Pré-Publicações do Departamento de Matemática
Universidade de Coimbra
Preprint Number 15–07

ANALYSIS OF ADAPTIVE FORWARD-BACKWARD DIFFUSION FLOWS WITH APPLICATIONS IN IMAGE PROCESSING

V. B. SURYA PRASATH, JOSÉ MIGUEL URBANO AND 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.
AMS Subject Classification (2010): 68U10, 35Q68, 65F22.

1. Introduction

The nonlinear diffusion model, introduced in image processing by Perona and Malik [35], involves solving the following initial-boundary value problem

$$\begin{cases} \dfrac{\partial u(x,t)}{\partial t} = \operatorname{div}\left(\mathcal{C}(|\nabla u(x,t)|^2)\nabla u(x,t)\right) & \text{in } \Omega\times(0,T), \\ \dfrac{\partial u(x,t)}{\partial \nu} = 0 & \text{on } \partial\Omega\times(0,T), \\ u(x,0) = u_0(x) & \text{in } \Omega, \end{cases} \tag{1.1}$$

where $u_0 : \Omega \to \mathbb{R}$ is the observed (noisy) image, $\Omega \subset \mathbb{R}^2$ is a bounded domain with Lipschitz boundary. The function $\mathcal{C} : \mathbb{R}_0^+ \to \mathbb{R}_0^+$ is non-increasing such that $\mathcal{C}(0) = 1$ and $\lim_{s\to+\infty}\mathcal{C}(s) = 0$. Note if $\mathcal{C}(s) \equiv 1$ then we recover

Received March 05, 2015.
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.

the heat equation. The diffusion coefficient function $\mathcal{C}(\cdot)$ in Eqn. (1.1) 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 [35] are

$$\mathcal{C}_1(s) = \exp\left(-\frac{s}{K^2}\right), \qquad \mathcal{C}_2(s) = \frac{1}{1+s/K^2}, \tag{1.2}$$

where $K > 0$ is a tunable parameter also known as the *contrast parameter* [37]. The motivation for the Perona-Malik nonlinear diffusivity is that inside the regions where the magnitude of the gradient of u is weak, equation (1.1) 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. (1.1),

$$\begin{aligned}&\operatorname{div}\left(\mathcal{C}(|\nabla u|^2)\nabla u\right)\\ &\quad = 2(u_x^2 u_{xx} + u_y^2 u_{yy} + 2u_x u_y u_{xy})\,\mathcal{C}'(|\nabla u|^2) + \mathcal{C}(|\nabla u|^2)\,(u_{xx}+u_{yy}).\end{aligned} \tag{1.3}$$

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

$$\mathcal{B}(s) = \mathcal{C}(s) + 2s\mathcal{C}'(s), \tag{1.4}$$

we infer

$$\operatorname{div}\left(\mathcal{C}(|\nabla u|^2)\nabla u\right) = \mathcal{C}(|\nabla u|^2)\,u_{\mathcal{T}\mathcal{T}} + \mathcal{B}(|\nabla u|^2)\,u_{\mathcal{N}\mathcal{N}}. \tag{1.5}$$

We thus see that the Perona-Malik diffusion (1.1) 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

$$\lim_{s\to\infty}\frac{\mathcal{B}(s)}{\mathcal{C}(s)} \leq 0, \text{ or equivalently, } \lim_{s\to\infty}\frac{s\,\mathcal{C}'(s)}{\mathcal{C}(s)} \leq -\frac{1}{2}. \tag{1.6}$$

For example, in the case of the power growth functions

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

V. B. Surya Prasath
Department of Computer Science, University of Missouri-Columbia, MO 65211 USA.
E-mail address: prasaths@missouri.edu

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

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

the above limit gives that

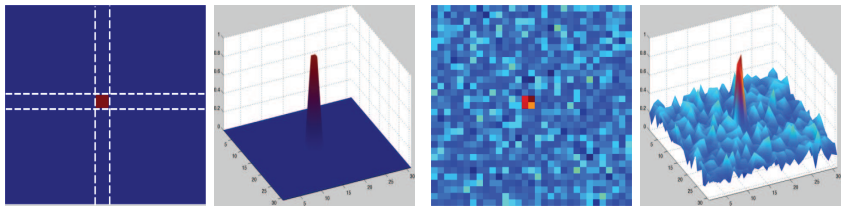
$$q \leq -\frac{1}{2}. \tag{1.7}$$

Note that for the original diffusion function $\mathcal{C}_2$ in (1.2) 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 threshhold value $2^{-1/2}K$. Thus, the PDE (1.1) promotes combined forward-backward diffusion flow, see Figure 1 for a comparison with the heat equation. The Perona-Malik anisotropic diffusion equation (PMADE for short) thus balances forward and backward diffusion regimes using a tunable K, the contrast parameter [29, 37]. These two competing requirements constitute a common theme in many PDE based image restoration models [52, 11, 9, 12]. Moreover, from Eqn. (1.5) we see that Eqn. (1.1) 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 partcular in edge preserving image restoration.

Despite impressive numerical results obtained in image processing using the Perona-Malik equation (1.1), it is known [26, 23] to be an ill-posed PDE due to the degenerate behavior (due to $\mathcal{C}(s) \approx 1/\sqrt{s}$ as $s \to \infty$, see Eqn. (1.6) above). To overcome this issue, Catté et al. [10] 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_σ denotes the two dimensional Gaussian kernel, $G_\sigma(x) = (2\pi\sigma)^{-1} \exp\left(-|x|^2/2\sigma^2\right)$ and $*$ denotes the convolution operation. By isotropic smoothing of scale σ, $|\nabla G_\sigma * u|$ might provide a better estimate of local gradients instead of the noise prone $|\nabla u|$. Moreover, we can relate the Gaussian smoothing $G_\sigma * u_0$ to the quadratic regularization function via linear diffusion (heat equation). Thus, the Catté et al. [10] Gaussian regularized anisotropic diffusion equation (GRADE for short) reads as

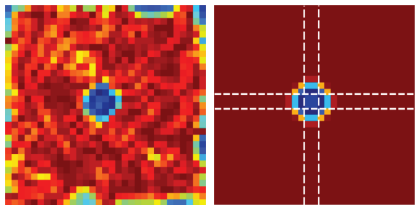
$$\frac{\partial u}{\partial t} = \operatorname{div}\left(\mathcal{C}(|\nabla G_\sigma * u|^2)\nabla u\right). \tag{1.8}$$

This modification of equation (1.1) is sufficient to prove the existence and uniqueness of solution to the initail-boundary value problem for GRADE (1.8). However, the space-invariant Gaussian smoothing inside the divergent term tends to push the edges away from their original locations, see

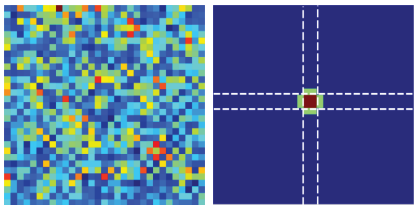


(a) Original

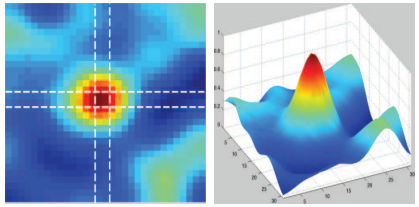
(b) Noisy



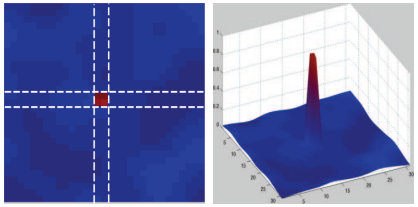
(c) Diffusion coefficient (left) initial (right) at iteration 20



(d) Flux function (left) initial (right) at iteration 20



(e) Heat equation result



(f) Perona and Malik anisotropic diffusion equation result

FIGURE 1. Diffusion process for a simple synthetic image. (a) Original synthetic image of size 31×31, a square (2×2, 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 (1.2), 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 (1.1) with 20 iterations in (left) image (right) surface format. The white dotted lines indicate the *influence* region at the center.

[32] P.-L. Lions. *Mathematical topics in fluid mechanics. Vol. 1*, volume 3 of *Oxford Lecture Series in Mathematics and its Applications*. The Clarendon Press Oxford University Press, New York, NY, USA, 1996. Incompressible models, Oxford Science Publications.

[33] 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.

[34] K. N. Nordstrom. Biased anisotropic diffusion: a unified regularization and diffusion approach to edge detection. *Image and Vision Computing*, 8(4):318–327, 1990.

[35] 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.

[36] 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.

[37] 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.

[38] V. B. S. Prasath and A. Singh. A hybrid convex variational model for image restoration. *Applied Mathematics and Computation*, 215(10):3655–3664, 2010.

[39] 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.

[40] 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.

[41] 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.

[42] 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.

[43] L. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60(1–4):259–268, 1992.

[44] J. Simon. Compact sets in the space $L^p(0,T;B)$. *Annali di Matematica Pura ed Applicata. Serie Quarta*, 146:65–96, 1987.

[45] D. Strong and T. Chan. Edge-preserving and scale-dependent properties of total variation regularization. *Inverse Problems*, 19(6):165–187, 2003.

[46] A.N. Tikhonov and V.Y. Arsenin. *Solutions of Ill-posed Problems*. John Wiley, New York, NY, USA, 1997.

[47] L. Vese. A study in the BV space of a denoising-deblurring variational problem. *Applied Mathematics & Optimization*, 44(2):131–161, 2001.

[48] D. Vorotnikov. Global generalized solutions for Maxwell-alpha and Euler-alpha equations. *Nonlinearity*, 25(2):309–327, 2012.

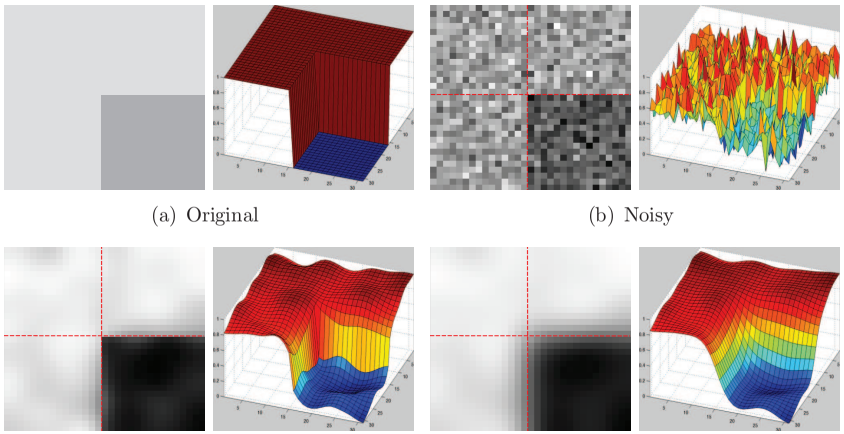
[49] D. A. Vorotnikov. Dissipative solutions for equations of viscoelastic diffusion in polymers. *Journal of Mathematical Analysis and Applications*, 339(2):876–888, 2008.

[50] 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.

[51] J. Wu. Analytic results related to magneto-hydrodynamic turbulence. *Physica D*, 136(3-4):353–372, 2000.

[52] 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.

[11] A. Chambolle and P. L. Lions. Image recovery via total variation minimization and related problems. *Numerische Mathematik*, 76(2):167–188, 1997.
[12] 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.
[13] 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.
[14] 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.
[15] 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.
[16] Emmanuele DiBenedetto. *Degenerate parabolic equations*. Universitext. Springer-Verlag, New York, 1993.
[17] L. C. Evans. *Partial differential equations*, volume 19 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, USA, second edition, 2010.
[18] 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.
[19] 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.
[20] 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.
[21] 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.
[22] 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.
[23] Patrick Guidotti. Anisotropic diffusions of image processing from Perona-Malik on. *Advanced Studies in Pure Mathematics, to appear*, 2015.
[24] 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.
[25] R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge University Press, Cambridge, UK, 1985.
[26] S. Kichenassamy. The Perona-Malik paradox. *SIAM Journal on Applied Mathematics*, 57(5):1328–1342, 1997.
[27] 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.
[28] O. A. Ladyženskaja, V. A. Solonnikov, and N. N. Ural′ceva. *Linear and quasilinear equations of parabolic type*. Translated from the Russian by S. Smith. Translations of Mathematical Monographs, Vol. 23. American Mathematical Society, Providence, R.I., USA, 1967.
[29] X. Li and T. Chen. Nonlinear diffusion with multiple edginess thresholds. *Pattern Recognition*, 27(8):1029–1037, 1994.
[30] J.-L. Lions. *Quelques méthodes de résolution des problèmes aux limites non linéaires*. Dunod, 1969.
[31] 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.



(a) Original (b) Noisy

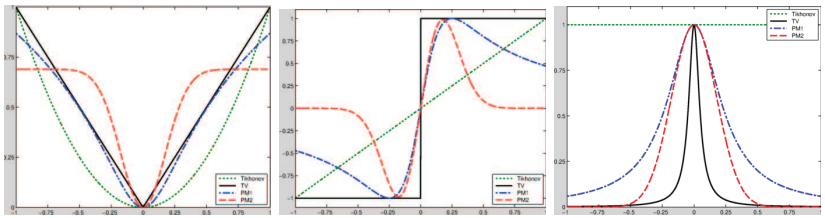
(c) PMADE (1.1) result (d) GRADE (1.8) result

FIGURE 2. Spatial regularization in the diffusion coefficient alters discontinuities in a given image. (a) Original synthetic image of size 31×31, 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 (1.1) with 20 iterations in (left) image (right) surface format (c), and GRADE (1.8) with 20 iterations in (left) image (right) surface format (d). The intersection of red dotted lines indicate the exact corner location of the square.

Figure 2 for an illustration of this effect on a synthetic corner image* Furthermore, the use of isotropic smoothing is against the principle of anisotropic diffusion which aims to smooth homogeneous regions without affecting edges.

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 [5, 38]. Following [47], consider the next

*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 higher order Sobolev space.



(a) ϕ (b) ϕ' (c) $\phi'(s)/s$

FIGURE 3. Well known functions and their qualitative shapes: (a) Regularization (b) Flux (c) Diffusion coefficients. Convex functions: Tikhonov, total variation (TV). Non-convex functions: Perona-Malik regularization, see Eqns. (1.11). Normalized to $[0, 1]$ for visualization.

minimization problem for image restoration:

$$\min_u E(u) = \frac{\beta}{2}\int_\Omega (u_0(x) - u(x))^2\,dx + \alpha\int_\Omega \phi(|\nabla u(x)|)\,dx. \tag{1.9}$$

Here $\alpha \geq 0$ is regularization parameter, $\beta \geq 0$ fidelity parameter, and $\phi : \mathbb{R} \to \mathbb{R}^+$ is an even function, which is sometimes 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 [33].

The formal gradient flow associated with the functional $E(u)$ looks like

$$\frac{\partial u}{\partial t} = \alpha\,\mathrm{div}\left(\frac{\phi'(|\nabla u|)}{|\nabla u|}\nabla u\right) - \beta\,(u - u_0). \tag{1.10}$$

It is easy to see that this equation almost coincides with the Perona-Malik anisotropic PDE (1.1) for $\phi'(s) = \mathcal{C}(s^2)s$, cf. [34], and the difference between the two equations is the lower order term coming from the data fidelity in the regularization functional (1.9). For edge preservation we need to work with functions ϕ with at most linear growth at infinity, cf. (1.7). For example, the Perona-Malik diffusion coefficients (1.2), up to multiplicative constants, correspond to the following non-convex regularization functions (see Figure 3(a),

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

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.

References

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

[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] 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.

[5] G. Aubert and P. Kornprobst. *Mathematical problems in image processing: Partial differential equation and calculus of variations*. Springer-Verlag, New York, USA, 2006.

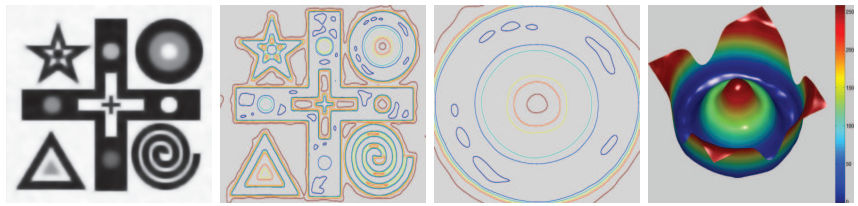
[6] C. A. Z. Barcelos and Y. Chen. Heat flows and related minimization problem in image restoration. *Computers & Mathematics with Applications*, 39(5–6):81–97, 2000.

[7] Michael Bildhauer and Martin Fuchs. A variational approach to the denoising of images based on different variants of the tv-regularization. *Applied Mathematics and Optimization*, page 31pp, 2012.

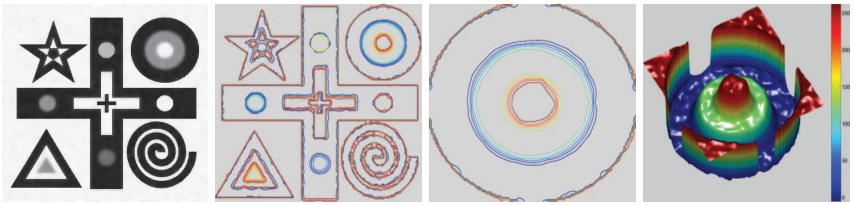
[8] Verena Bögelein and Frank Duzaar. Hölder estimates for parabolic $p(x,t)$-Laplacian systems. *Math. Ann.*, 354(3):907–938, 2012.

[9] 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.

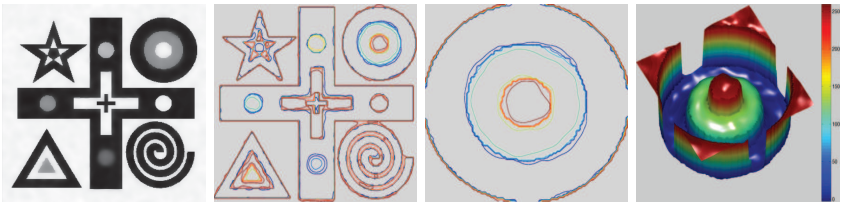
[10] V. Catte, P. L. Lions, J.-M. Morel, and T. Coll. Image selective smoothing and edge detection by nonlinear diffusion. *SIAM Journal on Numerical Analysis*, 29(1):182–193, 1992.



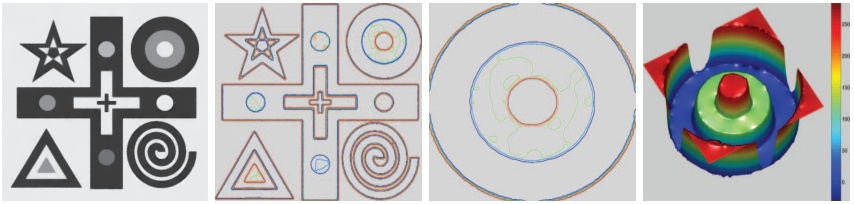
(a) GRADE



(b) Our-PM1



(c) Our-PM2



(d) Our-TV

FIGURE 9. GRADE Vs our approach on noisy synthetic *Shapes* image (noise standard deviation $\sigma_n = 30$, see Figure 6(a) right). (a) GRADE Our adaptive PDE with main diffusion function (b) PM1 ϕ_1 in (1.11), (c) PM2 ϕ_2 in (1.11) (d) TV ϕ. 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 $T = 200$.

denoted as PM1 and PM2),

$$\phi_1(|\nabla u(x)|) = 1 - \exp\left(-\left(\frac{|\nabla u(x)|}{K}\right)^2\right),$$

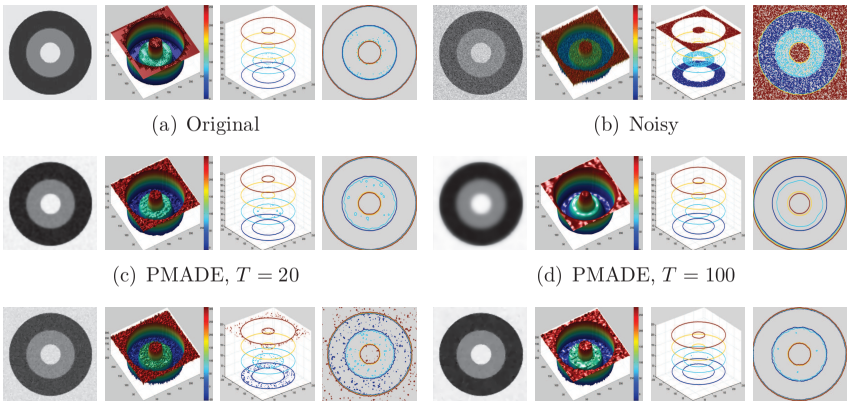
$$\phi_2(|\nabla u(x)|) = \log\left(1 + \left(\frac{|\nabla u(x)|}{K}\right)^2\right). \quad (1.11)$$

Several papers [12, 38, 7, 24] studied similar diffusion flows in the past, with an aim to have a regularizing function that will be linear near the edges, and quadratic away from them.

Motivated by the correspondence between the variational and PDE methods for imaging problems, in this paper we consider a Perona-Malik type PDE with the generic diffusion function inspired by the stationary regularization approach (1.9). Engrafting a mollified gradient within the diffusion function we obtain a general forward-backward diffusion PDE. Inspired by an adaptive regularization model [45], here we consider regularization functions of the type $\phi(x, |\nabla u|)$ along with a Catté et al's [10] mollified gradient approach. 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 reviews some classical regularization functions, and then motivates and presents our general forward-backward diffusion PDE. Section 3 examines the conditions needed for wellposedness of the proposed regularization-inspired forward-backward PDE in various scenarios. Section 4 we provide numerical experiments to prove the effectiveness of the proposed multi-scale scheme as well as examples for various cases.



(a) Original (b) Noisy

(c) PMADE, $T = 20$ (d) PMADE, $T = 100$

(e) TV-PDE, $T = 100$ (f) TV-PDE, $T = 200$

FIGURE 4. Advantages and disadvantages of TV-PDE and PMADE (with $\mathcal{C}_1$ as the diffusion coefficient and $K = 20$) models on a synthetic piecewise constant *Circles* image. Noisy image is obtained by adding Gaussian noise with variance $\sigma_n = 30$. $|i|ii|iii|iv|$: 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.

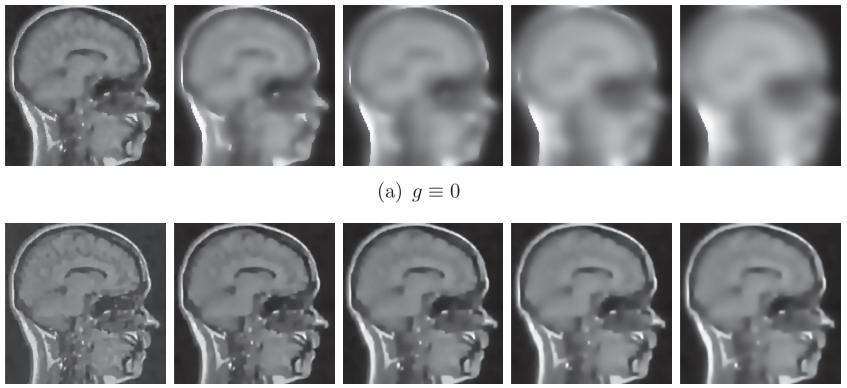
2. Forward and backward diffusion flows

2.1. Variational and PDE models for image restoration. We first 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 [46]. In this case the Euler-Lagrange equation (written with artificial time evolution, see also Eqn. (1.10)) is,

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

which is an isotropic diffusion equation and hence does not preserve edges, see Figure 1(e). This heat flow provides a linear scale space and



(a) $g \equiv 0$

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

FIGURE 8. Stabilizing property of the inverse mollification function when combined with power growth numerator. Solution of the PDE (2.7) 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) $g(|G_\sigma * \nabla u|) = |G_\sigma * \nabla u|^p$. In both cases we used $K = 10^{-4}$ and terminal time 100. It is clear visually that the inverse mollification has a stabilizing effect in keeping homogenous regions.

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 quantitative analysis of the numerical examples is deferred to an upcoming work and we refer to [38] for an earlier attempt in this direction with a strictly convex regularization.

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

4.3. GRADE vs. our approach. Finally, we provide a comparison of Catté et al. [10] GRADE (1.8) to illustrate qualitative differences in restoration with our inverse mollification term based PDE (2.7), with $g(|G_\sigma * \nabla u|) = |G_\sigma * \nabla u|^2$, and set $K = 10^{-4}$. In our PDE (2.7), we consider three cases for the diffusion function: Perona-Malik non-convex regularization functions (ϕ_1, ϕ_2 in (1.11)), and total variation regularization $\phi(s) = s$. Remember that the corresponding φ in (2.7) 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 φ_y from the Perona-Malik non-convex regularizations (PM1 for ϕ_1, PM2 for ϕ_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 (3.55)

has been widely used in various computer vision tasks such as feature point detection and object identification.

- $\phi(s) = s$: To reduce the smoothing when the magnitude of the gradient is high, Rudin et al., [43] introduced total variation (TV) based scheme by setting $\phi(s) = s$. In this case the Euler-Lagrange equation is written as (see Eqn. (1.10)),

$$\frac{\partial u}{\partial t} = \operatorname{div}\left(\frac{\nabla u}{|\nabla u|}\right) - \beta\,(u - u_0),$$
$$\text{or} \quad \frac{\partial u}{\partial t} = \operatorname{div}\left(\frac{\nabla u}{\sqrt{\epsilon + |\nabla u|^2}}\right) - \beta\,(u - u_0), \tag{2.1}$$

where $\epsilon > 0$ is a small number added to avoid numerical instabilities in discrete implementations. In [11], 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 [4], see also Remark 3.18 below. But this global TV model suffers from *staircasing* and blocky effects in the restored image [33]. Also, sharp corners will be rounded and thin features are removed under this regularization model. To see this, let χ_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 (1.9), $\int |\nabla \chi_B|$ is the perimeter of the set B. This shows that TV regularization penalizes edge lengths of an image. Note that this TV diffusion PDE (2.1) is a borderline case of anisotropic diffusion PMADE in Eqn. (1.1) with $\mathcal{C}(s) = s^{-1/2}$, a singular diffusion model.

Figure 4 shows an experimental analysis of PMADE (1.1) against TV PDE (2.1) 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 [45, 38, 39, 36, 41, 40, 42] 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., [45]),

$$\min_u E(u) = \beta \int_\Omega (u_0(x) - u(x))^2 \, dx + \int_\Omega \alpha(x) \, |\nabla u(x)| \, dx. \tag{2.2}$$

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 (1.9) 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 [45] is,

$$\alpha(x) = \frac{1}{\epsilon + |\nabla G_\sigma * u_0(x)|}, \qquad \epsilon > 0. \tag{2.3}$$

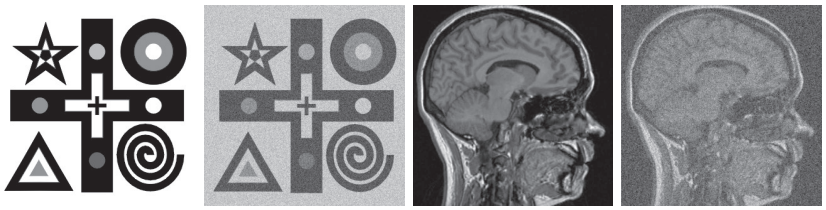
The term in the denominator provide an estimate of edges from the input image u_0 at scale σ using the Gaussian kernel G_σ filtered gradients. Introduction of such a spatially adaptive parameter, which self adjusts according to the smoothed gradient of the image, reduces the TV flow in homogenous regions thereby alleviating the *staircasing* problem. In [13], the existence and uniqueness of the functional satisfying (2.2) is proved under the weighted TV norm. An edge indicator function of the form (2.3) can also be introduced directly in the PDE of the form (1.10). For example, using it in the PMADE (1.1), we write adaptive PMADE as

$$\frac{\partial u(x,t)}{\partial t} = \operatorname{div}\left(\frac{\mathcal{C}(|\nabla u(x,t)|^2)}{\epsilon + |\nabla G_\sigma * u_0(x)|} \nabla u(x,t) \right). \tag{2.4}$$

It is advantageous to use the current estimate image u in the edge indicator $\alpha(x)$ in (2.4) instead of the initial noise image u_0, that is

$$\frac{\partial u(x,t)}{\partial t} = \operatorname{div}\left(\frac{\mathcal{C}(|\nabla u(x,t)|^2)}{\epsilon + |\nabla G_\sigma * u(x,t)|} \nabla u(x,t) \right). \tag{2.5}$$

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 σ) in one term, see Eqn. (2.6) below, can regularize the boundaries of the level set of u_0 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



(a) *Shapes* (b) *Brain*

FIGURE 6. Synthetic *Shapes* and real *Brain* images used in our experiments. (a) Noise-free (ground-truth) images. (b) Noisy images obtained by adding Gaussian noise level $\sigma_n = 30$.

4.1. Effect of inverse mollification, contrast parameter. We first consider the effect of inverse mollification function:

$$\frac{1}{1 + K \, g(|G_\sigma * \nabla u|)} \tag{4.1}$$

in edge detection under noise for different choices of g and K. Note that the diffusion coefficient acts like an edge detector within PDE based image restoration thereby guiding the diffusion smoothing process in and around edges. Figure 7 shows the computed inverse mollification function (4.1) for different choices of K and power growth in $g(\cdot)$ for the noisy *Shapes* image (noise standard deviation $\sigma_n = 30$, see 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 [37].

4.2. Effect of different exponent values in power growth diffusion. Next, we compare restoration results using PDE (2.7) with and without inverse mollification function (4.1). 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 stabilizes the final smoothing result when compared to the smoother results for higher exponent p values.

(3.48) with T replaced by any $t \in [0,T]$, u replaced by u_k, and $PV_{h(G*\nabla u)}(u)$ replaced by $\int_0^t (h(G*\nabla u_k(s)), |\nabla u_k(s)|)\, ds$. We have to prove that their limit (u,z) (in the weak-* topology of $L_\infty(0,T;L_2) \times L_\infty(0,T;L_\infty)$) is a weak solution. Estimates (3.31), (3.32) and (3.33) imply that the solutions u_k belong to a uniformly bounded set in the space (3.38) with $p=1$. Similarly to the proof of the previous theorem, $h(G*\nabla u_k) \to h(G*\nabla u)$ in $C([0,T]\times\overline{\Omega})$. Now we can pass to the limit in the first term of inequality (3.48) as in the previous subsection, cf. [32, 48], in the third term due to Lemma 3.13, and, quite straightforwardly, in (3.47), (3.5) and the remaining terms of (3.48). ∎

Remark 3.17. In view of lack of monotonicity/accretivity, uniqueness of weak solutions in Theorems 3.8 and 3.16 is an open problem.

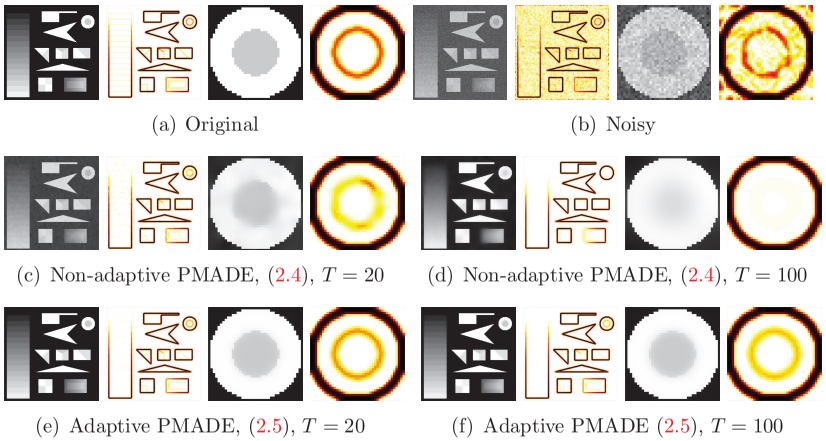
Remark 3.18. The total variation flow equation ($g \equiv 0$, $\varphi(x,y) = \ln y$) was treated in [4] via monotonicity/accretivity arguments such as the Crandall-Liggett theory. Since the operator

$$-\operatorname{div}\left(\frac{|\nabla u|^{-1}\nabla u}{1+K\, g(|G*\nabla u|)}\right) \tag{3.55}$$

is not accretive in any sense (for $g \neq const$), we had to use our (more straightforward and less elaborate) approach. The case $g \equiv 0$ is not excluded from our results: we have obtained existence of weak solutions for the total variation flow (with Dirichlet and Neumann boundary conditions) in a slightly different setting than in [4] and by a simpler method.

4. Experimental results and discussions

In what follows, we provide some experimental results using our adaptive forward-backward diffusion flows in image restoration. All the images are rescaled to $[0,1]$ for visualization and the PDE (2.7) is discretized using standard finite difference scheme via additive operator splitting [50] with time step $\triangle t = 0.2$. 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 (2.7) is fixed for additive Gaussian noise levels $\sigma_n = 30$ used here. Figure 6 shows the synthetic *Shapes* and real *Brain* images (ground truth) and its corresponding noisy versions which are utilized in our experiments.



(a) Original (b) Noisy

(c) Non-adaptive PMADE, (2.4), $T = 20$ (d) Non-adaptive PMADE, (2.4), $T = 100$

(e) Adaptive PMADE, (2.5), $T = 20$ (f) Adaptive PMADE (2.5), $T = 100$

FIGURE 5. Using updated edge indicator function results in better final restoration in PMADE (with $\mathcal{C}_1$ diffusion coefficient and $K = 20$). Shown here are the final restoration results at iterations $T = 20, 100$ for *non-adaptive* PMADE (2.4) and *adaptive* PMADE (2.5). $|i|ii|iii|iv|$: 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. (2.3). Better viewed online and zoomed in.

ϕ (see Eqn. (1.9)) to control the growth adaptively as mentioned. Thus the general multi-scale minimization problem now reads as

$$\min_u E(u) = \beta \int_\Omega (u(x) - u_0(x))^2\, dx + \int_\Omega \frac{\phi(\nabla u(x))}{\epsilon + |\nabla G_\sigma * u(x)|}\, dx. \tag{2.6}$$

A well known method to prove the existence and uniqueness of minimizer to this problem (2.6) is to obtain the lower semicontinuity of the functional E using the properties of regularizing function ϕ.

Remark 2.1. The functional (2.6) 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 [21], was considered in [24]. The arguments in [24] can also be extended to the general minimization problem (2.6).

2.2. General model. Motivated by the regularization functional (2.6) and previous discussions, we consider here a general forward-backward PDE of the following form[†],

$$\boxed{\frac{\partial u}{\partial t} = div\left(\frac{\varphi_y(x,|\nabla u|)\nabla u}{1+K\,g(|G_\sigma * \nabla u|)}\right).} \tag{2.7}$$

Here φ_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 φ which will be discussed at the end of this section and in Section 3. Observe also that, in the x- and u- independent case, ϕ and φ are related through $\phi'(s) = s\varphi'(s)$.

Remark 2.2. The parameter β in (2.6) can be adaptive to account for emphasize of fidelity depending on the problem at hand [22, 42], i.e,

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

The major questions we are now concerned with are:

(1) What are the conditions on φ to obtain existence of solutions for the PDE (2.7)?
(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 φ. If the derivative of g is sublinear 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.

[†]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.

(3.48), (3.5). Then

$$\begin{aligned} 2\int_0^T \left(\frac{d(u-w)}{ds}(s),(u-w)(s)\right) ds + 2\int_0^T \left(\frac{dw}{ds}(s),u(s)\right) ds \\ + 2\int_0^T \big(h(G*\nabla u(s)),|\nabla u(s)|\big)\,ds \\ \leq 2\int_0^T \left(\frac{dw}{ds}(s),w(s)\right) ds - 2\int_0^T \left(\frac{du}{ds}(s),w(s)\right) ds. \end{aligned} \tag{3.49}$$

Hence,

$$\int_0^T \big(h(G*\nabla u(s)),|\nabla u(s)|\big)\,ds \leq -\int_0^T \left(\frac{du}{ds}(s),u(s)\right) ds. \tag{3.50}$$

Therefore,

$$\int_0^T \big(h(G*\nabla u(s)),|\nabla u(s)|\big)\,ds \leq \int_0^T \big(h(G*\nabla u(s))z(s),\nabla u(s)\big)\,ds. \tag{3.51}$$

On the other hand,

$$h(G*\nabla u)|\nabla u| \geq h(G*\nabla u)z\cdot\nabla u. \tag{3.52}$$

All this can be true if and only if

$$|\nabla u| = z\cdot\nabla u, \tag{3.53}$$

i.e.

$$z = \frac{\nabla u}{|\nabla u|} = \varphi_y(\cdot,|\nabla u|)\nabla u. \tag{3.54}$$

But (3.47),(3.54) is a weak form of (3.1),(3.20).

Theorem 3.16. *Let $\varphi(x,y) = \ln y$, $u_0 \in L_2(\Omega)$. There exists a weak solution to (3.1), (3.20),(3.5).*

Proof: (*Sketch*) Let $\{(u_k, z_k)\}$ be a sequence of approximate solutions to (3.47),(3.54),(3.5) satisfying the a priori estimates (3.31), (3.32) and (3.33) (with $\varphi(x,y) = \ln y$). The approximate solutions can be chosen to satisfy

Lemma 3.13. *For any non-negative function $\Phi \in C([0,T]\times\overline{\Omega})$ and a weakly-* converging (in $\mathcal{M}$) sequence $\{v_m\} \subset BPV$, one has*

$$PV_\Phi(v) \leq \lim_{m\to+\infty} \inf PV_\Phi(v_m). \tag{3.45}$$

Using (3.45) with $\Phi \equiv 1$, we can show that BPV is a Banach space.

Definition 3.14. A function u from the class

$$u \in C_w(0,T;L_2(\Omega)) \cap BPV \cap W^1_\infty(0,T;[L_2 \cap W^1_1(\Omega)]^*) \tag{3.46}$$

is called a weak solution to problem (3.1), (3.20), (3.5) with $\varphi(x,y) = \ln y$ and $u_0 \in L_2(\Omega)$ if
i) there exists $z \in L_\infty((0,T)\times\Omega)$, $\|z\|_{L_\infty((0,T)\times\Omega)} \leq 1$, so that for all $v \in L_2 \cap W^1_1(\Omega)$,

$$\left\langle \frac{du}{dt}, v \right\rangle + \big(h(G*\nabla u)z, \nabla v\big) = 0, \tag{3.47}$$

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

$$\begin{aligned} &\|u(T)-w(T)\|^2 + 2\int_0^T \left(\frac{dw}{ds}(s), u(s)\right) ds + 2\,PV_{h(G*\nabla u)}(u) \\ &\leq \|u_0 - w(0)\|^2 + \|w(T)\|^2 - \|w(0)\|^2 + 2\int_0^T \big(h(G*\nabla u(s))z(s), \nabla w(s)\big)\, ds; \end{aligned} \tag{3.48}$$

iii) (3.5) holds in the space $L_2(\Omega)$.

Remark 3.15. The motivation for this definition is the following one. Consider, formally, a pair (u,z) of sufficiently smooth functions satisfying (3.47),

3. Existence of various types of solutions

3.1. Viscosity solutions. Here we study the equation

$$\frac{\partial u}{\partial t} = \operatorname{div}\left(\frac{\varphi_y(x,|\nabla u|)\nabla u}{1+K\,g(|G*\nabla u|)}\right), \tag{3.1}$$

which is slightly more general than (2.7) 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 $u : \mathcal{J} \to E$ is *weakly continuous* if for any linear continuous functional g on E the function $g(u(\cdot)) : \mathcal{J} \to \mathbb{R}$ is continuous.

We denote

$$a_{ij}(x,p) = \varphi_y(x,|p|)\delta_{ij} + \varphi_{yy}(x,|p|)\frac{p_i p_j}{|p|}, \tag{3.2}$$

$$h(q) = \frac{1}{1+Kg(|q|)}. \tag{3.3}$$

Here δ_{ij} is Kronecker's delta, and $x,p,q \in \mathbb{R}^n$. As usual, for the sake of simplification of the presentation, we consider the case of spatially periodic boundary conditions [2] for Eq. (3.1). Namely, we assume that there is an orthogonal basis $\{b_i\}$ in $\mathbb{R}^n$ so that

$$u(\cdot,x) = u(\cdot,x+b_i),\ x \in \mathbb{R}^n,\ i = 1,\dots,n. \tag{3.4}$$

The problem is complemented with the initial condition

$$u(0,x) = u_0(x), \tag{3.5}$$

where $x \in \mathbb{R}^n$, and u_0 is Lipschitz and satisfies (3.4). Of course, φ (and thus a) should also satisfy the same spatial periodicity restriction (with respect to x).

We also make the following assumptions:

$$\varphi_y, \varphi_{yx}, \varphi_{yxx} \text{ are continuous and bounded functions,} \tag{3.6}$$

$$\varphi_{yy}, \varphi_{yyx} \text{ are continuous for } y \neq 0, \text{ and satisfy} \tag{3.7}$$

$$\sup_{x\in\mathbb{R}^n} \lim_{y\to 0} |y|(|\varphi_{yy}(x,y)| + |\varphi_{yyx}(x,y)|) = 0. \tag{3.8}$$

$$a_{ij}(x,p)\xi_i\xi_j \geq C\left[mod\left(\frac{\partial a(x,p)}{\partial x_k}\right)\right]_{ij}\xi_i\xi_j,\ k=1,\dots,n,\ \xi,x,p\in\mathbb{R}^n, \quad (3.9)$$

$$\sqrt{h}\in W^1_\infty(\mathbb{R}^n),\ h\in W^2_\infty(\mathbb{R}^n),\ G\in W^3_1(\mathbb{R}^n). \quad (3.10)$$

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 [42]) maps any symmetric matrix to its suitably defined "positive-semidefinite part".

Note that we do not assume g, h and G to be space-periodic. Observe also that if

$$g\in W^2_{\infty,loc}(0,+\infty),\ g(s)\geq 0,\ |g'(s)|=O(s),$$
$$|g'(s)|\leq C(1+g(s))^{3/2},\ |g''(s)|+\frac{|g'(s)|}{s}\leq C(1+g(s))^2, \quad (3.11)$$

then the required conditions for h are satisfied.

Definition 3.1. A function u from the space

$$C([0,T]\times\mathbb{R}^n)\cap L_\infty(0,T,W^1_\infty(\mathbb{R}^n)) \quad (3.12)$$

is a viscosity sub-/supersolution to (3.1), (3.4), (3.5) 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

$$\frac{\partial\phi}{\partial t}-\operatorname{div}\left(\frac{\varphi_y(x,|\nabla\phi|)\nabla\phi}{1+K\,g(|G*\nabla u|)}\right)\leq 0\,/\,\geq 0, \quad (3.13)$$

and equalities (3.4), (3.5) hold in the classical sense. A viscosity solution is a function which is both a subsolution and a supersolution.

Theorem 3.2. *i) Problem* (3.1), (3.4), (3.5) *has a viscosity solution in class* (3.12) *for every positive* T. *Moreover,*

$$\inf_{\mathbb{R}^n}u_0\leq u(t,x)\leq\sup_{\mathbb{R}^n}u_0. \quad (3.14)$$

ii) Assume that

$$\left|\left(\sqrt{a(x,p)}-\sqrt{a(z,p)}\right)_{ij}\right|\leq C|x-z|,\ x,z,p\in\mathbb{R}^n. \quad (3.15)$$

Sketch of the proof. Equation (3.1) with $\varphi(x,y)=y^{p-1}$ can be written in the form

$$\frac{\partial u}{\partial t}=\operatorname{div}\left(\tilde{h}(t,x)|\nabla u|^{p-2}\nabla u\right), \quad (3.42)$$

where $\tilde{h}=(p-1)\,h(G*\nabla u)$. Similarly to the proof of Theorem 3.8, one shows that $G*\nabla u$ is continuous on $[0,T]\times\overline{\Omega}$. 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 [16, Chapter VIII, pp. 217–218] are fulfilled. The local regularity of the solution now follows from the general results of [16].

Concerning the optimal Lipschitz regularity of the weak solution we invoke the results of [8], with p constant. Indeed, condition (7) on page 912 of [8] 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 [16, Chapter III]). ∎

Now we treat the logarithmic growth case. Again, just to simplify the presentation, we merely consider $\varphi(x,y)=\ln y$. Let $\mathcal{M}$ be the Banach space of finite 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 $\Phi\in C([0,T]\times\overline{\Omega})$, $\Phi\geq 0$, we define the *weighted partial variation* of v as

$$PV_\Phi(v)=\sup_{\psi\in C_0^\infty((0,T)\times\Omega;\mathbb{R}^n):\,|\psi|\leq\Phi}\langle v,\operatorname{div}\psi\rangle_{\mathcal{M}\times C_0}. \quad (3.43)$$

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

$$PV(v)=PV_1(v). \quad (3.44)$$

Define the "bounded partial variation space"[‖] BPV as

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

Owing to lower semicontinuity of suprema, we have

[‖]We are not aware if anybody has introduced this space previously.

we may assume that $u_k \to u$ in $C([0,T];[W_2^1(\Omega)]^*)$. 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, $\tilde{u}_k \to \tilde{u}$ in $C([0,T];[W_2^1(\mathbb{R}^n)]^*)$. Thus, $G*\nabla u_k \to G*\nabla u$ in $C([0,T];L_\infty(\Omega))$. The operator $G*\nabla : L_2(\Omega) \to W_\infty^1(\Omega) \subset C(\overline{\Omega})$ is continuous. This implies that $G*\nabla u_k \to G*\nabla u$ in $C([0,T]\times\overline{\Omega}))$. Due to the continuity of the Nemytskii operator on the space of continuous of functions [27], we conclude that $h(G*\nabla u_k) \to h(G*\nabla u)$ uniformly on $[0,T]\times\overline{\Omega}$. Then we can manually proceed similarly to the classical monotonicity argument [17, 19, 30] but with necessary changes. We omit the further details. ∎

Remark 3.9. In a similar way, Theorem 3.8 may be generalized onto the case of more general $\varphi(x,y)$ with growth as $y \to \infty$ and decay as $y \to 0$ 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 3.10. *Assume that h is Lipschitz and u is a weak solution to (3.1), (3.5), (3.20), 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*

$$|u(t_1,x_1)-u(t_2,x_2)| \le M\left(|x_1-x_2|+\sqrt{|t_1-t_2|}\right), \quad (t_1,x_1),(t_2,x_2)\in\mathcal{K} \tag{3.40}$$

and

$$|\nabla u(t_1,x_1)-\nabla u(t_2,x_2)| \le M\left(|x_1-x_2|+\sqrt{|t_1-t_2|}\right)^{\alpha}, \quad (t_1,x_1),(t_2,x_2)\in\mathcal{K}. \tag{3.41}$$

Remark 3.11. 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 $\left(\frac{2n}{n+2},2\right)$ suffices though. Note that for applications in imaging $n=2$ and no extra assumption is needed.

Remark 3.12. 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 ∇u on $\mathcal{K}$, cf. [16, Chapter IX]. The constant α depends exclusively on p and dimension.

Here $\sqrt{\ }$ is the square root of a positive-semidefinite symmetric matrix [25]. Then the solution is unique. Moreover, for any two viscosity solutions u and v to (3.1), (3.4), the following estimate holds

$$\sup_{\mathbb{R}^n}|u(t,\cdot)-v(t,\cdot)| \le \Phi(t)\sup_{\mathbb{R}^n}|u(0,\cdot)-v(0,\cdot)| \tag{3.16}$$

with some non-decreasing continuous scalar function Φ dependent on u and v.

Proof: (*Sketch*) Note that (3.14) 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 $\delta \to +0$. Now, it suffices to formally establish a Bernstein estimate for $\sup_{\mathbb{R}^n}|\nabla u|$, for then we would be able to approximate our problem by a well-posed one in the sense of [28, Chapter 5], and the results would follow via a well-known technique [3, 2, 6, 14, 42]. Fix T. Differentiating (3.1) with respect to each x_k, $k=1,\dots,n$, multiplying by $2u_{x_k}$, and adding the results, we get

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

$$
\begin{aligned}
&= -2h(u*\nabla G)a_{ij}(x,\nabla u)u_{x_kx_i}u_{x_kx_j} \\
&\quad + 2\nabla h(u*\nabla G)\cdot\left(u*\frac{\partial\nabla G}{\partial x_k}\right)a_{ij}(x,\nabla u)u_{x_ix_j}u_{x_k} \\
&\quad + 2h(u*\nabla G)\frac{\partial a_{ij}(x,\nabla u)}{\partial x_k}u_{x_ix_j}u_{x_k} \\
&\quad + 2\nabla h(u*\nabla G)\cdot\left(u*\frac{\partial\nabla G}{\partial x_k}\right)\varphi_{yx_i}(x,|\nabla u|)u_{x_i}u_{x_k} \\
&\quad + 2h(u*\nabla G)\varphi_{yx_ix_k}(x,|\nabla u|)u_{x_i}u_{x_k} \\
&\quad + 2\frac{\partial^2 h}{\partial q_j\partial q_l}(u*\nabla G)\left(u*\frac{\partial^2 G}{\partial x_i\partial x_j}\right)\left(u*\frac{\partial^2 G}{\partial x_k\partial x_l}\right)\varphi_y(x,|\nabla u|)u_{x_i}u_{x_k} \\
&\quad + 2\nabla h(u*\nabla G)\cdot\left(u*\frac{\partial^2\nabla G}{\partial x_i\partial x_k}\right)\varphi_y(x,|\nabla u|)u_{x_i}u_{x_k} \\
&\quad + 2\nabla h(u*\nabla G)\cdot\left(u*\frac{\partial\nabla G}{\partial x_i}\right)\varphi_{yx_k}(x,|\nabla u|)u_{x_i}u_{x_k}. \quad (3.17)
\end{aligned}
$$

This eventually yields

$$\mathcal{L}(|\nabla u|^2)\le C(1+|\nabla u|^2), \tag{3.18}$$

and

$$|\nabla u|^2\le C. \tag{3.19}$$

We omit the details and the hints for they have much in common with the ones from [42, Section 3]. ∎

3.2. Dissipative solutions. The concept of *dissipative solution* (see [32, 31, 51, 49, 15, 20, 41] and an illustrative discussion in [48]) allows us to significantly relax the assumptions on φ 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 Ω be a bounded open domain in $\mathbb{R}^n$, $n\in\mathbb{N}$, with a regular boundary $\partial\Omega$. We thus consider (3.1), (3.5) to be coupled with

$$\frac{\partial u}{\partial\nu}=0,\ \ x\in\partial\Omega. \tag{3.20}$$

To prove the theorem, we can approximate our problem by a more regular one, and pass to the limit in inequality (3.24) maintaining its sign (cf. [32, 48]), since, due to the observations above, without loss of generality the sequence of corresponding solutions converges weakly in $L_1(0,T;W_1^1)$ and weakly-* in $L_\infty(0,T;L_2)$ for any $T>0$. ∎

3.3. Weak and strong solutions. For φ 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 Ω to be a bounded open domain in $\mathbb{R}^n$ with a regular boundary. We keep the Neumann boundary condition, but generalization to the Dirichlet case is straightforward. We assume that $g:[0,+\infty)\to[0,+\infty)$ is continuous[¶], and $G\in W_2^2(\mathbb{R}^n)$.

Firstly, let $\varphi(x,y)=c_1y^{p-1}$, $p>1$, $c_1>0$. To simplify the presentation, in the sequel we assume that $c_1=1$.

Definition 3.7. A function u from the class

$$u\in C_w([0,T];L_2(\Omega))\cap L_p(0,T;W_p^1(\Omega))\cap W^1_{p/p-1}(0,T;[L_2\cap W_p^1(\Omega)]^*) \tag{3.38}$$

is called a weak solution to problem (3.1), (3.20) if, for all $v\in L_2(\Omega)\times W_p^1(\Omega)$ and a.a. $t\in(0,T)$, one has

$$\left\langle\frac{du}{dt},v\right\rangle+\left(\frac{(p-1)|\nabla u|^{p-2}\nabla u}{1+K\,g(|G*\nabla u|)},\nabla v\right)=0. \tag{3.39}$$

Theorem 3.8. *Let $\varphi(x,y)=y^{p-1}$, $p>1$, $u_0\in L_2(\Omega)$. There exists a weak solution u to (3.1), (3.20) satisfying (3.5).*

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 3.8. The key point is to pass to the limit. Let $\{u_k\}$ be a sequence of approximate solutions satisfying the a priori estimates (3.31), (3.32) and (3.33) (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 (3.31), (3.32) and (3.33) imply that the solutions u_k belong to a uniformly bounded set in the space (3.38). Owing to [44, Corollary 4], without loss of generality

[¶]Thus, for weak solutions, g is not needed to be locally Lipschitz.

If (3.27) holds, then, due to (3.32), (3.26) and boundedness of Ω, we have

$$\begin{aligned}
&2\left([h(G*\nabla v)-h(G*\nabla u)]\varphi_y(x,|\nabla v|)\nabla v,\nabla(u-v)\right)\\
&\qquad -2\left(\frac{\varphi_y(x,|\nabla u|)\nabla u-\varphi_y(x,|\nabla v|)\nabla v}{1+K\,g(|G*\nabla u|)},\nabla(u-v)\right)\\
&\quad\le C\|\nabla h\|_{L_\infty}\|G*\nabla(u-v)\|_{L_\infty}\|\nabla(u-v)\|_{L_1}-C_1\|\nabla(u-v)\|^2\\
&\quad\le C\|\nabla G\|\|u-v\|\|\nabla(u-v)\|-C_1\|\nabla(u-v)\|^2\le C\|u-v\|^2,\quad(3.35)
\end{aligned}$$

where C_1 is the doubled constant from (3.27).

If (3.28) and (3.29) hold, then, by virtue of (3.21), we get

$$\begin{aligned}
&2\left([h(G*\nabla v)-h(G*\nabla u)]\varphi_y(x,|\nabla v|)\nabla v,\nabla(u-v)\right)\\
&\qquad -2\left(\frac{\varphi_y(x,|\nabla u|)\nabla u-\varphi_y(x,|\nabla v|)\nabla v}{1+K\,g(|G*\nabla u|)},\nabla(u-v)\right)\\
&\quad\le 2\left(\operatorname{div}\left[[h(\nabla G*\tilde u)-h(\nabla G*\tilde v)]\varphi_y(x,|\nabla v|)\nabla v\right],u-v\right)\\
&\quad= 2\left([h(\nabla G*\tilde u)-h(\nabla G*\tilde v)]\operatorname{div}[\varphi_y(x,|\nabla v|)\nabla v],u-v\right)\\
&\quad+2\Big(\left[\nabla h(\nabla G*\tilde u)\left(\frac{\partial\nabla G}{\partial x_i}*\tilde u\right)-\nabla h(\nabla G*\tilde v)\left(\frac{\partial\nabla G}{\partial x_i}*\tilde v\right)\right]\\
&\qquad\times\varphi_y(x,|\nabla v|)\frac{\partial v}{\partial x_i},u-v\Big)\\
&\quad\le C\|\nabla G*(\tilde u-\tilde v)\|_{L_\infty}\|\operatorname{div}[\varphi_y(x,|\nabla v|)\nabla v]\|\|u-v\|\\
&+2\left([\nabla h(\nabla G*\tilde u)-\nabla h(\nabla G*\tilde v)]\left(\frac{\partial\nabla G}{\partial x_i}*\tilde u\right)\varphi_y(x,|\nabla v|)\frac{\partial v}{\partial x_i},u-v\right)\\
&\quad+2\left(\nabla h(\nabla G*\tilde v)\left(\frac{\partial\nabla G}{\partial x_i}*(\tilde u-\tilde v)\right)\varphi_y(x,|\nabla v|)\frac{\partial v}{\partial x_i},u-v\right)\\
&\quad\le C\|\nabla G\|\|\operatorname{div}[\varphi_y(x,|\nabla v|)\nabla v]\|\|\tilde u-\tilde v\|\|u-v\|\\
&+C\|\nabla G\|\|\tilde u\|\|\nabla G\|_{W_2^1}\|\tilde u-\tilde v\|\|u-v\|+C\|\nabla G\|_{W_2^1}\|\tilde u-\tilde v\|\|u-v\|\\
&\quad\le C\|u-v\|^2.\quad(3.36)
\end{aligned}$$

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

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

$$\frac{d}{dt}\|u-v\|^2\le(\ln\gamma)\|u-v\|^2-2\left(\Phi(v),\nabla u-\nabla v\right)-2\left(\frac{\partial v}{\partial t},u-v\right).\quad(3.37)$$

By Gronwall's lemma, we infer (3.24).

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 φ,φ_y are continuous and bounded on $\Omega\times(1/N,N)$. The parabolicity conditions (3.6)–(3.9) are replaced by a weaker one:

$$(\varphi_y(x,|p_1|)p_1-\varphi_y(x,|p_2|)p_2)\cdot(p_1-p_2)\ge 0,\ x\in\Omega,\ p_1,p_2\in\mathbb{R}^n\backslash\{0\}.\quad(3.21)$$

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 $|g'|\le C(1+g)^2$, 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[‡] of u onto $\mathbb{R}^n$ which may depend on the boundary condition (cf. [10]).

Introduce the following formal expression

$$\Phi(v)(t,x)=\frac{\varphi_y(x,|\nabla v(t,x)|)\nabla v(t,x)}{1+K\,g(|G*\nabla v|\,(t,x))}.\quad(3.22)$$

Definition 3.3. Let $u_0\in L_2(\Omega)$. A function u from the class

$$u\in C_w([0,\infty);L_2(\Omega))\cap L_1(0,\infty;W_1^1(\Omega))\quad(3.23)$$

is called a dissipative solution to problem (3.1), (3.5), (3.20) if, for all regular[§] functions $v:[0,\infty)\times\overline{\Omega}\to\mathbb{R}$ satisfying the Neumann boundary condition (3.20) and all non-negative moments of time t, one has

$$\begin{aligned}
&\|u(t)-v(t,\cdot)\|^2\\
&\qquad\le\gamma^t\|u_0-v(0,\cdot)\|^2\\
&-\int_0^t 2\gamma^{t-s}\left[\left(\Phi(v(s,\cdot)),\nabla u(s)-\nabla v(s,\cdot)\right)+\left(\frac{\partial v(s,\cdot)}{\partial s},u(s)-v(s,\cdot)\right)\right]ds,
\end{aligned}\quad(3.24)$$

where γ is a certain constant depending on g, G, φ and v (in particular, γ=1 provided $g\equiv 0$).

[‡]The simplest possible extension is letting $\tilde u$ to be zero outside of Ω. Another option is to use Hestenes-Seeley-like extensions [1] which conserve the Sobolev class of u.

[§]Here "regular" means that v and ∇v are uniformly bounded and sufficiently smooth, and $|\nabla v|\ne 0$ a.e. in $(0,\infty)\times\Omega$.

Usual argument [48] shows that these dissipative solutions possess the weak-strong uniqueness property (any regular solution is a unique dissipative solution).

Theorem 3.4. *Assume*

$$\lim_{y\to+\infty}\inf_{x\in\Omega}\varphi_y(x,y)y=+\infty, \tag{3.25}$$

$$\lim_{y\to 0}\sup_{x\in\Omega}\varphi_y(x,y)y=0. \tag{3.26}$$

Assume also that either we have strong parabolicity, namely,

$$(\varphi_y(x,|p_1|)p_1-\varphi_y(x,|p_2|)p_2)\cdot(p_1-p_2)\geq C|p_1-p_2|^2,\ x\in\Omega,\ p_1,p_2\in\mathbb{R}^n\backslash\{0\}, \tag{3.27}$$

or better regularity of h and G,

$$h\in W^2_\infty(\mathbb{R}^n),\ G\in W^2_2(\mathbb{R}^n). \tag{3.28}$$

Let $u_0\in L_2(\Omega)$. Then there exists a dissipative solution to problem (3.1), (3.5), (3.20).

Remark 3.5. In the case when (3.28) but not (3.27) holds, the test functions for (3.24) should additionally satisfy the condition

$$\operatorname{div}\left(\varphi_y(x,|\nabla v|)\nabla v\right)\in L_\infty(0,+\infty;L_2(\Omega)), \tag{3.29}$$

which, by the way, automatically holds provided φ is more regular, e.g., satisfies the assumptions (3.6) – (3.8) of the previous subsection.

Remark 3.6. The total variation flow [4] corresponds to the case $g\equiv 0$, $\varphi(x,y)=\ln y$, so it satisfies (3.21) but is ruled out by (3.25) and (3.26). We will consider this particular form of φ (with generic g) in the next subsection (cf. Remark 3.18). The existence of dissipative solutions for the total variation flow is an open problem. We however believe that the hypotheses of Theorem 3.4 may be significantly weakened.

Proof of Theorem 3.4. Let us formally derive some a priori bounds and inequality (3.24) for the solutions to problem (3.1), (3.5), (3.20). Firstly, we formally take the $L_2(\Omega)$-scalar product of (3.1) with $2u(t)$, and integrate by parts:

$$\frac{d}{dt}\|u\|^2+\left(\frac{2\varphi_y(x,|\nabla u|)\nabla u}{1+K\,g(|G*\nabla u|)},\nabla u\right)=0. \tag{3.30}$$

Since the second term is non-negative due to (3.21) and (3.26), (3.30) a priori implies that

$$\|u\|_{L_\infty(0,+\infty;L_2)}\leq\|u_0\|. \tag{3.31}$$

Thus,

$$|G*\nabla u|\leq\|\nabla G\|\|\tilde{u}\|\leq C\|u\|\leq C. \tag{3.32}$$

Consider the scalar function $\Psi(y)=\inf_{x\in\Omega}\varphi_y(x,y)y^2$, $y>0$, $\Psi(0)=0$. From (3.30) and (3.32) we can conclude that

$$\int_0^{+\infty}\int_\Omega\Psi(|\nabla u|)\,dx\,dt\leq C\|u_0\|. \tag{3.33}$$

The function $\Psi(y)$ is non-negative, continuous (for $y=0$ this follows from (3.26), for positive y it can be derived from the compactness of $\overline{\Omega}$) and satisfies the condition $\lim_{y\to+\infty}\Psi(y)/y=+\infty$. By the Vallée-Poussin criterion [18], ∇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 $v:[0,\infty)\times\overline{\Omega}\to\mathbb{R}$ satisfying the Neumann boundary condition (3.20). Adding (3.1) 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

$$\begin{aligned}\frac{d}{dt}\|u-v\|^2+2\left(\frac{\varphi_y(x,|\nabla u|)\nabla u-\varphi_y(x,|\nabla v|)\nabla v}{1+K\,g(|G*\nabla u|)},\nabla(u-v)\right)\\ =2\left([h(G*\nabla v)-h(G*\nabla u)]\varphi_y(x,|\nabla v|)\nabla v,\nabla(u-v)\right)\\ -2\left(\Phi(v),\nabla u-\nabla v\right)-2\left(\frac{\partial v}{\partial t},u-v\right).\end{aligned} \tag{3.34}$$