

# Analysis of adaptive forward-backward diffusion flows with applications in image processing

V. B. Surya Prasath\*      José Miguel Urbano†      Dmitry Vorotnikov‡

## Abstract

The nonlinear diffusion model introduced by Perona and Malik in 1990 is well suited to preserve salient edges while restoring noisy images. This model overcomes well-known edge smearing effects of the heat equation by using a gradient dependent diffusion function. Despite providing better denoising results, the analysis of the PM scheme is difficult due to the forward-backward nature of the diffusion flow. We study a related adaptive forward-backward diffusion equation which uses a mollified inverse gradient term engrafted in the diffusion term of a general nonlinear parabolic equation. We prove a series of existence, uniqueness and regularity results for viscosity, weak and dissipative solutions for such forward-backward diffusion flows. In particular, we introduce a novel functional framework for wellposedness of flows of total variation type. A set of synthetic and real image processing examples are used to illustrate the properties and advantages of the proposed adaptive forward-backward diffusion flows.

**Keywords:** Anisotropic diffusion, forward-backward diffusion, wellposedness, regularization, Perona-Malik PDE, image restoration, total variation flow.

## 1 Introduction

The nonlinear diffusion model, introduced in image processing by Perona and Malik [55], involves solving the following initial-boundary value problem

$$\begin{cases} \frac{\partial u(x, t)}{\partial t} = \operatorname{div} \left( \mathcal{C}(|\nabla u(x, t)|^2) \nabla u(x, t) \right) & \text{in } \Omega \times (0, T), \\ \frac{\partial u(x, t)}{\partial \nu} = 0 & \text{on } \partial\Omega \times (0, T), \\ u(x, 0) = u_0(x) & \text{in } \Omega, \end{cases} \quad (1)$$

---

\*Corresponding author. Department of Computer Science, University of Missouri-Columbia, MO 65211 USA. E-mail: prasaths@missouri.edu

†CMUC, Department of Mathematics, University of Coimbra, 3001-501 Coimbra, Portugal. E-mail: jmurb@mat.uc.pt

‡CMUC, Department of Mathematics, University of Coimbra, 3001-501 Coimbra, Portugal. E-mail: mitvorot@mat.uc.pt

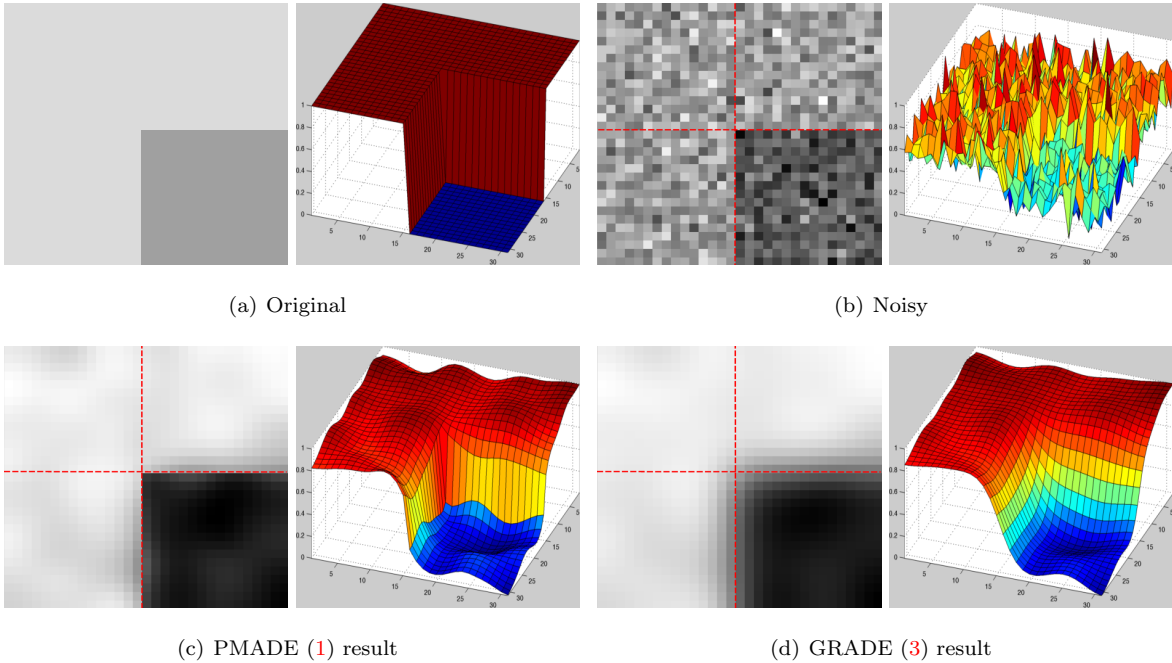


Figure 1: Spatial regularization in the diffusion coefficient alters discontinuities in a given image. (a) Original synthetic image of size  $31 \times 31$ , a square (gray value = 160) at the bottom right corner with uniform background (gray value = 219). (b) Input image obtained by adding Gaussian noise  $\sigma_n = 30$  to the original image. This noisy image is used as the initial value  $u_0$  for the nonlinear PDEs with  $\mathcal{C}_1$  diffusion coefficient and  $K = 20$ . Results of PMADE (1) with 20 iterations in (left) image (right) surface format (c), and GRADE (3) with 20 iterations in (left) image (right) surface format (d). The intersection of red dotted lines indicate the exact corner location of the square. (For interpretation of color in this figure, the reader is referred to the web version of this article.)

where  $u_0 : \Omega \rightarrow \mathbb{R}$  is the observed (noisy) image,  $\Omega \subset \mathbb{R}^2$  is a bounded domain with Lipschitz boundary. The function  $\mathcal{C} : \mathbb{R}_0^+ \rightarrow \mathbb{R}_0^+$  is non-increasing such that  $\mathcal{C}(0) = 1$  and  $\lim_{s \rightarrow +\infty} \mathcal{C}(s) = 0$ . Note if  $\mathcal{C}(s) \equiv 1$  then we recover the heat equation. The diffusion coefficient function  $\mathcal{C}(\cdot)$  in Eqn. (1) is an edge indicator function that reduces the amount of diffusion near edges and behaves locally as inverse heat equation. The original choices of  $\mathcal{C}(\cdot)$  by Perona and Malik [55] are

$$\mathcal{C}_1(s) = \exp\left(-\frac{s}{K^2}\right), \quad \mathcal{C}_2(s) = \frac{1}{1 + s/K^2}, \quad (2)$$

where  $K > 0$  is a tunable parameter also known as the *contrast parameter* [57].

Despite impressive numerical results obtained in image processing using the Perona-Malik equation (1) with  $\mathcal{C} = \mathcal{C}_1$ , it was shown [44, 43] to be an ill-posed PDE due to the degenerate behavior for large gradients. Thereby, although existence of infinitely many solutions in a relaxed setup involving Young measures was established in [77, 21], many authors have been looking for regularizations of the Perona-

Malik equations which inherit its usefulness in image restoration but have better mathematical behavior. One of the pioneering works of this kind [17] replaced the magnitude of the image gradient  $s = |\nabla u|$  used in the diffusivity functions by the spatially regularized gradient,  $s = |\nabla G_\sigma * u|$ , where  $G_\sigma$  denotes the two-dimensional Gaussian kernel,  $G_\sigma(x) = (2\pi\sigma)^{-1} \exp(-|x|^2/2\sigma^2)$  and  $*$  denotes the convolution operation. Thus, the Catté et al. [17] Gaussian regularized anisotropic diffusion equation (GRADE for short) reads as

$$\frac{\partial u}{\partial t} = \operatorname{div} \left( \mathcal{C}(|\nabla G_\sigma * u|^2) \nabla u \right). \quad (3)$$

This modification of equation (1) is sufficient to obtain the existence and uniqueness of solution to the initial-boundary value problem for GRADE in Eqn. (3). However, the space-invariant Gaussian smoothing inside the divergent term tends to push the edges away from their original locations, see Figure 1 for an illustration of this effect on a synthetic corner image. This effect, known as edge dislocation, can be detrimental to further image analysis. This can also be seen via the regularity of solution to GRADE, which belongs to a high order Sobolev space. Furthermore, the use of isotropic smoothing is against the principle of anisotropic diffusion which aims to smooth homogeneous regions without affecting edges. To remedy this, one can use time regularization instead of spatial regularization, or a related idea of decoupling the diffusion coefficient into a separate evolutionary PDE [53, 11, 12, 4, 63]. Another direction to improve the well-posedness of the Perona-Malik equation (1) is to approximate the nonlinear diffusion coefficient  $\mathcal{C}_1$  with less degenerate ones, see [34, 38, 35, 36]. Fourth order regularizations were introduced and studied in a bulk of papers, for instance, in [33, 13, 40, 39]. We refer to the recent paper [37] for a nice and detailed overview on anisotropic diffusions arising in image processing from the perspective of mathematical analysis.

Motivated by the correspondence between the variational and PDE methods for imaging problems, which we discuss in the next section, in this paper we consider a Perona-Malik type PDE with the generic diffusion function inspired by the stationary nonlinear regularization approach. Engrafting a mollified gradient within the adaptive diffusion function we obtain a general forward-backward diffusion PDE. We prove a series of existence, uniqueness and regularity results for viscosity, weak, strong and dissipative solutions for a wide class of the proposed generalized forward-backward diffusion models. Experimental results on synthetic, noisy standard test, and biomedical images are provided to illustrate different diffusion schemes considered here.

One of the highlights of the paper is the introduction of the concept of partial variation, which enables us to define and employ the Banach space of functions of bounded partial variation. Our approach appears to be relevant in the context of evolutionary problems which involve singular diffusion of 1-Laplacian kind or gradients of linear growth functionals, and holds promise for wide applicability.

The rest of the paper is organized as follows. Section 2 is devoted to the modelling issues. Section 3 ex-

amines the conditions needed for wellposedness of the proposed regularization-inspired forward-backward PDE in various scenarios. In Section 4, we provide numerical experiments to prove the effectiveness of the proposed multi-scale scheme as well as examples for various cases.

## 2 Preliminary observations and the proposed model

The motivation for the Perona-Malik nonlinear diffusivity is that inside the regions where the magnitude of the gradient of  $u$  is weak, equation (1) acts like a heat equation, resulting in isotropic smoothing, whereas near the edges, where the magnitude of the gradient is large, the diffusion is “stopped” and the edges are preserved.

To see the underlying details, we split the divergence term in Eqn. (1),

$$\operatorname{div} \left( \mathcal{C}(|\nabla u|^2) \nabla u \right) = 2(u_x^2 u_{xx} + u_y^2 u_{yy} + 2u_x u_y u_{xy}) \mathcal{C}'(|\nabla u|^2) + \mathcal{C}(|\nabla u|^2) (u_{xx} + u_{yy}). \quad (4)$$

Considering the tangent  $\mathcal{T}$  and normal  $\mathcal{N}$  directions of the isophote lines, and setting

$$\mathcal{B}(s) = \mathcal{C}(s) + 2s\mathcal{C}'(s), \quad (5)$$

we infer

$$\operatorname{div} \left( \mathcal{C}(|\nabla u|^2) \nabla u \right) = \mathcal{C}(|\nabla u|^2) u_{\mathcal{T}\mathcal{T}} + \mathcal{B}(|\nabla u|^2) u_{\mathcal{N}\mathcal{N}}. \quad (6)$$

We thus see that the Perona-Malik diffusion (1) is the sum of the tangential diffusion weighted by the function  $\mathcal{C}(\cdot)$  plus the normal (transverse) diffusion weighted by the function  $\mathcal{B}(\cdot)$ , resp. Since smoothing with edge preservation is of paramount importance in image processing, it is desirable to smooth more in the tangent direction than in the normal direction. This can be translated to the condition

$$\lim_{s \rightarrow \infty} \frac{\mathcal{B}(s)}{\mathcal{C}(s)} \leq 0, \quad \text{or equivalently,} \quad \lim_{s \rightarrow \infty} \frac{s\mathcal{C}'(s)}{\mathcal{C}(s)} \leq -\frac{1}{2}. \quad (7)$$

For example, in the case of the power growth functions

$$\mathcal{C}(s) \approx s^q$$

the above limit gives that

$$q \leq -\frac{1}{2}. \quad (8)$$

Note that for the original diffusion function  $\mathcal{C}_2$  in (2) we have  $\mathcal{B}_2(s) > 0$  if  $s < K^2$ , implying forward diffusion in the regions where the gradient magnitude of the image function is less than  $K$ , whereas  $\mathcal{B}_2(s) < 0$  if  $s > K^2$ , yielding backward diffusion in the area where absolute values of the gradient are larger than  $K$ . The same is true for  $\mathcal{C}_1$  with the threshold value  $2^{-1/2}K$ . Thus, the PDE (1) promotes combined forward-backward diffusion flow, see Figure 2 for a comparison with the heat equation. The

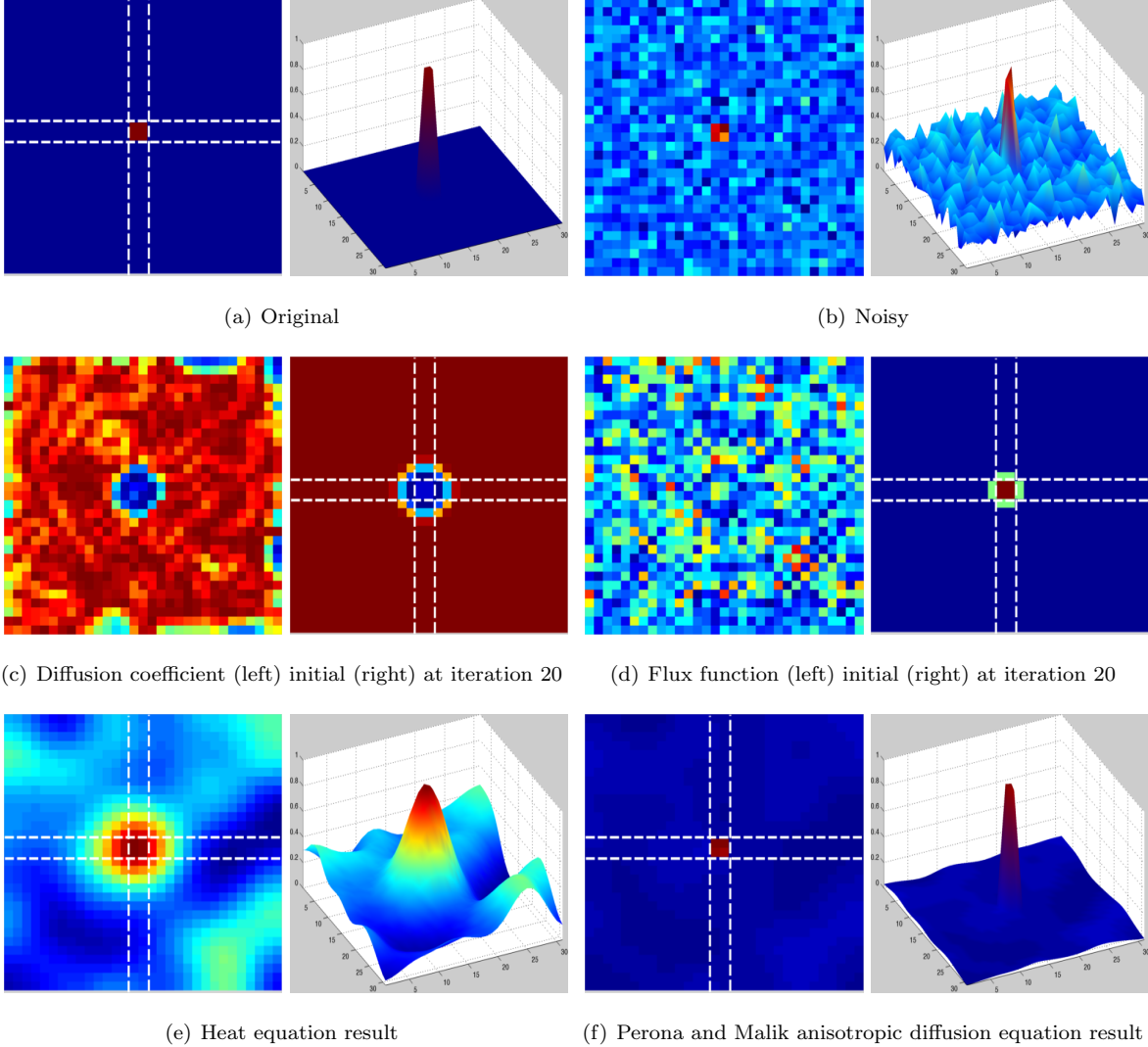


Figure 2: Diffusion process for a simple synthetic image. (a) Original synthetic image of size  $31 \times 31$ , a square ( $2 \times 2$ , gray value = 1) at the center with uniform background (gray value = 0). (b) Input image obtained by adding Gaussian noise  $\sigma_n = 30$  to the original image. This noisy image is used as the initial value  $u_0$ . (c) Diffusion coefficient  $\mathcal{C}_1$  in (2), with  $K = 20$ . This acts as a discontinuity detector and stops the diffusion spread across edges. (d) Flux function  $\mathcal{C}_1(|\nabla u|) \cdot |\nabla u|$ . (e) Result of heat equation with 20 iterations in (left) image (right) surface format. (f) Result of PMADE Eqn. (1) with 20 iterations in (left) image (right) surface format. The white dotted lines indicate the *influence* region at the center. (For interpretation of color in this figure, the reader is referred to the web version of this article.)

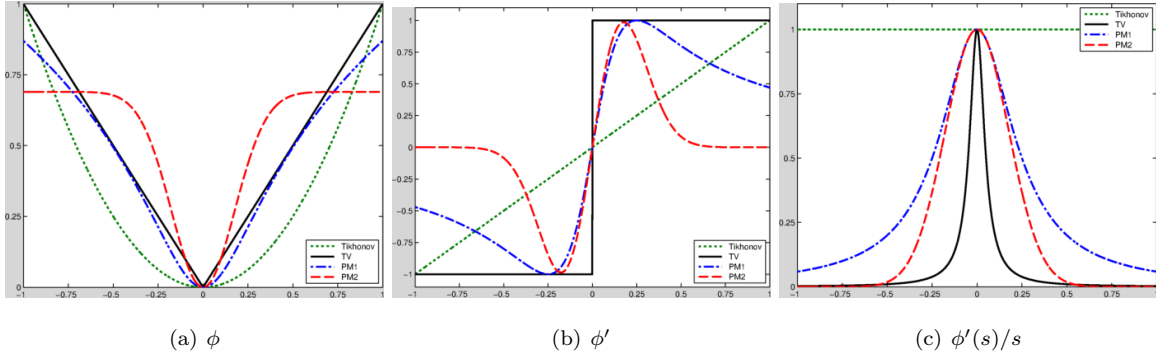


Figure 3: Well known regularization functions and their qualitative shapes: (a) Regularization (b) Flux (c) Diffusion coefficients. Convex functions: Tikhonov, total variation (TV). Non-convex functions: Perona-Malik regularizations  $\phi_1$ -PM1,  $\phi_2$ -PM2 see Eqns. (12). Normalized to  $[0, 1]$  for visualization.

Perona-Malik anisotropic diffusion equation (PMADE for short) thus balances forward and backward diffusion regimes using a tunable  $K$ , the contrast parameter [48, 57]. These two competing requirements constitute a common theme in many PDE based image restoration models [76, 18, 16, 19]. Moreover, from Eqn. (6) we see that Eqn. (1) is a time dependent nonlinear diffusion equation with preferential smoothing in the tangential direction  $\mathcal{T}$  than normal  $\mathcal{N}$  to edges. This property has been exploited in image processing and in particular in edge preserving image restoration [9].

The PDE models of Perona-Malik type have strong connections to variational energy minimization problems and this fact is exploited by many to design various diffusion functions [9, 60]. Following [71], consider the next minimization problem for image restoration:

$$\min_u E(u) = \frac{\beta}{2} \int_{\Omega} (u_0(x) - u(x))^2 dx + \alpha \int_{\Omega} \phi(|\nabla u(x)|) dx. \quad (9)$$

Here  $\alpha \geq 0$  is regularization parameter,  $\beta \geq 0$  fidelity parameter, and  $\phi : \mathbb{R} \rightarrow \mathbb{R}^+$  is an even function, which is called the *regularization function*. The *a priori* constraint on the solution is represented by the regularizing term  $\phi(|\nabla u|)$ , and the shape of the regularization determines the qualitative properties of solutions [52]. The formal gradient flow associated with the functional  $E(u)$  is given by,

$$\frac{\partial u}{\partial t} = \alpha \operatorname{div} \left( \frac{\phi'(|\nabla u|)}{|\nabla u|} \nabla u \right) - \beta (u - u_0). \quad (10)$$

We recall two primary choices used widely as regularization functions in various image processing tasks.

- $\phi(s) = s^2$ : This corresponds to the classical Tikhonov regularization method [70]. In this case the Euler-Lagrange equation (written with artificial time evolution, see Eqn. (10)) is,

$$\frac{\partial u}{\partial t} = \alpha \Delta u - \beta (u - u_0),$$

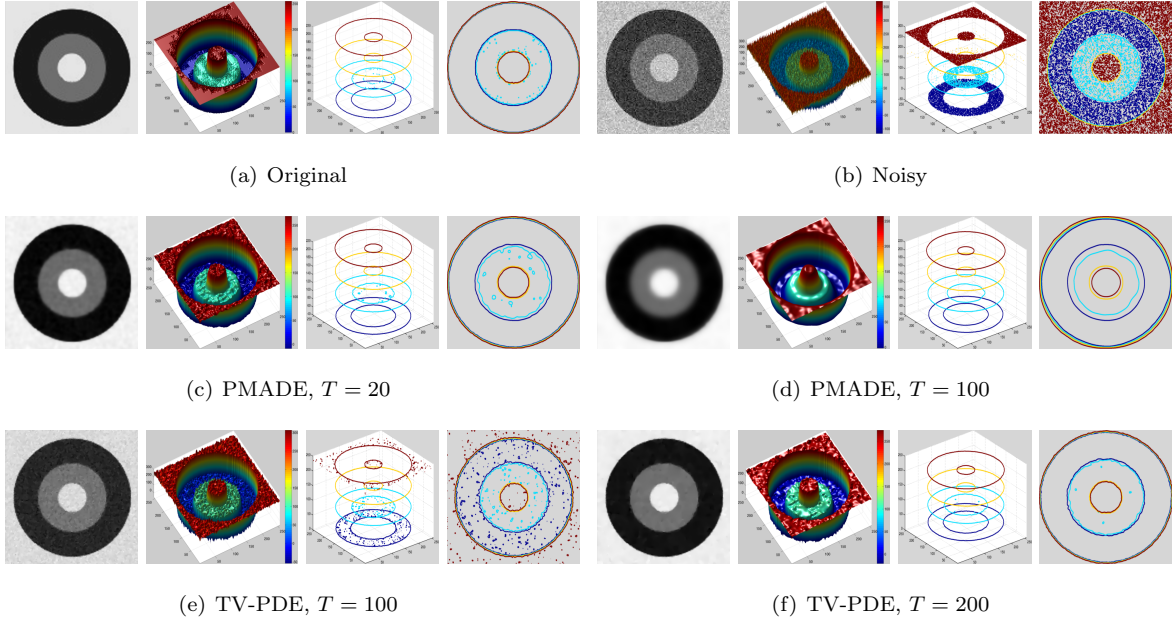


Figure 4: Advantages and disadvantages of TV-PDE and PMADE (with  $C_1$  as the diffusion coefficient and  $K = 20$ ) models on a synthetic piecewise constant *Circles* image. Noisy image is obtained by adding Gaussian noise  $\sigma_n = 30$ . *[i][ii][iii][iv]*: We show in each sub-figure (i) gray-scale image (ii) surface (pixel values as  $z$  values) (iii) level lines (only top 4 level lines are shown for clarity), and (iv) contour maps to highlight jaggedness of level lines and staircasing artifacts. Better viewed online and zoomed in.

which is an isotropic diffusion equation and hence does not preserve edges, see Figure 2(e). This heat flow provides a linear scale space and has been widely used in various computer vision tasks such as feature point detection and object identification [65].

- $\phi(s) = s$ : To reduce the smoothing when the magnitude of the gradient is high, Rudin et al., [66] introduced total variation (TV) based scheme by setting  $\phi(s) = s$ . In this case the Euler-Lagrange equation is written as (see Eqn. (10)),

$$\frac{\partial u}{\partial t} = \operatorname{div} \left( \frac{\nabla u}{|\nabla u|} \right) - \beta(u - u_0), \quad \text{or} \quad \frac{\partial u}{\partial t} = \operatorname{div} \left( \frac{\nabla u}{\sqrt{\epsilon + |\nabla u|^2}} \right) - \beta(u - u_0), \quad (11)$$

where  $\epsilon > 0$  is a small number added to avoid numerical instabilities in discrete implementations. In [18], the existence and uniqueness of the TV minimization is proved in the space of functions of bounded variation (BV), and the corresponding gradient flow is treated in [7], see also [27, 31, 5, 6, 15, 45] and Remark 10 below. But this global TV model suffers from *staircasing* and blocky effects in the restored image [52]. Also, sharp corners will be rounded and thin features are removed under this regularization model. To see this, let  $\chi_B$  be the indicator function of  $B \subset \mathbb{R}^N$  a bounded

set with Lipschitz boundary. Then the total variation term in the regularization functional (9),  $\int |\nabla \chi_B|$  is the perimeter of the set  $B$ . This shows that TV regularization penalizes edge lengths of an image. Note that this TV diffusion PDE (11) is a borderline case of anisotropic diffusion PMADE in Eqn. (1) with  $\mathcal{C}(s) = s^{-1/2}$ , a singular diffusion model.

It is easy to see that the Eqn. (10) almost coincides with the Perona-Malik anisotropic PDE (1) for  $\phi'(s) = \mathcal{C}(s^2)s$ , cf. [54], and the difference between the two equations is the lower order term coming from the data fidelity in the regularization functional (9). For edge preservation we need to work with functions  $\phi$  with at most linear growth at infinity, cf. (8). For example, the Perona-Malik diffusion coefficients (2), up to multiplicative constants, correspond to the following non-convex regularization functions (see Figure 3(a), denoted as PM1 and PM2),

$$\phi_1(|\nabla u(x)|) = 1 - \exp\left(-\left(\frac{|\nabla u(x)|}{K}\right)^2\right), \quad \phi_2(|\nabla u(x)|) = \log\left(1 + \left(\frac{|\nabla u(x)|}{K}\right)^2\right). \quad (12)$$

Figure 4 shows an experimental analysis of PMADE (1) against TV PDE (11) on a synthetic image which consist of various circles with constant pixel values. This piecewise constant image represents a near ideal scenario and both the PMADE ( $T = 20, 100$ ) and TV-PDE ( $T = 100, 200$ ) results indicate over-smoothing and staircasing artifacts.

Several studies [68, 60, 61, 56, 63, 62, 64] have introduced spatially adaptive regularization functions to reduce staircasing/blocky artifacts created by the classical TV and PMADE schemes. Such adaptive methods can be written as an energy minimization of the following form (see, e.g., [68]),

$$\min_u E(u) = \beta \int_{\Omega} (u_0(x) - u(x))^2 dx + \int_{\Omega} \alpha(x) |\nabla u(x)| dx. \quad (13)$$

where the function  $\alpha(\cdot)$  self adjusts itself according to an estimate of edge information from each pixel. Since we want to reduce the regularization/smoothing effect of (9) near edges, hence  $\alpha(x)$  is chosen to be inversely proportional to the likelihood of the presence of an edge. For example, the original function proposed in [68] is,

$$\alpha(x) = \frac{1}{\epsilon + |\nabla G_{\sigma} * u_0(x)|}, \quad \epsilon > 0. \quad (14)$$

The term in the denominator provide an estimate of edges from the input image  $u_0$  at scale  $\sigma$  using the Gaussian kernel  $G_{\sigma}$  filtered gradients. Introduction of such a spatially adaptive parameter, which self adjusts according to the smoothed gradient of the image, reduces the TV flow in homogenous regions thereby alleviating the *staircasing* problem. In [20], the existence and uniqueness of the functional satisfying (13) is proved under the weighted TV norm. An edge indicator function of the form (14) can also be introduced directly in the PDE of the form (10). For example, using it in the PMADE (1), we write adaptive PMADE as

$$\frac{\partial u(x, t)}{\partial t} = \operatorname{div} \left( \frac{\mathcal{C}(|\nabla u(x, t)|^2)}{\epsilon + |\nabla G_{\sigma} * u_0(x)|} \nabla u(x, t) \right). \quad (15)$$



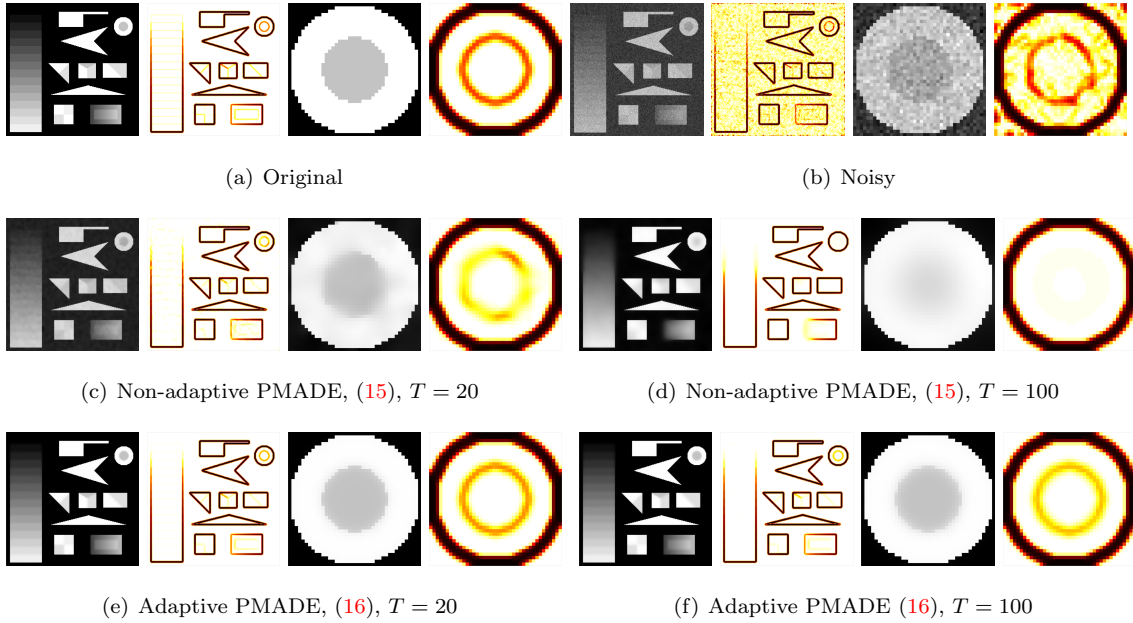


Figure 5: Using updated edge indicator function results in better final restoration in PMADE (with  $C_1$  diffusion coefficient and  $K = 20$ ). Shown here are the final restoration results at iterations  $T = 20, 100$  for *non-adaptive* PMADE (15) and *adaptive* PMADE (16).  $|i|ii|iii|iv|$ : We show in each sub-figure (i) gray-scale image (ii) edge map (heat colormap) (iii) close-up gray-scale image and (iv) close-up edge map. Note that the smoothness parameter  $\epsilon = 10^{-6}$ , is used in this example, see Eqn. (14). Better viewed online and zoomed in.

It is advantageous to use the current estimate image  $u$  in the edge indicator  $\alpha(x)$  in (15) instead of the initial noise image  $u_0$ , that is

$$\frac{\partial u(x, t)}{\partial t} = \operatorname{div} \left( \frac{C(|\nabla u(x, t)|^2)}{\epsilon + |\nabla G_\sigma * u(x, t)|} \nabla u(x, t) \right). \quad (16)$$

See Figure 5 for an illustration of using adaptive weight function in the final restoration results on a synthetic image with multiple objects. As can be seen using an adaptive edge indicator keeps the edges through higher iterations. Moreover, integration of two scales ( $|\nabla u|$  corresponds to scale  $\sigma = 0$  and  $|\nabla G_\sigma * u|$  to scale  $\sigma$ ) in one term, see Eqn. (17) below, can regularize the boundaries of the level set of  $u_0$  while at the same time keeping more features. Numerical experiments will support our claims about the advantage of interaction of two scales, see Section 4. Also we use the nonlinear function  $\phi$  (see Eqn. (9)) to control the growth adaptively as mentioned. Thus the general multi-scale minimization problem now reads as

$$\min_u E(u) = \beta \int_{\Omega} (u(x) - u_0(x))^2 dx + \int_{\Omega} \frac{\phi(|\nabla u(x)|)}{\epsilon + |\nabla G_\sigma * u(x)|} dx. \quad (17)$$

A well known method to prove the existence and uniqueness of minimizer to this problem (17) is to obtain the lower semicontinuity of the functional  $E$  using the properties of regularizing function  $\phi$ .

**Remark 1.** *The functional (17) with  $\phi(\nabla u)$  replaced by  $|G_\sigma * \nabla u|^2$ , and with additional quadratic regularization term ( $\eta \int_\Omega |\nabla u|^2 dx$ ), which is related to robust Geman-Mclure model [30], was considered in [41]. The arguments in [41] can also be extended to the general minimization problem (17).*

Motivated by the regularization functional (17) and previous discussions, we consider here a general forward-backward PDE of the following form<sup>1</sup>,

$$\frac{\partial u}{\partial t} = \operatorname{div} \left( \frac{\varphi_y(x, |\nabla u|) \nabla u}{1 + K g(|G_\sigma * \nabla u|)} \right). \quad (18)$$

Here  $\varphi_y$  is the partial derivative of  $\varphi(x, y)$  with respect to the second variable  $y = |\nabla u|$ . The variational problem could involve explicit dependence on the function  $u$  in the regularization term,  $\phi = \phi(x, u(x), \nabla u(x))$ , with the corresponding changes in the evolutionary problem, but we will not study this general case in this article. For the existence and uniqueness of a solution  $u$ , we need additional assumptions on  $\varphi$  which will be discussed at the end of this section and in Section 3. Observe also that, in the  $x$ - and  $u$ - independent case,  $\phi$  and  $\varphi$  are related through  $\phi'(s) = s\varphi'(s)$ .

**Remark 2.** *The fidelity parameter  $\beta$  in (17) can be tuned to fit the the problem at hand [32, 64], i.e.,*

$$D(u, u_0) = \int_\Omega \beta(x) (u(x) - u_0(x))^p dx, \quad p = 1, 2.$$

The major questions we are now concerned with are:

- (1) What are the conditions on  $\varphi$  to obtain existence of solutions for the PDE (18)?
- (2) What are the admissible inverse mollifier  $g$  functions?

The answer to the first question depends on the answer to the second. Let us briefly discuss this issue; more details will be given in the next section. If  $g$  is merely continuous, then we admit power growth, e.g.,  $\varphi(x, y) = y^p$ ,  $1 < p < +\infty$ , and logarithmic growth  $\varphi(x, y) = \ln y$ . If  $g$  is locally Lipschitz, then we need a sort of strong parabolicity condition involving  $\varphi$ . If the derivative of  $g$  is sub-linear near zero (that is,  $g$  may be of order  $s^p$ ,  $p \geq 2$ , for small  $s$ ), then  $\varphi(x, y)$  enjoys a wide range of possibilities with minor restrictions such as coercivity and weak parabolicity.

### 3 Existence of various types of solutions

Here we study the equation

$$\frac{\partial u}{\partial t} = \operatorname{div} \left( \frac{\varphi_y(x, |\nabla u|) \nabla u}{1 + K g(|G * \nabla u|)} \right), \quad (19)$$

---

<sup>1</sup>We will omit the image fidelity term for brevity, since this lower order term does not cause any special difficulties in the mathematical analysis of the model.

which is slightly more general than Eqn. (18) in the sense that we admit arbitrary space dimension  $n$  and generic convolution kernels  $G$ .

Throughout this section, we employ Einstein's summation convention. The inner product in  $\mathbb{R}^n$ ,  $n \in \mathbb{N}$ , is denoted by a dot. The symbols  $C(\mathcal{J}; E)$ ,  $C_w(\mathcal{J}; E)$ ,  $L_2(\mathcal{J}; E)$  etc. denote the spaces of continuous, weakly continuous, quadratically integrable etc. functions on an interval  $\mathcal{J} \subset \mathbb{R}$  with values in a Banach space  $E$ . We recall that a function  $u : \mathcal{J} \rightarrow E$  is *weakly continuous* if for any linear continuous functional  $g$  on  $E$  the function  $g(u(\cdot)) : \mathcal{J} \rightarrow \mathbb{R}$  is continuous.

We are going to introduce three different concepts of generalized solution to Eq. (19): viscosity, weak and dissipative. The relation between different kinds of solution is not an issue here, since our goal is to construct at least one kind of solution for the widest possible range of assumptions on  $\varphi$ . The interrelation question is very delicate, but it is true that a strong, regular solution, if it exists, is also a viscosity, dissipative, or weak solution. Moreover, in these circumstances, no other viscosity or dissipative solution may exist, but a weak solution might. On the other hand, if there are no strong solutions, then these classes of solutions intersect but do not coincide, e.g., there might exist weak solutions which are neither viscosity nor dissipative. Under additional assumptions, one can prove that a weak solution is also dissipative. Roughly speaking, dissipative is the weakest notion of the three, and viscosity is the strongest. Nonetheless, in the logarithmic case, only weak solutions are proven to exist, and for the infinity-Laplacian flow (which however does not fit into our framework here) we only know that viscosity solutions exist.

### 3.1 Viscosity solutions

We denote

$$a_{ij}(x, p) = \varphi_y(x, |p|)\delta_{ij} + \varphi_{yy}(x, |p|)\frac{p_i p_j}{|p|}, \quad (20)$$

$$h(q) = \frac{1}{1 + Kg(|q|)}. \quad (21)$$

Here  $\delta_{ij}$  is Kronecker's delta, and  $x, p, q \in \mathbb{R}^n$ . In this subsection we consider the case of spatially periodic boundary conditions [2] for Eq. (19). Namely, we assume that there is an orthogonal basis  $\{b_i\}$  in  $\mathbb{R}^n$  so that

$$u(\cdot, x) = u(\cdot, x + b_i), \quad x \in \mathbb{R}^n, \quad i = 1, \dots, n. \quad (22)$$

The problem is complemented with the initial condition

$$u(0, x) = u_0(x), \quad (23)$$

where  $x \in \mathbb{R}^n$ , and  $u_0$  is Lipschitz and satisfies (22). Of course,  $\varphi$  (and thus  $a$ ) should also satisfy the same spatial periodicity restriction (with respect to  $x$ ).

We also make the following assumptions:

$$\varphi_y, \varphi_{yx}, \varphi_{yxx} \text{ are continuous and bounded functions,} \quad (24)$$

$$\varphi_{yy}, \varphi_{yyx} \text{ are continuous for } y \neq 0, \text{ and satisfy} \quad (25)$$

$$\sup_{x \in \mathbb{R}^n} \lim_{y \rightarrow 0} |y|(|\varphi_{yy}(x, y)| + |\varphi_{yyx}(x, y)|) = 0. \quad (26)$$

$$a_{ij}(x, p)\xi_i\xi_j \geq C \left[ \text{mod} \left( \frac{\partial a(x, p)}{\partial x_k} \right) \right]_{ij} \xi_i\xi_j, \quad k = 1, \dots, n, \quad \xi, x, p \in \mathbb{R}^n, \quad (27)$$

$$\sqrt{h} \in W_\infty^1(\mathbb{R}^n), \quad h \in W_\infty^2(\mathbb{R}^n), \quad G \in W_1^3(\mathbb{R}^n). \quad (28)$$

Here and below  $C$  stands for a generic positive constant, which can take different values in different lines. The operator *mod* (see its exact definition in [64]) maps any symmetric matrix to its suitably defined “positive-semidefinite part”. Observe also that if

$$g \in W_{\infty,loc}^2(0, +\infty), \quad g(s) \geq 0, \quad |g'(s)| = O(s),$$

$$|g'(s)| \leq C(1 + g(s))^{3/2}, \quad |g''(s)| + \frac{|g'(s)|}{s} \leq C(1 + g(s))^2, \quad (29)$$

then the required conditions for  $h$  are satisfied.

**Definition 1.** A function  $u$  from the space

$$C([0, T] \times \mathbb{R}^n) \cap L_\infty(0, T, W_\infty^1(\mathbb{R}^n)) \quad (30)$$

is a viscosity sub-/supersolution to (19), (22), (23) if, for any  $\phi \in C^2([0, T] \times \mathbb{R}^n)$  and any point  $(t_0, x_0) \in (0, T] \times \mathbb{R}^n$  of local maximum/minimum of the function  $u - \phi$ , one has

$$\frac{\partial \phi}{\partial t} - \text{div} \left( \frac{\varphi_y(x, |\nabla \phi|) \nabla \phi}{1 + K g(|G * \nabla u|)} \right) \leq 0 / \geq 0, \quad (31)$$

and equalities (22), (23) hold in the classical sense. A viscosity solution is a function which is both a subsolution and a supersolution.

**Theorem 1.** *i) Problem (19), (22), (23) has a viscosity solution in class (30) for every positive  $T$ . Moreover,*

$$\inf_{\mathbb{R}^n} u_0 \leq u(t, x) \leq \sup_{\mathbb{R}^n} u_0. \quad (32)$$

*ii) Assume that*

$$\left| \left( \sqrt{a(x, p)} - \sqrt{a(z, p)} \right)_{ij} \right| \leq C|x - z|, \quad x, z, p \in \mathbb{R}^n. \quad (33)$$

Here  $\sqrt{\cdot}$  is the square root of a positive-semidefinite symmetric matrix [42]. Then the solution is unique. Moreover, for any two viscosity solutions  $u$  and  $v$  to (19), (22), the following estimate holds

$$\sup_{\mathbb{R}^n} |u(t, \cdot) - v(t, \cdot)| \leq \Phi(t) \sup_{\mathbb{R}^n} |u(0, \cdot) - v(0, \cdot)| \quad (34)$$

with some non-decreasing continuous scalar function  $\Phi$  dependent on  $u$  and  $v$ .

*Proof. (Sketch)* We follow the strategy of [64, Section 3], where we studied a related problem, and refer to it for further details. Note first that (32) is a direct consequence of the definition of viscosity solution: e.g., to get the second inequality, one can put  $\phi = \delta t$ , to derive that the function  $u(t, x) - \delta t$  attains its global maximum at  $t = 0$ , and to let  $\delta \rightarrow +0$ .

Now, we need to formally establish a Bernstein estimate for  $\sup_{\mathbb{R}^n} |\nabla u|$ . Consider the formal parabolic operator

$$\begin{aligned} \mathcal{L} = & \frac{\partial}{\partial t} - h(u * \nabla G) a_{ij}(x, \nabla u) \frac{\partial^2}{\partial x_i \partial x_j} \\ & - h(u * \nabla G) \frac{\partial a_{ij}(x, \nabla u)}{\partial p_l} u_{x_i x_j} \frac{\partial}{\partial x_l} - h(u * \nabla G) \varphi_{yx_l}(x, |\nabla u|) \frac{\partial}{\partial x_l} \\ & - h(u * \nabla G) \frac{\varphi_{yyx_i}(x, |\nabla u|) u_{x_l} u_{x_i}}{|\nabla u|} \frac{\partial}{\partial x_l} \\ & - \nabla h(u * \nabla G) \cdot \left( u * \frac{\partial \nabla G}{\partial x_i} \right) \varphi_y(x, |\nabla u|) \frac{\partial}{\partial x_i} \\ & - \nabla h(u * \nabla G) \cdot \left( u * \frac{\partial \nabla G}{\partial x_i} \right) \frac{\varphi_{yy}(x, |\nabla u|) u_{x_l} u_{x_i}}{|\nabla u|} \frac{\partial}{\partial x_l}. \end{aligned} \quad (35)$$

Fix  $T$ . Differentiating (19) with respect to each  $x_k$ ,  $k = 1, \dots, n$ , multiplying by  $2u_{x_k}$ , and adding the results, we get

$$\begin{aligned} \mathcal{L}(|\nabla u|^2) = & -2h(u * \nabla G) a_{ij}(x, \nabla u) u_{x_k x_i} u_{x_k x_j} + 2\nabla h(u * \nabla G) \cdot \left( u * \frac{\partial \nabla G}{\partial x_k} \right) a_{ij}(x, \nabla u) u_{x_i x_j} u_{x_k} \\ & + 2h(u * \nabla G) \frac{\partial a_{ij}(x, \nabla u)}{\partial x_k} u_{x_i x_j} u_{x_k} \\ & + 2\nabla h(u * \nabla G) \cdot \left( u * \frac{\partial \nabla G}{\partial x_k} \right) \varphi_{yx_i}(x, |\nabla u|) u_{x_i} u_{x_k} \\ & + 2h(u * \nabla G) \varphi_{yx_i x_k}(x, |\nabla u|) u_{x_i} u_{x_k} \\ & + 2 \frac{\partial^2 h}{\partial q_j \partial q_l} (u * \nabla G) \left( u * \frac{\partial^2 G}{\partial x_i \partial x_j} \right) \left( u * \frac{\partial^2 G}{\partial x_k \partial x_l} \right) \varphi_y(x, |\nabla u|) u_{x_i} u_{x_k} \\ & + 2\nabla h(u * \nabla G) \cdot \left( u * \frac{\partial^2 \nabla G}{\partial x_i \partial x_k} \right) \varphi_y(x, |\nabla u|) u_{x_i} u_{x_k} \\ & + 2\nabla h(u * \nabla G) \cdot \left( u * \frac{\partial \nabla G}{\partial x_i} \right) \varphi_{yx_k}(x, |\nabla u|) u_{x_i} u_{x_k}. \end{aligned} \quad (36)$$

Using [64, Lemma 1], it is possible to get rid of the second-order terms in the right-hand side, ending up with

$$\mathcal{L}(|\nabla u|^2) \leq C(1 + |\nabla u|^2), \quad (37)$$

which by maximum principle implies

$$|\nabla u|^2 \leq C. \quad (38)$$

We can now approximate our problem by well-posed uniformly parabolic problems in the sense of [47, Chapter 5], namely, we need to separate  $h$  and the eigenvalues of  $a(x, p)$  away from zero, and to make

$a(x, p)$  diagonal for large  $|p|$ , simultaneously keeping all involved constants uniformly bounded and the constant from (27) bounded away from zero. We obtain the required viscosity solution by passing to the limit in the viscosity sense, employing the general consistency/stability results from the viscosity solution theory [22] and the Arzelà-Ascoli compactness provided by the Bernstein estimate, cf. [3, 2, 10, 64] with similar considerations. The uniqueness of solutions follows from the stability inequality (34), which may be shown by mimicking the proofs of similar bounds in [3, 2, 64].  $\square$

### 3.2 Dissipative solutions

The concept of *dissipative solution* (see [51, 50, 75, 73, 23, 29, 63] and an illustrative discussion in [72]) allows us to significantly relax the assumptions on  $\varphi$  and  $g$  with respect to the viscosity solution case.

In this subsection we use Neumann's boundary condition. The Dirichlet boundary conditions can also be treated with some technical adjustments. Let  $\Omega$  be a bounded open domain in  $\mathbb{R}^n$ ,  $n \in \mathbb{N}$ , with a regular boundary  $\partial\Omega$ . We thus consider (19), (23) to be coupled with

$$\frac{\partial u}{\partial \nu} = 0, \quad x \in \partial\Omega. \quad (39)$$

The symbol  $\|\cdot\|$  will stand for the Euclidean norm in  $L_2(\Omega)$ . The corresponding scalar product will be denoted by parentheses  $(\cdot, \cdot)$ . We will also use this notation for duality between  $L_p(\Omega)$  and  $L_{p/p-1}(\Omega)$ .

We assume that for every natural number  $N$ , the functions  $\varphi, \varphi_y$  are continuous and bounded on  $\Omega \times (1/N, N)$ . The parabolicity conditions (24)–(27) are replaced by a weaker one:

$$(\varphi_y(x, |p_1|)p_1 - \varphi_y(x, |p_2|)p_2) \cdot (p_1 - p_2) \geq 0, \quad x \in \Omega, \quad p_1, p_2 \in \mathbb{R}^n \setminus \{0\}. \quad (40)$$

We also put weaker assumptions on  $g, h$  and  $G$ . Namely,  $h$  is merely needed to be Lipschitz, which holds, e.g., if  $g$  is non-negative, locally Lipschitz and  $|g'| \leq C(1 + g)^2$ , whereas  $G$  should be of class  $W_2^1(\mathbb{R}^n)$ .

We point out that  $G * \nabla u$  means the convolution  $\nabla G * \tilde{u}$ , where  $\tilde{u}$  is an appropriate linear and continuous extension<sup>2</sup> of  $u$  onto  $\mathbb{R}^n$  which may depend on the boundary condition (cf. [17]).

Introduce the following formal expression

$$\Phi(v)(t, x) = \frac{\varphi_y(x, |\nabla v(t, x)|) \nabla v(t, x)}{1 + K g(|G * \nabla v|(t, x))}. \quad (41)$$

**Definition 2.** Let  $u_0 \in L_2(\Omega)$ . A function  $u$  from the class

$$u \in C_w([0, \infty); L_2(\Omega)) \cap L_1(0, \infty; W_1^1(\Omega)) \quad (42)$$

---

<sup>2</sup>The simplest possible extension is letting  $\tilde{u}$  to be zero outside of  $\Omega$ . Another option is to use Hestenes-Seeley-like extensions [1] which conserve the Sobolev class of  $u$ .

is called a dissipative solution to problem (19), (23), (39) if, for all regular<sup>3</sup> functions  $v : [0, \infty) \times \overline{\Omega} \rightarrow \mathbb{R}$  satisfying the Neumann boundary condition (39) and all non-negative moments of time  $t$ , one has

$$\begin{aligned} & \|u(t) - v(t, \cdot)\|^2 \\ & \leq \gamma^t \|u_0 - v(0, \cdot)\|^2 - \int_0^t 2\gamma^{t-s} \left[ (\Phi(v(s, \cdot)), \nabla u(s) - \nabla v(s, \cdot)) + \left( \frac{\partial v(s, \cdot)}{\partial s}, u(s) - v(s, \cdot) \right) \right] ds, \end{aligned} \quad (43)$$

where  $\gamma$  is a certain constant<sup>4</sup> depending on  $g$ ,  $G$ ,  $\varphi$  and  $v$  (in particular,  $\gamma=1$  provided  $g \equiv 0$ ).

Usual argument [72] shows that these dissipative solutions possess the weak-strong uniqueness property (any regular solution is a unique dissipative solution).

**Theorem 2.** *Assume*

$$\lim_{y \rightarrow +\infty} \inf_{x \in \Omega} \varphi_y(x, y)y = +\infty, \quad (44)$$

$$\lim_{y \rightarrow 0} \sup_{x \in \Omega} \varphi_y(x, y)y = 0. \quad (45)$$

Assume also that either we have strong parabolicity, namely,

$$(\varphi_y(x, |p_1|)p_1 - \varphi_y(x, |p_2|)p_2) \cdot (p_1 - p_2) \geq C|p_1 - p_2|^2, \quad x \in \Omega, \quad p_1, p_2 \in \mathbb{R}^n \setminus \{0\}, \quad (46)$$

or better regularity of  $h$  and  $G$ ,

$$h \in W_\infty^2(\mathbb{R}^n), \quad G \in W_2^2(\mathbb{R}^n). \quad (47)$$

Let  $u_0 \in L_2(\Omega)$ . Then there exists a dissipative solution to problem (19), (23), (39).

**Remark 3.** *In the case when (47) but not (46) holds, the test functions for (43) should additionally satisfy the condition*

$$\operatorname{div} (\varphi_y(x, |\nabla v|)\nabla v) \in L_\infty(0, +\infty; L_2(\Omega)), \quad (48)$$

which, by the way, automatically holds provided  $\varphi$  is more regular, e.g., satisfies the assumptions (24) – (26) of the previous subsection.

**Remark 4.** *The total variation flow [7] corresponds to the case  $g \equiv 0$ ,  $\varphi(x, y) = \ln y$ , so it satisfies (40) but is ruled out by (44) and (45). We will consider this particular form of  $\varphi$  (with generic  $g$ ) in the next subsection (cf. Remark 10). The existence of dissipative solutions for the total variation flow is an open problem. We however believe that the hypotheses of Theorem 2 may be significantly weakened.*

*Proof of Theorem 2.* To begin with, we formally derive some a priori bounds and inequality (43) for the solutions to problem (19), (23), (39). Firstly, we formally take the  $L_2(\Omega)$ -scalar product of (19) with  $2u(t)$ , and integrate by parts:

$$\frac{d}{dt} \|u\|^2 + \left( \frac{2\varphi_y(x, |\nabla u|)\nabla u}{1 + K g(|G * \nabla u|)}, \nabla u \right) = 0. \quad (49)$$

<sup>3</sup>Here “regular” means that  $v$  and  $\nabla v$  are uniformly bounded and sufficiently smooth, and  $|\nabla v| \neq 0$  a.e. in  $(0, \infty) \times \Omega$ .

<sup>4</sup>The exact expression for  $\gamma$  follows from the proof below.

Since the second term is non-negative due to (40) and (45), (49) a priori implies that

$$\|u\|_{L_\infty(0,+\infty;L_2)} \leq \|u_0\|. \quad (50)$$

Thus,

$$|G * \nabla u| \leq \|\nabla G\| \|\tilde{u}\| \leq C \|u\| \leq C. \quad (51)$$

Consider the scalar function  $\Psi(y) = \inf_{x \in \Omega} \varphi_y(x, y) y^2$ ,  $y > 0$ ,  $\Psi(0) = 0$ . From (49) and (51) we can conclude that

$$\int_0^{+\infty} \int_\Omega \Psi(|\nabla u|) dx dt \leq C \|u_0\|. \quad (52)$$

The function  $\Psi(y)$  is non-negative, continuous (for  $y = 0$  this follows from (45), for positive  $y$  it can be derived from the compactness of  $\bar{\Omega}$ ) and satisfies the condition  $\lim_{y \rightarrow +\infty} \Psi(y)/y = +\infty$ . By the Vallée-Poussin criterion [26],  $\nabla u$  a priori belongs to a certain weakly compact set in  $L_1(0, T; L_1)$  for any  $T > 0$ .

Fix a regular test function  $v : [0, \infty) \times \bar{\Omega} \rightarrow \mathbb{R}$  satisfying the Neumann boundary condition (39). Adding (19) with the identity

$$-\frac{\partial v}{\partial t} = -\operatorname{div} \left( \frac{\varphi_y(x, |\nabla v|) \nabla v}{1 + K g(|G * \nabla v|)} \right) + \operatorname{div} \Phi(v) - \frac{\partial v}{\partial t},$$

which can be understood, e.g., in the sense of distributions, and formally multiplying by  $2[u(t) - v(t)]$  in  $L_2(\Omega)$ , we find

$$\begin{aligned} \frac{d}{dt} \|u - v\|^2 + 2 \left( \frac{\varphi_y(x, |\nabla u|) \nabla u - \varphi_y(x, |\nabla v|) \nabla v}{1 + K g(|G * \nabla u|)}, \nabla(u - v) \right) \\ = 2 ([h(G * \nabla v) - h(G * \nabla u)] \varphi_y(x, |\nabla v|) \nabla v, \nabla(u - v)) \\ - 2 (\Phi(v), \nabla u - \nabla v) - 2 \left( \frac{\partial v}{\partial t}, u - v \right). \end{aligned} \quad (53)$$

If (46) holds, then, due to (51), (45) and boundedness of  $\Omega$ , we have

$$\begin{aligned} 2 ([h(G * \nabla v) - h(G * \nabla u)] \varphi_y(x, |\nabla v|) \nabla v, \nabla(u - v)) - 2 \left( \frac{\varphi_y(x, |\nabla u|) \nabla u - \varphi_y(x, |\nabla v|) \nabla v}{1 + K g(|G * \nabla u|)}, \nabla(u - v) \right) \\ \leq C \|\nabla h\|_{L_\infty} \|G * \nabla(u - v)\|_{L_\infty} \|\nabla(u - v)\|_{L_1} - C_1 \|\nabla(u - v)\|^2 \\ \leq C \|\nabla G\| \|u - v\| \|\nabla(u - v)\| - C_1 \|\nabla(u - v)\|^2 \leq C \|u - v\|^2, \end{aligned} \quad (54)$$

where  $C_1$  is the doubled constant from (46).



If (47) and (48) hold, then, by virtue of (40), we get

$$\begin{aligned}
& 2([h(G * \nabla v) - h(G * \nabla u)]\varphi_y(x, |\nabla v|)\nabla v, \nabla(u - v)) - 2\left(\frac{\varphi_y(x, |\nabla u|)\nabla u - \varphi_y(x, |\nabla v|)\nabla v}{1 + Kg(|G * \nabla u|)}, \nabla(u - v)\right) \\
& \leq 2(\operatorname{div} [[h(\nabla G * \tilde{u}) - h(\nabla G * \tilde{v})]\varphi_y(x, |\nabla v|)\nabla v], u - v) \\
& = 2([h(\nabla G * \tilde{u}) - h(\nabla G * \tilde{v})] \operatorname{div}[\varphi_y(x, |\nabla v|)\nabla v], u - v) \\
& + 2\left(\left[\nabla h(\nabla G * \tilde{u})\left(\frac{\partial \nabla G}{\partial x_i} * \tilde{u}\right) - \nabla h(\nabla G * \tilde{v})\left(\frac{\partial \nabla G}{\partial x_i} * \tilde{v}\right)\right] \varphi_y(x, |\nabla v|) \frac{\partial v}{\partial x_i}, u - v\right) \\
& \leq C\|\nabla G * (\tilde{u} - \tilde{v})\|_{L^\infty} \|\operatorname{div}[\varphi_y(x, |\nabla v|)\nabla v]\| \|u - v\| \\
& + 2\left([\nabla h(\nabla G * \tilde{u}) - \nabla h(\nabla G * \tilde{v})]\left(\frac{\partial \nabla G}{\partial x_i} * \tilde{u}\right) \varphi_y(x, |\nabla v|) \frac{\partial v}{\partial x_i}, u - v\right) \\
& + 2\left(\nabla h(\nabla G * \tilde{v})\left(\frac{\partial \nabla G}{\partial x_i} * (\tilde{u} - \tilde{v})\right) \varphi_y(x, |\nabla v|) \frac{\partial v}{\partial x_i}, u - v\right) \\
& \leq C\|\nabla G\| \|\operatorname{div}[\varphi_y(x, |\nabla v|)\nabla v]\| \|\tilde{u} - \tilde{v}\| \|u - v\| \\
& + C\|\nabla G\| \|\tilde{u}\| \|\nabla G\|_{W_2^1} \|\tilde{u} - \tilde{v}\| \|u - v\| + C\|\nabla G\|_{W_2^1} \|\tilde{u} - \tilde{v}\| \|u - v\| \\
& \leq C\|u - v\|^2. \quad (55)
\end{aligned}$$

Note that this  $C$  can be set to be zero when  $g \equiv 0$ .

Thus, in both cases, there is  $\gamma > 0$  such that

$$\frac{d}{dt} \|u - v\|^2 \leq (\ln \gamma) \|u - v\|^2 - 2(\Phi(v), \nabla u - \nabla v) - 2\left(\frac{\partial v}{\partial t}, u - v\right). \quad (56)$$

By Gronwall's lemma, we infer (43).

We recall the following abstract observation [69, 78]. Assume that we have two Hilbert spaces,  $X \subset Y$ , with continuous embedding operator  $i : X \rightarrow Y$ , and  $i(X)$  is dense in  $Y$ . The adjoint operator  $i^* : Y^* \rightarrow X^*$  is continuous and, since  $i(X)$  is dense in  $Y$ , one-to-one. Since  $i$  is one-to-one,  $i^*(Y^*)$  is dense in  $X^*$ , and one may identify  $Y^*$  with a dense subspace of  $X^*$ . Due to the Riesz representation theorem, one may also identify  $Y$  with  $Y^*$ . We arrive at the chain of inclusions:

$$X \subset Y \equiv Y^* \subset X^*. \quad (57)$$

Both embeddings here are dense and continuous. Observe that in this situation, for  $f \in Y, u \in X$ , their scalar product in  $Y$  coincides with the value of the functional  $f$  from  $X^*$  on the element  $u \in X$ :

$$(f, u)_Y = \langle f, u \rangle. \quad (58)$$

Such triples  $(X, Y, X^*)$  are called Lions triples.

We will work with the Lions triple  $(H^m(\Omega), L_2(\Omega), (H^m)^*(\Omega))$  where  $m > 1 + \frac{n}{2}$  is a fixed number. Denote by  $A$  the Riesz bijection between the spaces  $H^m$  and  $(H^m)^*$  (which are not identified). We will employ the Sobolev embedding  $H^m \subset C^1$ , which is compact.

Consider the following approximate problem, where the first equality is understood in the sense of the space  $(H^m)^*$ , whereas the second one is in the sense of the space  $L_2$ :

$$\frac{du}{dt} + Q(u) + \epsilon Au = 0, \quad u|_{t=0} = u_0. \quad (59)$$

The operator  $Q : H^m \rightarrow (H^m)^*$ , which respects the Neumann boundary condition, is determined by the duality

$$\langle Q(u), w \rangle = \left( \frac{\varphi_y(\cdot, |\nabla u|) \nabla u}{1 + K g(|G * \nabla u|)}, \nabla w \right), \quad \forall w \in H^m.$$

We do not use a notation for partial time derivative since we treat (59) as an ODE in a Banach space.

The operator  $Q$  is bounded and continuous from  $C^1$  to  $(H^m)^*$ . Note that there is no loss of continuity as  $\nabla u \rightarrow 0$  due to (45). Therefore,  $Q : H^m \rightarrow (H^m)^*$  is a compact operator. This gives opportunity to secure existence of solutions to (59) in the class

$$L_2(0, +\infty; H^m) \cap W_2^1(0, +\infty; (H^m)^*) \cap C_b([0, +\infty); L_2) \quad (60)$$

by an application of the Leray-Schauder degree theory (a systematic approach to parabolic problems of kind (59) may be found, e.g., in [78]).

Repeating the arguments above, we deduce, for a fixed test function  $v$ , that the approximate solutions satisfy the inequality

$$\frac{d}{dt} \|u - v\|^2 \leq (\ln \gamma) \|u - v\|^2 - 2(\Phi(v), \nabla u - \nabla v) - 2 \left( \frac{\partial v}{\partial t}, u - v \right) - 2\epsilon \langle Au, u - v \rangle. \quad (61)$$

An application of Cauchy's inequality yields

$$-2\langle Au, u - v \rangle \leq \frac{1}{2} \langle Av, v \rangle.$$

Consequently, the approximate solutions satisfy inequality (43) up to a term of order  $O(\epsilon)$ .

Due to the observations above, without loss of generality the approximate solutions converge weakly in  $L_1(0, T; W_1^1)$  and weakly-\* in  $L_\infty(0, T; L_2)$  as  $\epsilon \rightarrow 0$  for any  $T > 0$ . This is enough for passing to the limit in inequality (43), since its left-hand side is the only term which is nonlinear in  $u$ , but this term is lower-semicontinuous (we refer to [51, 72] for detailed passages to the limit in some dissipative solution inequalities).  $\square$

### 3.3 Weak and strong solutions

For  $\varphi$  of power and logarithmic growth, we can show existence of weak solutions. Moreover, in the first case the solutions are locally Lipschitz and their gradients are Hölder continuous. We maintain  $\Omega$  to be a bounded open domain in  $\mathbb{R}^n$  with a regular boundary. We assume that  $g : [0, +\infty) \rightarrow [0, +\infty)$  is continuous<sup>5</sup>, and  $G \in W_2^2(\mathbb{R}^n)$ .

---

<sup>5</sup>Thus, for weak solutions,  $g$  is not needed to be locally Lipschitz.

Firstly, let  $\varphi(x, y) = c_1 y^{p-1}$ ,  $p > 1$ ,  $c_1 > 0$ . To simplify the presentation, in the sequel we assume that  $c_1 = 1$ . We keep the Neumann boundary condition, but generalization to the Dirichlet case is straightforward.

**Definition 3.** A function  $u$  from the class

$$u \in C_w([0, T]; L_2(\Omega)) \cap L_p(0, T; W_p^1(\Omega)) \cap W_{p/p-1}^1(0, T; [L_2 \cap W_p^1(\Omega)]^*) \quad (62)$$

is called a weak solution to problem (19), (39) if, for all  $v \in L_2(\Omega) \times W_p^1(\Omega)$  and a.a.  $t \in (0, T)$ , one has

$$\left\langle \frac{du}{dt}, v \right\rangle + \left\langle \frac{(p-1)|\nabla u|^{p-2} \nabla u}{1 + K g(|G * \nabla u|)}, \nabla v \right\rangle = 0. \quad (63)$$

**Theorem 3.** Let  $\varphi(x, y) = y^{p-1}$ ,  $p > 1$ ,  $u_0 \in L_2(\Omega)$ . There exists a weak solution  $u$  to (19), (39) satisfying (23).

*Proof. (Sketch)* Although the operator

$$-\operatorname{div} \left( \frac{|\nabla u|^{p-2} \nabla u}{1 + K g(|G * \nabla u|)} \right)$$

is not monotone, it is still possible to adapt the Minty-Browder technique to prove Theorem 3. The key point is to pass to the limit. Let  $\{u_k\}$  be a sequence of solutions to (59) with  $\epsilon = \epsilon_k \rightarrow 0$ . Since  $\langle Au_k, u_k \rangle \geq 0$ , every  $u_k$  satisfies the a priori estimates (50), (51) and (52) (with  $T$  instead of  $+\infty$ ). We have to prove that its limit  $u$  (in the weak-\* topology of  $L_\infty(0, T; L_2)$ ) is a solution. Estimates (50), (51) and (52) imply that the solutions  $u_k$  belong to a uniformly bounded set in the space (62). Owing to [67, Corollary 4], without loss of generality we may assume that  $u_k \rightarrow u$  in  $C([0, T]; [W_2^1(\Omega)]^*)$ , so the initial condition is preserved by the limit. One can check that the extension operators mentioned in the previous subsection are continuous from  $[W_2^1(\Omega)]^*$  to  $[W_2^1(\mathbb{R}^n)]^*$ . Therefore,  $\tilde{u}_k \rightarrow \tilde{u}$  in  $C([0, T]; [W_2^1(\mathbb{R}^n)]^*)$ . Thus,  $G * \nabla u_k \rightarrow G * \nabla u$  in  $C([0, T]; L_\infty(\Omega))$ . The operator  $G * \nabla : L_2(\Omega) \rightarrow W_\infty^1(\Omega) \subset C(\bar{\Omega})$  is continuous. This implies that  $G * \nabla u_k \rightarrow G * \nabla u$  in  $C([0, T] \times \bar{\Omega})$ . Due to the continuity of the Nemytskii operator on the space of continuous of functions [46], we conclude that  $h(G * \nabla u_k) \rightarrow h(G * \nabla u)$  uniformly on  $[0, T] \times \bar{\Omega}$ . Then we can proceed similarly to the classical monotonicity argument [25, 28, 49] but with necessary changes.  $\square$

**Remark 5.** In a similar way, Theorem 3 may be generalized onto the case of more general  $\varphi(x, y)$  with growth as  $y \rightarrow \infty$  and decay as  $y \rightarrow 0$  of order  $|y|^{p-1}$ .

We next obtain the local Lipschitz-regularity of a weak solution, as well as the local Hölder continuity of its gradient.

**Theorem 4.** Assume that  $h$  is Lipschitz and  $u$  is a weak solution to (19), (23), (39), which is locally bounded, together with its gradient. Then there exists  $\alpha \in (0, 1)$  such that for any compact set  $\mathcal{K} \subset$

$(0, T) \times \Omega$  there is  $M > 0$  so that

$$|u(t_1, x_1) - u(t_2, x_2)| \leq M \left( |x_1 - x_2| + \sqrt{|t_1 - t_2|} \right), \quad (t_1, x_1), (t_2, x_2) \in \mathcal{K} \quad (64)$$

and

$$|\nabla u(t_1, x_1) - \nabla u(t_2, x_2)| \leq M \left( |x_1 - x_2| + \sqrt{|t_1 - t_2|} \right)^\alpha, \quad (t_1, x_1), (t_2, x_2) \in \mathcal{K}. \quad (65)$$

**Remark 6.** For the degenerate case  $p > 2$  the behaviour of solutions is a purely local fact and the local boundedness follows for any local weak solution. On the contrary, in the singular case  $1 < p < 2$ , it must be derived from global information and may require extra assumptions. Restricting the values of  $p$  to the range  $(\frac{2n}{n+2}, 2)$  suffices though. Note that for applications in imaging  $n = 2$  and no extra assumption is needed.

**Remark 7.** The constant  $M$  is determined by the parabolic distance from  $\mathcal{K}$  to the parabolic boundary of  $(0, T) \times \Omega$  and by the supremum of  $u$  and  $\nabla u$  on  $\mathcal{K}$ , cf. [24, Chapter IX]. The constant  $\alpha$  depends exclusively on  $p$  and dimension.

*Proof of Theorem 4.* Equation (19) with  $\varphi(x, y) = y^{p-1}$  can be written in the form

$$\frac{\partial u}{\partial t} = \operatorname{div} \left( \tilde{h}(t, x) |\nabla u|^{p-2} \nabla u \right), \quad (66)$$

where  $\tilde{h} = (p-1)h(G * \nabla u)$ . Similarly to the proof of Theorem 3, one shows that  $G * \nabla u$  is continuous on  $[0, T] \times \bar{\Omega}$ . Therefore  $\tilde{h} = (p-1)h(G * \nabla u)$  is continuous and bounded from below by a positive constant. Moreover, since  $h$  is Lipschitz, the partial derivatives  $\frac{\partial \tilde{h}}{\partial x_i} = (p-1)\nabla h(G * \nabla u) \cdot \left( \frac{\partial \nabla G}{\partial x_i} * \tilde{u} \right)$  are bounded. Thus, the structure conditions of [24, Chapter VIII, pp. 217–218] are fulfilled. The local regularity of the solution now follows from the general results of [24].

Concerning the optimal Lipschitz regularity of the weak solution we invoke the results of [14], with  $p$  constant. Indeed, condition (7) on page 912 of [14] holds since  $\tilde{h}$  is bounded above and below by positive constants and Lipschitz in space. Moreover,  $\tilde{h}$  is Hölder continuous in time because the same holds for the nonlocal term  $G * \nabla u$ , due to the fact that  $u$  is Hölder continuous in time up to the lateral boundary  $\partial\Omega$  (see [24, Chapter III]).  $\square$

Now we treat the logarithmic growth case. Again, just to simplify the presentation, we merely consider  $\varphi(x, y) = \ln y$ . We restrict ourselves to the Neumann boundary condition. Let  $\mathcal{M}$  be the Banach space of finite signed Radon measures on  $(0, T) \times \Omega$ . It is the dual of the space  $C_0((0, T) \times \Omega)$  (the space of continuous functions on  $(0, T) \times \Omega$  that vanish at the boundary of this cylinder). For  $v \in \mathcal{M}$ , and  $\Phi \in C([0, T] \times \bar{\Omega})$ ,  $\Phi \geq 0$ , we define the *weighted partial variation* of  $v$  as

$$PV_\Phi(v) = \sup_{\psi \in C_0^\infty((0, T) \times \Omega; \mathbb{R}^n): |\psi| \leq \Phi} \langle v, \operatorname{div} \psi \rangle_{\mathcal{M} \times C_0}. \quad (67)$$

Observe that if  $v \in L_1(0, T; W_1^1(\Omega))$ , then  $PV_\Phi(v)$  is equal to  $\int_0^T \int_\Omega \Phi(t, x) |\nabla v(t, x)| dx dt$ .

The *partial variation* of  $v$  is

$$PV(v) = PV_1(v). \quad (68)$$

Define the “bounded partial variation space” *BPV* as

$$\left\{ v \in \mathcal{M} \mid \|v\|_{BPV} := \|v\|_{\mathcal{M}} + PV(v) < +\infty \right\}.$$

Owing to lower semicontinuity of suprema, we have

**Proposition 1.** *For any non-negative function  $\Phi \in C([0, T] \times \bar{\Omega})$  and a weakly- $*$  converging (in  $\mathcal{M}$ ) sequence  $\{v_m\} \subset BPV$ , one has*

$$PV_\Phi(v) \leq \liminf_{m \rightarrow +\infty} PV_\Phi(v_m). \quad (69)$$

Using (69) with  $\Phi \equiv 1$ , we can prove

**Proposition 2.** *BPV is a Banach space.*

**Definition 4.** *A function  $u$  from the class*

$$u \in C_w(0, T; L_2(\Omega)) \cap BPV \cap W_\infty^1(0, T; [L_2 \cap W_1^1(\Omega)]^*) \quad (70)$$

*is called a weak solution to problem (19), (39), (23) with  $\varphi(x, y) = \ln y$  and  $u_0 \in L_2(\Omega)$  if*

*i) there exists  $z \in L_\infty((0, T) \times \Omega; \mathbb{R}^n)$ ,  $\|z\|_{L_\infty((0, T) \times \Omega)} \leq 1$ , so that for all  $v \in L_2 \cap W_1^1(\Omega)$ ,*

$$\left\langle \frac{du}{dt}, v \right\rangle + (h(G * \nabla u)z, \nabla v) = 0, \quad (71)$$

*a.e. on  $(0, T)$ ;*

*ii) for all  $w \in L_1(0, T; W_1^1(\Omega)) \cap W_1^1(0, T; L_2(\Omega))$ , one has*

$$\begin{aligned} \|u(T) - w(T)\|^2 + 2 \int_0^T \left( \frac{dw}{ds}(s), u(s) \right) ds + 2 PV_{h(G * \nabla u)}(u) \\ \leq \|u_0 - w(0)\|^2 + \|w(T)\|^2 - \|w(0)\|^2 + 2 \int_0^T (h(G * \nabla u(s))z(s), \nabla w(s)) ds; \end{aligned} \quad (72)$$

*iii) (23) holds in the space  $L_2(\Omega)$ .*

**Remark 8.** *The motivation for this definition is the following one. Consider, formally, a pair  $(u, z)$  of sufficiently smooth functions satisfying (71), (72), (23). Then*

$$\begin{aligned} 2 \int_0^T \left( \frac{d(u-w)}{ds}(s), (u-w)(s) \right) ds + 2 \int_0^T \left( \frac{dw}{ds}(s), u(s) \right) ds + 2 \int_0^T (h(G * \nabla u(s)), |\nabla u(s)|) ds \\ \leq 2 \int_0^T \left( \frac{dw}{ds}(s), w(s) \right) ds - 2 \int_0^T \left( \frac{du}{ds}(s), w(s) \right) ds. \end{aligned} \quad (73)$$

Hence,

$$\int_0^T (h(G * \nabla u(s)), |\nabla u(s)|) ds \leq - \int_0^T \left( \frac{du}{ds}(s), u(s) \right) ds. \quad (74)$$

Therefore,

$$\int_0^T (h(G * \nabla u(s)), |\nabla u(s)|) ds \leq \int_0^T (h(G * \nabla u(s))z(s), \nabla u(s)) ds. \quad (75)$$

On the other hand,

$$h(G * \nabla u)|\nabla u| \geq h(G * \nabla u)z \cdot \nabla u. \quad (76)$$

All this can be true if and only if

$$|\nabla u| = z \cdot \nabla u, \quad (77)$$

i.e.

$$z = \frac{\nabla u}{|\nabla u|} = \varphi_y(\cdot, |\nabla u|)\nabla u. \quad (78)$$

Substituting expression (78) for  $z$  into (71) and integrating by parts, we infer

$$\int_{\Omega} \left[ \frac{\partial u}{\partial t} - \operatorname{div} (h(G * \nabla u)\varphi_y(x, |\nabla u|)\nabla u) \right] v dx + \int_{\partial\Omega} h(G * \nabla u)\varphi_y(\cdot, |\nabla u|)v \frac{\partial u}{\partial \nu} d\mathcal{H}^{n-1}. \quad (79)$$

Testing (79) by any  $v$  compactly supported in  $\Omega$ , we deduce (19). Consequently, both integrals in (79) are identically zero. By arbitrariness of  $v$ ,  $\mathcal{H}^{n-1}$ -a.e. in  $\partial\Omega$  it holds

$$h(G * \nabla u)\varphi_y(\cdot, |\nabla u|) \frac{\partial u}{\partial \nu} = 0. \quad (80)$$

Since both  $h$  and  $\varphi_y$  are non-vanishing, the Neumann boundary condition (39) is satisfied.

**Theorem 5.** Let  $\varphi(x, y) = \ln y$ ,  $u_0 \in L_2(\Omega)$ . There exists a weak solution to (19), (39), (23).

*Proof.* Consider the approximate problem

$$\frac{du}{dt} + Q_{\epsilon}(u) + \epsilon Au = 0, \quad u|_{t=0} = u_0. \quad (81)$$

The notation and functional framework are the same as in the previous subsection, see the proof of Theorem 2. The operator  $Q_{\epsilon} : H^m \rightarrow (H^m)^*$  is determined by the duality

$$\langle Q_{\epsilon}(u), w \rangle = \left( \frac{\nabla u}{(\epsilon + |\nabla u|)(1 + Kg(|G * \nabla u|))}, \nabla w \right), \quad \forall w \in H^m.$$

The existence of solutions to (81) in class (60) is similar to the existence of solutions to (59) in the previous subsection. Note that, since condition (45) is violated in the logarithmic case, we do not work directly with (59). However, the a priori bounds (50), (51) and

$$\int_0^T \int_{\Omega} \frac{|\nabla u|^2}{\epsilon + |\nabla u|} \leq C \quad (82)$$

(an analogue of (52)) still hold, with  $C$  independent of  $\epsilon$ . Estimate (82) easily yields

$$\int_0^T \int_{\Omega} |\nabla u| \leq C. \quad (83)$$

Let  $\{u_k\}$  be a sequence of solutions to (81) with  $\epsilon = \epsilon_k \rightarrow 0$ . Set

$$z_k = \frac{\nabla u_k}{\epsilon_k + |\nabla u_k|}.$$

Fixing a sufficiently smooth function  $w : [0, T] \times \bar{\Omega} \rightarrow \mathbb{R}$ , testing (81) with  $2(u_k - w)$ , and integrating in time, we deduce

$$\begin{aligned} \|u_k(t) - w(t)\|^2 + 2 \int_0^t \left( \frac{dw}{ds}(s), u_k(s) \right) ds + 2 \int_0^t (h(G * \nabla u_k(s)) z_k(s), \nabla u_k(s)) ds \\ = \|u_0 - w(0)\|^2 + \|w(t)\|^2 - \|w(0)\|^2 + 2 \int_0^t (h(G * \nabla u_k(s)) z_k(s), \nabla w(s)) ds \\ - 2\epsilon_k \int_0^t \langle Au_k(s), u_k(s) - w(s) \rangle ds, \quad t \in [0, T], \end{aligned} \quad (84)$$

whence by Cauchy's inequality

$$\begin{aligned} \|u_k(t) - w(t)\|^2 + 2 \int_0^t \left( \frac{dw}{ds}(s), u_k(s) \right) ds + 2 \int_0^t (h(G * \nabla u_k(s)), |\nabla u_k(s)|) ds \\ \leq \|u_0 - w(0)\|^2 + \|w(t)\|^2 - \|w(0)\|^2 + 2 \int_0^t (h(G * \nabla u_k(s)) z_k(s), \nabla w(s)) ds \\ + 2\epsilon_k \int_0^t \left( h(G * \nabla u_k(s)), \frac{|\nabla u_k(s)|}{\epsilon_k + |\nabla u_k(s)|} \right) ds + \frac{\epsilon_k}{2} \int_0^t \langle Aw(s), w(s) \rangle ds, \quad t \in [0, T]. \end{aligned} \quad (85)$$

Let us prove that the limit  $(u, z) = \lim_{k \rightarrow +\infty} (u_k, z_k)$ , which we without loss of generality assume to exist in the weak-\* topology of  $L_\infty(0, T; L_2) \times L_\infty(0, T; L_\infty)$ , is a weak solution. Estimates (50), (51) and (83) imply that the solutions  $u_k$  belong to a uniformly bounded set in the space (62) with  $p = 1$ . Similarly to the proof of Theorem 3, passing to a further subsequence if necessary, we have that  $u_k \rightarrow u$  in  $C([0, T]; [W_2^1(\Omega)]^*)$ , and  $h(G * \nabla u_k) \rightarrow h(G * \nabla u)$  in  $C([0, T] \times \bar{\Omega})$ . We can pass to the limit in (23) and in the fourth term of the right-hand side of (85). We can also pass to the limit in the first term of inequality (85) without affecting its sign, since this term is lower-semicontinuous (cf. [51, 72] and the proof of Theorem 2). We employ Proposition 1, triangle inequality and the above uniform convergence of  $h(G * \nabla u_k)$  in order to pass to the limit in the third term of (85). The remaining terms of (85) are either linear, or constants, or of order  $O(\epsilon_k)$ , so the passage to the limit is now straightforward, and we get (72).



Figure 6: Synthetic *Shapes* and real *Brain* images used in our experiments. Noise-free (ground-truth) images (left), and noisy images obtained by adding Gaussian noise level  $\sigma_n = 30$  (right).

Testing (81) by a sufficiently smooth  $v : [0, T] \times \bar{\Omega} \rightarrow \mathbb{R}$ , we derive

$$\left\langle \frac{du_k}{dt}, v \right\rangle + (h(G * \nabla u_k)z_k, \nabla v) + \epsilon_k \langle Av, u_k \rangle = 0. \quad (86)$$

Passing to the distributional limit, we obtain (71). Finally, by density, the test functions  $v$  and  $w$  in (71) and (72) can be taken from the spaces indicated in Definition 4.  $\square$

**Remark 9.** *In view of lack of monotonicity/accretivity, uniqueness of weak solutions in Theorems 3 and 5 is an open problem.*

**Remark 10.** *The total variation flow equation ( $g \equiv 0$ ,  $\varphi(x, y) = \ln y$ ) was studied in [7] via monotonicity/accretivity arguments such as the Crandall-Liggett theory. Since the operator*

$$- \operatorname{div} \left( \frac{|\nabla u|^{-1} \nabla u}{1 + K g(|G * \nabla u|)} \right) \quad (87)$$

*is not accretive in any sense (for  $g \neq \text{const}$ ), we had to use another approach. The case  $g \equiv 0$  is not excluded from our results: we have obtained existence of some weak solutions for the total variation flow (with Neumann boundary condition). However, the advantage of the solutions  $u$  obtained in [7] is that they belong to the space of functions of bounded variation  $BV(\Omega)$  for a.a.  $t$ , and it is thus possible to employ Anzellotti's approach [8] to give more direct meaning to the Neumann boundary condition (39). We conjecture that a theory of that kind (traces, Green formula, etc.) may be developed, for instance, in the space  $L_2((0, T) \times \Omega) \cap BPV$ , but for the purposes of this paper we opted to treat the boundary condition (39) in a more implicit way presented in Remark 8.*

## 4 Experimental results and discussions

In what follows, we provide some experimental results using adaptive forward-backward diffusion flows in image restoration. All the images are rescaled to  $[0, 1]$  for visualization and the PDE (18) is discretized



using standard finite difference scheme via additive operator splitting (AOS) [74] with time step  $\tau = 0.2$ , and grid size one. The discretized version of the PDE Eqn. (18) is utilized out using the following unconditionally stable semi-implicit scheme. Let  $h$  be the grid size,  $\tau > 0$  the time step and  $U_{ij}^t$  be the pixel value  $u(i, j)$  at iteration  $t$ . The AOS discretization in 1-D with matrix-vector notation is given by,

$$U^{t+1} = [1 - \tau A(U^t)]^{-1} U^t,$$

where  $A(U^t) = [a_{ij}(U^t)]$  with

$$a_{ij}(U^t) := \begin{cases} \frac{C_i^t + C_j^t}{2h^2}, & j \in \mathcal{N}_i, \\ -\sum_{k \in \mathcal{N}_i} \frac{C_i^t + C_k^t}{2h^2}, & j = i, \\ 0, & \text{otherwise.} \end{cases}$$

Here  $\mathcal{N}_i$  is the set of the two neighbors of location  $i$  (boundary locations have only one neighbor) The values  $C_i$  are the discrete values obtained by evaluating the overall diffusion function,  $\mathcal{C} = \varphi_y(x, |\nabla u|)/1 + K g(|G_\sigma * \nabla u|)$  at location  $i$ . The Gaussian convolution in the inverse mollification is carried out by the discrete version with rotationally symmetric Gaussian lowpass filter of size  $5 \times 5$  with variance 2. For n-D images the modified semi-implicit scheme (AOS) is written as,

$$U^{t+1} = \frac{1}{n} \sum_{l=1}^n [1 - n\tau A_l(U^t)]^{-1} U^t. \quad (88)$$

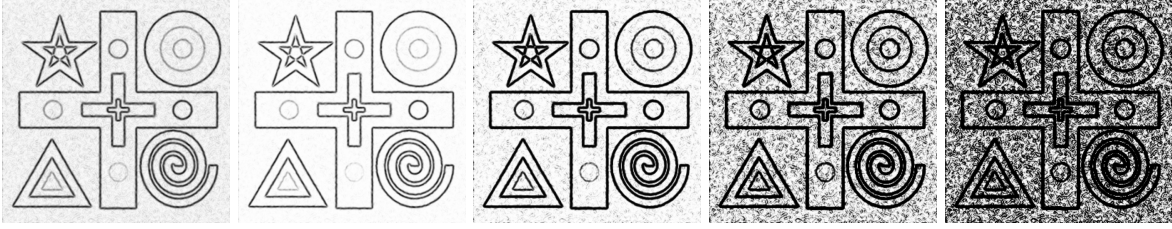
The matrix  $A_l = (a_{ijl})_{ij}$  corresponds to derivatives along the  $l$ -th coordinate axis. Note that this involves solving a linear system where the system matrix is tridiagonal and diagonally dominant, see [74, 59] for more details. To avoid the directional (x-y axis) smoothing bias we adapted a multi-direction based modification [58], and to be fair the same approach was used to compare all the PDEs here. We used a non-optimized MATLAB implementation on a 2.3 GHz Intel Core i7, 8GB 1600 MHz DDR3 Mac Book pro Laptop. The pre-smoothing parameter  $\sigma = 1$  in (18) is fixed for additive Gaussian noise level of standard deviation  $\sigma_n = 30$  used here. Figure 6 shows the synthetic *Shapes* and *Brain* images (noise-free/ground truth) and its corresponding noisy versions which are utilized in our experiments.

#### 4.1 Effect of inverse mollification, contrast parameter

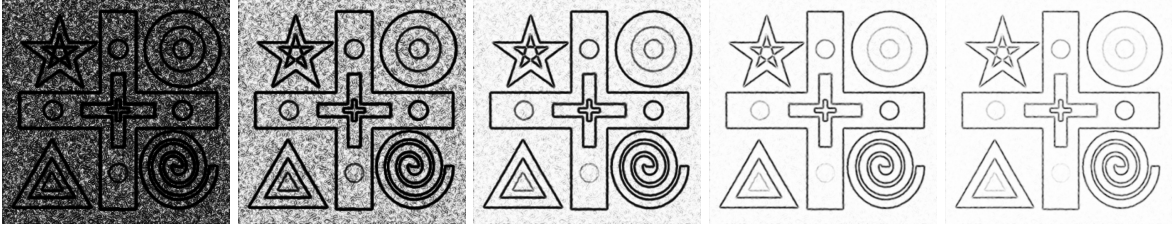
We first consider the effect of inverse mollification function:

$$M(x) = \frac{1}{1 + K g(|G_\sigma * \nabla u(x)|)} \quad (89)$$

in edge detection under noise for different choices of the power growth in  $g(|G_\sigma * \nabla u|) = |G_\sigma * \nabla u|^p$  and  $K$  values. Note that the overall diffusion coefficient acts like an edge detector within PDE based image restoration thereby guiding the smoothing process in and around edges. Figure 7 shows the computed



(a) Power growth for the function  $g(|G_\sigma * \nabla u|) = |G_\sigma * \nabla u|^p$  where  $p = 1, 2, 3, 4, 5$  and  $K = 10^{-4}$  fixed



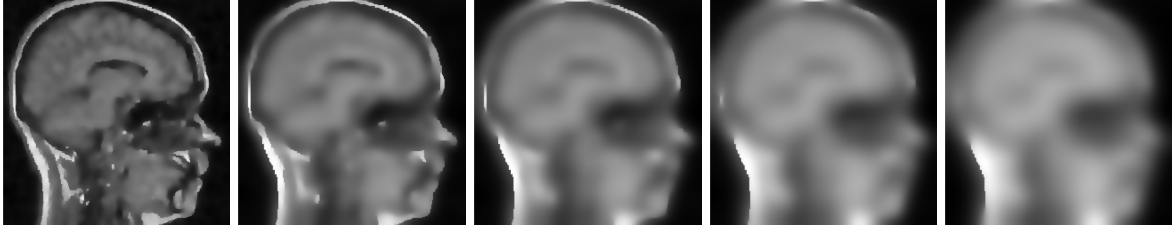
(b) Contrast parameter  $K = 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}$  with  $g(|G_\sigma * \nabla u|) = |G_\sigma * \nabla u|^2$  fixed

Figure 7: Inverse mollification function  $1/(1 + Kg(|G_\sigma * \nabla u|))$  with respect to  $g(\cdot)$  and contrast parameter  $K$  computed using noisy synthetic *Shapes* image (noise standard deviation  $\sigma_n = 30$ , see Figure 6(a) right). We show  $[0, 1]$  normalized  $1/(1 + Kg(|G_\sigma * \nabla u|))$  as an image when: (a)  $g(|G_\sigma * \nabla u|) = |G_\sigma * \nabla u|^p$  for  $p = 1, 2, 3, 4, 5$  with  $K = 10^{-4}$  fixed, and (b)  $K = 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}$  with  $g(|G_\sigma * \nabla u|) = |G_\sigma * \nabla u|^2$  fixed. Better viewed online and zoomed in.

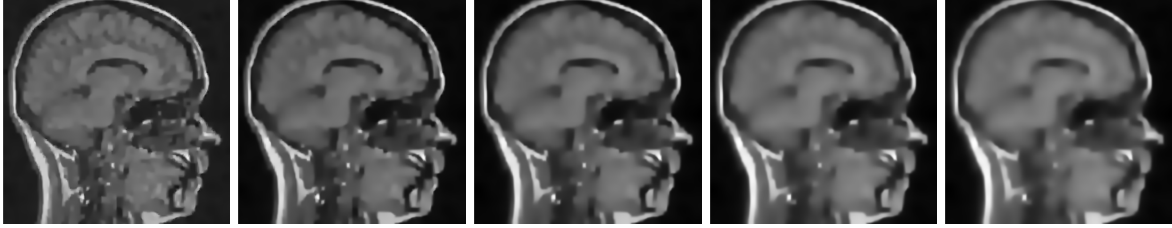
inverse mollification function  $M(x)$  in (89) as images for different values of  $K$  and power growth in  $g(\cdot)$  for the noisy *Shapes* image (noise standard deviation  $\sigma_n = 30$ , see Figure 6(a) right). The noisy image is shown in Figure 6(a) right. Higher growth in functions  $g(|G_\sigma * \nabla u|) = |G_\sigma * \nabla u|^p$ ,  $p > 2$  retains noisy edges and similarly lower  $K$  as well. Moreover, we see that the contrast parameter  $K$  controls the density of edges and can be chosen adaptively [57].

## 4.2 Effect of different exponent values in power growth diffusion

Next, we compare restoration results using PDE Eqn. (18) with and without inverse mollification function (89). We consider the power growth  $\varphi_y(x, |\nabla u|) = |\nabla u|^p$  for different exponent values  $p = 1, 2, 3, 4, 5$  as the diffusion function. Figure 8(a) shows restoration of noisy *Brain* image (noise standard deviation  $\sigma_n = 30$ , see Figure 6(b) right) without inverse mollification function, i.e. taking  $g \equiv 0$ , and Figure 8(b) with  $g(|G_\sigma * \nabla u|) = |G_\sigma * \nabla u|^2$ . As can be seen, taking the non-trivial inverse mollifier function preserves strong edges in the final result when compared to the smoother results for higher exponent  $p$  values.



(a)  $g \equiv 0$

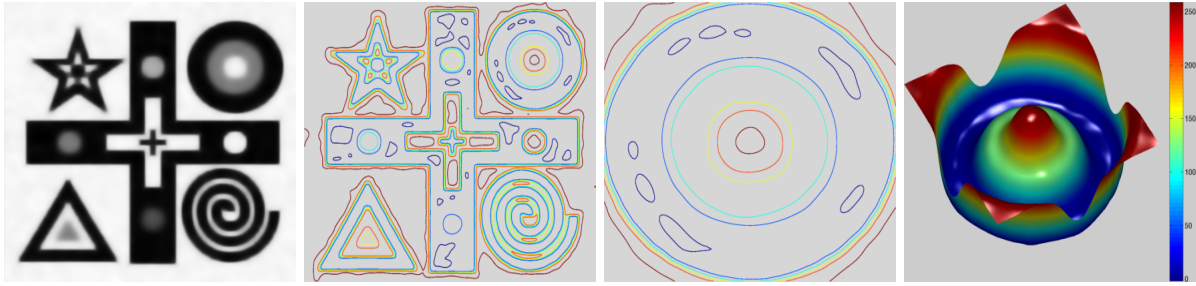


(b)  $g(|G_\sigma * \nabla u|) = |G_\sigma * \nabla u|^2$

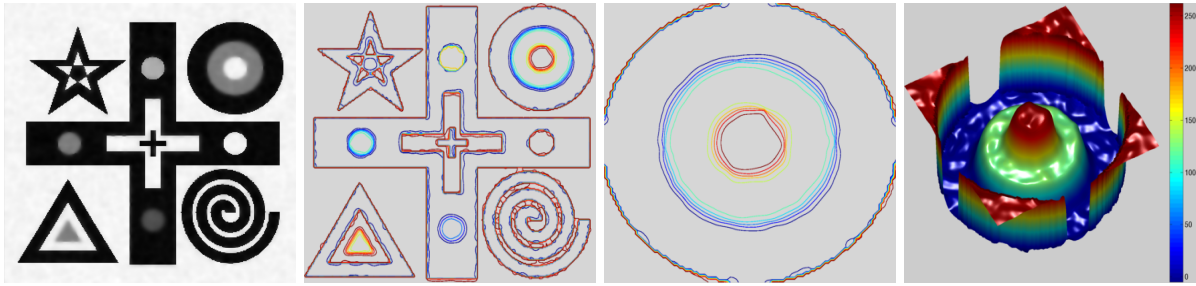
Figure 8: Inverse mollification function when combined with power growth numerator can stop diffusion across edges. Solution of the PDE (18) on noisy synthetic *Brain* image (noise standard deviation  $\sigma_n = 30$ , see Figure 6(b) right) with power growth  $\varphi_y(x, |\nabla u|) = |\nabla u|^p$ ,  $p = 1, 2, 3, 4, 5$  (left to right) without inverse mollification (a)  $g \equiv 0$ , and with (b)  $g(|G_\sigma * \nabla u|) = |G_\sigma * \nabla u|^2$ . In both cases we used  $K = 10^{-4}$  and terminal time 100. It is clear visually that the inverse mollification has an effect in keeping homogenous regions separated by strong edges and avoids leakage.

### 4.3 GRADE Vs our approach

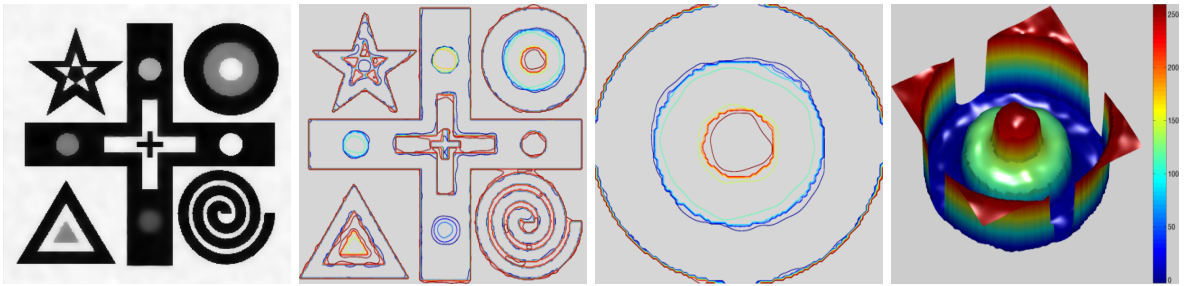
Finally, we provide a comparison of Catté et al. [17] GRADE (3) to illustrate qualitative differences in restoration with our inverse mollification term based PDE (18), with  $g(|G_\sigma * \nabla u|) = |G_\sigma * \nabla u|^2$ , and set  $K = 10^{-4}$ . In our PDE (18), we consider three cases for the diffusion function: Perona-Malik non-convex regularization functions ( $\phi_1, \phi_2$  in (12)), and total variation regularization  $\phi(s) = s$ . Remember that the corresponding  $\varphi$  in (18) may be recovered from the relation  $\varphi_y(x, y)y = \phi'(y)$ . In Figure 9 we show restoration results for the noisy *Shapes* image (noise standard deviation  $\sigma_n = 30$ , see Figure 6(a) right) corresponding to diffusion coefficients  $\varphi_y$  from the Perona-Malik non-convex regularizations (PM1 for  $\phi_1$ , PM2 for  $\phi_2$ ) and the total variation (TV) regularization  $\varphi_y(y) = 1/y$  (the corresponding PDE is  $\frac{\partial u}{\partial t}$  plus the diffusive term (87) equals zero; for technical reasons, we actually employ  $\varphi_y(y) = 1/\sqrt{\epsilon + y^2}$  with  $\epsilon = 10^{-6}$ ). As can be seen, by comparing the contour images, GRADE tends to smooth and displace the level lines of resultant image whereas our adaptive schemes obtain better preservation of level lines in both Perona-Malik and total variation diffusion. Table 1 lists well-known error metrics in the image processing literature for comparing GRADE against our adaptive PDE methods and supports visual comparison results that adaptive schemes are better at retaining structures while removing noise. Deeper



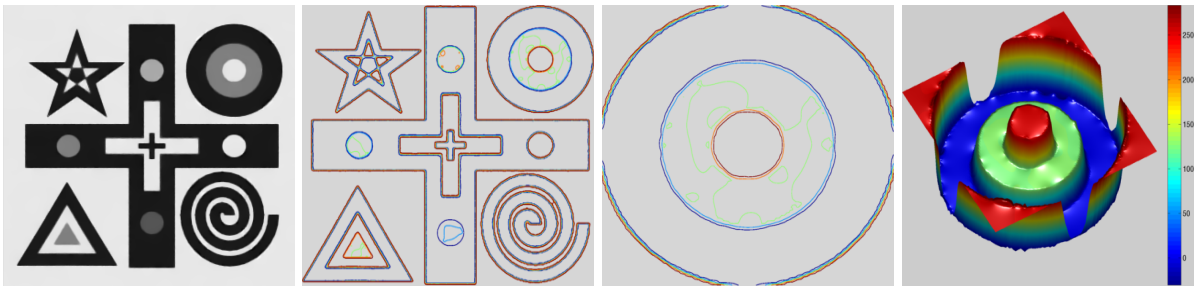
(a) GRADE



(b) Our-PM1



(c) Our-PM2



(d) Our-TV

Figure 9: GRADE versus our approach on noisy synthetic *Shapes* image (noise standard deviation  $\sigma_n = 30$ , see Figure 6(a) right). Results with: (a) GRADE [17], and our adaptive PDE with main diffusion function (b) PM1  $\phi_1$  in (12), (c) PM2  $\phi_2$  in (12) (d) TV. In each row (i|ii|iii|iv): we show (i) final denoised results, (ii) contours from final denoised results, (iii) close-up of the contour map, and (iv) close-up surface. In PM1, PM2 we used  $K = 10^{-4}$ , with terminal time  $T = 50$  and in TV terminal time  $T = 200$ .

Metric	Noisy	GRADE	Our-PM1	Our-PM2	Our-TV
ISNR	0	1.1120	4.6094	<b>8.4478</b>	5.824
SNR	16.1925	17.3044	24.3237	24.6402	<b>31.5589</b>
PSNR	18.5571	19.6691	26.6883	27.0049	<b>33.9235</b>
MSSIM	0.3333	0.7177	0.8776	0.8529	<b>0.9591</b>
MSE	906.4984	701.7305	139.395	129.5954	<b>26.3467</b>
RMSE	30.1081	26.4902	11.8066	11.384	<b>5.1329</b>
MAE	24.0084	14.7066	5.5361	6.2528	<b>2.2002</b>
MAX	132.8668	208.9502	150.7193	159.5855	<b>126.144</b>

Table 1: Error metrics comparison for PDE based smoothing results on noisy synthetic *Shapes* image obtained with noise standard deviation  $\sigma_n = 30$ . Higher improved signal to noise ratio (ISNR), signal to noise ratio (SNR), peak signal to noise ratio (PSNR), mean structural similarity (MSSIM) indicate better restoration whereas lower mean squared error (MSE), root mean square error (RMSE), maximum absolute error (MAE), maximum absolute difference (MAX) indicate better restoration.

quantitative analysis of the numerical examples is deferred to an upcoming work and we refer to [60] for an earlier attempt in this direction with a strictly convex regularization.

## Acknowledgment

JMU and DV partially supported by FCT project PTDC/MAT-CAL/0749/2012 and by CMUC, funded by the European Regional Development Fund, through the program COMPETE, and by the Portuguese Government, through the FCT, under the project PEst-C/MAT/UI0324/2011. DV was supported by the project 14-21-00066 of Russian Science Foundation. Part of this work was done while the first author VBSP was visiting the Institute for Pure and Applied Mathematics (IPAM), University of California, Los Angeles, CA, USA. The authors thank José Alberto Iglesias Martínez (Computational Science Center, University of Vienna, Austria) for discussions related to the regularization - diffusion flows. Authors thank the anonymous referees and the handling editor for their comments which helped to improve the presentation and content of the paper.

## References

- [1] R. Adams. *Sobolev spaces*, volume 65. Academic Press, New York, NY, USA, 1975. Pure and Applied Mathematics.

- [2] L. Alvarez and J. Esclarín. Image quantization using reaction-diffusion equations. *SIAM Journal on Applied Mathematics*, 57(1):153–175, 1997.
- [3] L. Alvarez, P. L. Lions, and J.-M. Morel. Image selective smoothing and edge detection by nonlinear diffusion II. *SIAM Journal on Numerical Analysis*, 29(3):845–866, 1992.
- [4] H. Amann. Time-delayed Perona-Malik type problems. *Acta Mathematica Universitatis Comenianae*, LXXVI(1):15–38, 2007.
- [5] F. Andreu, C. Ballester, V. Caselles, and J. M. Mazón. The Dirichlet problem for the total variation flow. *Journal of Functional Analysis*, 180(2):347–403, 2001.
- [6] F. Andreu, V. Caselles, J. I . Díaz, and J. M. Mazón. Some qualitative properties for the total variation flow. *Journal of functional analysis*, 188(2):516–547, 2002.
- [7] F. Andreu-Vaillo, V. Caselles, and J. M. Mazón. *Parabolic quasilinear equations minimizing linear growth functionals*, volume 223 of *Progress in Mathematics*. Birkhäuser Verlag, Basel, 2004.
- [8] G. Anzellotti. Pairings between measures and bounded functions and compensated compactness. *Annali di Matematica Pura ed Applicata*, 135(1):293–318, 1983.
- [9] G. Aubert and P. Kornprobst. *Mathematical problems in image processing: Partial differential equation and calculus of variations*. Springer-Verlag, New York, USA, 2006.
- [10] C. A. Z. Barcelos and Y. Chen. Heat flows and related minimization problem in image restoration. *Computers & Mathematics with Applications*, 39(5–6):81–97, 2000.
- [11] A. Belahmidi. Solvability of a coupled system arising in image and signal processing. *African Diaspora Journal of Mathematics*, 3(1):45–61, 2005.
- [12] A. Belahmidi and A. Chambolle. Time-delay regularization of anisotropic diffusion and image processing. *Mathematical Modelling and Numerical Analysis*, 39(2):231–251, 2005.
- [13] G. Bellettini and G. Fusco. The  $\Gamma$ -limit and the related gradient flow for singular perturbation functionals of Perona-Malik type. *Transactions of the American Mathematical Society*, 360(9):4929–4987, 2008.
- [14] Verena Bögelein and Frank Duzaar. Hölder estimates for parabolic  $p(x, t)$ -Laplacian systems. *Math. Ann.*, 354(3):907–938, 2012.
- [15] M. Bonforte and A. Figalli. Total variation flow and sign fast diffusion in one dimension. *Journal of Differential Equations*, 252(8):4455–4480, 2012.



- [16] A. Brook, R. Kimmel, and N. Sochen. Variational restoration and edge detection for color images. *Journal of Mathematical Imaging and Vision*, 18(3):247–268, 2003.
- [17] V. Catte, P. L. Lions, J.-M. Morel, and T. Coll. Image selective smoothing and edge detection by nonlinear diffusion. *SIAM Journal on Numerical Analysis*, 29(1):182–193, 1992.
- [18] A. Chambolle and P. L. Lions. Image recovery via total variation minimization and related problems. *Numerische Mathematik*, 76(2):167–188, 1997.
- [19] Y. Chen, S. Levine, and M. Rao. Variable exponent, linear growth functionals in image restoration. *SIAM Journal on Applied Mathematics*, 66(4):1383–1406, 2006.
- [20] Y. Chen and T. Wunderli. Adaptive total variation for image restoration in BV space. *Journal of Mathematical Analysis and Applications*, 272(3):117–137, 2002.
- [21] Y. Chen and K. W. Zhang. Young measure solutions of the two-dimensional Perona-Malik equation in image processing. *Communications on Pure and Applied Analysis*, 5(3):617–637, 2006.
- [22] M. G. Crandall, H. Ishii, and P.-L. Lions. User’s guide to viscosity solutions of second order partial differential equations. *American Mathematical Society Bulletin New Series*, 27(1):1–67, 1992.
- [23] C. De Lellis and L. Székelyhidi, Jr. On admissibility criteria for weak solutions of the Euler equations. *Arch. Ration. Mech. Anal.*, 195(1):225–260, 2010.
- [24] E. DiBenedetto. *Degenerate parabolic equations*. Universitext. Springer-Verlag, New York, 1993.
- [25] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, USA, second edition, 2010.
- [26] E. Feireisl and D. Pražák. *Asymptotic behavior of dynamical systems in fluid mechanics*, volume 4 of *AIMS Series on Applied Mathematics*. American Institute of Mathematical Sciences (AIMS), Springfield, MO, 2010.
- [27] T. Fukui and Y. Giga. Motion of a graph by nonsmooth weighted curvature. In *Proceedings of the First World Congress of Nonlinear Analysis*, volume 1, pages 47–56, 1996.
- [28] H. Gajewski, K. Gröger, and K. Zacharias. *Nichtlineare Operatorgleichungen und Operatordifferentialgleichungen*. Akademie-Verlag, Berlin, 1974. Mathematische Lehrbücher und Monographien, II. Abteilung, Mathematische Monographien, Band 38.
- [29] I. M. Gamba, V. Panferov, and C. Villani. Upper Maxwellian bounds for the spatially homogeneous Boltzmann equation. *Archive for Rational Mechanics and Analysis*, 194(1):253–282, 2009.

- [30] S. Geman and D. McClure. Statistical methods for tomographic image reconstruction. In *In Proceedings of the 46-th Session of the ISI, Bulletin of the ISI*, volume 52, pages 22–26, 1987.
- [31] M.-H. Giga and Y. Giga. Evolving graphs by singular weighted curvature. *Archive for rational mechanics and analysis*, 141(2):117–198, 1998.
- [32] G. Gilboa, N. Sochen, and Y. Y. Zeevi. Variational denoising of partly textured images by spatially varying constraints. *IEEE Transactions on Image Processing*, 15(8):2281–2289, 2006.
- [33] J.B. Greer and A.L. Bertozzi. Traveling wave solutions of fourth order PDEs for image processing. *SIAM Journal on Mathematical Analysis*, 36(1):38–68, 2004.
- [34] P. Guidotti. A new nonlocal nonlinear diffusion of image processing. *Journal of Differential Equations*, 246(12):4731–4742, 2009.
- [35] P. Guidotti. A new well-posed nonlinear nonlocal diffusion. *Nonlinear Analysis: Theory, Methods & Applications*, 72(12):4625–4637, 2010.
- [36] P. Guidotti. A backward–forward regularization of the Perona–Malik equation. *Journal of Differential Equations*, 252(4):3226–3244, 2012.
- [37] P. Guidotti. Anisotropic diffusions of image processing from Perona-Malik on. *Advanced Studies in Pure Mathematics*, to appear, 2015.
- [38] P. Guidotti and J. Lambers. Two new nonlinear nonlocal diffusions for noise reduction. *Journal of Mathematical Imaging Vision*, 33(1):25–37, 2009.
- [39] P. Guidotti and K. Longo. Two enhanced fourth order diffusion models for image denoising. *Journal of Mathematical Imaging Vision*, 40(2):188–198, 2011.
- [40] P. Guidotti and K. Longo. Well-posedness for a class of fourth order diffusions for image processing. *Nonlinear Differential Equations and Applications NoDEA*, 18(4):407–425, 2011.
- [41] P. Harjulehto, V. Latvala, and O. Toivanen. A variant of the Geman–McClure model for image restoration. *Journal of Mathematical Analysis and Applications*, 399(2):676–681, 2013.
- [42] R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge University Press, Cambridge, UK, 1985.
- [43] B. Kawohl and N. Kutev. Maximum and comparison principle for one-dimensional anisotropic diffusion. *Mathematische Annalen*, 311(1):107–123, 1998.
- [44] S. Kichenassamy. The Perona-Malik paradox. *SIAM Journal on Applied Mathematics*, 57(5):1328–1342, 1997.



- [45] K. Kielak, P. B. Mucha, and P. Rybka. Almost classical solutions to the total variation flow. *Journal of Evolution Equations*, 13(1):21–49, 2013.
- [46] M. A. Krasnosel'skii. *Topological methods in the theory of nonlinear integral equations*. Translated by A. H. Armstrong; translation edited by J. Burlak. A Pergamon Press Book. The Macmillan Co., New York, NY, USA, 1964.
- [47] O. A. Ladyženskaja, V. A. Solonnikov, and N. N. Ural'ceva. *Linear and quasilinear equations of parabolic type*, volume 23 of *Translated from the Russian by S. Smith. Translations of Mathematical Monographs*. American Mathematical Society, Providence, R.I., USA, 1967.
- [48] X. Li and T. Chen. Nonlinear diffusion with multiple edginess thresholds. *Pattern Recognition*, 27(8):1029–1037, 1994.
- [49] J.-L. Lions. *Quelques méthodes de résolution des problèmes aux limites non linéaires*. Dunod, 1969.
- [50] P.-L. Lions. Compactness in Boltzmann's equation via Fourier integral operators and applications. I, II. *Journal of Mathematics of Kyoto University*, 34(2):391–427, 429–461, 1994.
- [51] P.-L. Lions. *Mathematical topics in fluid mechanics*, volume 1 of *Oxford Lecture Series in Mathematics and its Applications*. Oxford University Press, New York, NY, USA, 1996.
- [52] M. Nikolova. Weakly constrained minimization: application to the estimation of images and signals involving constant regions. *Journal of Mathematical Imaging and Vision*, 21(2):155–175, 2004.
- [53] M. Nitzberg and T. Shiotu. Nonlinear image filtering with edge and corner enhancement. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(8):826–833, 1992.
- [54] K. N. Nordstrom. Biased anisotropic diffusion: A unified regularization and diffusion approach to edge detection. *Image and Vision Computing*, 8(4):318–327, 1990.
- [55] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7):629–639, 1990.
- [56] V. B. S. Prasath. A well-posed multiscale regularization scheme for digital image denoising. *International Journal of Applied Mathematics and Computer Science*, 21(4):769–777, 2011.
- [57] V. B. S. Prasath and R. Delhibabu. Automatic contrast parameter estimation in anisotropic diffusion for image restoration. In *Analysis of Images, Social Networks, and Texts*, Yekaterinburg, Russia, April 2014. Communications in Computer and Information Science (Springer) Vol. 436. Eds.: D. I. Ignatov, M. Y. Khachay, N. Konstantinova, A. Panchenko.

- [58] V. B. S. Prasath and J. C. Moreno. Feature preserving anisotropic diffusion for image restoration. In *Fourth National Conference on Computer Vision, Pattern Recognition, Image Processing and Graphics (NCVPRIPG 2013)*, India, December 2013.
- [59] V. B. S. Prasath and J. C. Moreno. On convergent finite difference schemes for variational - PDE based image processing. *Applied Mathematics and Computation*, to appear, 2015.
- [60] V. B. S. Prasath and A. Singh. A hybrid convex variational model for image restoration. *Applied Mathematics and Computation*, 215(10):3655–3664, 2010.
- [61] V. B. S. Prasath and A. Singh. Well-posed inhomogeneous nonlinear diffusion scheme for digital image denoising. *Journal of Applied Mathematics*, 2010:14pp, 2010. Article ID 763847.
- [62] V. B. S. Prasath and A. Singh. An adaptive anisotropic diffusion scheme for image restoration and selective smoothing. *International Journal of Image and Graphics*, 12(1):18pp, 2012.
- [63] V. B. S. Prasath and D. Vorotnikov. On a system of adaptive coupled PDEs for image restoration. *Journal of Mathematical Imaging and Vision*, 48(1):35–52, 2014.
- [64] V. B. S. Prasath and D. Vorotnikov. Weighted and well-balanced anisotropic diffusion scheme for image denoising and restoration. *Nonlinear Analysis: Real World Applications*, 17:33–46, 2014.
- [65] V. B. S. Prasath, D. Vorotnikov, R. Pelapur, S. Jose, G. Seetharaman, and K. Palaniappan. Multiscale Tikhonov-total variation image restoration using spatially varying edge coherence exponent. *IEEE Transactions on Image Processing*, to appear, 2015.
- [66] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60(1–4):259–268, 1992.
- [67] J. Simon. Compact sets in the space  $L^p(0, T; B)$ . *Annali di Matematica Pura ed Applicata. Serie Quarta*, 146:65–96, 1987.
- [68] D. Strong and T. Chan. Edge-preserving and scale-dependent properties of total variation regularization. *Inverse Problems*, 19(6):165–187, 2003.
- [69] R. Temam. *Navier-Stokes equations*, volume 2 of *Studies in Mathematics and its Applications*. North-Holland Publishing Co., Amsterdam, Revised edition, 1979. Theory and numerical analysis, With an appendix by F. Thomasset.
- [70] A.N. Tikhonov and V.Y. Arsenin. *Solutions of Ill-posed Problems*. John Wiley, New York, NY, USA, 1997.

- [71] L. Vese. A study in the BV space of a denoising-deblurring variational problem. *Applied Mathematics & Optimization*, 44(2):131–161, 2001.
- [72] D. Vortnikov. Global generalized solutions for Maxwell-alpha and Euler-alpha equations. *Nonlinearity*, 25(2):309–327, 2012.
- [73] D. A. Vortnikov. Dissipative solutions for equations of viscoelastic diffusion in polymers. *Journal of Mathematical Analysis and Applications*, 339(2):876–888, 2008.
- [74] J. Weickert, B. M. H. Romeny, and M. A. Viergever. Efficient and reliable schemes for nonlinear diffusion filtering. *IEEE Transactions on Image Processing*, 7(3):398–410, 1998.
- [75] J. Wu. Analytic results related to magneto-hydrodynamic turbulence. *Physica D*, 136(3-4):353–372, 2000.
- [76] Y.-L. You, W. Xu, A. Tannenbaum, and M. Kaveh. Behavioral analysis of anisotropic diffusion in image processing. *IEEE Transactions on Image Processing*, 5(11):1539–1553, 1996.
- [77] K. Zhang. Existence of infinitely many solutions for the one-dimensional Perona-Malik model. *Calculus of Variations and Partial Differential Equations*, 26(2):171–199, 2006.
- [78] V. G. Zvyagin and D. A. Vortnikov. *Topological approximation methods for evolutionary problems of nonlinear hydrodynamics*, volume 12 of *de Gruyter Series in Nonlinear Analysis and Applications*. Walter de Gruyter & Co., Berlin, Germany, 2008.