - \int_{\Omega} \frac{(u^{h,j} - u^{h,j-1})^2}{h} - \frac{\lambda}{2} \int_{\Omega} (u^{h,j} - u^{h,j-1})^2 dx \leq 0.$$

Thus

$$\mathcal{F}_h^{**}(u^{h,j}; u^{h,j-1}) \leq \mathcal{F}_h^{**}(u^{h,j-1}; u^{h,j-1}) \leq \mathcal{F}_h^{**}(u^{h,j-2}; u^{h,j-2}) \leq \dots \leq \mathcal{F}_h^{**}(u^{h,0}; u^{h,0}).$$

Hence

$$\mathcal{F}_h^{**}(u^{h,j}; u^{h,j-1}) \leq \int_{\Omega} \varphi_p^{**}(\nabla f_p) \leq \Lambda |\nabla f_p|^p + 1. \quad \square$$

Let $\nu^{h,j} = (\nu_x^{h,j})_{x \in \Omega}$ be the Young measure generated by $\{\nabla u^{h,j,k}\}_{k=1}^\infty$ in the proof of Lemma 5. By Lemma 3, $\nu^{h,j}$ is a $W^{1,p}$ -gradient Young measure. Then

$$\begin{aligned} \int_{\Omega} \langle \nu^{h,j}, \varphi_p^{**} \rangle dx &= \lim_{k \rightarrow \infty} \int_{\Omega} \varphi_p^{**}(\nabla u^{h,j,k}) dx \\ &= \lim_{k \rightarrow \infty} \int_{\Omega} \varphi_p(\nabla u^{h,j,k}) dx = \int_{\Omega} \langle \nu^{h,j}, \varphi_p \rangle dx. \end{aligned}$$

Noticing that $\varphi_p^{**} \leq \varphi_p$, we see that

$$\operatorname{supp} \nu_x^{h,j} \subset \{X \in R^N : \varphi_p(X) = \varphi_p^{**}(X)\}, \quad \text{a.e. } x \in \Omega. \quad (23)$$

Thus,

$$\langle v^{hj}, Z_p \rangle = \langle v^{hj}, Z_p^{**} \rangle \quad \text{a.e. in } \Omega,$$

$$\nabla u^{hj} = \langle v^{hj}, \text{id} \rangle \quad \text{a.e. in } \Omega.$$

Let χ^{hj} be the indicator function of $[hj, h(j+1))$ and

$$\beta^{hj}(t) = \begin{cases} \frac{t}{h} - j, & t \in [hj, h(j+1)), \\ 0, & \text{otherwise.} \end{cases}$$

Define

$$u^h(x, t) = \sum_{0 \leq j \leq T/h} \chi^{hj}(t) \{ u^{hj}(x) + \beta^{hj}(t) (u^{h(j+1)}(x) - u^{hj}(x)) \}, \quad (24)$$

$$w^h(x, t) = \sum_{0 \leq j \leq T/h} \chi^{hj}(t) u^{hj}(x), \quad (25)$$

$$v^h = (v_{x,t}^h)_{(x,t) \in Q_T} = \sum_{0 \leq j \leq T/h} \chi^{hj}(t) v_x^{hj}. \quad (26)$$

Then

$$u^h(x, t) \in L^\infty((0, T); W^{1,p}(\Omega)) \cap L^\infty(Q_T),$$

$$v^h \in L^1(Q_T; (\mathcal{E}_0^p(\mathbb{R}^N))') \cap L^\infty((0, T); (\mathcal{E}_0^p(\mathbb{R}^N))'),$$

$$u^h(x, 0) = f_p(x), w^h(x, 0) = f_p(x), u^h(x, T) = w^h(x, T) \quad \text{a.e. } x \in \Omega,$$

$$\sup_{0 \leq t \leq T} \|u^h\|_{W^{1,p}(\Omega)} + \|u^h\|_{L^\infty(Q_T)} \leq M_1,$$

$$\sup_{0 \leq t \leq T} \|w^h\|_{W^{1,p}(\Omega)} + \|w^h\|_{L^\infty(Q_T)} \leq M_1.$$

Based on the facts above, we can obtain the following lemma.

Lemma 6 *The functions u^h , w^h , and the Young measure v^h defined above satisfy*

$$\iint_{Q_T} \langle v^h, Z_p \rangle \cdot \nabla \zeta \, dx \, dt + \iint_{Q_T} \frac{\partial u^h}{\partial t} \zeta \, dx \, dt + \lambda \iint_{Q_T} (w^h - f_p) \zeta \, dx \, dt = 0, \quad (27)$$

for $\zeta \in C^\infty(\overline{Q_T})$, with $\zeta(x, 0) = \zeta(x, T) = 0$. Moreover,

$$\sup_{0 \leq t \leq T} \|u^h\|_{W^{1,p}(\Omega)} + \|u^h\|_{L^\infty(Q_T)} + \left\| \frac{\partial u^h}{\partial t} \right\|_{L^2(Q_T)} + \|\langle v_{x,t}^h, Z_p \rangle\|_{L^{p/(p-1)}(Q_T)} \leq M,$$

$$\sup_{0 \leq t \leq T} \|w^h\|_{W^{1,p}(\Omega)} + \|w^h\|_{L^\infty(Q_T)} \leq M,$$

where M are independent of p and h .

Proof Let $\xi \in C_0^\infty(\Omega)$, $-1 < \epsilon < 1$. Then there exists $C > 0$ such that

$$\varphi_p^{**}(X + \epsilon \nabla \xi) \leq \Lambda(1 + |X|^p).$$

By Lemma 2, we can see that

$$\begin{aligned} \mathcal{F}_h^{**}(u^{hj}; u^{hj-1}) &\leq \mathcal{F}_h^{**}(u^{hj} + \epsilon \xi; u^{hj-1}) \\ &= \int_{\Omega} \varphi_p^{**}(\nabla u^{hj} + \epsilon \nabla \xi) dx + \frac{1}{2h} \int_{\Omega} (u^{hj} + \epsilon \xi - u^{hj-1})^2 dx \\ &\quad + \frac{\lambda}{2} \int_{\Omega} (u^{hj} + \epsilon \xi - f^\delta)^2 dx \\ &\leq \lim_{k \rightarrow \infty} \int_{\Omega} \varphi_p^{**}(\nabla u^{hj,k} + \epsilon \nabla \xi) dx + \frac{1}{2h} \int_{\Omega} (u^{hj} + \epsilon \xi - u^{hj-1})^2 dx \\ &\quad + \frac{\lambda}{2} \int_{\Omega} (u^{hj} + \epsilon \xi - f^\delta)^2 dx \\ &= \int_{\Omega} \int_{\mathbb{R}^N} \varphi_p^{**}(X + \epsilon \nabla \xi) dv_x^{hj}(X) dx + \frac{1}{2h} \int_{\Omega} (u^{hj} + \epsilon \xi - u^{hj-1})^2 dx \\ &\quad + \frac{\lambda}{2} \int_{\Omega} (u^{hj} + \epsilon \xi - f^\delta)^2 dx, \end{aligned}$$

which implies the equilibrium equation

$$\int_{\Omega} \langle v^{hj}, Z_p^{**} \rangle \cdot \nabla \xi dx + \frac{1}{h} \int_{\Omega} (u^{hj} - u^{hj-1}) \xi dx + \lambda \int_{\Omega} (u^{hj} - f_p) \xi dx = 0,$$

$\forall \xi \in C_0^\infty(\Omega)$. At the minimizer u^{hj} , the Gâteaux derivative of \mathcal{F}_h^{**} is zero, and we obtain

$$\int_{\Omega} Z_p^{**}(\nabla u^{hj}(x)) \cdot \nabla \xi dx + \frac{1}{h} \int_{\Omega} (u^{hj} - u^{hj-1}) \xi dx + \lambda \int_{\Omega} (u^{hj} - f_p) \xi dx = 0,$$

$\forall \xi \in C_0^\infty(\Omega)$. Thus

$$Z_p(\nabla u^{hj}(x)) = \langle v^{hj}, Z_p \rangle = \langle v^{hj}, Z_p^{**} \rangle = Z_p^{**}(\nabla u^{hj}(x)), \quad \text{a.e. } x \in \Omega, \quad (28)$$

$$\int_{\Omega} \langle v^{hj}, Z_p \rangle \cdot \nabla \xi dx + \frac{1}{h} \int_{\Omega} (u^{hj} - u^{hj-1}) \xi dx + \lambda \int_{\Omega} (u^{hj} - f_p) \xi dx = 0, \quad (29)$$

and, by Lemma 5, we get the estimate

$$\begin{aligned} \|\langle v^{hj}, Z_p \rangle\|_{L^{p/(p-1)}(\Omega)} &= \|Z_p^{**}(\nabla u^{hj})\|_{L^{p/(p-1)}(\Omega)} \leq \Lambda \left(\int_{\Omega} |\nabla u^{hj}|^p dx \right)^{(p-1)/p} \\ &\leq \Lambda \|u^{hj}\|_{W^{1,p}(\Omega)}^\delta \leq \Lambda(M_1 + 1). \end{aligned}$$

From (24)–(26) and (29),

$$\int_{\Omega} \langle v^h, Z_p \rangle \cdot \nabla \xi dx + \frac{\partial u^h}{\partial t} \xi dx + \lambda \int_{\Omega} (u^h - f_p) \xi dx = 0, \quad \forall \xi \in C^\infty(\overline{\Omega}), \text{ a.e. } t \in [0, T],$$

which implies that

$$\iint_{Q_T} \langle v^h, Z_p \rangle \cdot \nabla \zeta \, dx \, dt + \iint_{Q_T} \frac{\partial u^h}{\partial t} \zeta \, dx \, dt + \lambda \iint_{Q_T} (w^h - f_p) \zeta \, dx \, dt = 0, \quad (30)$$

for $\zeta \in C^\infty(\overline{Q_T})$, with $\zeta(x, 0) = \zeta(x, T) = 0$. From the direct calculation, we see that

$$\|\langle v_{x,t}^h, Z_p \rangle\|_{L^{p/(p-1)}(Q_T)} \leq T \sup_{0 \leq j \leq T/h} \|\langle v_x^{h,j}, Z_p \rangle\|_{L^{p/(p-1)}(\Omega)} \leq T(M_1 + 1).$$

Because

$$\frac{\partial u^h(x, t)}{\partial t} = \frac{1}{h} \sum_{0 \leq j \leq T/h} \chi^{h,j}(t) (u^{h,j+1}(x) - u^{h,j}(x)) \in L^\infty((0, T); W^{1,p}(\Omega)) \cap L^\infty(Q_T),$$

$\partial_t u^h$ can be chosen as the test function in (27), and therefore

$$\begin{aligned} & \iint_{Q_T} \langle v^h, Z_p \rangle \cdot \nabla \frac{\partial u^h}{\partial t} \, dx \, dt + \iint_{Q_T} \left(\frac{\partial u^h}{\partial t} \right)^2 \, dx \, dt \\ & + \lambda \iint_{Q_T} |w^h - f_p|^{p-2} (w^h - f_p) \frac{\partial u^h}{\partial t} \, dx \, dt = 0. \end{aligned}$$

Then

$$\begin{aligned} & \iint_{Q_T} \left(\frac{\partial u^h}{\partial t} \right)^2 \, dx + \lambda \int_{\Omega} |u^h(x, T) - f_p|^2 \, dx \\ & = - \iint_{Q_T} \langle v^h, Z_p \rangle \cdot \nabla \frac{\partial u^h}{\partial t} \, dx \, dt + \lambda \int_{\Omega} (u^h(x, T) - f_p) f_p \, dx \\ & = - \sum_{0 \leq j \leq \frac{T}{h}} \iint_{Q_T} \chi^{h,j}(t) \langle v^h, Z_p \rangle \cdot \frac{\partial(\nabla u^h)}{\partial t} \, dx \, dt \\ & \quad + \lambda \int_{\Omega} (u^h(x, T) - f_p) f_p \, dx \\ & = - \sum_{0 \leq j \leq \frac{T}{h}} \iint_{Q_T} \chi^{h,j}(t) \frac{\partial(\langle v^h, Z_p \rangle \cdot \nabla u^h)}{\partial t} \, dx \, dt \\ & \quad + \lambda \int_{\Omega} (u^h(x, T) - f_p) f_p \, dx \\ & = - \int_{\Omega} \langle v_{x,T}^h, Z_p \rangle \cdot \nabla u^h(x, T) \, dx + \int_{\Omega} \langle v_{x,0}^h, Z_p \rangle \cdot \nabla u^h(x, 0) \, dx \\ & \quad + \lambda \int_{\Omega} (u^h(x, T) - f_p) f_p \, dx \\ & \leq \|\langle v_{x,T}^h, Z_p \rangle\|_{L^{p/(p-1)}(\Omega)} \|\nabla u^h(x, T)\|_{L^p(\Omega)} + \|\langle v_{x,0}^h, Z_p \rangle\|_{L^{p/(p-1)}(\Omega)} \|\nabla f_p\|_{L^p(\Omega)} \\ & \quad + \frac{\lambda}{2} \int_{\Omega} |u^h(x, T) - f_p|^2 \, dx + \frac{\lambda}{2} \int_{\Omega} f_p^2 \, dx, \end{aligned}$$

and therefore

$$\begin{aligned} & \iint_{Q_T} \left(\frac{\partial u^h}{\partial t} \right)^2 dx + \frac{\lambda}{2} \int_{\Omega} |u^h(x, T) - f_p|^2 dx \\ & \leq \| \langle v_{x,T}^h, Z_p \rangle \|_{L^{p/(p-1)}(\Omega)} \| \nabla u^h(x, T) \|_{L^p(\Omega)} \\ & \quad + \| \langle v_{x,0}^h, Z_p \rangle \|_{L^{p/(p-1)}(\Omega)} \| \nabla f_p \|_{L^p(\Omega)} + \frac{\lambda}{2} \int_{\Omega} f_p^2 dx, \\ & \leq 2\Lambda M_1(M_1 + 1) + \frac{\lambda}{2} \int_{\Omega} f_p^2 dx, \end{aligned}$$

which implies

$$\left\| \frac{\partial u^h}{\partial t} \right\|_{L^2(Q_T)} \leq (2\Lambda + \lambda^{1/2} + 1)(M_1 + 1).$$

□

Let

$$v^{h,k}(x, t) = \sum_{0 \leq j \leq \frac{T}{h}} \chi^{hj}(t) u^{h,j,k}(x),$$

for $k \geq 1$. By Lemma 5, $\{v^{h,k}\}_{k=1}^{\infty} \subset L^{\infty}((0, T); W^{1,p}(\Omega)) \cap L^{\infty}(Q_T)$, $\{v^{h,k}\}_{k=1}^{\infty}$ is the $W^{1,p}(Q_T)$ -gradient generating sequence of v^h and

$$\sup_{0 \leq t \leq T} \|v^{h,k}\|_{W^{1,p}(\Omega)} + \|v^{h,k}\|_{L^{\infty}(Q_T)} \leq M_1. \quad (31)$$

When $h \rightarrow 0$ and $k \rightarrow \infty$, we have the following lemma.

Lemma 7 *There exist $u, w \in L^{\infty}((0, T); W^{1,p}(\Omega)) \cap L^{\infty}(Q_T)$, $v \in L^{\infty}((0, T); (\mathcal{E}^p(\mathbb{R}^N))')$, $\frac{\partial u}{\partial t} \in L^2(Q_T)$, and subsequences of $\{u^h\}_{0 < h < 1}$, $\{w^h\}_{0 < h < 1}$, $\{v^{h,k}\}_{0 < h < 1, k=1}^{\infty}$, and $\{v^h\}_{0 < h < 1}$, denoted by $\{u^{h_m}\}_{m=1}^{\infty}$, $\{w^{h_m}\}_{m=1}^{\infty}$, $\{v^m = v^{h_m, k_m}\}_{m=1}^{\infty}$, and $\{v^{h_m}\}_{m=1}^{\infty}$, respectively, such that*

$$u^{h_m} \rightarrow u, \quad \text{weakly in } L^{\infty}((0, T); W^{1,p}(\Omega)) \cap L^{\infty}(Q_T) \text{ and strongly in } L^p(Q_T),$$

$$\frac{\partial u^{h_m}}{\partial t} \rightarrow \frac{\partial u}{\partial t}, \quad \text{weakly in } L^2(Q_T),$$

$$w^{h_m} \rightarrow u, \quad \text{weakly in } L^{\infty}((0, T); W^{1,p}(\Omega)) \cap L^{\infty}(Q_T) \text{ and strongly in } L^p(Q_T),$$

$$v^m \rightarrow v, \quad \text{weakly in } L^{\infty}((0, T); W^{1,p}(\Omega)) \cap L^{\infty}(Q_T) \text{ and strongly in } L^p(Q_T),$$

$$v^{h_m} \rightarrow v \quad \text{weakly in } L^1(Q_T; (\mathcal{E}^p(\mathbb{R}^N))'), \text{ weakly}^* \text{ in } L^{\infty}(Q_T; \mathcal{M}(\mathbb{R}^N)),$$

$$\text{and in } L^{\infty}(Q_T; (\mathcal{E}^p(\mathbb{R}^N))') \text{ in the biting sense}$$

and

$$u(x, t) = v(x, t), \quad \text{a.e. } (x, t) \in Q_T,$$

$$u(x, 0) = f_p(x), \quad x \in \Omega.$$

$$\nabla u = \nabla w = \nabla v = \langle v, \text{id} \rangle.$$

Proof From Lemma 6, we see that $\{u^h\}_{0 < h < 1}$ and $\{w^h\}_{0 < h < 1}$ are bounded in $L^\infty((0, T); W^{1,p}(\Omega)) \cap L^\infty(Q_T)$, $\{\partial_t u^h\}_{0 < h < 1}$ is bounded in $L^2(Q_T)$, $\{\langle v^h, Z_p \rangle\}_{0 < h < 1}$ is bounded in $L^{p/(p-1)}(Q_T)$, $\{v^h\}_{0 < h < 1}$ is bounded in $L^\infty((0, T); (\mathcal{E}_0^p(\mathbb{R}^N))')$, and the bounds are independent of δ and h . By weak compactness, we can obtain the convergence of the subsequences $\{u^{h_m}\}_{m=1}^\infty$, $\{w^{h_m}\}_{m=1}^\infty$, and $\{v^{h_m}\}_{m=1}^\infty$. From (26), we can see that (12) holds, namely

$$u(x, 0) = f_p(x), \quad x \in \Omega.$$

From (31), by Lemma 4 and the weak sequential precompactness of $L^\infty((0, T); W^{1,p}(\Omega)) \cap L^\infty(Q_T)$, there exist $w \in L^\infty((0, T); W^{1,p}(\Omega)) \cap L^\infty(Q_T)$ and a diagonal subsequence of $\{v^{h_m,k}\}_{m=1}^\infty$, denoted by $\{v^m = w^{h_m,k_m}\}_{m=1}^\infty$, such that $\{v^m\}_{m=1}^\infty$ converges to w weakly in $L^\infty((0, T); W^{1,p}(\Omega)) \cap L^\infty(Q_T)$ and strongly in $L^p(Q_T)$,

$$\sup_{1 \leq j \leq T/h_m} \|u^{h_m,j,k_m} - u^{h_m,j}\|_{L^2(\Omega)} \leq h_m,$$

$$\{v^m\}_{m=1}^\infty \subset L^\infty((0, T); W^{1,p}(\Omega)) \cap L^\infty(Q_T),$$

$$\sup_{0 \leq t \leq T} \|w^m\|_{W^{1,p}(\Omega)} + \|w^m\|_{L^\infty(Q_T)} \leq M_1,$$

and $\{\nabla v^m\}_{m=1}^\infty$ is the $W^{1,1}(Q_T)$ -bitting generating sequence of v , and then

$$\nabla v = \langle v, \text{id} \rangle.$$

Now let us prove that

$$u(x, t) = v(x, t), \quad \text{a.e. } (x, t) \in Q_T.$$

Since

$$\begin{aligned} u^h(x, t) &= \sum_{0 \leq j \leq T/h} \chi^{h,j}(t) \{u^{h,j}(x) + \beta^{h,j}(t)(u^{h,j+1}(x) - u^{h,j}(x))\}, \\ v^{h,k}(x, t) &= \sum_{0 \leq j \leq T/h} \chi^{h,j}(t) u^{h,j,k}(x), \end{aligned}$$

by Lemma 5, we have

$$\begin{aligned} & \iint_{Q_T} |u^h(x, t) - v^{h,k}(x, t)|^2 dx dt \\ & + 2 \sum_{0 \leq j \leq \frac{T}{h}} \left(\iint_{Q_T} |\chi^{h,j}(t) \lambda^{h,j}(t) (u^{h,j+1} - u^{h,j}(x))|^2 dx dt \right. \\ & \quad \left. + \iint_{Q_T} |\chi^{h,j}(t) (u^{h,j}(x) - u^{h,j,k}(x))|^2 dx dt \right) \\ & \leq 2h \sum_{0 \leq j \leq \frac{T}{h}} \left(\int_{\Omega} |u^{h,j+1}(x) - u^{h,j}(x)|^2 dx + \int_{\Omega} |u^{h,j}(x) - u^{h,j,k}(x)|^2 dx \right) \end{aligned}$$

$$\begin{aligned}
&\leq h^2 \sum_{0 \leq j \leq \frac{T}{h}} (\mathcal{F}_h(u^{h,j+1}; u^{h,j}) + \mathcal{F}_h(u^{h,j,k}; u^{h,j})) \\
&\leq 8hT \sum_{0 \leq j \leq \frac{T}{h}} \mathcal{F}_h(u^{h,j+1}; u^{h,j}) + \frac{4hT}{k} \\
&\leq 8hT \Lambda \int_{\Omega} |\nabla f_p| dx + 4hT \\
&\rightarrow 0 \quad (h \rightarrow 0).
\end{aligned}$$

Since $\{u^{h_m}\}_{m=1}^{\infty}$ converges to u in $L^p(Q_T)$ and $\{v^{h_m,k_m}\}_{m=1}^{\infty}$ converges to w in $L^p(Q_T)$, we see that

$$u(x, t) = v(x, t), \quad \text{a.e. } (x, t) \in Q_T.$$

Hence

$$\nabla u = \nabla v = \langle v, \text{id} \rangle,$$

which implies (7). Similar to the above arguments, we also obtain that $\{v^{h_m}\}_{m=1}^{\infty}$ converges to u in $L^p(Q_T)$. \square

Proof of Theorem 3 From Lemmas 6 and 7, we can obtain

$$\iint_{Q_T} \langle v, Z_p \rangle \cdot \nabla \zeta dx dt + \iint_{Q_T} \frac{\partial u}{\partial t} \zeta dx dt + \lambda \iint_{Q_T} (u - f_p) \zeta dx dt = 0, \quad (32)$$

for $\zeta \in C^\infty(\overline{Q_T})$, with $\zeta(x, 0) = \zeta(x, T) = 0$. Therefore, if we prove (8), we will obtain the weak solution of the problem (3)–(5). Let $\{u^{h_{j,k}}\}_{k=1}^{\infty} \subset W^{1,p}(\Omega)$ be a minimizing sequence of \mathcal{F}_h in the proof of Lemma 5. For all $\xi \in W^{1,p}(\Omega)$, we see that

$$\begin{aligned}
&\lim_{k \rightarrow \infty} \int_{\Omega} Z_p^{**}(\nabla u^{h_{j,k}}) \cdot \nabla \xi dx + \frac{1}{h} \int_{\Omega} (u^{h_{j,k}} - u^{h_{j-1}}) \xi dx + \lambda \int_{\Omega} (u^{h_{j,k}} - f_p) \xi dx \\
&= \int_{\Omega} \langle v^{h_j}, Z_p^{**} \rangle \cdot \nabla \xi dx + \frac{1}{h} \int_{\Omega} (u^{h_j} - u^{h_{j-1}}) \xi dx + \lambda \int_{\Omega} (u^{h_j} - f_p) \xi dx = 0
\end{aligned}$$

and

$$\begin{aligned}
&\lim_{k \rightarrow \infty} \left| \int_{\Omega} (Z_p^{**}(\nabla u^{h_{j,k}}) - \langle v^{h_j}, Z_p^{**} \rangle) \cdot (\nabla u^{h_{j,k}} - \nabla u^{h_j}) dx \right| \\
&= \lim_{k \rightarrow \infty} \left| \int_{\Omega} \frac{u^{h_{j,k}} - u^{h_j}}{h} (u^{h_{j,k}} - u^{h_j}) dx \right. \\
&\quad \left. + \lambda \int_{\Omega} ((u^{h_{j,k}} - f_p) - (u^{h_j} - f_p))(u^{h_{j,k}} - u^{h_j}) dx \right| \\
&\leq 1/h \lim_{k \rightarrow \infty} \|u^{h_{j,k}} - u^{h_j}\|_{L^2(\Omega)}^2 \\
&\quad + \lambda \lim_{k \rightarrow \infty} \left(\int_{\Omega} |u^{h_{j,k}} - f_p|^2 - |u^{h_j} - f_p|^2 dx \right)^{1/2} \left(\int_{\Omega} |u^{h_{j,k}} - u^{h_j}|^2 dx \right)^{1/2} \\
&= 0.
\end{aligned}$$

Since $Z_p^{**}(\nabla u^{h,j,k}) \cdot \nabla u^{h,j,k}$ converges weakly to $\langle v^{h,j}, Z_p^{**} \cdot \text{id} \rangle$ in $L^1(\Omega)$, $Z_p^{**}(\nabla u^{h,j,k})$ converges weakly to $\langle v^{h,j}, Z_p^{**} \rangle$ in $L^{p/(p-1)}(\Omega)$, $\nabla u^{h,j,k}$ converges weakly to $\langle v^{h,j}, \text{id} \rangle$ in $L^p(\Omega)$ as k tends to infinity, we get that

$$\langle v_x^{h,j}, Z_p^{**} \cdot \text{id} \rangle = \langle v_x^{h,j}, Z_p^{**} \rangle \cdot \langle v_x^{h,j}, \text{id} \rangle \quad \text{a.e. } x \in \Omega.$$

Thus (23) implies that

$$\langle v_x^{h,j}, Z_p \cdot \text{id} \rangle = \langle v_x^{h,j}, Z_p \rangle \cdot \langle v_x^{h,j}, \text{id} \rangle \quad \text{a.e. } x \in \Omega.$$

By the definition of v^h in (26), we see that

$$\langle v_{x,t}^h, Z_p \cdot \text{id} \rangle = \langle v_{x,t}^h, Z_p \rangle \cdot \langle v_{x,t}^h, \text{id} \rangle \quad \text{a.e. } x \in Q_T.$$

From Lemma 7, we can obtain that $\langle v_{x,t}^{h_m}, Z_p \cdot \text{id} \rangle$ converges weakly to $\langle v_{x,t}, Z_p \cdot \text{id} \rangle$ in the biting sense, $\langle v_{x,t}^{h_m}, Z_p \rangle$ converges weakly to $\langle v_{x,t}, Z_p \rangle$ in $L^{p/(p-1)}(Q_T)$, $\langle v_{x,t}^{h_m}, \text{id} \rangle$ converges weakly to $\langle v_{x,t}, \text{id} \rangle$ in $L^p(Q_T)$ as m tends to infinity. Thus for all $\eta \in C^\infty(\overline{Q_T})$, with $\eta(x, 0) = \eta(x, T) = 0$, (17) and (19) imply that

$$\begin{aligned} & \left| \iint_{Q_T} (\langle v^{h_m}, Z_p \rangle - \langle v, Z_p \rangle) \cdot \langle v^{h_m}, \text{id} \rangle \eta \, dx \, dt \right| \\ & \leq \left| \iint_{Q_T} (\langle v^{h_m}, Z_p \rangle - \langle v, Z_p \rangle) \cdot \nabla (w^{h_m} \eta) \, dx \, dt \right| \\ & \quad + \left| \iint_{Q_T} w^{h_m} (\langle v^{h_m}, Z_p \rangle - \langle v, Z_p \rangle) \cdot \nabla \eta \, dx \, dt \right| \\ & \leq \left| \iint_{Q_T} \left(\frac{\partial u^{h_m}}{\partial t} - \frac{\partial u}{\partial t} \right) w^{h_m} \eta \, dx \, dt \right| + \lambda \left| \iint_{Q_T} ((w^{h_m} - f_p) - (u - f_p)) w^{h_m} \eta \, dx \, dt \right| \\ & \quad + \left| \iint_{Q_T} w^{h_m} (\langle v^{h_m}, Z_p \rangle - \langle v, Z_p \rangle) \cdot \nabla \eta \, dx \, dt \right| \\ & \leq \left| \iint_{Q_T} \left(\frac{\partial u^{h_m}}{\partial t} - \frac{\partial u}{\partial t} \right) u \eta \, dx \, dt \right| + \left| \iint_{Q_T} \left(\frac{\partial u^{h_m}}{\partial t} - \frac{\partial u}{\partial t} \right) (w^{h_m} - u) \eta \, dx \, dt \right| \\ & \quad + \lambda \left| \iint_{Q_T} ((w^{h_m} - f_p) - (u - f_p)) w^{h_m} \eta \, dx \, dt \right| \\ & \quad + \left| \iint_{Q_T} u (\langle v^{h_m}, Z_p \rangle - \langle v, Z_p \rangle) \cdot \nabla \eta \, dx \, dt \right| \\ & \quad + \left| \iint_{Q_T} (w^{h_m} - u) (\langle v^{h_m}, Z_p \rangle - \langle v, Z_p \rangle) \cdot \nabla \eta \, dx \, dt \right| \\ & \rightarrow 0 \quad (m \rightarrow \infty). \end{aligned}$$

Hence,

$$\begin{aligned} & \left| \iint_{Q_T} (\langle v^{h_m}, Z_p \rangle \cdot \langle v^{h_m}, \text{id} \rangle - \langle v, Z_p \rangle \cdot \langle v, \text{id} \rangle) \eta \, dx \, dt \right| \\ & \leq \left| \iint_{Q_T} (\langle v^{h_m}, Z_p \rangle - \langle v, Z_p \rangle) \cdot \langle v^{h_m}, \text{id} \rangle \eta \, dx \, dt \right| \end{aligned}$$

$$+ \left| \iint_{Q_T} \langle v, Z_p \rangle \cdot (\langle v^{h_m}, \text{id} \rangle - \langle v, \text{id} \rangle) \eta \, dx \, dt \right| \\ \rightarrow 0 \quad (m \rightarrow \infty).$$

So $\langle v_{x,t}^{h_m}, Z_p \rangle \cdot \langle v_{x,t}^{h_m}, \text{id} \rangle$ weakly converges to $\langle v_{x,t}, Z_p \rangle \cdot \langle v_{x,t}, \text{id} \rangle$ in $L^1(Q_T)$. Since $\langle v_{x,t}^{h_m}, Z_p \cdot \text{id} \rangle$ converges to $\langle v_{x,t}, Z_p \cdot \text{id} \rangle$ in the biting sense, we obtain that

$$\langle v_{x,t}, Z_p \cdot \text{id} \rangle = \langle v_{x,t}, Z_p \rangle \cdot \langle v_{x,t}, \text{id} \rangle \quad \text{a.e. } (x, t) \in Q_T \setminus E_j, \forall j \geq 1.$$

Since $\langle v_{x,t}, Z_p \cdot \text{id} \rangle \in L^1(Q_T)$,

$$\langle v_{x,t}, Z_p \cdot \text{id} \rangle = \langle v_{x,t}, Z_p \rangle \cdot \langle v_{x,t}, \text{id} \rangle \quad \text{a.e. } (x, t) \in Q_T,$$

which implies (8). Hence, u is the desired Young measure solution of problem (3)–(5). The proof of Theorem 3 is complete. \square

Remark 2 Let u be the Young measure solution of problem (12)–(14) obtained in the proof of Theorem 3. Then from the proof we see that there exists a constant M depending only on $\|f_p\|_{W^{1,p}(\Omega)}$, $\|f_p\|_{L^\infty(\Omega)}$, Λ , and $\text{meas}(\Omega)$, but independent of p and T , such that

$$\|u\|_{L^\infty((0,T);W^{1,p}(\Omega))} + \|u\|_{L^\infty(Q_T)} \leq M, \\ \left\| \frac{\partial u}{\partial t} \right\|_{L^2(Q_T)} \leq M,$$

namely, $u \in L^\infty(\mathbb{R}^+; W^{1,p}(\Omega)) \cap L^\infty(Q_\infty)$, $\partial u / \partial t \in L^2(Q_\infty)$, where $Q_\infty = \Omega \times \mathbb{R}^+$.

3.4 Existence of solution to problem (3)–(5)

In this subsection, we consider the limit case of problem (3)–(5), namely, $p \rightarrow 1$.

Proof of Theorem 2 Let u_p be the Young measure solution of problem (12)–(14) with the initial data f_p with respect to the $W^{1,p}(Q_T)$ -gradient Young measures ν^p generated by the sequence $\{\nabla w^{p,k}\}_{k=1}^\infty$, which we obtained in the proof of Theorem 3. We see that

$$u_p \in L^\infty((0, T); W^{1,p}(\Omega)) \cap L^\infty(Q_T), \quad \frac{\partial u_p}{\partial t} \in L^2(Q_T), \\ \nu_p \in L^\infty((0, T); (\mathcal{E}_0^p(\mathbb{R}^N))'), \quad \langle \nu_p, Z_p \rangle \in L^{p/(p-1)}(Q_T), \\ w^{p,k} \in L^\infty((0, T); W^{1,p}(\Omega)) \cap L^\infty(Q_T),$$

and there exists a constant M_0 depending only on $\|f_p\|_{W^{1,p}(\Omega)}$, $\|f_p\|_{L^\infty(\Omega)}$, Λ , and $\text{meas}(\Omega)$, but independent of p , such that

$$\|u_p\|_{L^\infty((0,T);W^{1,p}(\Omega))} + \|u_p\|_{L^\infty(Q_T)} \leq M_0, \quad \left\| \frac{\partial u_p}{\partial t} \right\|_{L^2(Q_T)} \leq M_0, \\ \|\nu_p\|_{L^\infty((0,T);(\mathcal{E}^p(\mathbb{R}^N))')} \leq M_0, \quad \|\langle \nu_p, Z_p \rangle\|_{L^{p/(p-1)}(Q_T)} \leq M_0, \\ \sup_{0 \leq t \leq T} \|w^{p,k}\|_{W^{1,p}(\Omega)} + \|w^{p,k}\|_{L^\infty(Q_T)} \leq M_0.$$

So there exist $u \in L^\infty((0, T); \text{BV}(\Omega)) \cap L^\infty(Q_T)$ with $\partial u / \partial t \in L^2(Q_T)$ and a subsequence of $\{u_p\}_{1 \leq p < 2}$, denoted by $\{u_{p_m}\}_{m=1}^\infty$, such that

$$\begin{aligned} u_{p_m} &\rightarrow u, \quad \text{strongly in } L^1(Q_T), \\ \frac{\partial u_{p_m}}{\partial t} &\rightarrow \frac{\partial u}{\partial t}, \quad \text{weakly in } L^2(Q_T). \end{aligned}$$

By Lemma 4, there exist a $W^{1,1}(Q_T)$ -gradient Young measure $\nu \in L^\infty((0, T); (\mathcal{E}_0^1(\mathbb{R}^N))')$ and a subsequence of $\{\nu^{p_m}\}_{m=1}^\infty$, denoted the same, such that

$$\begin{aligned} \nu^{p_m} &\rightarrow \nu, \quad \text{weakly}^* \text{ in } L^\infty(Q_T; \mathcal{M}(\mathbb{R}^N)), \\ \nu^{p_m} &\rightarrow \nu, \quad \text{in } \mathcal{E}_0^1(\mathbb{R}^d) \text{ in the biting sense,} \end{aligned}$$

which implies that there is a decreasing sequence of subsets $E_{j+1} \subset E_j$ of Q_T with $\lim_{j \rightarrow \infty} \text{meas}(E_j) = 0$ such that $\langle \nu^{p_m}, \psi \rangle$ converges weakly to $\langle \nu, \psi \rangle$ in $L^1(Q_T \setminus E_j)$ for all $\psi \in \mathcal{E}_0^1(\mathbb{R}^d)$ and all $j \geq 1$. By (18) we get that

$$\text{supp } \nu_{x,t}^p \subset \{x \in \mathbb{R}^N : \varphi_p(X) = \varphi_p^{**}(X)\} \subset \{X \in \mathbb{R}^N : \varphi(X) = \varphi^{**}(X)\}$$

which implies (9). By Lemma 4, there exist $w \in L^\infty((0, T); \text{BV}(\Omega))$ and a subsequence of $\{w^{p_m,k}\}_{m,k=1}^\infty$, denoted by $\{w^k\}_{k=1}^\infty$, such that $\{w^k\}_{k=1}^\infty$ converges to w in $L^1(Q_T)$ and $\{\nabla w^k\}_{k=1}^\infty$ is the $W^{1,1}(Q_T)$ -biting generating sequence of ν , namely, there is a decreasing sequence of subsets $G_{j+1} \subset G_j$ of Q_T with $\lim_{j \rightarrow \infty} \text{meas}(G_j) = 0$ such that $\langle \nu^{p_m}, \psi \rangle$ converges weakly to $\langle \nu, \psi \rangle$ in $L^1(Q_T \setminus G_j)$ for all $\psi \in \mathcal{E}_0^1(\mathbb{R}^d)$ and all $j \geq 1$.

To prove (6), we first prove that $\{\langle \nu^{p_m}, Z_{p_m} \rangle\}_{m=1}^\infty$ converges weakly to $\langle \nu, Z \rangle$ in $L^1(Q_T)$. For $i \geq 1$, define

$$\theta^i(A) = \begin{cases} 1, & |A| \leq i, \\ i + 1 - |A|, & i \leq |A| \leq i + 1, \\ 0, & |A| \geq i + 1. \end{cases}$$

Let $\eta \in L^\infty(Q_T; \mathbb{R}^N)$. Then

$$\begin{aligned} &\left| \iint_{Q_T} (\langle \nu^{p_m}, Z_p \rangle - \langle \nu^{p_m}, Z \rangle) \cdot \eta \, dx \, dt \right| \\ &\leq \|\eta\|_{L^\infty(Q_T)} \iint_{Q_T} |\langle \nu^p, (1 - \theta^i)(Z_p - Z) \rangle| \, dx \, dt + \left| \iint_{Q_T} \langle \nu^p, \theta^i(Z_p - Z) \rangle \cdot \eta \, dx \, dt \right| \\ &= I + II, \\ I &\leq \|\eta\|_{L^\infty(Q_T)} \lim_{k \rightarrow \infty} \iint_{\{(x,t): |\nabla w^{p,k}| \geq i\}} |Z_p(\nabla w^{p,k}) - Z(\nabla w^{p,k})| \, dx \, dt \\ &\leq 2\Lambda \|\eta\|_{L^\infty(Q_T)} \lim_{k \rightarrow \infty} \iint_{\{(x,t): |\nabla w^{p,k}| \geq i\}} |\nabla w^{p,k}|^{p-1} \, dx \, dt. \end{aligned}$$

Noticing that

$$\lim_{i \rightarrow \infty} \text{meas}\{(x, t) : |\nabla w^{p,k}| \geq i\} = 0,$$

we see that I tends uniformly to 0 in p as $i \rightarrow \infty$. For II , we get that

$$\begin{aligned} II &\leq \|\eta\|_{L^\infty(Q_T)} \|\theta^i(Z_p - Z)\|_{L^\infty(Q_T)} \iint_{Q_T} \langle v^p, 1 \rangle dx dt \\ &= \|\eta\|_{L^\infty(Q_T)} \|\theta^i(Z_p - Z)\|_{L^\infty(Q_T)} \text{meas}(Q_T). \end{aligned}$$

So

$$\lim_{p \rightarrow 1^+} \iint_{Q_T} \langle v^p, Z_p - Z \rangle \cdot \eta dx dt = 0.$$

Therefore,

$$\begin{aligned} &\left| \iint_{Q_T} \langle v^{p_m}, Z_{p_m} \rangle \cdot \eta dx dt - \iint_{Q_T} \langle v, Z \rangle \cdot \eta dx dt \right| \\ &\leq \left| \iint_{Q_T} (\langle v^{p_m}, Z_{p_m} \rangle - \langle v^{p_m}, Z \rangle) \cdot \eta dx dt \right| + \left| \iint_{Q_T} (\langle v^{p_m}, Z \rangle - \langle v, Z \rangle) \cdot \eta dx dt \right| \\ &\rightarrow 0 \quad (m \rightarrow \infty). \end{aligned}$$

So $\{\langle v^{p_m}, Z_{p_m} \rangle\}_{m=1}^\infty$ converges weakly to $\langle v, Z \rangle$ in $L^1(Q_T)$. Thus (6) holds, namely,

$$\begin{aligned} &\iint_{Q_T} ((v, Z_p) \cdot \nabla \zeta + u_t \zeta - \lambda(u - f_p)\zeta) dx dt = 0, \\ &\forall \zeta \in C^\infty(\overline{Q_T}) \text{ with } \zeta(x, 0) = \zeta(x, T) = 0. \end{aligned}$$

Since $\{w^k\}_{k=1}^\infty$ converges to w in $L^1(Q_T)$, we get that for all $\eta \in C_0^\infty(Q_T \setminus G_j; \mathbb{R}^N)$,

$$\begin{aligned} &\lim_{k \rightarrow \infty} \left(\iint_{Q_T \setminus G_j} \eta \cdot \nabla w^k dx dt - \iint_{Q_T \setminus G_j} \eta \cdot \nabla w dx dt \right) \\ &= \lim_{k \rightarrow \infty} \iint_{Q_T \setminus G_j} (w - w^k) \text{div}(\eta) dx dt = 0. \end{aligned}$$

Thus for all $j \geq 1$,

$$\nabla w = \langle v, \text{id} \rangle \quad \text{in } Q_T \setminus G_j$$

in the sense of measure. Let $G = \bigcap_{j=1}^\infty G_j$. Then $\text{meas}(G) = 0$ and

$$\nabla w = \langle v, \text{id} \rangle \quad \text{in } Q_T \setminus G$$

in the sense of measure. Similar to the proof of Theorem 3, we get that

$$u(x, t) = w(x, t) \quad \text{a.e. } (x, t) \in Q_T.$$

Thus (7) holds, namely,

$$\nabla u(x, t) = \langle v_{x,t}, \text{id} \rangle \quad \text{a.e. } (x, t) \in Q_T$$

in the sense of measure.

We now prove (8). For all $\eta \in C_0^\infty(Q_T)$, we get that

$$\begin{aligned}
 & \left| \iint_{Q_T} (\langle v^{h_m}, Z_{p_m} \rangle - \langle v, Z \rangle) \cdot \langle v^{p_m}, \text{id} \rangle \eta \, dx \, dt \right| \\
 & \leq \left| \iint_{Q_T} (\langle v^{p_m}, Z_{p_m} \rangle - \langle v, Z \rangle) \cdot \nabla (u_{p_m} \eta) \, dx \, dt \right| \\
 & \quad + \left| \iint_{Q_T} u_{p_m} (\langle v^{p_m}, Z_{p_m} \rangle - \langle v, Z \rangle) \cdot \nabla \eta \, dx \, dt \right| \\
 & \leq \left| \iint_{Q_T} \left(\frac{\partial u_{p_m}}{\partial t} - \frac{\partial u}{\partial t} \right) u_{p_m} \eta \, dx \, dt \right| + \lambda \left| \iint_{Q_T} ((u_{p_m} - f_{p_m}) - (u - f)) u_{p_m} \eta \, dx \, dt \right| \\
 & \quad + \left| \iint_{Q_T} u_{p_m} (\langle v^{p_m}, Z_{p_m} \rangle - \langle v, Z \rangle) \cdot \nabla \eta \, dx \, dt \right| \\
 & \leq \left| \iint_{Q_T} \left(\frac{\partial u_{p_m}}{\partial t} - \frac{\partial u}{\partial t} \right) u \eta \, dx \, dt \right| + \left| \iint_{Q_T} \left(\frac{\partial u_{p_m}}{\partial t} - \frac{\partial u}{\partial t} \right) (u_{p_m} - u) \eta \, dx \, dt \right| \\
 & \quad + \lambda \left| \iint_{Q_T} (u_{p_m} - u) u_{p_m} \eta \, dx \, dt \right| + \lambda \left| \iint_{Q_T} (f_{p_m} - f) u_{p_m} \eta \, dx \, dt \right| \\
 & \quad + \left| \iint_{Q_T} u (\langle v^{p_m}, Z_{p_m} \rangle - \langle v, Z \rangle) \cdot \nabla \eta \, dx \, dt \right| \\
 & \quad + \left| \iint_{Q_T} (u^{p_m} - u) (\langle v^{h_m}, Z_{p_m} \rangle - \langle v, Z \rangle) \cdot \nabla \eta \, dx \, dt \right| \\
 & \rightarrow 0 \quad (m \rightarrow \infty).
 \end{aligned}$$

Since $\langle v_{x,t}^{p_m}, \text{id} \rangle$ converges to $\langle v_{x,t}, \text{id} \rangle$ in the biting sense, we see that for all $j \geq 1$ and all $\eta \in C_0^\infty(Q_T \setminus E_j)$,

$$\begin{aligned}
 & \left| \iint_{Q_T \setminus E_j} (\langle v^{p_m}, Z_{p_m} \rangle \cdot \langle v^{p_m}, \text{id} \rangle - \langle v, Z \rangle \cdot \langle v, \text{id} \rangle) \eta \, dx \, dt \right| \\
 & \leq \left| \iint_{Q_T \setminus E_j} (\langle v^{p_m}, Z_{p_m} \rangle - \langle v, Z \rangle) \cdot \langle v^{p_m}, \text{id} \rangle \eta \, dx \, dt \right| \\
 & \quad + \left| \iint_{Q_T \setminus E_j} (\langle v^{p_m}, \text{id} \rangle - \langle v, \text{id} \rangle) \cdot \langle v, Z \rangle \eta \, dx \, dt \right| \rightarrow 0 \quad (m \rightarrow \infty).
 \end{aligned}$$

Thus for all $j \geq 1$,

$$\lim_{m \rightarrow \infty} \langle v_{x,t}^{p_m}, Z_{p_m} \cdot \text{id} \rangle = \langle v_{x,t}, Z \rangle \cdot \langle v_{x,t}, \text{id} \rangle \quad \text{a.e. } (x, t) \in Q_T \setminus E_j.$$

Fix $i \geq C_0$. Let $0 \leq \eta \in C_0^\infty(Q_T \setminus E_j)$. Then

$$\begin{aligned}
 & \iint_{Q_T} \langle v^{p_m}, Z_{p_m} \cdot \text{id} \rangle \eta \, dx \, dt - \iint_{Q_T} \langle v, Z \cdot \text{id} \rangle \eta \, dx \, dt \\
 & = \left(\iint_{Q_T} \langle v^{p_m}, \theta^i(Z_{p_m} \cdot \text{id}) \rangle \eta \, dx \, dt - \iint_{Q_T} \langle v, \theta^i(Z \cdot \text{id}) \rangle \eta \, dx \, dt \right)
 \end{aligned}$$

$$\begin{aligned}
& + \left(\iint_{Q_T} \langle v^{p_m}, (1 - \theta^i)(Z_{p_m} \cdot \text{id}) \rangle \eta \, dx \, dt - \iint_{Q_T} \langle v, (1 - \theta^i)(Z \cdot \text{id}) \rangle \eta \, dx \, dt \right) \\
& = I + II.
\end{aligned}$$

Because $\theta^i(Z_{p_m} \cdot \text{id})$ converge uniformly to $\theta^i(Z \cdot \text{id})$ in \mathbb{R}^N , we see that

$$\begin{aligned}
& \left| \iint_{Q_T} \langle v^{p_m}, \theta^i(Z_{p_m} \cdot \text{id}) \rangle \eta \, dx \, dt - \iint_{Q_T} \langle v^{p_m}, \theta^i(Z \cdot \text{id}) \rangle \eta \, dx \, dt \right| \\
& \leq \|\eta\|_{L^\infty(Q_T)} \|\theta^i(Z_{p_m} \cdot \text{id}) - \theta^i(Z \cdot \text{id})\|_{L^\infty(Q_T)} \iint_{Q_T} \langle v^{p_m}, 1 \rangle \, dx \, dt \\
& = \|\eta\|_{L^\infty(Q_T)} \|\theta^i(Z_{p_m} \cdot \text{id}) - \theta^i(Z \cdot \text{id})\|_{L^\infty(Q_T)} \text{meas}(Q_T) \rightarrow 0 \quad (m \rightarrow \infty).
\end{aligned}$$

Hence

$$\begin{aligned}
\lim_{m \rightarrow \infty} I &= \lim_{m \rightarrow \infty} \left(\iint_{Q_T} \langle v^{p_m}, \theta^i(Z_{p_m} \cdot \text{id}) \rangle \eta \, dx \, dt - \iint_{Q_T} \langle v^{p_m}, \theta^i(Z \cdot \text{id}) \rangle \eta \, dx \, dt \right) \\
&+ \lim_{m \rightarrow \infty} \left(\iint_{Q_T} \langle v^{p_m}, \theta^i(Z \cdot \text{id}) \rangle \eta \, dx \, dt - \iint_{Q_T} \langle v, \theta^i(Z \cdot \text{id}) \rangle \eta \, dx \, dt \right) \\
&= 0.
\end{aligned}$$

From $(1 - \theta^i)(Z_{p_m} \cdot \text{id}) \geq (1 - \theta^i)(Z \cdot \text{id})$, we get that

$$\begin{aligned}
\lim_{m \rightarrow \infty} II &\geq \lim_{m \rightarrow \infty} \left(\iint_{Q_T} \langle v^{p_m}, (1 - \theta^i)(Z \cdot \text{id}) \rangle \eta \, dx \, dt - \iint_{Q_T} \langle v, (1 - \theta^i)(Z \cdot \text{id}) \rangle \eta \, dx \, dt \right) \\
&= 0.
\end{aligned}$$

Thus

$$\lim_{m \rightarrow \infty} \iint_{Q_T} \langle v^{p_m}, Z_{p_m} \cdot \text{id} \rangle \eta \, dx \, dt \geq \iint_{Q_T} \langle v, Z \cdot \text{id} \rangle \eta \, dx \, dt.$$

By the arbitrariness of $0 \leq \eta \in C_0^\infty(Q_T \setminus E_j)$, we see that

$$\lim_{m \rightarrow \infty} \langle v_{x,t}^{p_m}, Z_{p_m} \cdot \text{id} \rangle \geq \langle v_{x,t}, Z \cdot \text{id} \rangle, \quad \text{a.e. } (x, t) \in Q_T \setminus E_j.$$

Thus for all $j \geq 1$,

$$\langle v_{x,t}, Z \cdot \text{id} \rangle \leq \langle v_{x,t}, Z \rangle \cdot \langle v_{x,t}, \text{id} \rangle \quad \text{a.e. } (x, t) \in Q_T \setminus E_j.$$

Since $\langle v_{x,t}, Z \cdot \text{id} \rangle, \langle v_{x,t}, Z \rangle \cdot \langle v_{x,t}, \text{id} \rangle \in L^1(Q_T)$, we see that

$$\langle v_{x,t}, Z \cdot \text{id} \rangle \leq \langle v_{x,t}, Z \rangle \cdot \langle v_{x,t}, \text{id} \rangle \quad \text{a.e. } (x, t) \in Q_T.$$

So (8) holds, and therefore u is the Young measure solution of the problem (3)–(5). The proof of the theorem is complete. \square

Remark 3 Note that if the initial data $f \equiv C$, and then f is the Young measure of the problem (3)–(5).

4 Properties of the Young measure solution

Theorem 4 Let $u_1, u_2 \in \mathcal{B}$ be the Young measure solutions of the problem (3)–(5) with the initial data f_1, f_2 and satisfy (H6), respectively. Then for (x, t) a.e. in Q_T ,

$$-\operatorname{ess\,sup}_{x \in \Omega} (f_1 - f_2)^- \leq u_1(x, t) - u_2(x, t) \leq \operatorname{ess\,sup}_{x \in \Omega} (f_1 - f_2)^+. \quad (33)$$

In particular,

$$-\operatorname{ess\,sup}_{x \in \Omega} u_0^- \leq u(x, t) \leq \operatorname{ess\,sup}_{x \in \Omega} u_0^+, \quad \text{a.e. } (x, t) \in Q_T. \quad (34)$$

Proof Denote

$$K = \operatorname{ess\,sup}_{x \in \Omega} (f_1 - f_2)^+.$$

Let $G(t) \in C^1(\mathbb{R})$ be such that

$$G(t) = 0, \quad \forall t \in (-\infty, 0], \quad 0 < G'(t) \leq B \quad \forall t \in (0, \infty),$$

where B is a positive constant, and

$$\begin{aligned} H(t) &= \int_0^t G(s) \, ds, \\ \psi(t) &= \int_{\Omega} H(u_1(x, t) - u_2(x, t) - K) \, dx. \end{aligned}$$

Then $\psi \in C(\mathbb{R}^+) \cap H^1(\mathbb{R}^+)$, $\psi(0) = 0$, $\psi(t) \geq 0$, for $t \in \mathbb{R}^+$, and $G(u_1(x, t) - u_2(x, t) - K) \in \text{BV}(\Omega)$. Note that $G(u_1^k(x, t) - u_2^k(x, t) - K) \in W^{1,1}(\Omega)$ can be chosen as the test function and, taking Q_s for $s \in [0, T]$ as the domain of integration, we obtain

$$\begin{aligned} & \iint_{Q_s} \langle v_1, Z \rangle \cdot (\nabla u_1^k - \nabla u_2^k) G'(u_1^k - u_2^k - K) \, dx \, dt + \iint_{Q_s} \frac{\partial u_1}{\partial t} G(u_1^k - u_2^k - K) \, dx \, dt \\ & + \lambda \iint_{Q_s} (u_1 - f_1) G(u_1^k - u_2^k - K) \, dx \, dt = 0, \\ & \iint_{Q_s} \langle v_2, Z \rangle \cdot (\nabla u_1^k - \nabla u_2^k) G'(u_1^k - u_2^k - K) \, dx \, dt + \iint_{Q_s} \frac{\partial u_2}{\partial t} G(u_1^k - u_2^k - K) \, dx \, dt \\ & + \lambda \iint_{Q_s} (u_2 - f_2) G(u_1^k - u_2^k - K) \, dx \, dt = 0. \end{aligned}$$

From the proof of Theorem 4 and due to $0 < G' \leq B$,

$$\begin{aligned} \psi(s) &= \int_0^s \psi'(t) \, dt \\ &= \lim_{k \rightarrow \infty} \iint_{Q_s} G(u_1^k - u_2^k - K) \frac{\partial(u_1 - u_2)}{\partial t} \, dx \, dt \\ &= - \lim_{k \rightarrow \infty} \iint_{Q_s} (\langle v_1, Z \rangle - \langle v_2, Z \rangle) \cdot (\nabla u_1^k - \nabla u_2^k) G'(u_1^k - u_2^k - K) \, dx \, dt \end{aligned}$$

$$\begin{aligned}
& -\lambda \lim_{k \rightarrow \infty} \iint_{Q_s} (u_1 - f_1 - u_2 + f_2) G(u_1^k - u_2^k - K) \, dx \, dt \\
& \leq -\lambda \iint_{Q_s} (u_1 - f_1 - u_2 + f_2) G(u_1 - u_2 - K) \, dx \, dt \\
& \leq 0.
\end{aligned}$$

Hence

$$\psi(s) = 0, \quad \text{a.e. } s \in [0, T].$$

And therefore

$$H(u_1(x, t) - u_2(x, t) - K) = 0, \quad \text{a.e. } t \in [0, T].$$

Then

$$u_1(x, t) - u_2(x, t) - K \leq 0, \quad \text{a.e. } (x, t) \in Q_T,$$

which implies the right-hand side inequality of (33). Changing the position of u_1 and u_2 will yield the left-hand side inequality of (33). When the initial data $f \equiv \text{ess sup}_{x \in \Omega} u_0^+$, f is the Young measure solution of (3)–(5); when the initial data $f \equiv -\text{ess sup}_{x \in \Omega} u_0^-$, f is also the Young measure solution of (3)–(5), which completes the proof of the theorem. \square

Corollary 1 *Let $u_1, u_2 \in \mathcal{B}$ be the Young measure solutions of the problem (3)–(5) with the initial data f_1, f_2 and satisfy (H1), respectively. Assume*

$$f_1(x) \leq f_2(x), \quad \text{a.e. } x \in \Omega.$$

Then

$$u_1(x, t) \leq u_2(x, t), \quad \text{a.e. } (x, t) \in Q_T.$$

5 Numerical scheme

In this section, numerical results are presented to demonstrate the performance of our proposed algorithm for image restoration involving white Gaussian noise. The results are compared with those obtained by the PM method in [23] and the TV method in [8]. In the next two subsections, two numerical discrete schemes, the PM scheme (PMS) and the AOS scheme, will be proposed.

5.1 The AOS scheme

Let

$$C(X) = \frac{\mu_1}{\sqrt{1 + |X|^2}} + \frac{\mu_2}{1 + |X|^2}.$$

Using the scheme in [43], the problem (3)–(5) can be discretized as

$$\lambda^0 = 0,$$

$$\begin{aligned}
u^{n+1} &= \frac{1}{m} \sum_{l=1}^m [I - m\tau A_l(u^k)]^{-1} [u^n + \lambda\tau(f - u^n)], \\
\operatorname{div}^n &= (u^{n+1} - u^n)/\tau, \\
\lambda^n &= \frac{1}{\sigma^2 MN} (u - f) \operatorname{div}^n, \\
u_{i,j}^0 &= f_{i,j} = f(ih, jh), \\
u_{i,0}^n &= u_{i,1}^n, \quad u_{0,j}^n = u_{1,j}^n, \quad u_{I,i}^n = u_{I-1,i}^n, \quad u_{i,J}^n = u_{i,J-1}^n,
\end{aligned}$$

where $A_l(u^n) = [a_{i,j}(u^n)]$,

$$a_{i,j}(u^n) := \begin{cases} \frac{C_i^n + C_j^n}{2h^2}, & j \in \mathcal{N}(i), \\ -\sum_{n \in \mathcal{N}(i)} \frac{C_i^n + C_N^n}{2h^2}, & j = i, \\ 0, & \text{else,} \end{cases}$$

and

$$C_i^n := C(|\nabla u_{i,j}^n|),$$

where

$$|\nabla u_{i,j}^n| = \frac{1}{2} \sum_{p,q \in \mathcal{N}(i)} \frac{|u_p^n - u_q^n|}{2h},$$

with $\mathcal{N}(i)$ being the set of the two neighbors of pixel i (boundary pixels have only one neighbor).

5.2 The PM scheme

Similar to the original PM method, the discrete explicit scheme of the problem (3)–(5) is as follows:

$$\begin{aligned}
C_{Ni,j}^n &= C(|\nabla_N u_{i,j}|), \quad C_{Si,j}^n = C(|\nabla_S u_{i,j}|), \\
C_{Ei,j}^n &= C(|\nabla_E u_{i,j}|), \quad C_{Wi,j}^n = C(|\nabla_W u_{i,j}|), \\
\operatorname{div}_{i,j}^n &= (C_{Ni,j}^n \nabla_N u_{i,j} + C_{Si,j}^n \nabla_S u_{i,j} \\
&\quad + C_{Ei,j}^n \nabla_E u_{i,j} + C_{Wi,j}^n \nabla_W u_{i,j}), \\
\lambda^n &= \frac{1}{\sigma^2 MN} \sum_{i,j} \operatorname{div}_{i,j}^n (u_{i,j} - f_{i,j}), \\
u_{i,j}^{n+1} &= u_{i,j}^n + \tau \operatorname{div}_{i,j}^n - \lambda^n \tau (u_{i,j} - f_{i,j}), \\
u_{i,j}^0 &= f_{i,j} = f(ih, jh), \\
u_{i,0}^n &= u_{i,1}^n, \quad u_{0,j}^n = u_{1,j}^n, \quad u_{I,i}^n = u_{I-1,i}^n, \quad u_{i,J}^n = u_{i,J-1}^n,
\end{aligned}$$

where

$$\begin{aligned}\nabla_N u_{i,j} &= u_{i-1,j} - u_{i,j}, & \nabla_S u_{i,j} &= u_{i+1,j} - u_{i,j}, \\ \nabla_E u_{i,j} &= u_{i,j+1} - u_{i,j}, & \nabla_W u_{i,j} &= u_{i,j-1} - u_{i,j},\end{aligned}$$

for $0 \leq i \leq I, 0 \leq j \leq J$.

6 Denoising performance

The denoising algorithms are tested on four images: a synthetic image (128×128 pixels), Lena image (300×300 pixels), and a tower image (500×500 pixels). For each image, a noisy observation is generated by adding Gaussian noise with the standard deviation $\sigma \in \{30, 50\}$ to the original image.

Peak signal-to-noise ratio (PSNR) and the mean absolute deviation error (MAE) are used to measure the quality of the restoration results. They are defined as

$$\begin{aligned}\text{PSNR} &= 10 \log_{10} (255^2 MN / \|u_O - u\|_2^2) \text{ dB}, \\ \text{MAE} &= \|u_O - u\|_1 / (MN),\end{aligned}$$

where u_O and u are the original and restored image, respectively. The stopping criterion of all methods is set to achieve the maximal PSNR or the best MAE. All methods are implemented in Matlab R2007b on a 2.8 GHz Pentium 4 processor.

6.1 Measure of similarity of edges

The PSNR does not always give a clear guide as to whether one image is less staircased than another, so the authors [44] take into account the value of $\text{PSNR}_{\text{grad}}$ which they define as $1/2(\text{PSNR}(\partial_x u, \partial_x u_O) + \text{PSNR}(\partial_y u, \partial_y u_O))$, and this should measure how well the derivatives of reconstruction match those of the true image.

The edge maps are defined as

$$\text{EM}(u) = 2 - \frac{2}{1 + k |\nabla G_\sigma * u|^2}, \quad k > 0, \sigma > 0, \quad (35)$$

where $G_\sigma(x) = \frac{1}{4\pi\sigma} \exp(-\frac{|x|^2}{4\sigma})$. If all images are normalized, their gray-scale is in the range $[0, 255]$. The authors [45] find that a value of $0.025 \leq c \leq 0.0025$ and $\sigma = 0.5$ give the best edge map. In [46], the authors propose the following PSNR:

$$\widetilde{\text{PSNR}} = 10 \log_{10} \frac{MN |\max u_O - \min u_O|^2}{\|u - u_O\|_2^2} \text{ dB}, \quad (36)$$

where $|\max u_O - \min u_O|$ gives the gray-scale range of the original image. It is noticed that $\widetilde{\text{PSNR}}$ can measure how well the reconstruction data match the true data, and the data need not to be an image. Combining (35) and (36), in [47], we define the following PSNR_E in order to measure the similarity of edges:

$$\text{PSNR}_E = \widetilde{\text{PSNR}}(\text{EM}(u), \text{EM}(u_O)) \text{ dB}.$$

If there are wrong edges in the restored image by some method, then the PSNR_E will be positive.

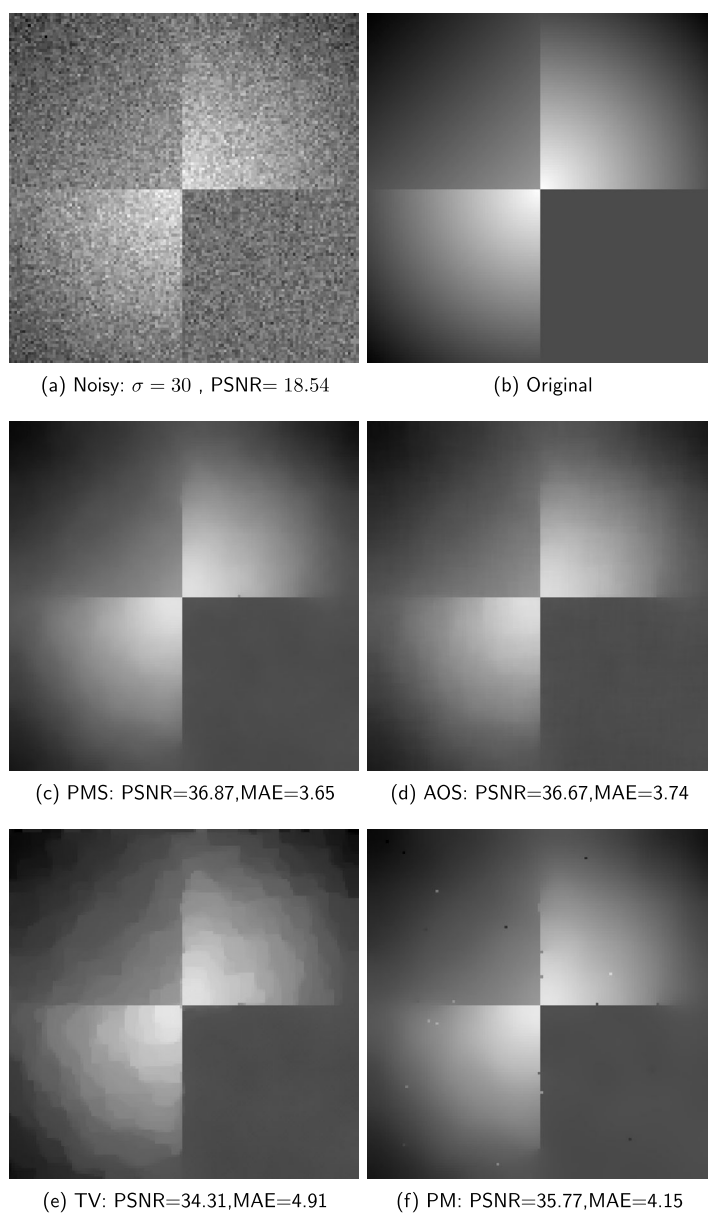


Figure 1 Synthetic image (128×128): (a) noisy image corrupted by Gaussian noise for $\sigma = 30$; (b) original image; (c) our algorithm by PMS, $\mu_1 = 0.5$, $\mu_2 = 0.5$, and $\tau = 0.2$; (d) our algorithm by AOS, $\mu_1 = 0.9$, $\mu_2 = 0.1$, and $\tau = 2$; (e) TV algorithm, $\tau = 0.2$; and (f) PM algorithm, $\tau = 0.2$ and $K = 7$

6.2 Comparison with other methods

The results are depicted in Figs. 1–3 for the synthetic image and Figs. 4–5 for Lena image. Our methods do a good job in restoring faint geometrical structures of the images even for high values of σ , for instance, the results on Lena image for $\sigma = 50$. Our algorithm distinguishes itself from its competitors most of the time both visually and quantitatively, as revealed by the PSNR and MAE values. For TV method, the number of iterations which is necessary to satisfy the stopping rule increases rapidly when σ increases. For PM method, the appropriate parameter K is indispensable to get the best result.

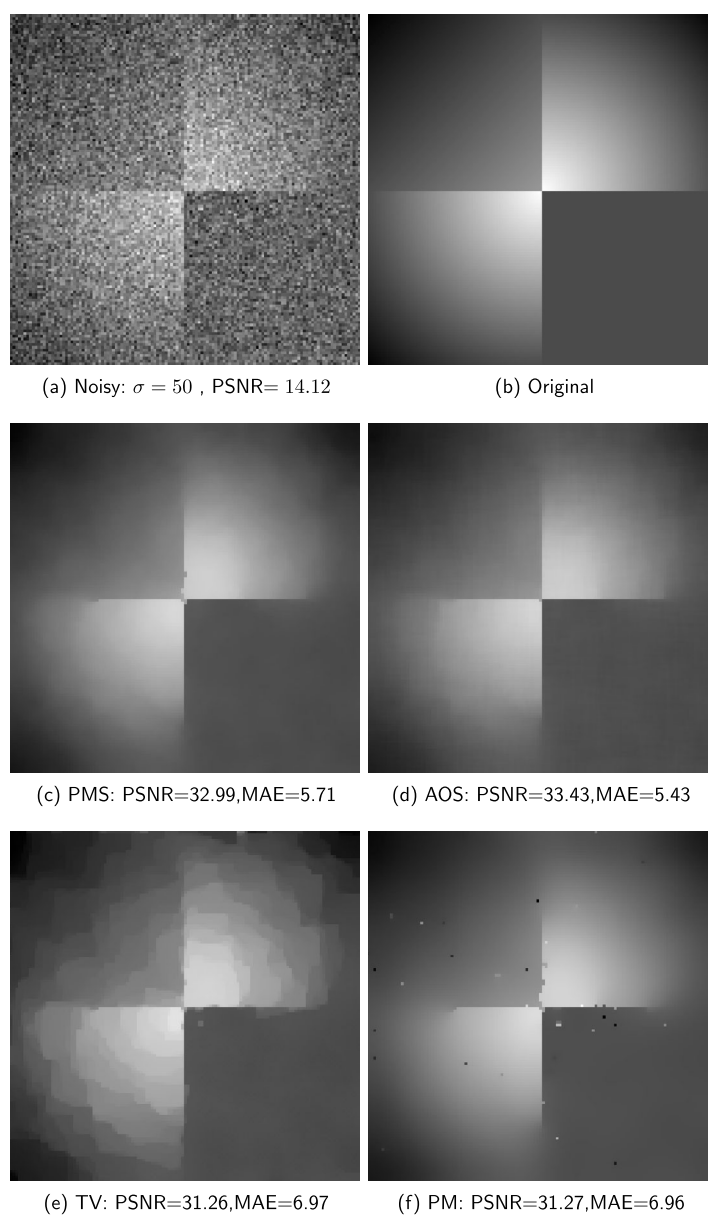


Figure 2 Synthetic image (128×128): (a) noisy image corrupted by Gaussian noise for $\sigma = 50$; (b) original image; (c) our algorithm by PMS, $\mu_1 = 0.3$, $\mu_2 = 0.7$, and $\tau = 0.2$; (d) our algorithm by AOS, $\mu_1 = 0.9$, $\mu_2 = 0.1$, and $\tau = 2$; (e) TV algorithm, $\tau = 0.2$; and (f) PM algorithm, $\tau = 0.2$ and $K = 8$

Figures 1–3 illustrate that the proposed model is able to reconstruct sharp edges and nonuniform regions while avoiding staircasing. TV-based diffusion reconstructs sharp edges, but the staircasing effect is obvious. PM-based diffusion also reconstructs sharp edges, but it creates isolated black and white speckles in the restored image. The proposed model reconstructs sharp edges as effectively as PM-based diffusion and meanwhile recovers smooth regions as effectively as pure isotropic diffusion (in particular, without staircasing). Figure 3 shows the edge functions when the smoothed images by the new methods, TV, and PM attain the largest PSNR, respectively. Note that PM and TV methods create many new edges in the restored images.

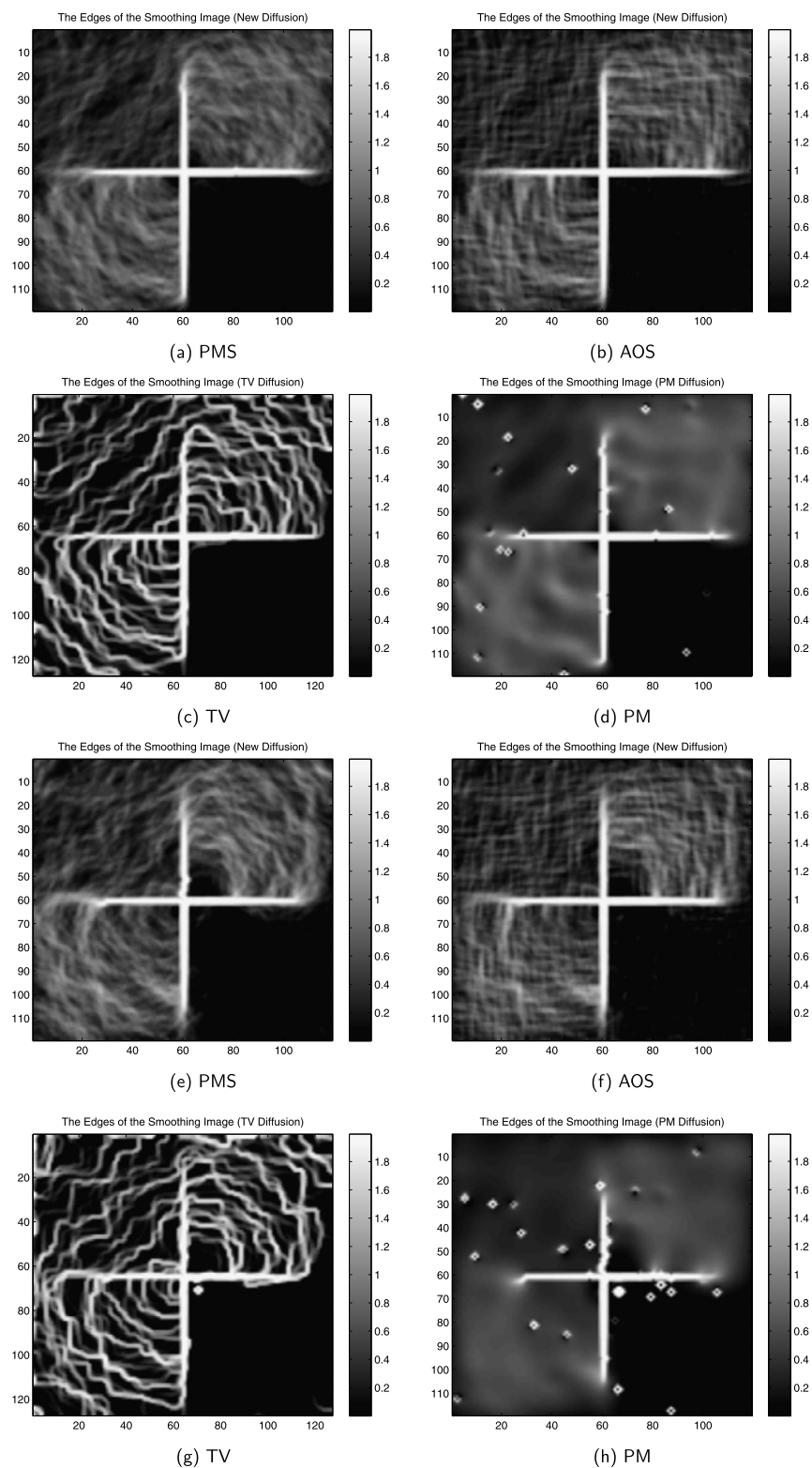


Figure 3 The edge indicator function. **(a)–(d)** The edge indicator functions for the denoising results of every methods when $\sigma = 30$; **(e)–(h)** The edge indicator functions for the denoising results of every methods when $\sigma = 50$



Figure 4 Lena image (300×300): (a) noisy image corrupted by Gaussian noise for $\sigma = 20$; (b) original image; (c) our algorithm by PMS, $\mu_1 = 0.3$, $\mu_2 = 0.7$, and $\tau = 0.2$; (d) our algorithm by AOS, $\mu_1 = 0.9$, $\mu_2 = 0.1$, and $\tau = 2$; (e) TV algorithm, $\tau = 0.2$; and (f) PM algorithm, $\tau = 0.2$ and $K = 5$

Figures 4–5 show the restored Lena images produced by the Perona–Malik equation, the TV method, PMS, and AOS. Figures 4(e) and 5(e) show the processed images by the PM diffusion with isolated black and white speckles. However, in Figs. 4(c)–4(d), the processed images using the new methods are very clear. Inside the regions, the new diffusion acts as Gaussian smoothing, so our method can effectively avoid staircase effect. In Tables 1 and 2, we observe that both PSNRs and MAEs of the restored images in our methods are better than those in the PM and TV methods. And the increase in PSNR_E is obvious with the new diffusion operators.



Figure 5 Lena image (300×300): (a) noisy image corrupted by Gaussian noise for $\sigma = 30$; (b) original image; (c) our algorithm by PMS, $\mu_1 = 0.5$, $\mu_2 = 0.5$, and $\tau = 0.2$; (d) our algorithm by AOS, $\mu_1 = 0.9$, $\mu_2 = 0.1$, and $\tau = 2$; (e) TV algorithm, $\tau = 0.2$; and (f) PM algorithm, $\tau = 0.2$ and $K = 7$

The denoising performance results are tabulated in Tables 1–2 where the best PSNR, MAE, PSNR_E, and CPU time are shown in boldface. The PSNR improvement brought by our approach can be quite high, particularly for $\sigma = 50$ (see, e.g., Figs. 1–2), and the visual resolution is quite remarkable. For $\sigma = 30$, the PSNRs of our algorithm also can be higher than that of PM and TV methods. Moreover, the new algorithm by AOS scheme shows high PSNRs on real images (Figs. 3–4). Note that for big size images (Figs. 3–4, 300×300), the new methods take less time than TV and PM methods.

Table 1 PSNR, MAE, PSNR_E, and CPU time (seconds) of all methods for the synthetic image (128 × 128)

	σ							
	PSNR		MAE		PSNR _E		CPU(s)	
	30	50	30	50	30	50	30	50
PMS	36.87	32.99	3.65	5.71	21.20	18.73	1.67	1.94
AOS	36.67	33.43	3.74	5.43	20.83	19.11	0.33	0.59
PM	35.78	31.27	4.14	6.97	20.58	18.36	0.15	0.29
TV	34.31	31.26	4.91	6.97	13.81	12.58	1.12	1.79

Table 2 PSNR, MAE, PSNR_E, and CPU time (seconds) of all methods for the Lena image (300 × 300)

	σ							
	PSNR		MAE		PSNR _E		CPU(s)	
	20	30	30	50	30	50	30	50
PMS	29.64	27.97	8.40	10.36	11.35	10.09	1.69	4.86
AOS	29.94	28.13	8.11	10.00	11.75	10.64	0.65	0.83
PM	28.49	26.36	9.58	12.25	10.77	9.53	0.80	0.92
TV	28.62	26.84	9.42	11.33	10.87	9.71	0.73	4.02

7 Conclusion

In this paper, based on convex and nonconvex linear growth functionals, we proposed a class of singular diffusion equations for noise removal. In our method, the convex and nonconvex functionals are combined into a new functional. It is hard to consider the analysis of the new functional. However, the new singular forward–backward diffusion equation is introduced from the functional. And the existence, uniqueness, and long-time behavior of solutions for the new equation are investigated. Finally, experimental results illustrate the effectiveness of the model in noise reduction.

This work proposes quite an original and efficient method for noise removal. Noise removal is a difficult problem that arises in various applications relevant to active imaging system. The main ingredients of our method are as follows: (1) The new framework is based on a combination between convex and nonconvex functions; (2) The new equation is a forward–backward diffusion and it is singular; (3) The Young measure solution is obtained, and some useful properties of the solution are considered, such as long-time behavior, stability, maximum principle, and so on; (4) The new diffusion can be simulated by the efficient AOS scheme.

The obtained numerical results are really encouraging since they outperform the most recent methods in this field.

Acknowledgements

Not applicable.

Funding

The authors would like to express their sincere thanks to the referees for their valuable suggestions in the revision of the paper which contributed greatly to this work. This work was partially supported by the National Science Foundation of China (Grant no. 11271100) and the 985 project of Harbin Institute of Technology, and the Natural Science Foundation of Heilongjiang Province (Grant no. LH2020A015).

Abbreviations

AOS, additive operator splitting; PSNR, peak signal-to-noise ratio; MAE, mean absolute deviation error; TV, total variation; PM, Perona–Malik; CPU, central processing unit.

Availability of data and materials

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

The authors read and approved the final manuscript.

Author details

¹Department of Applied Mathematics, Harbin University of Science and Technology, GA, 150001 Harbin, China. ²School of Mathematics, Harbin Institute of Technology, GA, 150001 Harbin, China.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 27 March 2020 Accepted: 3 January 2021 Published online: 19 January 2021

References

- Rudin, L.I., Osher, S., Fatemi, E.: Nonlinear total variation based noise removal algorithms. *Phys. D: Nonlinear Phenom.* **60**(1–4), 259–268 (1992)
- Vese, L.: *Problèmes variationnels et edp pour l'analyse d'images et l'évolution de courbes*. PhD thesis, Nice (1996)
- Meyer, Y.: *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations: The Fifteenth Dean Jacqueline B. Lewis Memorial Lectures*, vol. 22. Am. Math. Soc., Providence (2001)
- Strong, D., Chan, T.: Edge-preserving and scale-dependent properties of total variation regularization. *Inverse Probl.* **19**(6), 165 (2003)
- Belletti, G., Caselles, V., Novaga, M.: The total variation ow in \mathbb{R}^N . *J. Differ. Equ.* **184**(2), 475–525 (2002)
- Chan, T., Esedoglu, S., Park, F.: Total variation image restoration: overview and recent developments. In: *Handbook of Mathematical Models in Computer Vision*. Springer, Berlin (2006)
- Strong, D.M., Chan, T.F. (eds.): *Spatially and scale adaptive total variation based regularization and anisotropic diffusion in image processing*. Diffusion in image processing. UCLA Math Department CAM report (1996)
- Chambolle, A., Lions, P.L.: Image recovery via total variation minimization and related problems. *Numer. Math.* **76**, 167–188 (1997)
- Chen, Y.M., Levine, S., Rao, M.: Variable exponent, linear growth functionals in image restoration. *SIAM J. Appl. Math.* **66**(4), 1383–1406 (2006)
- Chan, T., Marquina, A., Mulet, P.: High-order total variation-based image restoration. *SIAM J. Sci. Comput.* **22**, 503–516 (2000)
- Lysaker, M., Lundervold, A., Tai, X.C.: Noise removal using fourth-order partial differential equation with applications to medical magnetic resonance images in space and time. *IEEE Trans. Image Process.* **12**(12), 1579–1590 (2003)
- Chan, T., Marquina, A., Mulet, P.: High-order total variation-based image restoration. PhD thesis (2000)
- You, Y.L., Kaveh, M.: Fourth-order partial differential equations for noise removal. *IEEE Trans. Image Process.* **9**(10), 1723–1730 (2000)
- Theljani, A., Belhachmi, Z., Moakher, M.: High-order anisotropic diffusion operators in spaces of variable exponents and application to image inpainting and restoration problems. *Nonlinear Anal., Real World Appl.* **47**, 251–271 (2019)
- Mei, J.-J., Huang, T.-Z.: Primal–dual splitting method for high-order model with application to image restoration. *Appl. Math. Model.* **40**(3), 2322–2332 (2016)
- Liu, C., Jin, M.: Some properties of solutions of a fourth-order parabolic equation for image processing. *Bull. Malays. Math. Sci. Soc.* **43**(1), 333–353 (2020)
- Rathish Kumar, B.V., Halim, A.: A linear fourth-order PDE-based gray-scale image inpainting model. *Comput. Appl. Math.* **38**(1), 6–21 (2019)
- Aubert, G., Kornprobst, P.: *Mathematical Problems in Image Processing* (2002)
- Hebert, T., Leahy, R.: A generalized em algorithm for 3-d Bayesian reconstruction from Poisson data using Gibbs priors. *IEEE Trans. Med. Imaging* **8**, 194–202 (1990)
- Geman, S., Reynolds, G.: Constrained restoration and the recovery of discontinuities. *IEEE Trans. Pattern Anal. Mach. Intell.* **14**, 367–383 (1992)
- Geman, D., Yang, C.: Nonlinear image recovery with half-quadratic regularization. *IEEE Trans. Image Process.* **4**, 932–946 (1995)
- Chipot, M., March, R., Rosati, M., Caarelli, G.V.: Analysis of a nonconvex problem related to signal selective smoothing. *Math. Models Methods Appl. Sci.* **7**, 313–328 (1997)
- Perona, P., Malik, J.: Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Anal. Mach. Intell.* **12**, 629–639 (1990)
- Blake, A.: Comparison of the efficiency of deterministic and stochastic algorithms for visual reconstruction. *IEEE Trans. Pattern Anal. Mach. Intell.* **11**(1), 2–12 (1989)
- Charbonnier, B.-F.L.A.G.P., Barlaud, M.: Deterministic edge-preserving regularization in computed imaging. *IEEE Trans. Image Process.* **6**(2), 298–311 (1997)
- Nikolova, M., Michael, K.N.: Analysis of half-quadratic minimization methods for signal and image recovery. *SIAM J. Sci. Comput.* **27**(3), 937–966 (2005)
- Xiong, K., Lu, H.H.C., Wang, S.: Kernel correntropy conjugate gradient algorithms based on half-quadratic optimization. *IEEE Trans. Cybern.*, 1–14 (2020)
- Bergmann, R., Chan, R.H., Hielscher, R., Persch, J., Steidl, G.: Restoration of manifold-valued images by half-quadratic minimization. *Inverse Probl. Imaging* **10**(2), 281–304 (2016)
- Robini, M.C., Zhu, Y.: Generic half-quadratic optimization for image reconstruction. *SIAM J. Imaging Sci.* **8**(3), 1752–1797 (2015)

30. Weickert, J., Romeny, B.M.T.H., Viergever, M.A.: Efficient and reliable schemes for nonlinear diffusion filtering. *IEEE Trans. Image Process.* **7**(3), 398–410 (1998)
31. Charbonnier, B.-F.L.A.G.P., Barlaud, M. (eds.): In: *Proceedings of 1st International Conference on Image Processing*, Nov. 1994, Austin, TX, USA, pp. 168–172. IEEE, Austin (1994)
32. Weickert, J.: *Anisotropic Diffusion in Image Processing*. Teubner, Stuttgart (1998)
33. Catte, F., Lions, P.-L., Morel, J.-M., Coll, T.: Image selective smoothing and edge detection by nonlinear diffusion. *SIAM J. Numer. Anal.* **29**(1), 182–193 (1992)
34. Keeling, S.L., Stollberger, R.: Nonlinear anisotropic diffusion filtering for multiscale edge enhancement. *Inverse Probl.* **18**(1), 175–190 (2002)
35. Gilboa, G., Sochen, N., Zeevi, Y.Y.: Forward-and-backward diffusion processes for adaptive image enhancement and denoising. *IEEE Trans. Image Process.* **11**(7), 689–703 (2002)
36. Smolka, B.: Combined forward and backward anisotropic diffusion filtering of color images. In: *Joint Pattern Recognition Symposium*, pp. 314–322. Springer, Berlin (2002)
37. Kinderlehrer, D., Pedregal, P.: Remarks about the analysis of gradient young measures. *J. Geom. Anal.* **4**(1), 59–90 (1994)
38. Demoulini, S.: Young measure solutions for a nonlinear parabolic equation of forward–backward type. *SIAM J. Math. Anal.* **27**(2), 376–403 (1996)
39. Kinderlehrer, D., Pedregal, P.: Weak convergence of integrands and the young measure representation. *SIAM J. Math. Anal.* **23**(1), 1–19 (1992)
40. Yin, J., Wang, C.: Young measure solutions of a class of forward–backward diffusion equations. *J. Math. Anal. Appl.* **279**(2), 659–683 (2003)
41. Guidotti, P.: A backward–forward regularization of the Perona–Malik equation. *J. Differ. Equ.* **252**, 3226–3244 (2012)
42. Guidotti, P., Kim, Y., Lambers, J.: Image restoration with a new class of forward–backward–forward diffusion equations of Perona–Malik type with applications to satellite image enhancement. *SIAM J. Imaging Sci.* **6**(3), 1416–1444 (2013)
43. Dacorogna, B.: *Direct Methods in the Calculus of Variations*, vol. 78. Springer, Berlin (2007)
44. Tai, X.C., Lie, K.A., Chan, T.F., Osher, S.: *Image Processing Based on Partial Differential Equations*. Springer, Heidelberg (2006)
45. Levine, S., Chen, Y., Stanich, J.: Image restoration via nonstandard diffusion. Technical Report 04-01, Department of Mathematics and Computer Science, Duquesne University (2004)
46. Durand, S., Fadili, J., Nikolova, M.: Multiplicative noise removal using l_1 fidelity on frame coefficients. *J. Math. Imaging Vis.* **36**, 201–226 (2010)
47. Guo, Z., Sun, J., Zhang, D., Wu, B.: Adaptive Perona–Malik model based on the variable exponent for image denoising. *IEEE Trans. Image Process.* **21**(3), 958–967 (2012)

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)