An Efficient and Stable Two-Pixel Scheme for 2D Forward-and-Backward Diffusion

Martin Welk[1] and Joachim Weickert[2]

[1] Institute of Biomedical Image Analysis
Private University for Health Sciences, Medical Informatics and Technology
Eduard-Wallnöfer-Zentrum 1, 6060 Hall/Tyrol, Austria
martin.welk@umit.at

[2] Mathematical Image Analysis Group, Saarland University
Campus E1.7, 66041 Saarbrücken, Germany
weickert@mia.uni-saarland.de

Abstract. Image enhancement with forward-and-backward (FAB) diffusion is numerically very challenging due to its negative diffusivities. As a remedy, we first extend the explicit nonstandard scheme by Welk et al. (2009) from the 1D scenario to the practically relevant two-dimensional setting. We prove that under a fairly severe time step restriction, this 2D scheme preserves a maximum–minimum principle. Moreover, we find an interesting Lyapunov sequence which guarantees convergence to a flat steady state. Since a global application of the time step size restriction leads to very slow algorithms and is more restrictive than necessary for most pixels, we introduce a much more efficient scheme with locally adapted time step sizes. It applies diffusive two-pixel interactions in a randomised order and adapts the time step size to the specific pixel pair. These space-variant time steps are synchronised at sync times. Our experiments show that our novel two-pixel scheme allows to compute FAB diffusion with guaranteed L^∞-stability at a speed that can be three orders of magnitude larger than its explicit counterpart with a global time step size.

1 Introduction

Partial differential equations (PDEs) and variational approaches for enhancing digital images have been investigated intensively in the last thirty years. An overview can be found e.g. in [1, 12]. As continuous frameworks, these approaches excel by their concise and transparent formulation and their natural representation of rotational invariance.

However, some highly interesting models are affected by well-posedness problems, making their analysis in a continuous setting difficult. Well-posedness properties of space-discrete and fully-discrete formulations therefore receive increasing attention.

Regarding the Perona–Malik filter, a space-discrete and fully discrete theory for smooth nonnegative diffusivities was established by Weickert [12]. The corresponding explicit scheme was proven in [13] to preserve monotonicity in 1D. An extension of this analysis to singular nonnegative diffusivities was accomplished by Pollak et al. [10] who verified the well-posedness of dynamical systems with discontinuous right hand sides arising from a space-discrete Perona-Malik model.

In F. Lauze, Y. Dong, A. B. Dahl, eds., Scale Space and Variational Methods in Computer Vision, Lecture Notes in Computer Science, vol. 10302, pp. 94–106,

For the stabilised inverse linear diffusion process introduced by Osher and Rudin, a continuous well-posedness theory is lacking, but a stable minmod discretisation could be devised [7]. For shock filtering [5, 8] which, too, is difficult to analyse in the continuous setting, discrete well-posedness results are found in [15], including an analytic solution of the corresponding dynamical system.

The *forward-and-backward (FAB) diffusion* model of Gilboa et al. [4] is another example for these difficulties. Designed for the sharpening of images, it is basically a Perona–Malik type PDE filter. However, its diffusivities take positive values in some regions and negative values in others. The absence of well-posedness results in the continuous setting is plausible given that inverse diffusion in its pure form, with negative diffusivity, is a standard example of an ill-posed problem. Experiments with standard explicit discretisations show violations of a maximum–minimum principle.

For space-discrete and fully discrete FAB diffusion some analytical results have been obtained in [14]. It was shown that space-discrete FAB diffusion is well-posed and satisfies a maximum–minimum principle if a specific nonstandard discretisation is applied at extrema. For a fully discrete 1D FAB diffusion framework with an explicit time discretisation, a maximum–minimum principle and a total variation reduction property were established. However, results for higher dimensions are still missing.

Our contribution. In this paper, we consider the practically relevant fully discrete 2D case. We prove a maximum–minimum principle for an explicit time discretisation. By introducing a novel Lyapunov sequence – which is also interesting by itself – we prove convergence to a flat steady state. Our theoretical findings allow to devise an explicit finite difference scheme for FAB diffusion which is L^∞-stable for small positive time step sizes. Unfortunately, the time step bound is extremely small which limits the practical applicability of that scheme. However, as our analysis reveals, the small time step size is actually needed only at few locations in typical images. This motivates us to adapt the time step size locally. To obtain maximal locality and keep the diffusion interaction process as simple and transparent as possible, we split it into a sequence of two-pixel interactions. Selecting the pixel pairs randomly averages out any directional bias. Our experiments show that this gives a stable scheme for 2D FAB diffusion that is three orders of magnitude more efficient than an explicit scheme with global time step size restriction.

Structure of the paper. In Section 2, we briefly review the basic ideas behind FAB diffusion. Our theoretical results for an explicit scheme with nonstandard discretisation in a two-dimensional setting are established in Section 3. In Section 4 we introduce our novel two-pixel scheme with randomised updates. Its performance is evaluated experimentally in Section 5. Our paper is concluded with a summary in Section 6.

2 Forward-and-Backward Diffusion Filtering

Forward-and-backward (FAB) diffusion filtering has been proposed by Gilboa, Sochen and Zeevi in 2002 [4]. Let a greyscale image $f : \Omega \to \mathbb{R}$ on a rectangular image domain

$\Omega \in \mathbb{R}^2$ be given. To sharpen this image, filtered versions $u(\boldsymbol{x}, t)$ of $f(\boldsymbol{x})$ are created by solving an initial–boundary value problem for the Perona-Malik type [9] PDE

$$\partial_t u \;=\; \operatorname{div}\left(g(|\boldsymbol{\nabla} u|^2)\,\boldsymbol{\nabla} u\right) \tag{1}$$

with initial condition $u(\boldsymbol{x}, 0) = f(\boldsymbol{x})$, and homogeneous Neumann boundary conditions, $\partial_{\boldsymbol{n}} u = 0$, where $\boldsymbol{n}$ denotes a vector normal to the image boundary $\partial\Omega$. Here $\boldsymbol{x}$ stands for $(x, y)^\top$. Writing partial derivatives by subscripts, we denote by $\boldsymbol{\nabla} := (\partial_x, \partial_y)^\top$ the spatial gradient and by div its corresponding divergence operator.

For the diffusivity g different models have been proposed, see for example [3, 11]. Constitutive for a FAB diffusivity is that it is positive for small image gradients, while it becomes negative for larger ones. An example, adapted from [11], is

$$g(s^2) \;=\; 2\exp\left(-\frac{\kappa^2 \ln 2}{\kappa^2 - 1} \cdot \frac{s^2}{\lambda^2}\right) - \exp\left(-\frac{\ln 2}{\kappa^2 - 1} \cdot \frac{s^2}{\lambda^2}\right) \tag{2}$$

with admissible parameters $\lambda > 0$ and $\kappa > 1$. In contrast to this diffusivity, some other FAB diffusivities can also become positive again for large gradient magnitudes; see e.g. [3]) In [3] FAB diffusion has been interpreted as an energy minimisation process of a nonmonotone potential in the shape of a triple-well. FAB diffusion has also been put into relation with wavelet methods for image enhancement [6].

Beyond these works, there is not much theoretical analysis of the fully continuous FAB process documented in the literature. In particular, no existence, uniqueness and stability results have been proven. It was conjectured [4] that FAB diffusion violates a maximum–minimum principle due to the effect of negative diffusivities. In numerical experiments based on standard numerical methods, such violations were indeed observed. However, [14] brought out that using a more sophisticated space discretisation, the space-discrete process obeys the maximum–minimum principle and useful theoretical results on the space-discrete process could be established. Stability properties of fully discrete FAB diffusion were considered in [14], too, but limited to the 1D case. We will present analytical results for the 2D case in the next section.

3 Analysis of Fully Discrete FAB Diffusion in 2D

To study FAB diffusion in the fully discrete 2D case, we consider the discrete image domain $J := \{1, 2, \ldots, m\} \times \{1, 2, \ldots, n\}$. Following [14] we use a simple explicit scheme for (1) with time step size τ and grid sizes h_1 and h_2 in x- and y-direction:

$$\begin{aligned} u_{i,j}^{k+1} = u_{i,j}^k + \tau \cdot \Bigg(& \frac{g_{i+1,j}^k + g_{i,j}^k}{2} \cdot \frac{u_{i+1,j}^k - u_{i,j}^k}{h_1^2} - \frac{g_{i,j}^k + g_{i-1,j}^k}{2} \cdot \frac{u_{i,j}^k - u_{i-1,j}^k}{h_1^2} \\ & + \frac{g_{i,j+1}^k + g_{i,j}^k}{2} \cdot \frac{u_{i,j+1}^k - u_{i,j}^k}{h_2^2} - \frac{g_{i,j}^k + g_{i,j-1}^k}{2} \cdot \frac{u_{i,j}^k - u_{i,j-1}^k}{h_2^2} \Bigg) . \end{aligned} \tag{3}$$

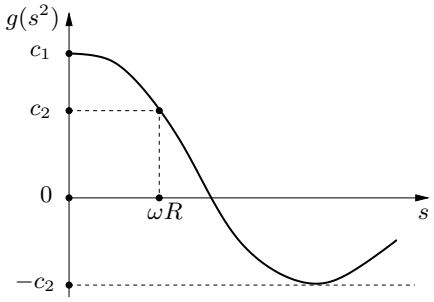

Fig. 1. Schematic view of a FAB diffusivity satisfying the conditions of Proposition 1.

Here, $u^k_{i,j}$ approximates u at location $\big((i-\frac{1}{2})h_1,(j-\frac{1}{2})h_2\big)$ and time $k\tau$. Further following [14], we use a nonstandard approximation for the FAB diffusivity,

$$g^k_{i,j} = g\Bigg(\max\bigg(\frac{u^k_{i+1,j}-u^k_{i,j}}{h_1}\cdot\frac{u^k_{i,j}-u^k_{i-1,j}}{h_1},0\bigg) + \max\bigg(\frac{u^k_{i,j+1}-u^k_{i,j}}{h_2}\cdot\frac{u^k_{i,j}-u^k_{i,j-1}}{h_2},0\bigg)\Bigg)\,. \tag{4}$$

In contrast to the standard approximation

$$g^k_{i,j} = g\left(\left(\frac{u^k_{i+1,j}-u^k_{i-1,j}}{2h_1}\right)^2+\left(\frac{u^k_{i,j+1}-u^k_{i,j-1}}{2h_2}\right)^2\right) \tag{5}$$

it offers the advantage that at extrema the gradient approximation is zero which leads to a positive FAB diffusivity. This will be essential for guaranteeing stability.
To implement homogeneous Neumann boundary conditions, we set

$$u^k_{0,j} := u^k_{1,j},\quad u^k_{m+1,j} := u^k_{m,j},\quad u^k_{i,0} := u^k_{i,1},\quad u^k_{i,n+1} := u^k_{i,n} \tag{6}$$

for all indices i and j. Then (3) can be used verbatim also at boundary pixels.

3.1 Maximum–Minimum Principle

Our first result is a 2D analogue for the first statement of [14, Prop. 4], i.e. the maximum–minimum principle. The hypotheses on the grey-value range and shape of the diffusivity function g are the same as there. Our bound for τ is adapted to the 2D grid geometry.

Proposition 1. *Let an initial 2D image $\boldsymbol{f} = (f_{i,j})$ on $J = \{1,2,\dots,m\}\times\{1,2,\dots,n\}$ be given, and let the sequence of images $\boldsymbol{u}^k = (u^k_{i,j})$ evolve according to (3), (4) with the initial condition $\boldsymbol{u}^0 = \boldsymbol{f}$. Let the grey-values $f_{i,j}$ be restricted to a finite interval $[a,b]$ of length $R := b-a$. Assume that there are two constants $c_1 > c_2 > 0$ such that the diffusivity g fulfils $g(0) = c_1$, and $g(z) \in [-c_2, c_1]$ for all $z > 0$, compare Fig. 1. Assume further that there exists an $\omega > 0$ such that $g(s^2) > c_2$ holds for all s with $0 < s < \omega R$.*

Let the time step τ satisfy the inequality

$$\tau \leq \vartheta := \frac{\omega^2 h_1^4 h_2^4}{2c_1(h_1^2+h_2^2)(\omega^2 h_1^2 h_2^2 + h_1^2 + h_2^2)} \; . \tag{7}$$

Then $\boldsymbol{u}$ obeys the following maximum–minimum principle: If the initial signal is bounded by $f_{i,j} \in [a,b]$ for all $(i,j) \in J$, then $u_{i,j}^k \in [a,b]$ for all $(i,j) \in J$, $k \geq 0$.

The proof of this statement relies on two local properties.

Lemma 1. *Under the assumptions of Proposition 1, local maxima do not increase.*

Proof. Let $u_{i,j}^k$ be a local maximum of $\boldsymbol{u}^k$. Then $g_{i,j}^k = c_1$, and we have that $g_{i+1,j}^k + g_{i,j}^k$ etc. are positive while $u_{i+1,j}^k - u_{i,j}^k$ etc. are negative such that all summands in the bracket on the r.h.s. of (3) are negative. This holds independent of τ.

The r.h.s. of (3) is a convex combination of $u_{i,j}^k$, $u_{i\pm1,j}^k$, $u_{i,j\pm1}^k$ if

$$1 - \frac{\tau}{2h_1^2}\left(g_{i+1,j}^k + 2g_{i,j}^k + g_{i-1,j}^k\right) - \frac{\tau}{2h_2^2}\left(g_{i,j+1}^k + 2g_{i,j}^k + g_{i,j-1}^k\right) \geq 0 \; , \tag{8}$$

which is certainly the case if $\tau \leq h_1^2 h_2^2/(2c_1(h_1^2+h_2^2))$ (the well-known time-step limit for the standard explicit scheme in the case of nonnegative diffusivity). □

Lemma 2. *Under the assumptions of Proposition 1, a non-maximal pixel does not grow in excess of its greatest adjacent pixel within one time step.*

Proof. Let $u_{i,j}^k$ be the non-maximal pixel under consideration. Assume first that the largest grey-value among its neighbours is attained by a horizontal neighbour, say $u_{i-1,j}^k$. So $u_{i-1,j}^k$ is greater than $u_{i,j}^k$ and not less than each other neighbour of $u_{i,j}^k$.

To outline the proof first, notice that basically two situations can happen: If $u_{i,j}^k$ is only slightly smaller than $u_{i-1,j}^k$, it turns out that $u_{i,j}^k$ is "not far from maximality". In this case, its diffusivity $g_{i,j}^k$ will be positive and large enough to ensure forward diffusion (Case 1 in the following), such that again the time step limit of explicit forward diffusion applies. Otherwise, negative diffusivity may occur but at the same time there is quite some way to go before $u_{i,j}^{k+1}$ could exceed $u_{i-1,j}^k$. The overall bounds on the image contrast limit the diffusion flow and thereby the "speed" of the pixel. So one can state also in this case a time step size limit that ensures the desired inequality (Case 2). In the following, these two cases are treated exactly.

Case 1: $\frac{1}{h_1^2}(u_{i,j}^k - u_{i-1,j}^k)(u_{i+1,j}^k - u_{i,j}^k) + \frac{1}{h_2^2}(u_{i,j}^k - u_{i,j-1}^k)(u_{i,j+1}^k - u_{i,j}^k) \leq \omega^2 R^2$.

Then $g_{i,j}^k \geq c_2$, and thus $g_{i,j}^k + g_{i+1,j}^k$, $g_{i,j}^k + g_{i,j\pm1}^k \geq 0$. As a consequence, $u_{i,j}^{k+1}$ is a convex combination of $u_{i,j}^k$, $u_{i\pm1,j}^k$, $u_{i,j\pm1}^k$ if

$$1 - \frac{\tau}{2h_1^2}(g_{i+1,j}^k + 2g_{i,j}^k + g_{i-1,j}^k) - \frac{\tau}{2h_2^2}(g_{i,j+1}^k + 2g_{i,j}^k + g_{i,j-1}^k) \geq 0 \; , \tag{9}$$

which is certainly fulfilled if $1 - \frac{4\tau c_1}{2h_1^2} - \frac{4\tau c_1}{2h_2^2} \geq 0$, i.e. $\tau \leq h_1^2 h_2^2/(2c_1(h_1^2+h_2^2))$.

Case 2: $\frac{1}{h_1^2}(u_{i,j}^k - u_{i-1,j}^k)(u_{i+1,j}^k - u_{i,j}^k) + \frac{1}{h_2^2}(u_{i,j}^k - u_{i,j-1}^k)(u_{i,j+1}^k - u_{i,j}^k) > \omega^2 R^2$.

The difference between pixel $u_{i,j}^k$ and its greatest neighbour fulfils the inequality $u_{i-1,j}^k - u_{i,j}^k > \omega^2 R/(h_1^{-2} + h_2^{-2})$ because the contrary would imply the hypothesis of Case 1. Further, we have by the hypothesis of Case 2 that $-c_2 \le g_{i,j}^k \le c_2$. Together with the hypotheses from Proposition 1 on the range of g and the image range, one has

$$\begin{aligned} -2c_2 \le g_{i+1,j}^k + g_{i,j}^k \le c_1 + c_2\,, \qquad & -R \le u_{i+1,j}^k - u_{i,j}^k \le R\,, \\ -2c_2 \le g_{i-1,j}^k + g_{i,j}^k \le c_1 + c_2\,, \qquad & 0 \le u_{i-1,j}^k - u_{i,j}^k \le R\,, \\ -2c_2 < g_{i,j\pm1}^k + g_{i,j}^k \le c_1 + c_2\,, \qquad & -R \le u_{i,j\pm1}^k - u_{i,j}^k \le u_{i-1,j}^k - u_{i,j}^k. \end{aligned} \tag{10}$$

Inserting this into (3) gives

$$u_{i,j}^{k+1} \le u_{i,j}^k + \tau\left(\frac{c_1+c_2}{2h_1^2}R + \frac{c_1+c_2}{2h_1^2}(u_{i-1,j}^k - u_{i,j}^k) + 2\,\frac{c_1+c_2}{2h_2^2}R\right), \tag{11}$$

$$u_{i-1,j}^k - u_{i,j}^{k+1} \ge (u_{i-1,j}^k - u_{i,j}^k)\left(1 - \frac{\tau(c_1+c_2)}{2h_1^2}\right) - \frac{\tau(c_1+c_2)R}{2}\left(\frac{1}{h_1^2} + \frac{2}{h_2^2}\right). \tag{12}$$

The r.h.s. of (12) is certainly nonnegative if

$$\tau \le \frac{2h_1^2}{c_1+c_2} \cdot \frac{u_{i-1,j}^k - u_{i,j}^k}{u_{i-1,j}^k - u_{i,j}^k + R\,(1 + 2\,h_1^2/h_2^2)}\,, \tag{13}$$

for which by our initial estimate for $u_{i-1,j}^k - u_{i,j}^k$ and monotonicity it suffices that

$$\tau \le \frac{2\omega^2 h_1^4 h_2^4}{(c_1+c_2)(\omega^2 h_1^2 h_2^4 + (h_2^2 + 2h_1^2)(h_1^2 + h_2^2))}\,. \tag{14}$$

This limit on τ ensures that the pixel under consideration cannot grow in excess of its greatest neighbour if this neighbour is a horizontal neighbour. If $u_{i,j}^k$ has its greatest neighbour in vertical direction, analogous considerations lead to a similar constraint, with h_1 and h_2 exchanged. Both bounds on τ are larger than the one of Proposition 1. □

Proof (of the Proposition). The maximum–minimum principle follows immediately from Lemmas 1 and 2 and analogous statements for local minima. □

Remark 1. Unlike in the 1D case [14, Prop. 4], the statement that an extremum may not split into two does not hold. Similar to ordinary homogeneous diffusion, a "dumbbell" configuration with a narrow ridge between two more extended plateaus can serve as a counterexample. Note that for sufficiently small grey-value differences between the ridge and the adjacent plateaus all diffusivities in this region are positive.

3.2 Strict Lyapunov Condition

The maximum–minimum principle from Proposition 1 suggests the use of the difference between global maximum and global minimum of the image as a Lyapunov function to investigate the possible convergence of discrete FAB diffusion. Our proofs from

the previous subsection, however, still leave the possibility that the global maximum and minimum of the image stay constant, and different from each other, forever.

To rule out this possibility, we will refine our analysis and construct a strictly decreasing Lyapunov function by incorporating multiplicities of maxima and minima as additional information. From now on, we require that τ is strictly smaller than the bound from Proposition 1. We introduce notations for the global extremal grey-values of images with their multiplicities, and an ordering for pairs of values with multiplicities.

Definition 1. *For any image $\boldsymbol{u} = (u_{i,j})_{(i,j)\in J}$, let $u_{\max} := \max\limits_{i,j} u_{i,j}$, $u_{\min} := \min\limits_{i,j} u_{i,j}$ denote its maximal and minimal grey-value, respectively, and $n_{\max} := \#\{(i,j) \mid u_{i,j} = u_{\max}\}$, $n_{\min} := \#\{(i,j) \mid u_{i,j} = u_{\min}\}$ their multiplicities.*

Definition 2. *Let the relation $\prec$ on $\mathbb{R} \times \mathbb{N}$ be given by*

$$(u_1, n_1) \prec (u_2, n_2) \quad :\Longleftrightarrow \quad (u_1 < u_2) \;\; \textit{or} \;\; (u_1 = u_2 \;\; \textit{and} \;\; n_1 < n_2)\,. \tag{15}$$

Clearly, $\prec$ is a strict total order. We can now establish the maximum–minimum difference with multiplicities as Lyapunov function for discrete FAB diffusion.

Proposition 2. *Consider the fully discrete FAB diffusion (3), (4). If the time step size is chosen as $\tau < \vartheta$ with ϑ as in Proposition 1, then*

$$(u_{\max}^{k+1} - u_{\min}^{k+1},\; n_{\max}^{k+1} + n_{\min}^{k+1}) \;\prec\; (u_{\max}^{k} - u_{\min}^{k},\; n_{\max}^{k} + n_{\min}^{k}) \tag{16}$$

holds, unless $u_{\max}^k = u_{\min}^k$.

Proof. Let $u_{i,j}^k$ be a local maximum of $\boldsymbol{u}^k$. As in the proof of Lemma 1, we have $g_{i,j}^k = c_1$. Thus, the new value $u_{i,j}^{k+1}$ of that pixel will be a convex combination of the old grey-values of pixel (i,j) and its neighbours, with all neighbours having positive weights. Therefore, $u_{i,j}^{k+1} = u_{i,j}^k$ can happen only if all neighbours have the same value as $u_{i,j}$ in time step k. As long as not all pixels of the image have the same value, there will be at least one pixel $u_{i,j}^k = u_{\max}^k$ with a neighbour of smaller grey-value.

Following the proof of Lemma 2 we see that for $\tau < \vartheta$ the new pixel value $u_{i,j}^{k+1}$ remains strictly below the old value of its largest neighbour $u_{i-1,j}^k$. Thus, $u_{i,j}^{k+1} = u_{\max}^k$ cannot hold for a pixel with $u_{i,j}^k < u_{\max}^k$.

Combining both arguments, we see that the number of pixels attaining the value $u_{\max}^k$ decreases in time step $k+1$. As a consequence, one has $u_{\max}^{k+1} < u_{\max}^k$ (if no pixel with that value remains), or $n_{\max}^{k+1} < n_{\max}^k$ (if the maximal value remains equal).

Analogous reasoning for minima completes the proof. □

3.3 Convergence to a Flat Steady State

An immediate consequence of Proposition 2 is the following statement.

Corollary 1. *The only fixed points of the discrete FAB diffusion process (3), (4) are the flat images given for each $\mu \in \mathbb{R}$ by*

$$u_{i,j} = \mu \quad \textit{for all } (i,j) \in J\,. \tag{17}$$

By average grey-value invariance, the only steady state that could be reached from a given initial image $\boldsymbol{f}$ is that for which μ is the average grey value of $\boldsymbol{f}$. We will now prove convergence to this steady state.

Proposition 3. *Fully discrete FAB diffusion (3), (4) with $\boldsymbol{u}^0 = \boldsymbol{f}$ and time step size $\tau < \vartheta$ converges to the fixed point (17) where μ is the average grey value of $\boldsymbol{f}$.*

Proof. Consider the strictly decreasing (w.r.t. $\prec$) sequences $\big((u^k_{\max}, n^k_{\max})\big)_{k\in\mathbb{N}}$ and $\big((-u^k_{\min}, n^k_{\min})\big)_{k\in\mathbb{N}}$ from Proposition 2. These sequences are bounded from below by (a, N), $(-b, N)$, respectively, where $a := f_{\min}$, and $b := f_{\max}$. By an easy adaptation of the standard argument for sequences in $\mathbb{R}$ to sequences in $\mathbb{R} \times \mathbb{N}$ it follows that the sequences $(u^k_{\max})$, $(u^k_{\min})$ converge. Denote by $\bar{u}$, $\underline{u}$ their respective limits.

Assume $\underline{u} < \bar{u}$. By the maximum–minimum principle, $\boldsymbol{u}^k \in [a, b]^N$ holds for all k. Since $[a, b]^N$ is compact, the sequence $(\boldsymbol{u}^k)$ has a cumulation point. Because of the monotonicity of $(u^k_{\max})$, $(u^k_{\min})$ each cumulation point satisfies $u^*_{\max} = \bar{u}$, $u^*_{\min} = \underline{u}$.

We choose one cumulation point $\boldsymbol{u}^*$ and consider the FAB evolution $(\tilde{\boldsymbol{u}}^k)_{k\in\mathbb{N}_0}$ with initial condition $\tilde{\boldsymbol{u}}^0 = \boldsymbol{u}^*$. By Proposition 2, there exists a natural number K such that $\tilde{u}^K_{\max} = \max \tilde{\boldsymbol{u}} < \bar{u}$. Let therefore $\delta := \bar{u} - \tilde{u}^K_{\max} > 0$.

Moreover, the evolution (3), (4) satisfies a Lipschitz condition on $[a, b]^N$ with respect to the maximum norm $\|\cdot\|$, i.e.

$$\|u^k - \hat{u}^k\| < B \quad \Rightarrow \quad \|u^{k+1} - \hat{u}^{k+1}\| < LB \tag{18}$$

with some Lipschitz constant $L > 0$. Since $\boldsymbol{u}^*$ is a cumulation point of $(\boldsymbol{u}^k)$, we can choose k such that $\|\boldsymbol{u}^k - \boldsymbol{u}^*\| < \delta/L^K$. Consequently, $\|\boldsymbol{u}^{k+K} - \tilde{\boldsymbol{u}}^K\| < \delta$, and by the triangle inequality it follows that

$$u^{k+K}_{\max} \;\le\; \tilde{u}^K_{\max} + \|\boldsymbol{u}^{k+K} - \tilde{\boldsymbol{u}}^K\| \;<\; \bar{u} - \delta + \delta = \bar{u}\;, \tag{19}$$

contradicting the convergence of $(u^k_{\max})$ to $\bar{u}$. Thus, our assumption $\underline{u} < \bar{u}$ must be wrong, and we have $\underline{u} = \bar{u}$, i.e. convergence to a flat steady state. □

4 An Efficient and Stable Two-Pixel Scheme

Based on the stability result from Proposition 1 it is possible to compute FAB diffusion by a stable explicit scheme. The time step size limit imposed by (7), however, is way too small for practical purposes. In fact, (7) is an *a priori* estimate for the time step size resulting from several worst-case estimates which in general apply only to a few pixel locations. The theory in Section 3.1 can be used to derive instead *a posteriori* estimates which allow to steer the time stepping in an adaptive way. The resulting time steps will almost always be considerably larger than (7).

To make this precise, note that (3) can be written as $u^{k+1}_{i,j} = u^k_{i,j} + \tau \dot{u}^k_{i,j}$ where the finite-difference approximation $\dot{u}^k_{i,j}$ of $\operatorname{div}(g\,\boldsymbol{\nabla} u)$ can be computed independent of the time step size τ. To ensure that the maximum–minimum property is not violated within a time step, bounds on τ can then be derived directly from the criteria formulated in Lemma 1 and Lemma 2 for each pixel location. First, non-enhancement of extrema is

warranted for $\tau \leq \tau_{\max} := h_1^2 h_2^2/(2c_1(h_1^2 + h_2^2))$ as stated in the proof of Lemma 1. Second, local monotonicity preservation as stated in Lemma 2 can be achieved if in each time step, each non-maximal pixel is prevented from growing in excess of its largest neighbour, and each non-minimal pixel from decreasing below its smallest neighbour. Appropriate time step bounds can be computed for each pixel individually from $\dot{u}_{i,j}^k$.

Whereas these conditions could be evaluated to derive a *global* time step τ in each iteration, we obtain a higher effiency by following a radically *localised* approach: The diffusion flow $\dot{u}_{i,j}^k$ in (3) is composed of four two-pixel flows, each between the location (i,j) and one of its neighbours. Each of these flows appears with opposite signs in the flows of its two participating pixels, which ensures that the average grey-value conservation property of the continuous diffusion process is exactly fulfilled also in its discretisation (3). In order to maintain this conservation property, we choose the two-pixel flows as the elementary units for our scheme; compare also the proceeding in [2]. Starting from an approximation $\boldsymbol{u}^k$ pertaining to evolution time $k\tau_{\max}$, a global time step of size $\tau_{\max}$ is carried out by updating two-pixel flows in random order, each one with an appropriate time step. In the case of forward diffusion the time step is chosen so that the two interacting pixels preserve their monotonicity order. In the case of backward diffusion it is chosen such that they are prevented from growing above the maximum, or decreasing below the minimum of their respective neighbourhoods. Updated values $u_{i,j}$ enter immediately the computation of other pixels. This is repeated until all flows have reached the new time level, yielding the new approximation $\boldsymbol{u}^{k+1}$.

The time step $k \mapsto k+1$ thus starts by initialising an evolution time account for each pair of neighbouring pixels, $\{(i,j),(i',j')\}$ with $(i',j') = (i+1,j)$ or $(i',j') = (i,j+1)$, as $T_{i,j;i',j'} := \tau_{\max}$. Then, the following steps are repeated until all $T_{i,j;i',j'}$ reach zero, implying that we have progressed from sync time $k\tau_{\max}$ to $(k+1)\tau_{\max}$:

1. **Random selection:** Select a two-pixel pair $\{(i,j),(i',j')\}$ randomly, with the probability of each pair to be selected being proportional to $T_{i,j;i',j'}$.
2. **Diffusivity computation:** Compute $g_{i,j}$ and $g_{i',j'}$ as in (4); let $g := \frac{1}{2}(g_{i,j} + g_{i',j'})$.
3. **Flow computation:** Compute the flow $\dot{u} := g \cdot (u_{i',j'} - u_{i,j})/h^2$, using $h = h_1$ for horizontal or $h = h_2$ for vertical neighbours.
4. **Step-size determination:** Let $\tau^* := T_{i,j;i',j'}$.
   If $g > 0$ and $\tau^* > h^2/(2g)$, reduce τ^* to $h^2/(2g)$.
   If $g < 0$ and (i,j) is not a discrete local maximum, let (i^*,j^*) be its maximal neighbour. If $u_{i,j} + \tau^* \dot{u} > u_{i^*,j^*}$, reduce τ^* to $\tau^* := (u_{i^*,j^*} - u_{i,j})/\dot{u}$.
   If $g < 0$ and (i,j) is not a discrete local minimum, let (i^*,j^*) be its minimal neighbour. If $u_{i,j} + \tau^* \dot{u} < u_{i^*,j^*}$, reduce τ^* to $\tau^* := (u_{i^*,j^*} - u_{i,j})/\dot{u}$.
   If $g < 0$ and (i',j') is not a discrete local maximum, let (i^*,j^*) be its maximal neighbour. If $u_{i',j'} - \tau^* \dot{u} > u_{i^*,j^*}$, reduce τ^* to $\tau^* := (u_{i^*,j^*} - u_{i',j'})/(-\dot{u})$.
   If $g < 0$ and (i',j') is not a discrete local minimum, let (i^*,j^*) be its minimal neighbour. If $u_{i',j'} - \tau^* \dot{u} < u_{i^*,j^*}$, reduce τ^* to $\tau^* := (u_{i^*,j^*} - u_{i',j'})/(-\dot{u})$.
5. **Two-pixel flow update:** Update $u_{i,j}$ and $u_{i',j'}$ by replacing them with the new values $\tilde{u}_{i,j} := u_{i,j} + \tau^* \dot{u}$ and $\tilde{u}_{i',j'} := u_{i',j'} - \tau^* \dot{u}$. Decrease $T_{i,j;i',j'}$ by τ^*.

This process terminates, since τ^* cannot fall below the global positive limit (7). The approximation error of our scheme can be shown to be $\mathcal{O}(\tau + h^2 + \tau/h)$. This

ensures a conditionally consistent approximation to the FAB diffusion PDE if τ/h^2 is bounded by a constant when $\tau, h \to 0$. However, this is always satisfied by the stability conditions of our automatic time step size adaptation. For an efficient implementation of the algorithm, the performance of the selection in Step 1 is crucial. To this end, the bookkeeping of time step accounts $T_{i,j;i',j'}$ is done within a binary tree structure. The selection then requires logarithmic time w.r.t. the total number of pixels.

5 Experiments

Let us now evaluate our numerical algorithms. To this end we have implemented them in ANSI C and compiled the code with a GNU gcc compiler. We report runtimes on a single core of an Acer P 645 Laptop with an Intel® Core™ i5-5200U CPU running at 2.20 GHz. No advanced code optimisations took place.

In our first experiment we compare an implementation of our explicit scheme (3) with standard (5) or nonstandard discretisation (4) and its two-pixel variant (with nonstandard discretisation). We use the diffusivity function (2). Fig. 2 shows the results for $\lambda = 4$, $\kappa = 2.5$ and stopping time $t = 10$. First, we observe that an explicit scheme with standard discretisation is unstable, even for very small time step sizes. Thus, it is not considered any further. Second, we see that an explicit scheme with nonstandard discretisation is stable for small time steps and visually equivalent to its two-pixel counterpart. However, the runtimes of both schemes differ enormously:

- **Explicit scheme (with nonstandard discretisation):**
  For the diffusivity (2) with $\lambda = 4$ and $\kappa = 2.5$, the constant ω in the time step size limit (7) is given by $\omega = 0.009568$. Together with the grid sizes $h_1 = h_2 = 1$ this yields a time step size restriction of $\tau \leq 1.14 \cdot 10^{-5}$. Choosing $\tau := 1 \cdot 10^{-5}$ requires as many as 1 million iterations to reach a stopping time of $t = 10$. The corresponding CPU time was 70 minutes 13 seconds.
- **Two-pixel scheme:**
  For the two-pixel variant of the explicit scheme with nonstandard discretisation we used a sync step size of $\tau_{\max} = 0.1$. Thus, only 100 sync steps are necessary to reach a stopping time of $t = 10$. This requires a CPU time of 7.73 seconds.

We observe that our two-pixel scheme gives a speed-up by a factor 544 !

Since the two-pixel algorithm is highly efficient, it can also be used for long term computations, arising e.g. in scale-space analysis. Fig. 3 depicts the scale-space behaviour of a noisy test image when the scale-space is governed by FAB diffusion. We observe the high robustness of the FAB scale-space in spite of the fact that it uses negative diffusities. Moreover, the experiment confirms convergence to a flat steady state for $t \to \infty$. This is in full accordance with our theoretical results established in Section 3.

6 Summary and Conclusions

While backward diffusion suffers from an extremely bad reputation of being a terribly ill-posed process, we have seen in our paper that it can be turned into a highly stable evolution, provided that some essential requirements are met:

Fig. 2. Influence of the numerical scheme on the result of FAB diffusion ($\lambda = 4$, $\kappa = 2.5$, $t = 10$). **From left to right: (a)** Test image, 256×256 pixels. **(b)** Explicit scheme (3) with standard discretisation (5). Computing 1 million iterations with time step size $\tau = 10^{-5}$ requires 70 minutes and 13 seconds, while the result is unstable. **(c)** Explicit scheme (3) with nonstandard discretisation (5). Performing 1 million iterations with $\tau = 10^{-5}$ takes 66 minutes and 34 seconds. **(d)** Corresponding two-pixel scheme. We used 100 iterations with synchronisation step size $\tau_{\max} = 0.1$, leading to a runtime of 7.73 seconds. The average time step size was 0.0991.

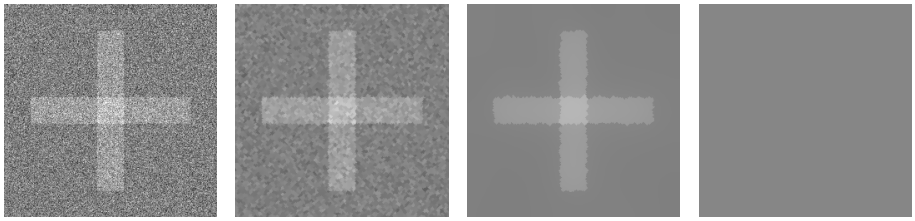

Fig. 3. Scale-space behaviour of FAB diffusion ($\lambda = 2$, $\kappa = 2.5$). **From left to right: (a)** Test image, 256×256 pixels. **(b)** After a diffusion time of $t = 10$. **(c)** $t = 300$. **(d)** $t = 10\,000$. All computations have been done with a two-pixel scheme with sync step size $\tau_{\max} = 0.1$.

First of all, it must be stabilised at extrema in order to avoid under- and overshoots. The FAB diffusion paradigm does take care of this. Our discrete analysis is based on a FAB diffusivity that attains a positive diffusivity in zero which is larger than the moduli of all negative diffusivities. Under this mild model assumption we were able to establish a maximum–minimum principle for an explicit 2D scheme with nonstandard discretisation as well as convergence to a flat steady state, if one adheres to a very restrictive time step size limit. In order to make this concept practically viable, we came up with a novel scheme that combines several unconventional features:

- By splitting the diffusion process into a sequence of **two-pixel interactions**, we ended up with the most local scheme that is possible. The simplicity of two-pixel interactions allowed to adapt the time step size locally and use only small time steps at those locations where this is unavoidable. Although **local time step size adaptations** are uncommon in PDE-based image analysis, we have shown that they may lead to speed-ups by three orders of magnitude.
- In contrast to other splittings in the PDE-based image analysis literature which are usually synchronous, our splitting is **asynchronous**: Generating a sequence of simple two-pixel interactions turns out to be attractive, because their stability follows trivially from the stability of each two-pixel interaction.

– Introducing a **randomisation** in the order of the two-pixel diffusions removes any directional bias that is characteristic of sequential splittings which are carried out in a deterministic order.

It is our hope that our results may help to improve the reputation of backward parabolic processes, since they can offer some very attractive image enhancement properties that have hardly been explored so far, mainly because of the lack of stable numerical schemes. In our ongoing work we are also looking into extensions to other time discretisations such as (semi-)implicit schemes.

References

1. Aubert, G., Kornprobst, P.: Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations. Second edn. Volume 147 of Applied Mathematical Sciences. Springer, New York (2006)
2. Burgeth, B., Weickert, J., Tari, S.: Minimally stochastic schemes for singular diffusion equations. In Tai, X.C., Lie, K.A., Chan, T.F., Osher, S., eds.: Image Processing Based on Partial Differential Equations. Springer, Berlin (2007) 325–339
3. Gilboa, G., Sochen, N., Zeevi, Y.Y.: Image sharpening by flows based on triple well potentials. Journal of Mathematical Imaging and Vision **20** (2004) 121–131
4. Gilboa, G., Sochen, N.A., Zeevi, Y.Y.: Forward-and-backward diffusion processes for adaptive image enhancement and denoising. IEEE Transactions on Image Processing **11** (2002) 689–703
5. Kramer, H.P., Bruckner, J.B.: Iterations of a non-linear transformation for enhancement of digital images. Pattern Recognition **7** (1975) 53–58
6. Mrázek, P., Weickert, J., Steidl, G.: Diffusion-inspired shrinkage functions and stability results for wavelet denoising. International Journal of Computer Vision **64** (2005) 171–186
7. Osher, S., Rudin, L.: Shocks and other nonlinear filtering applied to image processing. In Tescher, A.G., ed.: Applications of Digital Image Processing XIV. Volume 1567 of Proceedings of SPIE. SPIE Press, Bellingham (1991) 414–431
8. Osher, S., Rudin, L.I.: Feature-oriented image enhancement using shock filters. SIAM Journal on Numerical Analysis **27** (1990) 919–940
9. Perona, P., Malik, J.: Scale space and edge detection using anisotropic diffusion. IEEE Transactions on Pattern Analysis and Machine Intelligence **12** (1990) 629–639
10. Pollak, I., Willsky, A.S., Krim, H.: Image segmentation and edge enhancement with stabilized inverse diffusion equations. IEEE Transactions on Image Processing **9** (2000) 256–266
11. Smolka, B.: Combined forward and backward anisotropic diffusion filtering of color images. In Van Gool, L., ed.: Pattern Recognition. Volume 2449 of Lecture Notes in Computer Science. Springer, Berlin (2002) 314–320
12. Weickert, J.: Anisotropic Diffusion in Image Processing. Teubner, Stuttgart (1998)
13. Weickert, J., Benhamouda, B.: A semidiscrete nonlinear scale-space theory and its relation to the Perona–Malik paradox. In Solina, F., Kropatsch, W.G., Klette, R., Bajcsy, R., eds.: Advances in Computer Vision. Springer, Wien (1997) 1–10
14. Welk, M., Gilboa, G., Weickert, J.: Theoretical foundations for discrete forward-and-backward diffusion filtering. In Tai, X.C., Mørken, K., Lysaker, M., Lie, K.A., eds.: Scale Space and Variational Methods in Computer Vision. Volume 5567 of Lecture Notes in Computer Science. Springer, Berlin (2009) 527–538
15. Welk, M., Weickert, J., Galić, I.: Theoretical foundations for spatially discrete 1-D shock filtering. Image and Vision Computing **25** (2007) 455–463