Stable Models and Algorithms for Backward Diffusion Evolutions

Joachim Weickert
Mathematical Image Analysis Group
Saarland University, Saarbrücken, Germany

joint work with
Martin Welk (UMIT, Hall, Austria)
Leif Bergerhoff (Saarland University)
Marcelo Cárdenas (Saarland University)
Guy Gilboa (Technion, Haifa, Israel)

partial funding: DFG Leibniz Award and ERC Advanced Grant INCOVID

Introduction (1)

Introduction

♦ Forward diffusion equations blur or smooth images.
   → attempts to invert these evolutions for deblurring or sharpening images
Problems

- Backward diffusion is typically regarded as ill-posed:
  - Solution does not exist for non-smooth initial data.
  - If it exists, it is highly sensitive w.r.t. perturbations.
- Thus, many researchers refrain from using backward diffusion.

Goals

- show how these problems can be handled by sophisticated numerics or circumvented by smart modelling
- demonstrate these principles with two prototypical applications:
  - advanced numerics for the FAB diffusion of Gilboa et al. 2002
  - novel convex model for backward diffusion (Bergerhoff et al. 2018)

Outline

- FAB Diffusion
  - Continuous Model
  - Explicit Scheme
  - Efficient Numerics
  - Experiments
- Backward Diffusion with Convex Energy
  - Model and Theory
  - Numerical Algorithm
  - Experiment
- Conclusions
FAB Diffusion: Continuous Model

The Perona–Malik Filter (1990)

Consider open image domain $\Omega \subset \mathbb{R}^2$ and some bounded image $f : \Omega \to \mathbb{R}$.

Create family of filtered versions $u(x, t)$ of $f(x)$ as solution of

$$
\partial_t u = \text{div} \left( g(|\nabla u|^2) \nabla u \right) \quad \text{on} \quad \Omega \times (0, \infty),
$$

$$
u(x, 0) = f(x) \quad \text{on} \quad \Omega,
$$

$$
n^\top \nabla u = 0 \quad \text{on} \quad \partial \Omega \times (0, \infty),
$$

where $n$ denotes the outer normal vector to the image boundary $\partial \Omega$.

Diffusivity $g$ is monotonically decreasing positive function of $|\nabla u|^2$.

Smoothes within flat regions and enhances edges between them.

Gradient descent of a possibly nonconvex but monotone energy

$$
E(u) = \int_{\Omega} \Psi(|\nabla u|^2) \, dx
$$

where the penaliser (potential) $\Psi(|\nabla u|^2)$ satisfies $\Psi'(|\nabla u|^2) = g(|\nabla u|^2)$.  

FAB Diffusion: Continuous Model (2)

Forward-and-Backward (FAB) Diffusion

(Gilboa / Sochen / Zeevi 2002)

- goal: stronger sharpening than classical Perona-Malik filters
- equip Perona-Malik diffusion

\[
\partial_t u = \text{div} \left( g(|\nabla u|^2) \nabla u \right)
\]

with a diffusivity that takes positive and \textit{negative} values.

- fairly mild assumptions in this talk:

\[
g \in C^1[0, \infty), \quad g(0) = c_1 > 0, \quad g(.) \geq -c_2 \quad \text{with} \quad c_1 > c_2 \geq 0.
\]

- corresponds to \textit{nonconvex and nonmonotone} potential \( \Psi(|\nabla u|^2) \)

FAB Diffusion: Continuous Model (3)

How Unpleasant can this Become?

- Diffusivity \( g(s^2) \)
- Diffusivity \( g(s^2) \)
- Diffusivity \( g(s^2) \)
- Potential \( \Psi(s^2) \)
- Potential \( \Psi(s^2) \)
- Potential \( \Psi(s^2) \)
Theoretical Results so Far

- cannot be covered by standard theory for diffusion filters (W. 1998)
- no continuous well-posedness theory
- Gilboa / Sochen / Zeevi (JMIV 2004): experimental stabilisation with a fidelity term and biharmonic regularisation

Can we establish a fully discrete theory?
Does this lead to practical algorithms for images?

Outline

- FAB Diffusion
  - Continuous Model
  - Explicit Scheme
  - Efficient Numerics
  - Experiments
- Backward Diffusion with Convex Energy
  - Model and Theory
  - Numerical Algorithm
  - Experiment
- Conclusions
FAB Diffusion: Explicit Scheme

Goal
- establish comprehensive theory for an explicit discretisation of FAB diffusion

Explicit Scheme

Explicit finite difference discretisation of diffusion equation

\[ \partial_t u = \partial_x \left( g(|\nabla u|^2) \partial_x u \right) + \partial_y \left( g(|\nabla u|^2) \partial_y u \right) \]

in some inner pixel \((i,j)\) at time level \(k\) yields the scheme

\[ \frac{u_{i,j}^{k+1} - u_{i,j}^k}{\tau} = \frac{1}{h_1} \left( \frac{g_{i+1,j}^k + g_{i,j}^k}{2} \frac{u_{i+1,j}^k - u_{i,j}^k}{h_1} - \frac{g_{i,j}^k + g_{i-1,j}^k}{2} \frac{u_{i,j}^k - u_{i-1,j}^k}{h_1} \right) \]

\[ + \frac{1}{h_2} \left( \frac{g_{i,j+1}^k + g_{i,j}^k}{2} \frac{u_{i,j+1}^k - u_{i,j}^k}{h_2} - \frac{g_{i,j}^k + g_{i,j-1}^k}{2} \frac{u_{i,j}^k - u_{i,j-1}^k}{h_2} \right) \]

with grid sizes \(h_1, h_2\) and time step size \(\tau\).

Where Do Problems Arise?
- The standard discretisation of \(g(|\nabla u|^2)\) is given by

\[ g_{i,j}^k := g \left( \frac{u_{i+1,j}^k - u_{i-1,j}^k}{2h_1} \right)^2 + \left( \frac{u_{i,j+1}^k - u_{i,j-1}^k}{2h_2} \right)^2. \]

- It can be positive at extrema

- It can create negative diffusivities in extrema, which give rise to instabilities.

Is There a Remedy?
- A nonstandard discretisation produces a vanishing gradient in extrema:

\[ g_{i,j}^k := g \left( \max \left( \frac{u_{i+1,j}^k - u_{i,j}^k}{h_1}, \frac{u_{i,j}^k - u_{i-1,j}^k}{h_1}, 0 \right) \right) \]

\[ + \max \left( \frac{u_{i,j+1}^k - u_{i,j}^k}{h_2}, \frac{u_{i,j}^k - u_{i,j-1}^k}{h_2}, 0 \right) \].

- same quadratic consistency order as standard discretisation
Two Technical Definitions

• The grey values of $f = (f_i) \in \mathbb{R}^N$ are restricted to a finite interval of length

$$R := \max_i f_i - \min_i f_i.$$ 

• Since $g$ is continuous and $c_1 > c_2 \geq 0$, there exists a constant $\omega > 0$ such that

$$g(s^2) > c_2 \quad \forall s \in (0, \omega R).$$

Theorem [Theory for the Explicit FAB Scheme]

With the preceding assumptions and definitions, consider the explicit scheme for FAB diffusion with nonstandard discretisation. If the time step size $\tau$ satisfies

$$\tau \leq \frac{\omega^2 h_1^4 h_2^4}{2 c_1 \cdot (h_1^2 + h_2^2) \cdot \left(\omega^2 h_1^2 h_2^2 + h_1^2 + h_2^2\right)},$$

then this scheme has the following properties:

• Well-Posedness

For every $k \in \mathbb{N}_0$, the solution $u^{k+1}$ depends in a continuous way on perturbations of the initial image $f$.

• Average Grey Value Invariance

$$\frac{1}{N} \sum_{j=1}^N u_j^k = \frac{1}{N} \sum_{j=1}^N f_j =: \mu \quad \forall k \in \mathbb{N}_0.$$
Maximum-Minimum Principle

\[ \min_j f_j \leq u^k_i \leq \max_j f_j \quad \forall i, \forall k \in \mathbb{N}_0. \]

Lyapunov Sequence

\[ V^k := \max_j u^k_j - \min_j u^k_j \]

is a Lyapunov sequence: decreasing in \( k \) and bounded from below.

Convergence to a Constant Steady State

\[ \lim_{k \to \infty} u^k_i = \mu \quad \forall i. \]
FAB Diffusion: Efficient Numerics

Our explicit FAB scheme with nonstandard discretisation has a clean theory.

However, it must satisfy a severe time step size restriction:
This can lead to impractically small time steps: $10^{-6}, \ldots, 10^{-5}$

Reason:

- Based on worst case \textit{a priori} estimates
- Restrictions are not needed everywhere and at all time steps

Remedy:

- Replace pessimistic \textit{a priori} estimates by realistic \textit{a posteriori} estimates.
- Act as locally/adaptive as possible in space and time.
- Realisation: two-pixel interactions.

Motivation for Two-Pixel Interactions

- Most local diffusion interaction that respects a conservation law.

The explicit scheme performs four two-pixel interactions simultaneously:

$$u_{i,j}^{k+1} = u_{i,j}^k + \tau \left(\frac{g_{i+1,j}^k + g_{i,j}^k}{2} \cdot \frac{u_{i+1,j}^k - u_{i,j}^k}{h_1^2} - \frac{g_{i,j}^k + g_{i-1,j}^k}{2} \cdot \frac{u_{i,j}^k - u_{i-1,j}^k}{h_1^2}\right) + \frac{g_{i,j+1}^k + g_{i,j}^k}{2} \cdot \frac{u_{i,j+1}^k - u_{i,j}^k}{h_2^2} - \frac{g_{i,j}^k + g_{i,j-1}^k}{2} \cdot \frac{u_{i,j}^k - u_{i,j-1}^k}{h_2^2}$$
Basic Idea behind Two-Pixel Scheme

- Decouple explicit scheme into sequential (asynchronous) two-pixel interactions. Then stability follows trivially from the stability of each interaction.

- In each interaction, choose the largest time step ensuring two stability criteria:
  - For a positive diffusivity, the order of grey values must not be flipped.
  - For a negative diffusivity, we have a non-extremal pixel. It must not become larger/smaller than its largest/smallest neighbour.

- This gives highly localised time step size restrictions in space and time.

- To avoid directional bias, randomise the order of two-pixel interactions.

- Introduce sync times at which each all pixels reach the same time level. Use e.g. the stability bounds of an explicit forward diffusion scheme.

- selection probability for two-pixel interaction: proportional to remaining time until synchronisation.

Outline

- FAB Diffusion
  - Continuous Model
  - Explicit Scheme
  - Efficient Numerics
  - Experiments

- Backward Diffusion with Convex Energy
  - Model and Theory
  - Numerical Algorithm
  - Experiment

- Conclusions
FAB Diffusion: Experiments

Technical Details

We use the FAB diffusivity

\[ g(s^2) = 2 \exp \left( -\frac{\kappa^2 \ln 2 \cdot s^2}{\kappa^2 - 1} \right) - \exp \left( -\frac{\ln 2 \cdot s^2}{\kappa^2 - 1} \right) \]

with contrast parameter \( \lambda > 0 \) and stretching parameter \( \kappa > 1 \).

- Run times refer to a C implementation on a single core.
- Hardware: Intel Core i5–5200U CPU at 2.20 GHz.

Stability: Standard versus Nonstandard Discretisation

Left: Original image, 256 × 256 pixels. Middle: Using the standard discretisation within FAB diffusion creates instabilities. Diffusivity parameters \( \lambda = 4 \) and \( \kappa = 2.5 \), time step size \( \tau = 10^{-5} \), and stopping time \( t = 10 \). Values outside \([0, 255]\) have been cropped in the visualisation. Right: The same experiment with nonstandard discretisation does not give rise to instabilities.
**FAB Diffusion: Experiments (3)**

**Efficiency: Explicit Scheme versus Two-Pixel Scheme**

![Explicit scheme](image1) ![Two-pixel scheme](image2)

- **10^6 iterations with** $\tau = 10^{-5}$
- **run time:** 66 min
- **100 sync steps with** $\tau_{\text{max}} = 0.1$
- **average time step size:** 0.0991
- **run time:** 7.73 s

- Both schemes give results of comparable visual quality.
- The two-pixel scheme is 544 times faster.

**FAB Diffusion: Experiments (4)**

**Scale-Space Behaviour**

- $t = 0$
- $t = 6$
- $t = 40$
- $t = 200$
- $t = 800$
- $t = 3000$

Scale-space behaviour of FAB diffusion with $\lambda = 2$ and $\kappa = 2.5$. All computations use the two-pixel scheme with sync time step size $\tau_{\text{max}} = 0.1$. 
Robustness under Noise

test image *barbara*, 512 × 512 pixels

Gaussian noise with $\sigma = 50$, truncated outside $[0, 255]$

FAB diffusion, 2-pixel scheme ($\lambda = 2$, $\kappa = 2.5$, $t = 120$)

Outline

- FAB Diffusion
  - Continuous Model
  - Explicit Scheme
  - Efficient Numerics
  - Experiments

- Backward Diffusion with Convex Energy
  - Model and Theory
  - Numerical Algorithm
  - Experiments

- Conclusions
Backward Diffusion with Convex Energy: Model and Theory

Problem

- common stabilisation constraints for backward diffusion:
  - forward or zero diffusion at extrema: \( \implies \) requires sophisticated numerical schemes
  - fidelity term: \( \implies \) dependence on fidelity weights and input data

Remedy: Smarter Modelling

- novel backward diffusion model with globally negative diffusivities
- results from gradient descent of a convex (!) energy
- stabilisation through reflecting boundary conditions in the co-domain
- will allow standard numerics

Model and Theory

- vector \( \mathbf{v} = (v_1, \ldots, v_N)^\top \in (0,1)^N \) with \( N \) distinct 1D particle positions \( v_i \)
- extend \( \mathbf{v} \) with additional particles \( v_{N+1}, \ldots, v_{2N} \):
  mirror all positions \( v_1, \ldots, v_N \) at the right domain boundary 1

Exemplary setup for \( N = 7 \) particles.

- As our baseline model, we consider an energy with nonlocal interactions:

\[
E(\mathbf{v}) = \frac{1}{2} \sum_{i=1}^{2N} \sum_{j=1}^{2N} \Psi((v_j - v_i)^2),
\]

where \( \Psi : \mathbb{R}_0^+ \rightarrow \mathbb{R} \) is a specific repulsive penaliser function.
**Which Repulsive Penaliser $\Psi(s^2)$ Do We Choose?**

Penaliser $\Psi(s^2) = (s - 1)^2 - 1$ for $s \in [0, 1]$, extended to $[-1, 1]$ by symmetry and to $[-1, 3]$ by periodicity.

- decreasing and *strictly convex* for $s \in [0, 1]$
- extension to $[-1, 1]$ by symmetry
- extension to $\mathbb{R}$ by periodicity
- differentiable everywhere except at even integers

**What is the Gradient Descent Evolution of Our Model?**

Our discrete energy function

$$E(v) = \frac{1}{2} \cdot 2^N \sum_{i=1}^{2N} \sum_{j=1}^{2N} \Psi((v_j - v_i)^2)$$

yields the gradient descent evolution

$$\partial_t v_i = \sum_{j=1 \atop j \neq i}^{2N} \Psi'((v_j - v_i)^2) (v_j - v_i) =: \sum_{j=1 \atop j \neq i}^{2N} \Phi(v_j - v_i) .$$

This is a *space-discrete nonlocal diffusion* model with

- *diffusivity function* $\Psi'(s^2)$
- *flux function* $\Phi(s) := \Psi'(s^2) s$
Diffusivity and Flux Functions for $\Psi(s^2) = (s-1)^2 - 1$

- The diffusivity function $\Psi'(s^2)$ is a shifted backward TV diffusivity in $(0, 1]$:

![Graph showing diffusivity function]

For $s \in (0, 1]$, the diffusivity satisfies $\Psi'(s^2) = 1 - \frac{1}{s}$.

- The flux function $\Phi(s) := \Psi'(s^2) s$ is negative everywhere in $(0, 1)$.

![Graph showing flux function]

For $s \in (0, 2)$, the flux function is given by $\Phi(s) = s - 1$.

Which Properties can be Proven for the Baseline Model?

- Well-posedness:
  The strictly convex energy has a unique minimiser.
  The gradient descent evolution depends continuously on the input data.

- Particles can never reach the domain boundaries.

- Particles cannot occupy the same position.

- Strict global minimum of $E(v)$:
  Convergence to equilibrium point $v^*$ for $t \to \infty$

- Steady-state solution $v^*$ explicitly known:
  Particles are distributed equidistantly in $(0, 1)$.

![Steady state for $N = 7$ particles]

Can we change the model such that we get a more interesting steady state?
Generalised Model with Weights

- We can assign fixed nonnegative weights \( w_1, \ldots, w_N \) to our \( N \) particles.
- They are mirrored and periodically extended like the particles.
- With \( p_i := \sqrt{w_i} \cdot v_i \) the energy for the **generalised model** reads

\[
E(p, w) = \frac{1}{2} \cdot \sum_{i=1}^{2N} \sum_{j=1}^{2N} w_i \cdot w_j \cdot \Psi \left( \left( \frac{p_j}{\sqrt{w_j}} - \frac{p_i}{\sqrt{w_i}} \right)^2 \right).
\]

Exemplary setup for \( N = 7 \) particles with \( w_i = \frac{1}{i} \).

Which Properties can be Proven for the Generalised Model?

- same as for the baseline model
- Only difference: The steady-state \( p_i^* \) is no longer equidistantly distributed:

\[
p_i^* = \sqrt{w_i} \cdot \frac{\sum_{j=1}^{i} w_j - \frac{1}{2} w_i}{\sum_{j=1}^{N} w_j}, \quad i = 1, \ldots, N
\]

Steady state for \( N = 7 \) particles with \( w_i = \frac{1}{i} \).
Backward Diffusion with Convex Energy: Numerical Algorithm

A simple explicit time discretisation of the $N$ particle evolution works well!

time step size restriction involves Lipschitz constant $L_\Phi$ of flux $\Phi(s), s \in (0, 2)$:

$$0 < \tau < \frac{1}{2 L_\Phi \sum_{i=1}^{N} w_i}$$

algorithm reproduces the stability properties of time-continuous evolution:

- Particles cannot reach the domain boundary.
- Particles do not change their order.

no problems due to negative diffusivities
Backward Diffusion with Convex Energy: Experiment

Goal

- enhance global contrast of a digital greyscale image
  \[ f : \{1, \ldots, n_x\} \times \{1, \ldots, n_y\} \rightarrow (0, 1) \]

Algorithm Using our Generalised Model

- If a grey value \( i \) appears \( n \) times, set its weight to \( w_i := n \).
- Evolve the explicit scheme for some given time \( t \),
  or use the known steady state solution for \( t \rightarrow \infty \).
- Map the original grey values to the processed ones to get the enhanced image.
Outline

**FAB Diffusion**
- Continuous Model
- Explicit Scheme
- Efficient Numerics
- Experiments

**Backward Diffusion with Convex Energy**
- Model and Theory
- Numerical Algorithm
- Experiments

**Conclusions**
Conclusions

- Backward diffusion can be tamed by sophisticated numerics or smart models.

- important components of sophisticated numerics:
  - nonstandard discretisations to preserve continuous qualities
  - two-pixel interactions to achieve highest locality
  - local time step size adaptations to increase efficiency
  - asynchronous splittings for simple stability guarantees
  - randomisation to avoid directional bias

- features of well-posed smart models:
  - gradient descent of strictly convex energies
  - stabilisation through range constraints
  - allow simple numerical schemes

References


Download:  www.mia.uni-saarland.de/weickert/publications.shtml

Postdoc Opening in Inpainting-Based Image Compression

- funded by ERC Advanced Grant
- required: expertise in optimisation, variational methods, or PDEs
- e-mail CV: weickert@mia.uni-saarland.de

Thank you!