Calc. Var. (2006) 26(2): 171–199
DOI 10.1007/s00526-005-0363-4

Calculus of Variations

Kewei Zhang

Existence of infinitely many solutions for the one-dimensional Perona-Malik model

Received: 4 November 2004 / Accepted: 10 October 2005 / Published online: 8 March 2006

Abstract We establish the existence of infinitely many weak solutions for the the one-dimensional version of the well-known and widely used Perona-Malik anisotropic diffusion equation model in image processing. We establish the existence result under the homogeneous Neumann condition with smooth non-constant initial values. Our method is to convert the problem into a partial differential inclusion problem.

Keywords Perona-Malik model · One-dimensional · Infinitely many solutions · Differential inclusion · Relaxation property

1 Introduction

In this paper we establish the existence of infinitely many weak solutions for the one-dimensional Perona-Malik anisotropic diffusion equation under the homogeneous Neumann boundary condition:

$$\begin{cases} u_t = \sigma(u_x)_x, & (t, x) \in (0, T) \times (0, l) := Q_T, \\ u(0, x) = u_0(x), & 0 \le x \le l, \\ u_x(t, 0) = u_x(t, l) = 0, & 0 \le t \le T, \end{cases} \tag{1.1}$$

where $\sigma(s) = s/(1 + s^2)$.

This work can be viewed as an unexpected application of the variational techniques originated by Ball and James [6, 7] for the study of material microstructure to forward-backward diffusion equations. The main idea is to rephrase (1.1) into a

K. Zhang (✉)
Department of Mathematics, University of Sussex, Brighton, BN1 9RF, UK
E-mail: k.zhang@sussex.ac.uk

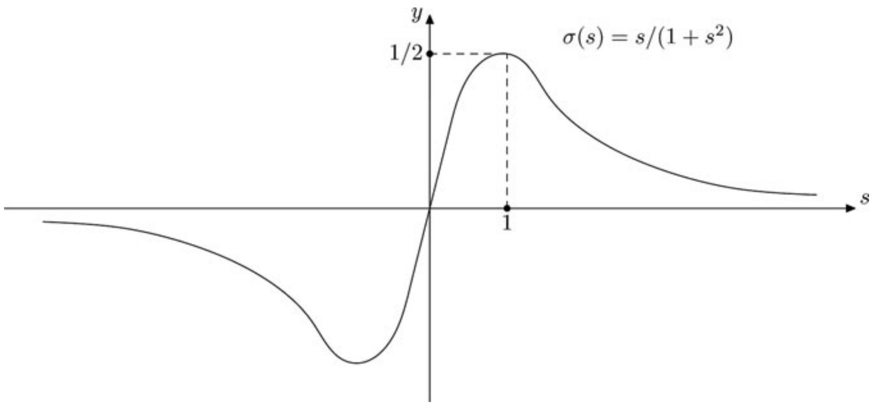

Fig. 1 1-d Peronal-Malik model $\sigma(s) = s/(1+s^2)$ for $k = 1$ with scaling

first order partial differential inclusion problem and apply the variational methods for such problems developed in the last decade [9, 10, 18, 23–27].

The Perona-Malik anisotropic diffusion equation [28] was introduced in 1990 as an edge enhancement model in image processing. The model has a great impact in the study of image enhancement and edge detection by using evolutionary partial differential equations. It has motivated many new models and methods (see e.g. [3, 8, 29, 32] and references therein).

The original Perona-Malik equation is a two dimensional forward-backward diffusion equation in the form

$$u_t(t,x) = \operatorname{div}(\rho(|Du(t,x)|)Du(t,x)), \quad \text{in } (0,\ T)\times\Omega$$

under the homogeneous Neumann boundary condition $\frac{\partial u}{\partial \nu} = 0$ on $\partial\Omega$, where $\Omega \subset \mathbb{R}^2$ is a square. Perona and Malik proposed two models for $\rho(s)$, that is

$$\text{either} \quad \rho(s) = \frac{1}{1+\left(\frac{s}{k}\right)^2}, \quad \text{or} \quad \rho(s) = \exp\left(-\left(\frac{s}{k}\right)^2\right), \quad (k>0).$$

In both cases, the Perona-Malik equations are diffusion equations of non-coercive and of forward-backward type. Let $k = 1$ for simplicity and let $\sigma(s) = s/(1+s^2)$, we see that $\sigma(s)$ reaches its maximum at $s = 1$.

The main mathematical results on Perona-Malik model is mostly for the one-dimensional version (1.1). So far, there were no existence results when the initial datum u_0 has large derivative $|(u_0)_x|$. In [19] it was proved that for small $|(u_0)_x|$, (1.1) has a global smooth solution, as one only needs to use the increasing part of $\sigma(\cdot)$ and the maximum principle. Also in [19], it was reported that an attempt to use the vanishing viscosity argument did not seem to produce a solution. It was shown in [17] that for large initial data, there are no C^1 solutions. In [5], a one-dimensional steady-state model of Perona-Malik equation was studied by using regularization and Γ-convergence, showing the formation of staircase function as a possible limit. A connection between Perona-Malik model and the Mumford-Shah functional was established via Γ-convergence in [22]. In a famous work on

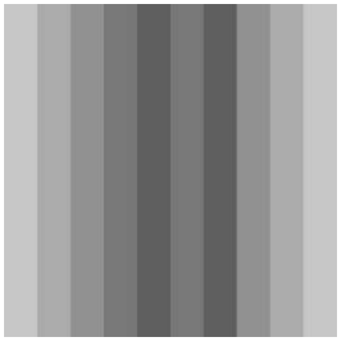

Fig. 2 Illustration of a 'one-dimensional image'

evolutionary partial differential equation models for image processing [4], it was suggested that the solutions of Perona-Malik model might not be stable.

The one-dimensional Perona-Malik model (1.1) can be viewed as a reduced two-dimensional problem when the initial value depends only on one variable (see Fig. 2 for an illustration).

A one-dimensional equation deduced from Perona-Malik model is also used in the continuum modelling for movements of biological organisms by Horstmann, Painter and Othmer [16], where the equation is

$$v_t = \sigma(v)_{xx}.$$

Obviously, if u is a weak solution of the Perona-Malik equation (1.1), $v = u_x$ is a solution of the biological model above. For large initial data, numerical computations in [16] indicate that the approximate solutions may oscillate very fast in certain regions.

There are many numerical schemes [8, 12, 16, 29, 32] to simulate the solution. The schemes intuitively work well except that the staircase phenomenon may occur [8, 29, 32]. However, there is no proof that any numerical schemes devised so far converge to a solution of the original Peronal-Malik equation except when the initial value has small gradient.

As for the existence problem of general coercive forward-backward diffusion equations, it was shown in the pioneering work of K. Höllig [15] that for a special piecewise affine flux function $\sigma(\cdot)$ as illustrated in Fig. 3, infinitely many weak solutions for (1.1) can be constructed.

The construction of solutions by Höllig depends heavily on the fact that $\sigma(\cdot)$ is piecewise affine and satisfies the coercivity condition $\sigma(s)s \geq c|s|^2$ for some constant $c > 0$. We observe that neither requirements hold for the one-dimensional Perona-Malik model $\sigma(s) = s/(1+s^2)$.

In this paper, we rephrase the one dimensional Perona-Malik model (1.1) as a partial differential inclusion problem and establish the existence of infinitely many weak solutions in $W^{1,\infty}(Q_T)$ for smooth (say $C^{3,1}$) initial values u_0 satisfying the boundary condition $(u_0)_x(0) = (u_0)_x(l) = 0$, whose derivative $(u_0)_x$ is not identically zero. In other words, u_0 is not identically a constant.

As a by product, our main result also implies that the following one-dimensional model for movements of biological organisms [16] has infinitely

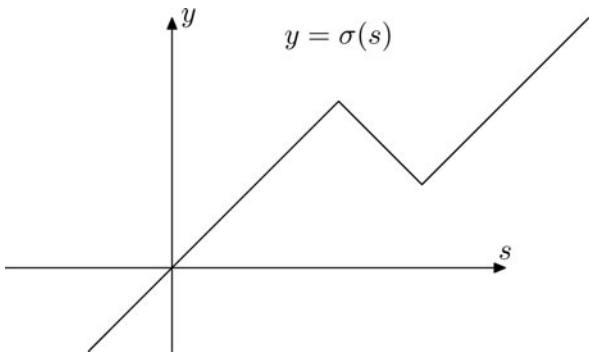

Fig. 3 Höllig's $\sigma(s)$ in Eq. (1.1)

many weak solutions in $L^\infty(Q_T)$ if the initial value v_0 satisfies $v_0(0) = v_0(l) = 0$ and is not identically zero.

$$\begin{cases} v_t = \sigma(v)_{xx}, & \text{in } Q_T, \\ v(0,x) = v_0(x), & 0 \le x \le l, \\ v(t,0) = v(t,l) = 0, & 0 \le t \le T, \end{cases} \tag{1.2}$$

where $\sigma(s) = s/(1+s^2)$. It is clear that a $W^{1,\infty}$ solution for (1.1) gives rise to an L^∞ solution for (1.2) if we let $v = u_x$.

When $s^* := \max_{[0,\,l]} |(u_0)_x(x)| < 1$, we denote by $u_K(t,x)$ the smooth solution of (1.1) obtained in [19]. This solution was obtained by modifying $\sigma(s)$ for $|s| \ge s^*$ by straight lines so that the resulting flux function is strictly increasing and agrees with $\sigma(s)$ when $|s| \le s^*$. As the maximum principle applies to $(u_K)_x$, we have $|(u_K)_x(t,x)| \le s^*$.

The following are our main results.

Theorem 1.1 *Suppose $\sigma(s) = s/(1+s^2)$ for $s \in \mathbb{R}$. Let $u_0 \in C^{3,\alpha}([0,l])$ $(0 < \alpha \le 1)$, with*

$$(u_0)_x(0) = (u_0)_x(l) = 0, \quad \text{and} \quad \max_{[0,\,l]} |(u_0)_x(x)| \ne 0.$$

Then the Neumann problem

$$\begin{cases} u_t - \sigma(u_x)_x = 0, & (t,x) \in Q_T, \\ u(0,x) = u_0(x), & x \in [0,\, l], \\ u_x(t,0) = u_x(t,l) = 0, & 0 \le t \le T \end{cases} \tag{1.3}$$

has infinitely many weak solutions $u \in W^{1,\infty}(Q_T)$ satisfying that

(a) for every $\psi \in C_0^1(Q_T)$,

$$\int_{Q_T} [u_t \psi + \sigma(u_x)\psi_x]\, dx\, dt = 0; \tag{1.4}$$

(b) for some $\delta > 0$,

$$u \in C^{1,2}((0,T)\times\{(0,\delta)\cup(l-\delta,l)\}), \quad u_t - \sigma(u_x)_x = 0,$$
$$(t,x) \in (0,T)\times\{(0,\delta)\cup(l-\delta,l)\}; \tag{1.5}$$

(c) the initial condition holds:

$$u(0,x) = u_0(x), \quad x \in [0,l]; \tag{1.6}$$

(d) and the boundary condition is satisfied:

$$u_x(t,0) = u_x(t,l) = 0, \quad 0 \le t \le T. \tag{1.7}$$

In case $\max_{[0,l]} |(u_0)_x(x)| < 1$*, the solutions* u *above are different from the smooth solution* u_K *obtained in* [19] *in the sense that* $u_x(t,x) \neq (u_K)_x(t,x)$ *on a set of positive measure in* Q_T.

Items (a)–(d) in Theorem 1.1 simply say that our solutions are classical solutions of the original problem (1.3) near the boundary, satisfying the Neumann boundary condition in the classical sense.

Remark 1.1 I will not speculate the importance of Theorem 1.1 in anisotropic diffusion modelling of image enhancement. However, Theorem 1.1 is the first non-trivial existence result for the Perona-Malik equation. It simply says that for any non-constant initial value, the one-dimensional Perona-Malik model will generate infinitely many non-smooth weak solutions no matter whether the derivative of the initial value is large or small. This is in contrast with the result in [19] that for initial value with small derivative, the smooth solution exists and is unique. Theorem 1.1 gives a new multiplicity result even under the assumptions of [19].

As for the initial condition and the solution, I have claimed in Theorem 1.1 that for smooth initial values, there are solutions u in $W^{1,\infty}(Q_T)$ hence are continuous. Some people might argue that both initial value u_0 and the solutions must be full of jump discontinuities. My answer to that is that before this work the only known existence result Perona-Malik model (1.1) is for initial values with small derivative [19]. Also it is necessary to show that there are solutions in the first place under reasonable mathematical assumptions before an attempt to generalize the existence theory to 'bad' initial data and to 'bad' solutions.

Similarly, for the homogeneous Dirichlet problem (1.2), we have

Corollary 1.2 *Suppose* $\sigma(s) = s/(1+s^2)$, $s \in \mathbb{R}$. *Let* $v_0 \in C^{2,\alpha}([0,l])$ *with* $v_0(0) = v_0(l) = 0$, *and* $\max_{[0,l]} |v_0(x)| \neq 0$. *Then the Dirichlet problem*

$$\begin{cases} v_t + \sigma(v)_{xx} = 0, & (t,x) \in Q_T, \\ v(0,x) = v_0(x), & x \in [0,l], \\ v(t,0) = v(t,l) = 0, & t \ge 0, \end{cases} \tag{1.8}$$

has infinitely many solutions in $L^\infty(Q_T)$ *in the sense that*

$$\int_{Q_T} [v\psi_t - \sigma(v)\psi_{xx}]\,dx\,dt = 0,$$

for every $\psi \in C_0^{1,2}(Q_T)$ (see notation below). *Furthermore, every weak solution* v *is smooth* (C^2) *near the boundary* $\{0, 1\} \times [0, T)$, *satisfying Eq. (1.8) and the corresponding boundary conditions.*

Our approach to problem (1.3) (or (1.4)–(1.7)) is based on a completely different idea from that of [15]. Instead, we rephrase the weak form (1.4) of equation (1.3) into a partial differential inclusion problem as we did earlier in [31] for Young measure solutions. We write $D = (\partial_t, \partial_x)$ as the gradient in $\mathbb{R}^2 = \{(t, x),\ t,\ x \in \mathbb{R}\}$. Our question can be stated alternatively as:

Find $\Psi \in W^{1,\infty}(Q_T, \mathbb{R}^2)$ with $\Psi(t, x) = (\psi(t, x),\ u(t, x))$ such that

$$D\Psi(t, x) \in K(u(t, x)), \quad \text{a.e. } (t, x) \in Q_T, \tag{1.9}$$

with $K(u(t, x))$ an appropriate subset of

$$\Gamma(u(t, x)) := \left\{ \begin{pmatrix} \sigma(X) & u(t, x) \\ Y & X \end{pmatrix} \in M^{2\times 2}, \quad X,\ Y \in \mathbb{R} \right\}.$$

We can easily see this. Suppose $\Psi = (\psi, u)$ is a $W^{1,\infty}$ solution of the differential inclusion problem (1.9) above, we obtain $\psi_t = \sigma(u_x)$, $\psi_x = u$. Since $\operatorname{curl} D\psi = 0$ in the sense of distributions, we have $\sigma(u_x)_x - u_t = 0$ in the sense of distributions. Thus u is a $W^{1,\infty}$ weak solution for (1.3).

Let $M^{N\times n}$ be the space of $N \times n$ real matrices. The study of systems of homogeneous partial differential inclusion problem under affine Dirichlet boundary condition

$$Df(x) \in K \subset M^{N\times n} \quad \text{a.e. } x \in \Omega \subset \mathbb{R}^n \quad u(x) = Ax \quad x \in \partial\Omega$$

and its inhomogeneous counterpart $Df(x) \in K(x, f(x))$ has been a very active area of research [9–11, 18, 23–27, 30]. The original problem was motivated from the variational approach to material microstructure using nonlinear elasticity models [6, 7]. A particular problem that is directly connected to homogeneous partial differential inclusions is the so-called attainment problem for the double-well model in two dimension [23]. For classical ordinary differential inclusions ($n = 1$), we refer to [2]. For systems of partial differential inclusions ($n,\ N > 1$), there are two main approaches. One approach [23–27, 30] uses the idea of convex integrals introduced by Gromov [14]. The other applies the Baire category theory [9–11]. B. Kirchheim [18] attempted to unify the two approaches. As shown in [18] (also see [27] for the in-approximation approaches), both approaches require the so-called Relaxation Property or Reduction Property. We use the homogeneous case as an example to explain what is required.

Let $K,\ E \subset M^{N\times n}$ be two bounded subsets with K compact. We say that E has the *relaxation property* with respect to K (or E can be reduced to K) if for every bounded open set $\Omega \subset \mathbb{R}^n$, any affine function $u_A(x) = Ax + b$ with $A \in E$, and any $\epsilon > 0$, there is a piecewise affine function $u_\epsilon \in W^{1,\infty}(\Omega, \mathbb{R}^N)$ (or piecewise C^1 function in the inhomogeneous case [9–11]) such that $u_\epsilon = u_A$ on $\partial\Omega$, $Du_\epsilon \in K \cup E$ a.e. in Ω and $\int_\Omega \operatorname{dist}(Du_\epsilon(x), K)dx < \epsilon$.

Note that in the above definition, piecewise affine (piecewise C^1 respectively) in $\bar{\Omega}$ means that the function u_ϵ is allowed to be affine (C^1) on countably many 'pieces' with complement of their union in $\bar{\Omega}$ to be of measure zero. For the inhomogeneous problems, so far for the issue of general existence, it is assumed that the corresponding set E above must be either open in the Baire-category approach [10, 11], or in the convex-integral approach, certain continuity with respect to the parameters is required for E [27] and the piecewise affine boundary values are assumed.

In our case, we cannot apply any of the general theorems [10, 11, 18, 27] directly. However, we may view the problem as a differential inclusion problem with constraints which is somehow related to that in [25] where the constraint is on the Jacobian: $\det Df(x) = 1$. If we let $V \subset M^{2\times 2}$ be the subspace of lower triangular matrices and P_V the orthogonal projection to V, then, we may rewrite our differential inclusion problem as

$$P_V(D\Psi(t,x)) \in K(0), \quad \text{a.e. in } Q_T \text{ subject to } \psi_x = u \text{ in } Q_T.$$

We will take the in-approximation like approach [23–27] streamlined by B. Kirchheim [18]. In order to apply the method, we need to establish the relaxation property (Lemma 3.2 below). In Sect. 2, we give some preliminary results which are needed for establishing Theorem 1.1. We prove Theorem 1.1 in Sect. 3 accepting the technical result Lemma 3.1 and Lemma 3.2 first. We then establish Lemma 3.2 - the relaxation property, followed by some comments on the long time behaviour of the solutions. The last section is devoted to the proof of Lemma 3.1.

2 Notation and preliminaries

For a measurable set $\Omega \subset \mathbb{R}^n$, we denote by $|\Omega|$ its Lebesgue measure. The norm $|X|$ of a matrix $X \in M^{N\times n}$ - the space of $N \times n$ matrices is identified with its standard Euclidean norm in $\mathbb{R}^{Nn}$. Let $Q_T = (0, T) \times (0, l)$ with $T > 0$, $l > 0$. We denote by $C^{k+\alpha/2,2k+\alpha}(\bar{Q}_T)$ the parabolic Hölder space on $\bar{Q}_T$ - the closure of Q_T, where $k \geq 0$ is an integer and $0 \leq \alpha < 1$ [13, 20, 21]. We write, for a function u defined on Q_T, the partial derivatives as u_t, u_{xx} etc. However, for an interval or a finite union of intervals $I \subset \mathbb{R}$, we write $C^{k,1}(I)$ as C^k functions on I whose k-th order derivatives are Lipschitz functions on I. The Sobolev space $W^{1,\infty}(Q_T)$ is defined as usual [1].

Let $K \subset \mathbb{R}^n$ be bounded and let $\text{dist}(X, K)$ be the Euclidean distance function from a point $X \in \mathbb{R}^n$ to K, given by $\text{dist}(X, K) = \min\{|X - Q|,\ Q \in K\}$.

The following one-dimensional version of the theorem concerning the existence, uniqueness and regularity result for one-dimensional parabolic equations is well known [13, 20, 21].

Lemma 2.1 *Suppose $\sigma^* \in C^{2,1}(\mathbb{R})$ satisfies $0 < \lambda \leq (\sigma^*)'(s) \leq \Lambda$ for some constants $0 < \lambda < \Lambda$. Let $u_0 \in C^{3,\alpha}[0, l]$, $(0 < \alpha < 1)$ be such that $(u_0)_x(0) = (u_0)_x(l) = 0$. Then the problem*

$$\begin{cases} u_t - \sigma^*(u_x)_x = 0, & (t,x) \in Q_T, \\ u(0,x) = u_0(x), & x \in [0, l], \\ u_x(t,0) = u_x(t,l) = 0, & t \geq 0. \end{cases} \tag{2.1}$$

has a unique solution $u^* \in C^{1+\alpha/2,2+\alpha}(\bar{Q}_T)$ *satisfying*

$$\|Du^*\|_{C^0(\bar{Q}_T)} \le C\|(u_0)_x\|_{C^0[0,l]}, \qquad \|u^*\|_{C^0(\bar{Q}_T)} \le C_T. \tag{2.2}$$

In particular, by the maximum principle,

$$\max_{(t,x)\in\bar{Q}_T} |u^*_x(t,x)| = \max_{0\le x\le l} |(u_0)_x(x)|.$$

The following construction of two simple piecewise affine functions is crucial for the proof of our main result.

Given $a > 0,\ b > 0, \delta > 0$, we define in the triangular domain

$$\Delta = \{(t,x) \in \mathbb{R}^2, \quad 0 \le x \le \delta(t+1), \quad -1 \le t \le 0\}$$

two piecewise affine functions.

$$g_+(t,x) = \begin{cases} bx, & 0 \le x \le \dfrac{a\delta(t+1)}{a+b}, & -1 \le t \le 0,\\ a\delta(t+1) - ax, & \delta(t+1) \ge x \ge \dfrac{a\delta(t+1)}{a+b}, & -1 \le t \le 0. \end{cases}$$

$$g_-(t,x) = \begin{cases} -ax, & 0 \le x \le \dfrac{b\delta(t+1)}{a+b}, & -1 \le t \le 0,\\ bx - a\delta(t+1), & \delta(t+1) \ge x \ge \dfrac{b\delta(t+1)}{a+b}, & -1 \le t \le 0. \end{cases}$$

We denote by Θ^+_T and Θ^-_T respectively the union of the finitely many line segments contained in $T(0,1,\delta)$ on which g_+ and g_- are not differentiable. We extend $g+$ and g_- respectively in the x-direction as odd functions: $g_\pm(-x,t) = -g_\pm(x,t)$ then extend the resulting function along the t-direction as even functions in t respectively. Let the diamond-shaped domain thus obtain for $g_\pm$ as $T(0;1,\delta)$ representing the centre at the origin 0, the horizontal length 1 from the centre to the right and the left vertices, and vertical length to the top and bottom vertices. We define $g_\pm = 0$ outside $T(0;1,\delta)$. We also call the length between the top and bottom vertices as the hight of $T(0;1,\delta)$ and we denote it by h. We see that $h = 2\delta$. We parameterize the bottom part of the boundary $\partial T(0,1,\delta)$ by $(t, x_T(t), -1 \le t \le 1$, where

$$x_T(t) = \begin{cases} -\delta(1+t), & -1 \le t \le 0,\\ -\delta(1-t), & 0 \le t \le 1. \end{cases}$$

We call $T(0;1,\delta)$ the *standard tile*. When we need to be more precise about the parameters, we write $g = g_\pm(-a,b,\delta,t,x)$. Fig. 3 below shows the tile $T(0;1,\delta)$ and the domains on which $Dg_\pm(t,x)$ equal a constant. The following are some properties of $g_\pm(t,x)$ whose proofs are easy as g is odd in x and the integrals of $(g_+)_t$ and $(g_-)_t$ against x cancel each other.

Lemma 2.2 *The piecewise affine functions* $g_\pm(t,x)$ *defined above satisfy*

(i) *Along any vertical line across* $T(0, 1, \delta)$*, the integral of* $g_\pm$ *against* x *are both zero:*

$$\int_{-\delta(1-|t|)}^{\delta(1-|t|)} g_\pm(t, x)dx = 0, \quad |t| \le 1,$$

and for $(x, t) \in T(0, 1, \delta) \setminus \Theta_T^\pm$*,*

$$\frac{\partial}{\partial t}\int_{x_T(t)}^{x} g_\pm(t, s)ds = \int_{x_T(t)}^{x} (g_\pm)_t(t, s)ds,$$

due to the fact that $g_\pm(t, x_T(t)) = 0$ *for* $-1 \le t \le 1$*.*

(ii) *The partial derivatives* $(g_+)_t$ *and* $(g_-)_t$ *satisfy*

$$\int_{-\delta(1-|t|)}^{\delta(1-|t|)} (g_+)_t(t, x)dx = \frac{2ab(1-|t|)\delta^2}{a+b},$$
$$\int_{-\delta(|t|+1)}^{\delta(1-|t|)} (g_-)_t(t, x)dx = -\frac{2ab(1-|t|)\delta^2}{a+b},$$

when $-1 < t < 0$ *and*

$$\int_{-\delta(1-|t|)}^{\delta(1-|t|)} (g_+)_t(t, x)dx = -\frac{2ab(1-|t|)\delta^2}{a+b},$$
$$\int_{-\delta(|t|+1)}^{\delta(1-|t|)} (g_-)_t(t, x)dx = \frac{2ab(1-|t|)\delta^2}{a+b},$$

when $0 < t < 1$*.*

(iii) *The gradient* $Dg_\pm(t, x) = ((g_\pm)_t(t, x), (g_\pm)_x(t, x))$ *take values*

$$Dg_+(t, x) \in \{(-a\delta, -a), (a\delta, -a), (0, b)\}$$

$$Dg_-(t, x) \in \{(-b\delta, b), (b\delta, b), (0, -a)\}.$$

(iv) *Furthermore,*

$$|g_\pm(t, x)| \le \frac{(a+b)}{4}\delta. \tag{2.3}$$

Remark 2.3 Later we need to scale and translate the construction of the above standard tile $T(0, 1, \delta)$ to be centred at a general point $p = (t_0, x_0)$ with horizonal width 2μ and height $2\mu\delta$. We denote such a tile by $\mathcal{T}(p, \mu, \mu\delta)$. Similarly by translation and scaling we may define functions by translating and scaling $g_\pm(t, x)$ on $\mathcal{T}(p, \mu, \delta\mu)$. In fact we may define

$$g_\pm^{\mathcal{T}}(t, x) = g_\pm(-a, b, \mu, \mu\delta, p, t, x) = \mu g_\pm\left(\frac{t-t_0}{\mu}, \frac{x-x_0}{\mu}\right).$$

We see that similar properties as Lemma 2.2 hold for $g_\pm^{\mathcal{T}}$ on $\mathcal{T}(p, \mu, \delta\mu)$. Let $(t, x_{\mathcal{T}}(t))$ be the parameterization of the bottom part of $\mathcal{T}(p, \mu, \mu\delta)$ obtained from that of $T(0, 1, \delta)$ and $\Theta_{\mathcal{T}}^\pm$ be the union of finitely many line segments contained in $\mathcal{T}(p, \mu, \mu\delta)$ on which $g_\pm^{\mathcal{T}}(t, x)$ are respectively not differentiable. We have

(a) The diamond shaped tile $\mathcal{T}(p, \mu, \mu\delta)$ is centred at p with width 2μ and height $2\mu\delta$.
(b) The functions $g_{\pm}^{\mathcal{T}}$ is supported on $\mathcal{T}(p, \mu, \mu\delta)$ and along almost every vertical sections of $\mathcal{T}(p, \mu, \mu\delta)$, the integrals of $g_{\pm}^{\mathcal{T}}$ and $g_{+}^{\mathcal{T}} + g_{-}^{\mathcal{T}}$ across $\mathcal{T}(p, \mu, \mu\delta)$ against x equal to zero.
(c) $Dg_{+}^{\mathcal{T}}(t, x) \in \{(-a\delta, -a), (a\delta, -a), (0, b)\}$, $Dg_{-}^{\mathcal{T}}(t, x)) \in \{(-b\delta, b), (b\delta, b), (0, -a)\}$ a.e. in $T(p, \mu, \mu\delta)$ and

$$\left|g_{\pm}^{\mathcal{T}}(t, x)\right| \leq \frac{(a+b)}{4}\mu\delta.$$

(d) For $(t, x) \in \mathcal{T}(p, \mu, \mu\delta) \setminus \Theta_{\mathcal{T}}^{\pm}$,

$$\frac{\partial}{\partial t}\int_{x_{\mathcal{T}}(t)}^{x} g_{\pm}^{\mathcal{T}}(t, s)ds = \int_{x_{\mathcal{T}}(t)}^{x} (g_{\pm}^{\mathcal{T}})_t(t, s)ds.$$

Remark 2.4 The advantages of the constructions of the piecewise affine functions $g_{\pm}(t, x)$ over simpler functions with four affine pieces is that the integral across the tile $T(0, 1, \delta)$ against x-variable is zero, so that in our later proofs, we may localize the approximate solutions to a single tile (see the proof of Lemma 3.2) and 'semi-localize' the partial derivative against t of these approximate solutions.

The following result was established by B. Kirchheim [18, Lemma 3.27].

Lemma 2.4 *Let $\Omega \subset \mathbb{R}^m$ be bounded and open. For a Lipschitz mapping $f : \Omega \to \mathbb{R}^n$ and $k \in \mathbb{N}$, let $r(f, k)$ be the supremum of all $r > 0$ such that there is a compact set $K \subset \Omega$ with $|\Omega \setminus K| < 2^{-k}$ and*

$$|f(x+y) - f(x) - \langle Df(x), y\rangle| \leq \frac{1}{k}|y| \quad \text{if } x \in K \text{ and } |y| \leq kr.$$

By Rademacher's Theorem, $r(f, k) > 0$. Consider a sequence $f_k : \Omega \to \mathbb{R}^n$ of uniformly Lipschitz mappings and suppose $0 < r_k < \min\{1/k^2, r(f_k, k)\}$ for all k. If $f \in \cap_k B_\infty(f_k, r_k)$, then $\lim_{k\to\infty} Df_k(x) \to Df(x)$ for a.e. $x \in \Omega$.

We need the following well-known covering result for a bounded open set of $\mathbb{R}^n$ which is called the Vitali covering principle [27] or simply the exhaustion argument [18].

Lemma 2.5 *Let $U \subset \mathbb{R}^n$ be a bounded open set satisfying the regularity condition $|\partial U| = 0$ and let $V \subset \mathbb{R}^n$ be another bounded open set. Then there is a sequence $(x_i, r_i) \in \mathbb{R}^n \times (0, \infty)$, $i = 1, 2, \ldots$, such that*

(a) $U_i = x_i + r_i U$ is contained in V, where $x_i + r_i U = \{x_i + r_i x,\ x \in U\}$;
(b) $U_i \cap U_j = \emptyset$ if $i \neq j$;
(c) $|V \setminus \bigcup_{i=1} U_i| = 0$.

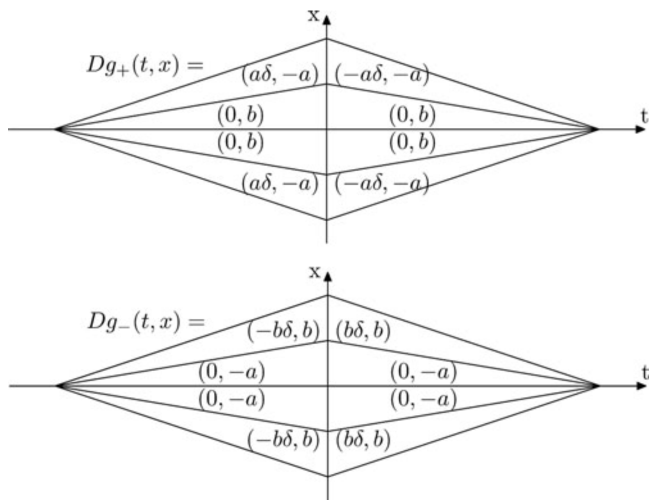

Fig. 4 Tile $T(0; 1, \delta)$ and values of $Dg_+(t, x)$ and $Dg_-(t, x)$ in each affine pieces

3 Proof of Theorem 1.1

In this section we establish our main result Theorem 1.1 and the crucial relaxation lemma (Lemma 3.2). We will leave the proof of the elementary technical Lemma 3.1 to next section.

Proof of Theorem 1.1 As we will convert our original problem (1.3)–(1.7) to an inhomogeneous differential inclusion problem similar to (1.9), we need to define two subsets $K(u(t, x))$ and $E(u(t, x))$ in $M^{2\times 2}$ and establish an relaxation property [10] or reduction property [18]. Then we apply Lemma 2.4 to find a solution.

We define bounded subsets $K(u)$ and $E(u) \subset M^{2\times 2}$ as follows. First note that we may identify diagonal matrices by vectors in $\mathbb{R}^2$ through a simple isometry $(x, y) \to \mathrm{diag}(x, y)$. So we define certain sets in $\mathbb{R}^2$ first, then identify them with 2×2 diagonal matrices. This means that for $\tilde{K} \subset \mathbb{R}^2$, we define

$$K = \left\{ \begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix}, \quad (x, y) \in \tilde{K} \right\}.$$

Given initial value u_0, without loss of generality, we may assume that

$$\max_{x\in[0,l]} |(u_0)_x(x)| = \max_{x\in[0,l]} (u_0)_x(x) := s^* > 0$$

and set $m^* = \min_{0\le x\le l}(u_0)_x(x) \le 0$, then obviously, $-s^* \le -m^*$. Let $0 < y_- < y_+ < \sigma(s^*)$ and let $< x_-^{(1)} < x_-^{(2)} < s^* < x_+(2) < x_+^{(1)}$ be such

$$\sigma\big(x_\pm^{(1)}\big) = y_-, \quad \sigma\big(x_\pm^{(2)}\big) = y_+.$$

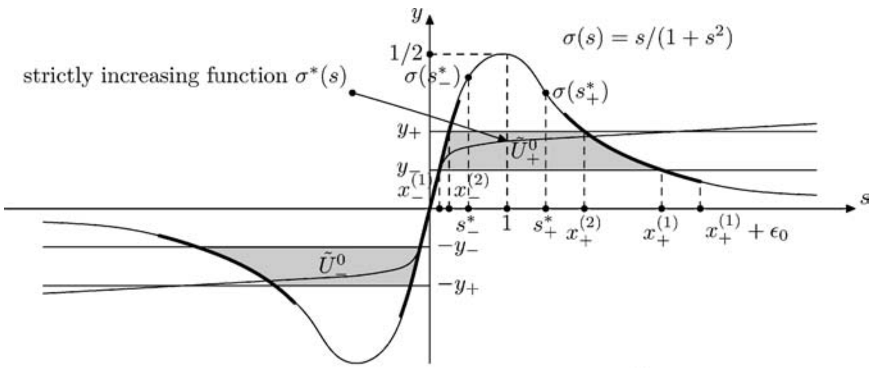

Fig. 5 $\tilde{K}^0$, $\tilde{U}^0$ and the strictly increasing function $\sigma^*(s)$ in Lemma 3.3

Now for some $\epsilon_0 > 0$ sufficiently small, we define (see Fig. 5 above) $\tilde{K}^0 = \tilde{K}_0^0 \cup \tilde{K}_-^0 \cup \tilde{K}_+^0$, with

$$\begin{aligned} \tilde{K}_0^0 &= \{(\sigma(s), s),\ |s| \le x_-^{(2)} + \epsilon_0\}, \\ \tilde{K}_+^0 &= \{(\sigma(s), s),\ x_+^{(2)} \le s \le x_+^{(1)} + \epsilon_0\}, \\ \tilde{K}_-^0 &= \{(\sigma(s), s),\ -x_+^{(2)} \ge s \ge -x_+^{(1)} - \epsilon_0\}. \end{aligned} \tag{3.1}$$

Next we define $\tilde{E}_0 = \tilde{U}_- \cup \tilde{U}_+$ with

$$\begin{aligned} \tilde{U}_+ &= \{(t, s),\ x^{(1)} < s < x_+^{(2)},\ y_- < t < \min\{\sigma(s),\ y_+\}\}, \\ \tilde{U}_- &= -\tilde{U}_+ = \{(t, s),\ -x^{(1)} > s > -x_+^{(2)},\ -y_- > t > \max\{\sigma(s),\ -y_+\}\}. \end{aligned} \tag{3.2}$$

Note on Fig. 5. In Fig. 5 above, $\tilde{K}_0^0$ is the thicker curve in the middle of the picture while $\tilde{K}_-^0$ and $\tilde{K}_+^0$ are the thicker curves on the left and right respectively. $\tilde{U}_+^0$ and $\tilde{U}_-^0$ the two shaded open domains respectively. If $s^* \le 1$, we denote it by $s^* = s_-^*$. If $s^* > 1$, we denote it by $s^* = s_-^*$ in the sketch.

Let K^0, K_0^0, K_+^0, K_-^0, U_-, U_+ and $E_0 = U_- \cup U_+$ be the corresponding sets of 2×2 diagonal matrices. We also define intervals

$$I(t) = \{s \in \mathbb{R},\ (t, s) \in \tilde{E}_0\} := (\alpha(t),\ \beta(t))$$

and for $\delta > 0$, we set

$$I_{-\delta}(t) = (\alpha(t) + \delta,\ \beta(t) - \delta).$$

It is easy to see that both $\alpha(\cdot)$ and $\beta(\cdot)$ are Lipschitz functions with $|\alpha'(t)| \le M$, $|\beta'(t)| \le M$ a.e. for some absolute constant $M > 0$. Let $e_{ij} \in M^{2\times 2}$ be the matrix with (i, j)-entry 1 and other entries zero. Now we define, for fixed $u \in \mathbb{R}$ and $m > 0$

$$\begin{aligned}
\left(K_0^0\right)_m(0) &= \left\{ \begin{pmatrix} s & u \\ r & t \end{pmatrix}, -2m \le r \le 2m, \ (s,t) \in \tilde{K}_0^0 \right\}, \\
(K_+^0)_m(0) &= \left\{ \begin{pmatrix} s & u \\ r & t \end{pmatrix}, -2m \le r \le 2m, \ (s,t) \in \tilde{K}_+^0 \right\}, \\
(K_-^0)_m(0) &= \left\{ \begin{pmatrix} s & u \\ r & t \end{pmatrix}, -2m \le r \le 2m, \ (s,t) \in \tilde{K}_-^0 \right\}, \\
K_m(0) &= \left(K_0^0\right)_m(0) \cup (K_+^0)_m(0) \cup (K_-^0)_m(0), \quad K_m(u) = K_m(0) + ue_{12}, \\
E_m(u) &= \left\{ \begin{pmatrix} s & u \\ r & t \end{pmatrix}, -m \le r \le m, \ (s,t) \in U_0 \right\}.
\end{aligned} \tag{3.3}$$

We see that

$$K_m(u) = \left(K_0^0\right)_m(u) \cup (K_+^0)_m(u) \cup (K_-^0)_m(u) = K_m(0) + ue_{12}$$

with $K_m(0) \subset V$ compact and $E_m(u) = E_m(0) + ue_{12}$ with $E_m(0) \subset V$ open in V. From now on we take $m = \|u_t^*\|_{C^0(\bar{Q}_T)} + 1$ where u^* is given by Lemma 2.1 and restrict u to the interval $[-a, a]$ with $a = \|u^*\|_{C^0(\bar{Q}_T)} + 1$.

Later, we need to consider two pairs of points $y_- < y_+$ and $y'_- < y_+\prime$, we denote by $(x_+^{(1)})'$, $(\tilde{K}^0)'$, $(K^0)'$ etc. for the corresponding parameters and sets for the pair $y'_- < y_+\prime$.

Now we define a strictly increasing function $\sigma^* : \mathbb{R} \to \mathbb{R}$ by the following (see Fig. 5 above)

Lemma 3.1 *Given $x_-^{(i)}, x_+^{(i)}$, $i = 1, 2$, y_-, y_+ as above, there is a strictly increasing function $\sigma^* : \mathbb{R} \to \mathbb{R}$ satisfying*

(i) $\sigma^*(s) = \sigma(s)$ *for* $|s| \le x_-^{(1)}$ *and*

$$y_- < \sigma^*(s^*) < y_+, \quad -y_- > \sigma^*(-s^*) > -y_+.$$

(ii) there exists $c > 0$ depending on $x_-^{(i)}$, $x_+^{(i)}$, $i = 1, 2$, y_-, y_+, and $\sigma'(\cdot)$, such that $(\sigma^*)'(s) > c$ for $s \in \mathbb{R}$;

(iii) $\sigma \in C^{2,1}(\mathbb{R})$.

We prove Lemma 3.1 in Sect. 4.

Proof of Theorem 1.1 (continued). By using the above $\sigma^*(\cdot)$ and Lemma 3.1, we see that the corresponding solution u^* of (2.1) satisfies $\|u_x^*\|_{C^0(\bar{Q}_T)} = \|(u_0)_x\|_{C^0[0,l]}$. From now on we take $m = \|(u_0)_x\|_{C^0[0,l]} + 1$ drop the subscript m. We restrict u to the interval $[-a, a]$ with $a = C_T + 1$, where C_T is the bound of the solution u^* given by Lemma 2.1.

We observe that for each fixed $u \in \mathbb{R}$, $K(u) = K(0) + e_{12}u$ with $K(0) \subset V$ compact in the subspace of lower-triangular matrices

$$V = \left\{ \begin{pmatrix} s & 0 \\ r & t \end{pmatrix}, \ s, r, t \in \mathbb{R} \right\} \subset M^{2\times 2}.$$

and $E(u) = E(0) + e_{12}u$ with $E(0)$ open in V. We denote by P_V the orthogonal projection from $M^{2\times 2}$ to V.

Remark 3.2 Let $X \in M^{2\times 2}$. We denote by $\mathrm{diag}(X)$ the 2×2 diagonal matrix consists of the diagonal entries of X. Given $X \in E(0) \cup K(0) \subset V$, we have

$$\begin{aligned}\mathrm{dist}(X,\ K(0)) &= \min_{Y\in K_+^0\cup K_-^0,\ -2m\le z\le 2m} (|\,\mathrm{diag}(X) - Y|^2 + |x_{21} - z|^2)^{1/2}\\ &= \min_{Y\in K_+^0\cup K_-^0} |\,\mathrm{diag}(X) - Y| = \mathrm{dist}(\mathrm{diag}(X),\ K_+^0 \cup K_-^0).\end{aligned}$$

Now since u_x^* is uniformly continuous in $\bar{Q}_T$ and $u_x^*(t,0) = u_x^*(t,l) = 0$ $(0 \le t \le T)$, there is some $\delta > 0$, such that $|u_x^*(t,x)| \le x_-^{(1)}/2$ whenever $(t,x) \in Q_T^\delta := ([0,T] \times [0,\delta]) \cup ([0,T] \times [l-\delta, l])$, where $x_-^{(1)} < x_-^{(2)} < 1$ are given by Lemma 3.1 (see Fig. 5). Let $Q^* = Q_T \setminus Q_T^\delta$.

Since u^* is smooth, and $\mathrm{curl}(\sigma^*(u_x^*),\ u^*) = 0$ in Q_T, there is some $\psi^* \in C^1(\bar{Q}_T)$ such that $(\sigma^*(u_x^*),\ u^*) = (\psi_t^*,\ \psi_x^*)$ in Q_T. If we let $\Phi^* = (\psi^*,\ u^*)$, we see that

$$\begin{aligned}&D\Phi^*(t,x) \in K(u^*) \cup E(u^*), \quad (t,x) \in Q^*,\\ &D\Phi^*(t,x) \in K(u^*), \quad (t,x) \in Q_T^\delta.\end{aligned}$$

Now we try to solve the following inhomogeneous differential inclusion problem:

$$\text{Find } \Psi = (\psi, u) \in W^{1,\infty}(Q^*, \mathbb{R}^2), \quad \begin{cases} D\Psi(t,x) \in K(u), \quad (t,x) \in Q^* \text{ a.e.}\\ \Psi|_{\partial Q^*} = \Phi^*. \end{cases} \tag{3.4}$$

Suppose Ψ is a solution of (3.4), then we extend Ψ to $\bar{Q}_T$ by Φ^*. Thus $\Psi \in W^{1\infty}(Q_T, \mathbb{R}^2)$ remains a solution of the differential inclusion problem $D\Psi(t,x) \in K(u(t,x))$, $(t,x) \in Q_T$ a.e. hence $D\psi = (\sigma(u_x), u)$. This implies that u is a solution of the equation $u_t - (\sigma(u_x))_x = 0$ in the weak sense. As on Q_T^δ, $u = u^*$, we see that u satisfies the homogeneous Neumann boundary condition $u_x(t,0) = u_x(t,l) = 0$ as well.

We will show that even in the case $0 < s^* \le 1$, solutions we obtained are different from those u_K obtained in [19] in the last paragraph of the proof.

We define a piecewise C^1 function f on an bounded open set $\Omega \subset \mathbb{R}^2$ to be such that there are at most countably many disjoint open triangular-shaped open domains $G_i \subset \Omega$, $f \in C^1(\bar{G}_i)$ is continuous and $|\Omega \setminus \cup_{i=1}^\infty G_i| = 0$. Obviously, in the definition, if G_i is not a triangular shaped domain, we can further divide it into at most countably many triangular-shaped sub-domains G_i^k satisfying $G_i^k \subset G_i$, $|G_i \setminus \cup_{k=1}^\infty G_i^k| = 0$ and $f \in C^1(\bar{G}_i^k)$.

We denote the set of all piecewise C^1 function on Ω as $C_{pw}^1(\Omega)$.

Let

$$\begin{aligned}\mathcal{P} = \Big\{&\Psi = (\psi,\ u) \in C_{pw}^1(Q^*, \mathbb{R}^2),\\ &D\Psi(t,x) \in K(u) \cup E(u),\ \text{a.e. } \Psi|_{\partial Q^*} = \Phi^*\Big\}.\end{aligned}$$

Clearly $\mathcal{P} \neq \emptyset$ as $\Phi^* = (\psi^*, u^*) \in \mathcal{P}$. Let $\bar{\mathcal{P}}^\infty$ be the closure of $\mathcal{P}$ under L^∞ norm. Firstly, we have

Lemma 3.2 *For any $\epsilon > 0$,*

$$\mathcal{P}_\epsilon = \left\{ \Psi \in \mathcal{P}, \quad \int_{Q^*} \mathrm{dist}(D\Psi(t,x),\ K(u))dtdx < \epsilon |Q^*| \right\}$$

is dense in $\mathcal{P}$ under the L^∞ norm.

We will establish Lemma 3.2 after the proof of Theorem 1.1. Accepting the conclusion of Lemma 3.2 for the moment, we can prove Theorem 1.1 following roughly the general approach of B. Kirchheim [18, Theorem 3.28].

Proof of Theorem 1.1 (continued) Suppose we are given any ball $B_\infty(\Psi_{2k-1}, r_{2k-1}) \subset \bar{\mathcal{P}}^\infty$ with $k \geq 1$. Since the ball intersects the set $\mathcal{P}$ itself, we may use Lemma 3.2 to find

$$\Psi_{2k} \in B_\infty(\Psi_{2k-1},\ r_{2k-1}/2) \cap \mathcal{P} \quad \text{with} \quad \Psi_{2k} = (\psi_{2k}, u_{2k}) \quad \text{satisfying}$$
$$\int_{Q^*} \mathrm{dist}(D\Psi_{2k}(t,x),\ K(u_{2k}(t,x))dt\,dx < \frac{1}{2^k}.$$

By Lemma 2.4, we take $R_{2k} = r(\Psi_{2k}, k)$. We then take our new radius as

$$r_{2k} = \min\left\{ R_{2k},\ r_{2k-1}/3,\ \frac{1}{k^2} \right\}.$$

Now we consider $B_\infty(\Psi_{2k}, r_{2k})$ which is included in the ball we were given. Since $r_{2k} \to 0_+$ as $k \to \infty$, we see that (Ψ_{2k}) is a Cauchy sequence in $\bar{\mathcal{P}}^\infty$. Let $\Psi = \lim_{k\to\infty} \Psi_{2k}$. Write $\Psi = (\psi, u)$. By Lemma 2.4,

$$\lim_{k\to\infty} D\Psi_{2k}(t,x) = D\Psi(t,x), \quad \text{a.e.}$$

Next we show that $D\Psi(t,x) \in K(u(t,x))$ a.e. in Q^*. We have, by our choice of Ψ_{2k} that $\Psi_{2k} \in \mathcal{P}$,

$$\int_{Q^*} \mathrm{dist}(D\Psi_{2k}(t,x),\ K(u_{2k}(t,x))))dt\,dx < \frac{1}{2^k}.$$

We also have

$$\begin{aligned} \mathrm{dist}(D\Psi_{2k}(t,x),\ K(u_{2k}(t,x))) &= \mathrm{dist}\left(\begin{pmatrix} (\psi_{2k})_t & 0 \\ 0 & (u_{2k})_x \end{pmatrix},\ K^0_+ \cup K^0_- \right) \\ &= \mathrm{dist}[((\psi_{2k})_t,\ (u_{2k})_x), \tilde{K}^0]. \end{aligned}$$

Since $\tilde{K}^0 \subset \mathbb{R}^2$ is compact, the distance function $X \to \mathrm{dist}(X, \tilde{K}^0)$ is Lipschitz, hence continuous, we have, by passing to the limit $k \to \infty$ that

$$\begin{aligned} &\lim_{k\to\infty} \int_{Q^*} \mathrm{dist}(D\Psi_{2k}(t,x),\ K(u_{2k}(t,x)))dt\,dx \\ &= \int_{Q^*} \mathrm{dist}[((\psi_{2k})_t,\ (u_{2k})_x), \tilde{K}^0]dt\,dx \\ &= \int_{Q^*} \mathrm{dist}[(\psi_t,\ u_x), \tilde{K}^0]dt\,dx = 0. \end{aligned}$$

Since $(\psi_{2k})_x = u_{2k}$ a.e., we also see that, as $k \to \infty$, we obtain $\psi_x = u$ a.e. Consequently,

$$0 = \int_{Q^*} \mathrm{dist}[(\psi_t,\ u_x), \tilde{K}^0]dt\,dx = \int_{Q^*} \mathrm{dist}(D\Psi(t,x),\ K(u(t,x)))dt\,dx.$$

Thus $Psi \in \bar{\mathcal{P}}_\infty$ satisfies

$$D\Psi \in K(u), \quad \text{a.e. in } Q^*, \tag{3.5}$$

and Ψ is a solution of (3.4).

Clearly, if we extend the solution $\Psi = (\psi, u)$ of (3.4) to Q_T by $\Phi^* = (\psi^*, u^*)$ outside Q^*, then the extended mapping Ψ is still Lipschitz and satisfies (3.5) a.e. in Q_T. Hence the extended u is a weak solution of the Neumann problem (1.4)–(1.7) in $W^{1,\infty}(Q_T)$.

Now we give a geometric proof that the Neumann problem (1.4)–(1.7) has infinitely many solutions in $W^{1,\infty}(Q_T)$. An alternative proof is easily obtained following the proof of Lemma 3.2 and Remark 3.4

Let $0 < y_- < y_+ < y'_- < y'_+ < \sigma(s^*)$. Let $K(0)$ and $K'(0)$ be the compact set $K(0)$ defined early corresponding to the pairs (y_-, y_+) and (y'_-, y'_+) respectively (see (3.1), (3.2), (3.3) for notation). We let $K(u) = K(0) + ue_{12}$, $K'(u) = K'(0) + ue_{12}$ and require the corresponding ϵ_0 and ϵ'_0 in the definition of $K'(0)$ and $K'(0)$ to be sufficiently small so that

$$(\tilde{K}^0_+ \cup \tilde{K}^0_-) \cap ((\tilde{K}')^0_+ \cup (\tilde{K}')^0_-) = \emptyset, \qquad \tilde{K}^0_0 \subset (\tilde{K}')^0_0.$$

Now we have two solutions Ψ_1 and Ψ_2 such that

$$D\Psi_1(t,x) \in K(u_1(t,x)),\ \ (t,x) \in Q^*,$$
$$D\Psi_2(t,x) \in K'(u_2(t,x)),\ \ (t,x) \in (Q')^*.$$

By extending Ψ_1 and Ψ_2 to be defined in Q_T as we did earlier we obtain two solutions of (1.4)–(1.7). Now we show that these two solutions cannot be the same. Otherwise, $\Psi_1 = \Psi_2 := \Psi$ in Q_T so that $u_1 = u_2 := u$. From our construction of $K(u)$ and $K'(u)$ we see that

$$D\Psi(t,x) \in K(u(t,x)) \cap K'(u(t,x)), \quad (t,x) \in Q^*,$$

hence

$$(\sigma(u_x(t,x)),\ u_x(t,x)) \in \tilde{K}^0 \cap (\tilde{K}')^0 = \tilde{K}^0_0.$$

This implies that

$$|u_x(t,x)| \le x_-^{(2)} + \epsilon_0, \tag{3.6}$$

a.e. in $\bar{Q}_T$.

We consider two different cases.

Case (i) $s^* > 1$.

In this case we see that (3.6) implies that

$$|u_x(t,x)| \le x_-^{(2)} + \epsilon_0 < 1 < s^*,$$

a.e. in $\bar{Q}_T$. By Lemma 3.1, we may construct yet another strictly monotone function (still denoted by $\sigma^*(\cdot)$) such that $\sigma^* \in C^{2,1}(\mathbb{R})$, $\sigma(s) = \sigma^*(s)$ if $|s| \leq x_-^{(2)} + \epsilon_0$, $(\sigma^*)'(s) \geq c_0 > 0$ for $s \in \mathbb{R}$ and $\sigma^*(s)$ is affine when $|s|$ is large.

Thus u is a weak solution of (2.1) given by Lemma 2.1, hence $u \in C^1(\bar{Q}_T)$. Now since u satisfies the initial condition $u(0,x) = u_0(x)$ so that $u_x(0,x) = (u_0)_x(x)$ in $(0,\ l)$. By our assumption for Case (i),

$$\max_{[0,\ l]} |(u_0)_x(x)| = \max_{[0,\ l]} (u_0)_x(x) = s^* > 1,$$

and we may find some $x_0 \in (0,\ l)$, $u_x(0,x_0) = (u_0)_x(x_0) = s^* > 1$. Now we see that for small $t > 0$, $u_x(t,x_0) > 1$. This contradicts to (3.6). Thus $\Psi_1 \neq \Psi_2$.

Case (ii) $s^* \leq 1$.

In this case we see that (3.6) implies that

$$|u_x(t,x)| \leq x_-^{(2)} + \epsilon_0 < s^* \leq 1,$$

a.e. in $\bar{Q}_T$. Again by Lemma 3.1, we may construct yet another strictly monotone function (still denoted by $\sigma^*(\cdot)$) such that $\sigma^* \in C^{2,1}(\mathbb{R})$, $\sigma(s) = \sigma^*(s)$ if $|s| \leq x_-^{(2)} + \epsilon_0$, $(\sigma^*)'(s) \geq c_0 > 0$ for $s \in \mathbb{R}$ and $\sigma^*(s)$ is affine when $|s|$ is large. Again u is a weak solution of (2.1) given by Lemma 2.1, satisfying the initial condition, hence $u \in C^1(\bar{Q}_T)$.

As $\sigma^*(x_-^{(2)} + \epsilon_0) < \sigma^*(s^*)$ and at the maximal point $(0,x_0)$, $u_x(0,x_0) = s^*$, for small $t > 0$, we see that

$$\sigma^*(u_x(t,x_0)) > \sigma^*\left(x_-^{(2)} + \epsilon_0\right) \geq \sigma^*(u_x(t,x))$$

for all $(t,x) \in Q_T$. This is a contradiction.

Now we show that in the interval $[0,\ \sigma(s^*)]$ we can easily find infinitely many pairs of disjoint intervals $\{[y_-(i),\ y_+(i)] \cup [-y_+(i),\ -y_-(i)]\}_{i=1}^{\infty}$ with $y_+(i) < y_-(i+1)$, hence the Neumann problem (1.4)–(1.7) has infinitely many solutions in $W^{1,\infty}(Q_T)$.

Finally we show that in case that $0 < s^* \leq 1$, our solutions above are different from the solution u_K obtained in [19]. For any solution u obtained above, we always have $D\Psi(t,x) \in K(u(t,x))$. Let $y_\pm$ and $x_\pm^{(i)}$ $(i = 1, 2)$ are the points in the definition of $K(0)$, we see that our solution satisfies either $u_x(t,x)) \leq x_-^{(2)} + \epsilon_0 < s^* \leq 1$ or $u_x(t,x)) \geq x_+^{(2)} - \epsilon_0 > 1 \geq s^*$ a.e. in Q_T (see Fig. 5, where our s^* in the present case is represented by s_-^*).

In [19] the monotone function $\sigma^{**}(\cdot)$ modified from $\sigma(\cdot)$ agrees with $\sigma(\cdot)$ in the interval $[-s^*,\ s^*]$. Let u_K be the unique smooth solution obtained in [19]. By the maximum principle, $(u_K)_x(t,x)$ reaches its maximum value $s^* > 0$ at some point $(0,x_0)$, so that $(u_K)_x(0,x_0) = (u_0)_x(x_0) = s^*$, by continuity, we see that for small $\tau > 0$, on $Q_\tau := (0,\delta] \times [x_0 - \delta,\ x_0 + \delta] \subset Q_T$, we may have

$$x_-^{(2)} + \epsilon_0 < (u_K)_x(t,x)) < x_+^{(2)} - \epsilon_0.$$

Thus $(u_K)_x(t,x) \neq u_x(t,x)$ for $(t,x) \in Q_\tau$. Geometrically, in Fig. 5, we have that the value $(\sigma(u_x), u_x))$ of our solution u stays in the set $\tilde{K}^0$ which consists of three pieces of thick curves. The value of u_x is given by the projection of the thick curves to the s-axis. The image of this projection has a positive distance to our s^* which is s^*_- in Fig. 5. As $s^* = s^*_-$ is reached by $(u_K)_x$ at some initial point $(0, x_0)$, by continuity of $(u_K)_x$, we see that in a neighbourhood of $(0, x_0)$, the values of $(u_K)_x$ is close to s^*_- hence in this small neighbourhood, $(u_K)_x \neq u_x$. The last claim in Theorem 1.1 is proved. □

Proof of Lemma 3.2 Before we proceed, let us notice that for any $\Psi \in \mathcal{P}$, we have, as $\Psi = (\psi, u)$ satisfies $\psi_x = u$ a.e. in Q^*,

$$\operatorname{dist}(D\Psi(t,x), K(u(t,x))) = \operatorname{dist}(P_V(D\Psi(t,x)),\ K(0)).$$

Also if we let $\partial|_V E(u(t,x)) = u(t,x)e_{12} + \partial|_V E(0)$, where $\partial|_V E(0)$ is the boundary of $E(0)$ in V, we have

$$\begin{aligned}&\operatorname{dist}[D\Psi(t,x),\ K(u(t,x)) \cup \partial|_V E(u(t,x))]\\ &\quad = \operatorname{dist}[P_V(D\Psi(t,x)),\ K(0) \cup \partial_V E(0)].\end{aligned}$$

These simple observations help us to simplify $K(u)$ and $E(u)$. So under the constraint $\psi_x = u$, we only need to consider a 'homogeneous' problem $P_V(D\Psi(t,x)) \in K(0)$. However, the difficulty remains that in any modification of Ψ, the constraint $\psi_x = u$ must be kept.

Given $\Psi \in \mathcal{P}$ and let $0 < \eta < 1$, we need to find some $\Psi_\eta \in \mathcal{P}_\epsilon$ such that $\|\Psi_\eta - \Psi\|_\infty < \eta$.

As $D\Psi(t,x) \in K(u(t,x)) \cup E(u(t,x))$ a.e. in Q^*, for each open set G_i on which $\Psi \in C^1(\bar{G}_i)$ we may find $\delta_i > 0$ such that the closed set

$$K_i = \{(t,x) \in \bar{G}_i,\ \operatorname{dist}[D\Psi(t,x),\ K(u) \cup \partial|_V E(u(t,x))] \le \delta_i\}, \quad (\delta_i < 1)$$

satisfies

$$\int_{K_i} \operatorname{dist}(D\Psi(t,x),\ K(u(t,x))dtdx < \frac{\epsilon}{2^{i+2}}|Q^*|.$$

This is due to the fact that near $K(u(t,x)$, the distance itself is small while near

$$\partial|_V E(u(t,x)) \setminus K(u(t,x)) = u(t,x)e_{12} + \partial|_V E(0) \setminus K(0),$$

the integral is small because $D\psi(t,x) \in K(u(t,x) \cup E(u(t,x)$ a.e. in Q^*.

Furthermore we may require that the boundary of $\hat{G}_i := G_i \setminus K_i$ has measure zero. This can be easily achieved as the function

$$\operatorname{dist}(D\Psi(t,x),\ K(u(t,x)) = \operatorname{dist}(P_V(D\Psi(t,x)),\ K(0))$$

is continuous on $\bar{G}_i$, the set on which the function equal to a constant $c \ge 0$ must be of measure zero except on possibly countably many c's.

Since K_i is closed, $\hat{G}^i \setminus K_i \subset G_i$ is open and on each $\tilde{G}^i$,

$$\operatorname{dist}(D\Psi(t,x),\ K(u) \cup \partial|_V E(u(t,x)) > \delta_i.$$

As we need to defined piecewise C^1 functions, we consider

$$\hat{K}_i = \{(t,x) \in \bar{G}_i,\ \operatorname{dist}[D\Psi(t,x),\ K(u) \cup \partial|_V E(u(t,x))] < \delta_i\}$$

which is an open set. Notice that

$$\partial \hat{K}_i \cup \partial \hat{G}_i \subset \partial G_i \cup \{(t,x) \in G_i, \quad \text{dist}[P_V(D\Psi(t,x)),\ K(0) \cup \partial|_V E(0)] = \delta_i,$$

so we define $\Psi_\eta = \Psi$ on $\cup_i \hat{K}_i$ as implicitly we have decomposed $\cup_i \hat{K}_i$ into countably many triangular shaped domains.

Now we cover each $\hat{G}_i$, by at most countably many squares $\{D_i^k\}_{k=1}^\infty$ whose sides parallel to the coordinate axes with disjoint interiors. Let $p_i^k \in D_i^k$ be the centre of D_i^k. By continuity of $D\Psi$ on $\bar{G}_i$, there is $\eta_i > 0$, such that $|D\Psi(t,x) - D\Psi(t',x')| < \rho\delta_i$, if $(t,x),\ (t',x') \in \bar{G}_i$, $|(t,x) - (t',x')| < \eta_i$, where $\rho > 0$ is to be determined (see (3.10) below). We further divide each D_i^k if the side length $l_i^k \geq \eta_i$. Thus we may assume that all D_i^k satisfy $l_i^k < \eta_i$. Consequently, on each D_i^k, $|D\Psi(t,x) - D\Psi(p_i^k)| < \rho\delta_i$.

Note that

$$\text{dist}\big[D\psi\big(p_i^k\big), K\big(u\big(p_i^k\big)\big) \cup \partial_V E\big(u\big(p_i^k\big)\big)\big] > \delta_i,$$

so that $D\psi(p_i^k) \in E(u_x(p_i^k)$ which implies $(\psi_t(p_i^k),\ u_x(p_i^k)) \in \tilde{U}_-^0 \cup \tilde{U}_+^0$. Without loss of generality, we may assume that $u_x(p_i^k) > 0$ so that $(\psi_t(p_i^k),\ u_x(p_i^k)) \in \tilde{U}_+^0$. Therefore

$$\psi_t\big(p_i^k\big) \in I_{-\delta_i}\big(u_x\big(p_i^k\big) = \big(\alpha\big(u_x\big(p_i^k\big) + \delta_i,\ \beta\big(u_x\big(p_i^k\big) - \delta_i\big).$$

Now we define $a_i^k > 0$, $b_i^k > 0$ be such that

$$\text{dist}\left(\begin{pmatrix} \psi_t(p_i^k) & 0 \\ 0 & u_x(p_i^k) - a_i^k \end{pmatrix}, K_0^0\right) = \frac{\delta_i}{2}$$
$$\text{dist}\left(\begin{pmatrix} \psi_t(p_i^k) & 0 \\ 0 & u_x(p_i^k) + b_i^k \end{pmatrix}, K_+^0\right) = \frac{\delta_i}{2}.$$

Consequently,

$$\psi_t\big(p_i^k\big) \in I_{-\delta_i/2}\big(u_x\big(p_i^k\big) - a_i^k\big), \quad \psi_t\big(p_i^k\big) \in I_{-\delta_i/2}\big(u_x\big(p_i^k\big) + b_i^k\big). \tag{3.7}$$

If $(\psi_t(p_i^k),\ u_x(p_i^k)) \in \tilde{U}_-^0$, by symmetry, we may define a_i^k, $-b_i^k$ using U_-^0, K_0^0 and K_-^0.

Next we construct ϕ on each D_i^k. We we decompose D_i^k into countably may symmetric tiles $\mathcal{T}_{i,s}^k$ centred at $p_{i,s}^k$ as described in Remark 2.3 immediately after Lemma 2.2 and define functions $g_{i,s,+}^k(t,x)$ supported on $\mathcal{T}_{i,s}^k$ by

$$g_{i,s,+}^k(t,x) := g_+\big(-a_i^k, b_i^k, \xi_{i,s}^k, \xi_{i,s}^k\delta_i^*, p_{i,s}^k, t, x\big)$$

with $\bar{\mathcal{T}}_{i,s}^k = \overline{\mathcal{T}(p_{i,s}^k, \xi_{i,s}^k,\ \xi_{i,s}^k\rho\delta_i)}$ as its support, where $\xi_{i,s}^k > 0$ is the scaling factor for g_+ as Remark 2.3 with $\rho > 0$ to be defined. Note that $\max_{i,k}\{a_i^k,\ b_k^i\} \leq x_+^{(1)}$, we have

$$\big|\big(g_{i,s,+}^k\big)_t(t,x)\big| \leq \rho\delta_i x_+^{(1)}, \quad (t,x) \in \mathcal{T}_{i,s}^k \setminus \Theta_{\mathcal{T}_{i,s}^k}^+.$$

Now we define $\phi_i^k(t,x) = \sum_{s=1}^{\infty} g_{i,s,+}^k(t,x)$. Let $\phi = \sum_{i,k} \phi_i^k$, and define

$$\psi_0(t,x) = \int_{\delta_0}^{x} \phi(t,\tau)\,d\tau, \quad x \in [\delta_0,\ l-\delta_0], \quad t \in [0,\ T].$$

Next we show that $\Psi_\eta = (\psi_\eta,\ u_\eta)$ with $\psi_\eta = \psi + \psi_0$, $u_\eta = u + \phi$ satisfies $\Psi_\eta \in P_\epsilon$ and $\|\Psi - \Psi_\eta\|_{L^\infty} < \eta$ if we further require that the largest vertical dimension of $T_{i,s}^k$ to be small enough.

We need to prove that

(i) $\Psi_\eta = \Phi^*$ on ∂Q^*;
(ii) $\Psi_\eta \in C_{pw}^1(Q^*)$;
(iii) $(\psi_\eta)_x = u_\eta$ a.e. in Q^*, that is, $\psi_x(t,x) + (\psi_0)_x(t,x) = u(t,x) + \phi(t,x)$ for a.e.$(t,x) \in Q^*$;
(iv) $\psi_t(t,x) + (\psi_0)_t(t,x) \in I(u_x(t,x) + \phi_x(t,x))$.
(v) $|u_t + \phi_t| < m$ so that together with (iv), one has $D\Psi_\eta \in K(u_\eta) \cup E(u_\eta)$;
(vi) $\int_{Q^*} \mathrm{dist}(D\Psi_\eta, K(u_\eta))dtdx < \epsilon|Q^*|$.

Assertion (i) is easy to prove. Obviously $\phi = 0$ on ∂Q^* by construction. Since

$$\psi_0(t,x) = \int_{\delta_0}^{x} \phi(t,s)ds, \tag{3.8}$$

we have $\psi_0(t,\delta_0) = 0$. For each fixed $t \in (0,T)$, the set $\{x \in [\delta_0, l - \delta_0], \phi(t,x) \neq 0\}$ is a countable union of open intervals. Each of such intervals is the intersection of the vertical line $l_t := \{(t,x),\ \delta_0 < x < l - \delta_0\}$ and some $\mathcal{T}_{i,s}^k$. The integral of $\phi(t,\cdot)$ over such an interval against x is zero, hence $\psi_0(t, l-\delta_0) = \int_{\delta_0}^{l-\delta_0} \phi(t,s)ds = 0$. By our construction of ϕ we see that $\psi_0(0,x) = \psi_0(T,x) = 0$ for $x \in [\delta_0, T-\delta_0]$. Thus $\Psi_0 = \Phi^*$ on ∂Q^*.

Remark 3.3 We observe that the values of the approximation ψ_0 defined by (3.8) is localized by ϕ as it depends only on the values of ϕ in each individual $\mathcal{T}_{i,s}^k$.

Next we prove (ii). Obviously ϕ is piecewise affine on $\cup_{i,k,s} T_{i,s}^k$. To show that $\phi,\ \psi_0 \in C_{pw}^1(\bar{Q}^*)$, we only need to prove that $\psi_0 \in C_{pw}^1(\bar{Q}^*)$ as it is simpler to establish the same property for ϕ. Due to the cancellation property of the integral of $\phi(t,\cdot)$ across each $\mathcal{T}_{i,s}^k$, we see that $\psi_0 \neq 0$ only in $\cup \mathcal{T}_{i,s}^k)$ and $\psi_0 = 0$ on $Q^* \setminus [\cup \mathcal{T}_{i,s}^k)]$. The boundary $\partial(\cup_s \mathcal{T}_{i,s,}^k) \subset \bar{D}_i^k$ is of measure zero. Also on the open set $\cup \hat{K}_i$, $\psi_0 = 0, \phi = 0$ with

$$\left|Q^* \setminus \left[(\cup_i \hat{K}_i) \cup \left(\cup_{i,k,s} T_{i,s}^k\right)\right]\right| = 0.$$

Thus both ψ_0 and ϕ are piecewise C^1 in Q^*.

Item (iii) is easy to prove as $\psi_x(t,x) = u(t,x)$ and $\psi_0(t,x) = \int_{\delta_0}^{x} \phi(t,\tau)d\tau$.

Now we prove (iv).

For each $t \in (0,T)$ the line $l_t := \{(t,x),\ \delta_0 < x < l - \delta_0)$ intersects at most countably many $\mathcal{T}_{i,s}^k$'s. Outside these $T_{i,s}^k$'s, $\phi(t,x) = 0$. The value of $\phi_t(t,x) = (g_{i,s,+}^k)_t(t,x)$ in $\mathcal{T}_{i,s}^k \setminus \Theta_{\mathcal{T}_{i,s}^k}^+$ will be in the set $\{-\rho\delta_i a_k^i,\ \rho\delta_i a_k^i,\ 0\}$ and integrating ϕ across each $\mathcal{T}^k_{i,s}$ against x is zero due to Remark 2.3 (b). We also have, for $(t,x) \in \mathcal{T}^k_{i,s}$,

$$\psi_0(t,x) = \int_{x_{\mathcal{T}^k_{i,s}}(t)}^{x} g^k_{i,s,+}(t,s)ds, \quad \text{so that}$$

$$(\psi_0)_t(t,x) = \int_{x_{\mathcal{T}^k_{i,s}}(t)}^{x} \big(g^k_{i,s,+}\big)_t(t,s)ds, \quad (t,x) \in \mathcal{T}^k_{i,s} \setminus \Theta^+_{\mathcal{T}^k_{i,s}}.$$

Therefore we have the bound

$$|(\psi_0)_t(t,x)| \le \eta^i_k \rho \delta_i x^{(1)}_- \quad \text{a.e. in } \mathcal{T}^k_{i,s}.$$

Since $\max\{a^i_k,\ b^i_k\} \le x^{(1)}_-$ for all $i,\ k$, where η^i_k is the side-length of D^i_k. By further dividing D^i_k if necessary, we may assume that $\eta^i_k < \eta/2 < 1/2$ so that

$$|(\psi_0)_t(t,x)| \le \frac{\eta}{2}\rho\delta_i x^{(1)}_- < \rho\delta_i x^{(1)}_- \quad \text{a.e. in } D^i_k. \tag{3.9}$$

Thus, if $(t,x) \in T^k_{i,s}$ and $\phi_x(t,x) = -a^k_i$,

$$\begin{aligned}\psi_t(t,x) + (\psi_0)_t(t,x) &= \psi_t\big(p^k_i\big) + \big[\psi_t(t,x) - \psi_t\big(p^k_i\big)\big] + (\psi_0)_t(t,x)\\ &\in I_{-\delta_i/2}\big(u_x\big(p^k_i\big) - a^k_i\big) - \big[\psi_t\big(p^k_i\big) - \psi_t(t,x)\big] + (\psi_0)_t(t,x).\end{aligned}$$

Here we have assumed as before that $(\psi_t(p^k_i),\ u_x(p^k_i)) \in \tilde{U}^0_+$ (see Fig. 5). As the symmetrical case $(\psi_t(p^k_i),\ u_x(p^k_i)) \in \tilde{U}^0_-$ can be treated similarly. Therefore

$$\begin{aligned}\psi_t(t,x)+(\psi_0)_t(t,x) &\ge \alpha\big(u_x\big(p^k_i - a^k_i\big) + \frac{\delta_i}{2} - \big[\psi_t\big(p^k_i\big) - \psi_t(t,x)\big] + (\psi_0)_t(t,x)\\ &= \alpha\big(u_x(t,x) - a^k_i\big) + \frac{\delta_i}{2} + \big[\alpha\big(u_x(t,x) - a^k_i\big) - \alpha\big(u_x\big(p^k_i\big) - a^k_i\big)\big]\\ &\quad + \big[\psi_t\big(p^k_i\big) - \psi_t(t,x)\big] + (\psi_0)_t(t,x)\\ &\ge \alpha(u_x(t,x) + \phi_x(t,x)) + \frac{\delta_i}{2} - M\big|u_x(t,x) - u_x\big(p^k_i\big)\big| - \big|\psi_t\big(p^k_i\big)\\ &\quad - \psi_t(t,x)\big| - |(\psi_0)_t(t,x)|\\ \alpha(u_x(t,x) + \phi_x(t,x)) &+ \frac{\delta_i}{2} - M\rho\delta_i - \rho\delta_i - \rho\delta_i x^{(1)}_-\\ &\ge \alpha(u_x(t,x) + \phi_x(t,x)) + \frac{\delta_i}{4},\end{aligned}$$

if we require

$$\rho < \frac{1}{4\big(M + x^{(1)}_- + 1\big)}. \tag{3.10}$$

Similarly, we have

$$\psi_t(t,x) + (\psi_0)_t(t,x) \le \beta(u_x(t,x) + \phi_x(t,x)) - \frac{\delta_i}{4},$$

so that

$$\psi_t(t,x) + (\psi_0)_t(t,x) \in I_{-\delta_i/4}(u_x(t,x) + \phi_x(t,x)).$$

If $\phi_x(t,x) = b_i^k$, we can prove the same assertion. Thus (iv) is proved. □

Remark 3.4 A different pair (ϕ, ψ_0) can be constructed by using $g_{i,s,-}^k$ instead of $g_{i,s,+}^k$ on $\mathcal{T}_{i,s}^k$ where $g_{i,s,-}^k$ is defined by using the functiong_-. Thus on each $\mathcal{T}_{i,s}^k$ we have two choices of defining ϕ and ψ_0. Therefore we may construct infinitely many weak solutions of (1.3) satisfying (1.4)–(1.7).

Item (v) is easy to prove as on each $T_{i,s}^k$, $|u_t| < m - \delta_i$ so that

$$|u_t + \phi_t| \le m - \delta_i + |\phi_t| \le m - \delta_i + \rho\delta_i x_-^{(1)} \le m - \frac{\delta_i}{2} < m$$

as (3.10) implies that $\rho < 1/(2x_-^{(1)})$.

Now we prove (vi). We have

$$\begin{aligned}
&\int_{Q^*} \text{dist}(D\Psi_\eta(t,x),\ K(u(t,x)+\phi(t,x)))dtdx \\
&\quad = \sum_{i=1}^{\infty}\int_{K_i} \text{dist}(D\Psi_\eta(t,x),\ K(u(t,x)+\phi(t,x)))dtdx \\
&\qquad + \sum_{i=1}^{\infty}\int_{G_i} \text{dist}(D\Psi_\eta(t,x),\ K(u(t,x)+\phi(t,x)))dtdx \\
&\quad \le \frac{\epsilon}{2}|Q^*| + \sum_{i=1}^{\infty}\sum_{k,s}\int_{T_{i,s}^k} \text{dist}(D\Psi_\eta(t,x),\ K(u(t,x)+\phi(t,x)))dtdx.
\end{aligned}$$

Given any $\mathcal{T}_{k,s}^i$, we assume as before that $(\psi_t(p_i^k),\ u_x(p_i^k)) \in \tilde{U}_+^0)$. Let $(t,x) \in T_{k,s}^i$ such that $\phi_x(t,x) = -a_i^k$, then by (3.9) and (3.10),

$$\begin{aligned}
&\text{dist}(D\Psi_\eta(t,x),\ K(u(t,x)+\phi(t,x))) \\
&\quad \le \text{dist}\left[\begin{pmatrix} \psi_t(t,x) + (\psi_0)_t(t,x) & 0 \\ 0 & u_x(t,x) + \phi_x(t,x) \end{pmatrix}, K_0^0 \cup K_+^0\right] \\
&\quad \le \text{dist}\left[\begin{pmatrix} \psi_t(p_i^k) & 0 \\ 0 & u_x(p_i^k) - a_i^k \end{pmatrix}, K_0^0 \cup K_+^0\right] \\
&\qquad + \left|\psi_t(t,x) - \psi_t(p_i^k)\right| + |(\psi_0)_t(t,x)| + \left|u_x(p_i^k) - u_x(t,x)\right| \\
&\quad \le \frac{\delta_i}{2} + 2\rho\delta_i + h_{i,s}^k\rho\delta_i s_2^* \le \left(\frac{1}{2} + 2\rho + x_-^{(1)}\rho\right)\delta_i < \delta_i < \frac{1}{4}\epsilon,
\end{aligned}$$

if we require that $\delta_i < \epsilon/4$ for all $i > 0$. The proof for the case when $(t, x) \in T^i_{k,s}$ with $\phi_x(t, x) = b^k_i$ is similar.

Consequently,

$$\begin{aligned}
&\sum_{i=1}^{\infty} \sum_{k,s} \int_{\mathcal{T}^k_{i,s}} \operatorname{dist}(D\Psi_\eta(t, x),\ K(u(t, x) + \phi(t, x)))dt\,dx \\
&\quad \le \sum_{i=1}^{\infty} \sum_{k,s} \frac{1}{4}\epsilon \big|\mathcal{T}^k_{i,s}\big| = \frac{1}{4}\epsilon \big| \cup_{i,k,s} \mathcal{T}^k_{i,s} \big| \\
&\quad \le \frac{1}{4}\epsilon |Q^*|.
\end{aligned}$$

Thus (vi) is proved.

Finally, by adjusting $h^k_{i,s}$ in the definition of $T^k_{i,s}$, we may have

$$\|\Psi_\eta - \Psi\|_{L^\infty(Q^*)} < \eta. \qquad \square$$

Remark 3.5 In Theorem 1.1 we see that for any $T > 0$, problem (1.3) has infinitely many solutions satisfying (1.4)–(1.7). To construct solutions for all $t > 0$, a plausible approach is to take a sequence of $T_j \to \infty$ and see whether a corresponding sequence of solutions u_j converges, in some sense to a solution on $Q_\infty = (0, +\infty) \times (0, l)$. However, this is potentially very difficult as the sequence, if not carefully selected might converge in some weak sense to a Young measure-valued solution.

Alternative constructions are the following:

(A) We take a fixed increasing function $\sigma^* : \mathbb{R} \to \mathbb{R}$ obtained by Lemma 3.1 and let $T_j = j \ge 0$ $j = 1, 2, \dots$. Let u^* be the smooth solution of (2.1) obtained by Lemma 2.1 using this particular $\sigma^*(\cdot)$. We may consider in $Q_j = (T_j, T_{j+1}$ problem (1.3) for initial value $u(T_j, x) = u^*(T_j, x)$. We denote by $u_j \in W^{1,\infty}(Q_j)$ the corresponding solution of (1.4)–(1.7) on Q_j and let $u_\infty \in W^{1,\infty}(Q_\infty)$ the function defined by $u_\infty(t, x) = u_j(t, x)$ if $(t, x) \in \bar{Q}_j$, $j = 1, 2, \dots$. Then from our construction of weak solutions in Theorem 1.1, we see that u_∞ is Lipschitz continuous in Q_∞ and satisfies (1.4)–(1.7) for $T = \infty$ because $\Psi_\infty = (\psi_\infty, u_\infty)$ satisfies

$$D\Psi_\infty(t, x) \in \Gamma(u(t, x)), \quad \text{a.e. } (t, x) \in Q_\infty,$$

where

$$\Gamma(u(t, x)) := \left\{ \begin{pmatrix} \sigma(X) & u(t, x) \\ Y & X \end{pmatrix} \in M^{2\times 2}, \quad X, Y \in \mathbb{R} \right\},$$

and u_∞ satisfies the boundary condition $(u_\infty)_x(t, 0) = (u_\infty)_x(t, l) = 0$ in the classical sense.

Another interesting observation is that when the smooth solution u^* given by Lemma 2.1 satisfies

$$\lim_{t\to+\infty} \max_{x\in[0,l]} |u^*_x(t, x)| = 0,$$

by our construction of $\sigma^*(\cdot)$ in Lemma 3.1, we have that for large $t > 0$ then $\sigma^*(u_x^*(t,x)) = \sigma(u_x^*(t,x))$ hence for large $t > 0$ u^* is a smooth solution of the Perona-Malik equation, satisfying the homogeneous Neumann condition hence $u_\infty(t,x) = u^*(t,x)$. This solution will eventually become smooth and decay to zero.

(B) If in Q_j we solve (1.3) by using a modified $\sigma_j^*(\cdot)$ as we did in the proof of Theorem for initial values with small derivative, we may let $s_j^* = \max_{x\in[0,l]} |(u_{j-1}^*)_x(T_j,x)|$ where u_{j-1}^* is the solution obtained in Q_{j-1}. When the maximum s_j^* is small, we may define $\sigma_j^*(\cdot)$ in Lemma 3.1 to have very 'flat' slope, that is, to have very small derivative $(\sigma_j^*)'(s)$ for when $|s|$ is, say greater that $s_j^*/2$. If we write our solution in Q_j as u_j satisfying the initial condition $u_j(T_j,x) = (u_{j-1})(T_j,x)$ and let $u_\infty(t,x) = u_j(t,x)$ when $(t,x) \in \bar{Q}_j$, we see that $u_\infty \in W^{1,\infty}_{loc}(Q_\infty)$. As the value $-y_+^j \le \sigma((u_j)_x(t,x) \le y_+^j$ and y_+^j can be very small if s_j^* is small (see Fig. 5), we see that $(u_j)_x(t,x)$ oscillates between points very close to zero and points close to infinity. We may then view this as the so-called 'stair-case' phenomenon described in the image processing literature [29, 32].

4 Proof of Lemma 3.1

Now we prove the technical result Lemma 3.1 whose proof involves only elementary calculus. We prove the lemma in two steps. We first establish the following

Claim Let $h > 0$ and $\sigma \in C^{2,1}[0,\ h]$ be such that

(i) $\sigma'(s) \ge \alpha > 0$ for $s \in [0,\ h]$ and for some fixed $\alpha > 0$;
(ii) $|\sigma''(s)| \le M$ for $s \in [0,\ h]$ and $|\sigma''(s) - \sigma''(t)| \le M|s-t|$ for $s,\ t \in [0,\ h]$, where $M > 0$ is a fixed constant.

Then for $\epsilon > 0$ sufficiently small and every $0 < \delta < \alpha/2$, there is a function $\sigma_{\epsilon,\delta} \in C^{2,1}[0,\infty)$ such that

(a) $\sigma_{\epsilon,\delta}(0) = \sigma(0)$, $\sigma'_{\epsilon,\delta}(0) = \sigma'(0)$ and $\sigma''_{\epsilon,\delta}(0) = \sigma''(0)$;
(b) $\sigma'_{\epsilon,\delta}(s) > \delta/2$ for $s \in [0,\infty)$, and $\sigma'_{\epsilon,\delta}(s) = \delta$ for $s \in [2\epsilon,\ h]$;
(c) $\sigma_{\epsilon,\delta}(s) < \sigma(s)$ for $s \in (0,\ h]$.

Proof of Claim We define, for some $m > 0$ to be determined later,

$$\sigma''_{\epsilon,\delta}(s) = \begin{cases} \sigma''(0) - \frac{1}{\epsilon}(\sigma''(0)+m)s, & s \in [0,\ \epsilon], \\ \dfrac{m}{\epsilon}(s-2\epsilon), & s \in [\epsilon,\ 2\epsilon], \\ 0, & s \in [2\epsilon,\ \infty). \end{cases} \tag{4.1}$$

Clearly, $\sigma''_{\epsilon,\delta}(\cdot)$ thus defined is a Lipschitz function. Next by requiring $\sigma'_{\epsilon,\delta}(0) = \sigma'(0)$ and $\sigma'_{\epsilon,\delta}(s) = \delta$ when $s \ge 2\epsilon$, we have

$$\sigma'_{\epsilon,\delta}(s) = \begin{cases} \sigma'(0) + \sigma''(0)s - \dfrac{\epsilon}{2}(\sigma''(0)+m)s^2, & s \in [0,\ \epsilon], \\ \sigma'(0) + \dfrac{\epsilon}{2}\sigma''(0) - \epsilon m + \dfrac{m}{2\epsilon}(s-2\epsilon)^2, & s \in [\epsilon,\ 2\epsilon], \\ \delta, & s \in [2\epsilon,\ \infty). \end{cases} \tag{4.2}$$

In order to satisfy $\sigma'_{\epsilon,\delta}(2\epsilon) = \delta$, we have to satisfy

$$\sigma'(0) + \frac{\epsilon}{2}\sigma''(0) - \epsilon m = \delta,$$

hence we choose

$$m = \frac{1}{\epsilon}\left(\sigma'(0) - \delta\right) + \frac{1}{2}\sigma''(0). \tag{4.3}$$

Clearly, $m > 0$ if $\epsilon > 0$ is small enough.

Before we proceed, we notice that

$$\sigma''(s) > \sigma''_{\epsilon,\delta}(s), \quad s \in (0, \epsilon].$$

This is due to the fact that

$$\begin{aligned}\sigma''(s) - \sigma''_{\epsilon,\delta}(s) &\geq \sigma''(0) - Ms - \left(\sigma''(0) - \frac{1}{\epsilon}(\sigma''(0) + m)s\right)\\ &\geq \frac{s}{\epsilon^2}(\sigma'(0) - \delta - M\epsilon - M\epsilon^2) \geq \frac{s}{\epsilon^2}\left(\frac{\alpha}{2} - 2M\epsilon\right) > 0,\end{aligned}$$

if $\epsilon < \min\{1,\ \alpha/(2M)\}$. This implies that

$$\sigma'(s) > \sigma'_{\epsilon,\delta}(s), \quad s \in (0,\ \epsilon].$$

By integrating $\sigma'_{\epsilon,\delta}(s)$ with the initial condition $\sigma_{\epsilon,\delta}(0) = \sigma(0)$, we obtain

$$\sigma_{\epsilon,\delta}(s) = \begin{cases} \sigma(0) + \sigma'(0)s + \dfrac{\sigma''(0)}{2}s^2 - \dfrac{1}{6\epsilon}(\sigma''(0) + m)s^3, \\ \quad s \in [0,\ \epsilon], \\ \sigma(0) + \sigma'(0)s + \dfrac{\epsilon}{2}\sigma''(0)s^2 - m\epsilon(s - \epsilon) - \dfrac{1}{6}\sigma''(0)\epsilon^2 + \dfrac{m}{6\epsilon}(s - 2\epsilon)^3, \\ \quad s \in [\epsilon\ 2\epsilon], \\ \sigma(0) + 2\sigma'(0)\epsilon - \dfrac{5\sigma''(0)}{6}\epsilon^2 - \dfrac{m}{2}\epsilon^2 + \delta(s - 2\epsilon), \\ \quad s \in [2\epsilon,\ \infty). \end{cases} \tag{4.4}$$

From our construction of $\sigma_{\epsilon,\delta}(s)$ it is clear that item (a) holds. Now we prove (b). We have, by recalling (4.3), that for $s \in (0,\ \epsilon]$,

$$\begin{aligned}\sigma'_{\epsilon,\delta}(s) &= \sigma'(0) + \sigma''(0)s - \frac{1}{2\epsilon}(\sigma''(0) + m)s^2\\ &= \sigma'(0) + \sigma''(0)s - \frac{s^2}{2\epsilon}\sigma''(0) - \frac{1}{2\epsilon}\left(\frac{\sigma''(0) - \delta}{\epsilon} + \frac{\sigma''(0)}{2}\right)\\ &= \left(1 - \frac{s^2}{2\epsilon^2}\right)\sigma'(0) + \left(s - \frac{3s^2}{4\epsilon^2}\right)\sigma''(0) + \frac{s^2}{2\epsilon^2}\delta\\ &\geq \frac{\sigma'(0)}{2} - M\epsilon \geq \frac{\alpha}{2} - M\epsilon > \frac{\alpha}{4} > \frac{\delta}{2}\end{aligned}$$

if $\epsilon < \alpha/(4M)$. Next, for $s \in (\epsilon, 2\epsilon)$, we have

$$\begin{aligned}\sigma'_{\epsilon,\delta}(s) &= \sigma'(0) + \frac{\epsilon}{2}\sigma''(0) - \epsilon m + \frac{m}{2\epsilon}(s-2\epsilon)^2\\ &= \frac{(s-2\epsilon)^2}{2\epsilon^2}\left(\sigma'(0)+\sigma''(0)\epsilon\right) + \left(1-\frac{(s-2\epsilon)^2}{2\epsilon^2}\right)\delta\\ &\geq \frac{(s-2\epsilon)^2}{2\epsilon^2}(\alpha-\delta-M\epsilon)+\delta > \delta > \frac{\delta}{2},\end{aligned}$$

as $s \in (\epsilon, 2\epsilon)$. We also have

$$\sigma'_{\epsilon,\delta}(s) = \delta \geq \frac{\delta}{2}$$

when $s \in [2\epsilon, \infty)$. Thus item (b) is proved.

Finally we prove item (c). As $\sigma_{\epsilon,\delta}(0) = \sigma(0)\ \sigma'_{\epsilon,\delta}(s) < \sigma'(s)$ in $(0, \epsilon)$, we see that

$$\sigma_{\epsilon,\delta}(s) < \sigma(s) \quad \text{in } (0, \epsilon].$$

Now for $s \in (\epsilon, 2\epsilon]$, we have

$$\begin{aligned}&\sigma(s) - \sigma_{\epsilon,\delta}(s)\\ &\quad= \sigma(s) - \Bigg(\sigma(0)+\sigma'(0)s+\frac{\epsilon}{2}\sigma''(0)s - m\epsilon(s-\epsilon)\\ &\qquad -\frac{1}{6}\sigma''(0)\epsilon^2 + \frac{m}{6\epsilon}(s-2\epsilon)^3\Bigg)\\ &\quad\geq \sigma(0)+\sigma'(0)s+\frac{\epsilon}{2}\sigma''(0)s^2 - \frac{Ms^3}{6}\\ &\qquad - \left(\sigma(0)+\sigma'(0)s+\frac{\epsilon}{2}\sigma''(0)s - m\epsilon(s-\epsilon) - \frac{1}{6}\sigma''(0)\epsilon^2 + \frac{m}{6\epsilon}(s-2\epsilon)^3\right)\\ &\quad= m\epsilon(s-\epsilon) + \frac{m}{6\epsilon}(2\epsilon-s)^3 + \frac{\sigma''(0)}{2}s(s-\epsilon) + \frac{\sigma''(0)}{6}\epsilon^2 - \frac{M}{6}\epsilon^3\\ &\quad\geq \left(\frac{\sigma''(0)}{2}(s+\epsilon)+\sigma'(0)-\delta\right)(s-\epsilon)\\ &\qquad + \left(\sigma'(0)-\delta+\frac{\sigma''(0)}{2}\epsilon\right)\frac{(2\epsilon-s)^3}{\epsilon^2} - \frac{M}{6}\epsilon^2 - \frac{2M}{3}\epsilon^3\\ &\quad\geq \left(\frac{\alpha}{2}-\frac{3M}{2}\epsilon\right)(s-\epsilon) + \frac{(2\epsilon-s)^3}{\epsilon^2}\left(\frac{\alpha}{2}-\frac{M}{2}\epsilon\right) - \frac{5M}{6}\epsilon^2\\ &\quad\geq \frac{1}{2}(\alpha-3M)\left((s-\epsilon)+\frac{(2\epsilon-s)^3}{\epsilon^2}\right)\frac{5M}{6}\epsilon^2\\ &\quad\geq \frac{1}{2}(\alpha-3M)\frac{\epsilon}{6} - \frac{5M}{6}\epsilon^2 = \frac{\epsilon}{12}(\alpha-13M\epsilon) > 0,\end{aligned}$$

if we further require that $\epsilon < \alpha/(13M)$. Here we have used the fact that

$$(s-\epsilon)+\frac{(2\epsilon-s)^3}{\epsilon^2}\geq\frac{\epsilon}{6}$$

as $s\in[\epsilon,\ 2\epsilon]$. In particular, we have the bounds for $\sigma_{\epsilon,\delta}(2\epsilon)$:

$$\frac{1}{2}\left(\sigma'(0)-\frac{11}{3}M\epsilon\right)\epsilon\leq\sigma_{\epsilon,\delta}(2\epsilon)\leq\left(\sigma'(0)+\frac{11}{6}M\epsilon\right)\epsilon,$$

hence

$$\sigma_{\epsilon,\delta}(2\epsilon)=o(\epsilon)\text{ as }\epsilon\to 0_+.$$

Now when $s\in(2\epsilon,\ h]$, we notice that $\sigma_{\epsilon,\delta}(s)=\sigma_{\epsilon,\delta}(2\epsilon)+\delta(s-2\epsilon)$, hence

$$\begin{aligned}\sigma(s)-\sigma_{\epsilon,\delta}(s)&=[\sigma(s)-\sigma(2\epsilon)]+[\sigma(2\epsilon)-\sigma_{\epsilon,\delta}(2\epsilon)]-\delta(s-2\epsilon)\\&>\sigma(s)-\sigma_{\epsilon,\delta}(s)-\delta(s-2\epsilon)>\left(\frac{\alpha}{2}-\delta\right)(s-2\epsilon)>0.\end{aligned}$$ □

Proof of Lemma 3.1 Let $0<y_-<y_+<1$ with y_+-y_- small, we can easily check that that for $s,\ r\in\{t\in(0,\ x_-^{(2)}),$ one has, for some $\alpha>0$ and $M>0$ that

$$\begin{aligned}&\sigma'(s)\geq\alpha>0,\\&|\sigma''(s)|\leq M,\\&|\sigma''(s)-\sigma''(r)|\leq M|s-r|.\end{aligned}$$

Now we define

$$\sigma^*_{\epsilon,\delta}(s)=\sigma_{\epsilon,\delta}(s-x_-^{(1)}),$$

where $\sigma_{\epsilon,\delta}(\cdot)$ is defined in the above Claim. As $\epsilon>0$ is sufficiently small, we may claim that the extended function

$$\sigma^*(s)=\begin{cases}\sigma(s), & s\in\left(0,\ x_-^{(1)}\right],\\ \sigma^*_{\epsilon,\delta}(s), & s\in\left[x_-^{(1)},\ \infty\right],\end{cases}$$

is of $C^{2,1}((0,\ \infty])$, strictly increasing and $\sigma^*(x_-^{(1)}+2\epsilon)<y_+$, for $0<\delta<\alpha/2$.

Notice that the function

$$\delta\to\sigma^*_{\epsilon,\delta}\left(x_-^{(2)}+\right)=\sigma_{\epsilon,\delta}\left(x_+^{(2)}-x_-^{(1)}\right)$$

is continuous for $\delta\in[0,\ \alpha/2]$. When $\delta=0$,

$$\sigma^*_{\epsilon,0}(x_-^{(2)}+)=\sigma_{\epsilon,0}(x_+^{(2)}-x_-^{(1)})=\sigma_{\epsilon,0}(2\epsilon)<y_+.$$

Therefore, for sufficiently small $\delta>0$, $\sigma^*_{\epsilon,\delta}(x_-^{(2)}+)<y_+$, so that

$$\sigma^*(s):=\sigma^*_{\epsilon,\delta}(s),\qquad s\in[0,\infty),\quad \epsilon>0,\ \delta>0\text{ small}$$

satisfies the requirement of Lemma 3.1.

Finally, by reflecting the graph of $\sigma^*(\cdot)$ with respect to the origin, we may define $\sigma^*(s)$ for $s\leq 0$. The proof is finished. □

Acknowledgements I wish to thank Ming Jiang and the referee for some helpful suggestions.

References

1. Adams, R.A.: Sobolev spaces. Academic Press (1975)
2. Aubin, J.P., Cellina, A.: Differential inclusions: set-valued maps and viability theory. Springer-Verlag (1984)
3. Aubert, G., Kornprobst, P.: Mathematical problems in image processing. Partial differential equations and the calculus of variations. Applied Mathematical Sciences, vol. 147. New York, Springer (2002)
4. Alvarez, L., Guichard, F., Lions, P.L., Morel, S.M.: Axioms and fundamental equations of image processing. Arch. Rational Mech. Anal **123**, 199–257 (1993)
5. Bellettini, G., Fusco, G.: A regularized Perona-Malik functional: some aspects of the gradient dynamics. Preprint (1993)
6. Ball, J.M., James, R.D.: Fine phase mixtures as minimizers of energy. Arch. Rational Mech. Anal. **100**, 13–52 (1987)
7. Ball, J.M., James, R.D.: Proposed experimental tests of a theory of fine microstructures and the two-well problem. Phil. Royal Soc. Lon. **338A**, 389–450 (1992)
8. Caselles, V., Morel, J. (eds): Special issue on partial differential equations and geometry-driven diffusion in image processing and analysis. IEEE Trans. Image Processing **7**(3) (1998)
9. Dacorogna, B., Marcellini, P.: General existence theorems for Hamilton-Jacobi Equations in the scalar and vectorial cases. Acta Mathematica **178**, 1–37 (1997)
10. Dacorogna, B., Marcellini, P.: Implicit partial differential equations. Progress in Nonlinear Differential Equations and their Applications, vol. 37. Birkhäuser (1999)
11. Dacorogna, B., Pisante, G.: A general existence theorem for differential inclusions in the vector valued case. Preprint (2004)
12. Esedoglu, S.: An analysis of the Perona-Malik scheme. Comm. Pure Appl. Math. **54**, 1442–1487 (2001)
13. Friedman, A.: Partial differential equations of parabolic type. Prentice-Hall (1964)
14. Gromov, M.: Partial differential relations. Springer-Verlag (1986)
15. Höllig, K.: Existence of infinitely many solutions for a forward backward heat equation. Trans. Amer. Math. Soc. **278**, 299–316 (1983)
16. Horstmann, D., Painter, K.J., Othmer, H.G.: Aggregation under local reinforcement, from lattice to continuum. Eur. J. Appl. Math. **15**, 545–576 (2004)
17. Kichenassamy, S.: The Pernoa-Malik paradox. SIAM J. Appl. Math **57**, 1328–1342 (1997)
18. Kirchheim, B.: Rigidity and Geometry of Microstructures. MPI for Mathematics in the Sciences Leipzig, Lecture notes (http://www.mis.mpg.de/preprints/ln/lecturenote-1603.pdf) (2003)
19. Kawohl, B., Kutev, N.: Maximum and comparision principle for one-dimensional anisotropic diffusion. Math. Ann **311**, 107–123 (1998)
20. Lieberman, G.M.: Second order parabolic differential equations. Singapore, London: World Scientific (1996)
21. Ladyzenskaya, O.A., Solonnikov, V.A., Uralceva, N.N.: Linear and quasilinear equations of parabolic type. Nauka, Moscow (1967)
22. Morini, M., Negri, M.: Mumford-Shah functional as Γ-limit of discrete Perona-Malik energies. Math. Models Methods Appl. Sci **13**, 785–805 (2003)
23. Müller, S., Šverák, V.: Attainment results for the two-well problem by convex integration. In: Jost, J. (ed.) Geometric analysis and the calculus of variations, pp. 239–251. International Press (1996)
24. Müller, S., Šverák, V.: Unexpected solutions of first and second order partial differential equations. Doc. Math. J. DMV, Extra Vol. ICM 98, pp. 691–702
25. Müller, S., Šverák, V.: Convex integration with constraints and applications to phase transitions and partial differential equations. J. Eur. Math. Soc. **1**, 393–422 (1999)
26. Müller, S., Šverák, V.: Convex integration for Lipschitz mappings and counterexamples to regularity. Ann. Math **157**, 715–742 (2003)
27. Müller, S., Sychev, M.A.: Optimal existence theorems for nonhomogeneous differential inclusions. J. Funct. Anal. **181**, 447–475 (2001)
28. Perona, P., Malik, J.: Scale space and edge detection using anisotropic diffusion. IEEE Trans. Pattern Anal. Mach. Intell **12**, 629–639 (1990)

29. Sapiro, G.: Geometric partial differential equations and image analysis. Cambridge: Cambridge University Press (2001)
30. Sychev, M.A.: Comparing two methods of resolving homogeneous differential inclusions. Calc. Var. PDEs **13**, 213–229 (2001)
31. Taheri, S., Tang, Q., Zhang, K.: Young measure solutions and instability of the one-dimensional Perona-Malik equation. J. Math. Anal. Appl. **308**, 467–490 (2005)
32. Weickert, J.: Anisotropic diffusion in image processing. ECMI Series, Teubner, Stuttgart (1998)