Stable Backward Diffusion Models that Minimise Convex Energies

Leif Bergerhoff\textsuperscript{1}, Marcelo Cárdenas\textsuperscript{2}, Joachim Weickert\textsuperscript{1}, and Martin Welk\textsuperscript{2}

\textsuperscript{1} Mathematical Image Analysis Group, Faculty of Mathematics and Computer Science, Saarland University, Campus E1.7, 66041 Saarbrücken, Germany
{bergerhoff,cardenas,weickert}@mia.uni-saarland.de

\textsuperscript{2} 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

Abstract. Backward diffusion processes appear naturally in image enhancement and de-blurring applications. However, the inverse problem of backward diffusion is known to be ill-posed and straightforward numerical algorithms are unstable. So far, existing stabilisation strategies in the literature require sophisticated numerics to solve the underlying initial value problem. Therefore, it is desirable to establish a backward diffusion model which implements a smart stabilisation approach that can be used in combination with a simple numerical scheme. We derive a class of space-discrete one-dimensional backward diffusion as gradient descent of energies where we gain stability by imposing range constraints. Interestingly, these energies are even convex. Furthermore, we establish a comprehensive theory for the time-continuous evolution and we show that stability carries over to a simple explicit time discretisation of our model. Finally, we confirm the stability and usefulness of our technique in experiments in which we enhance the contrast of digital greyscale and colour images.

Keywords: Inverse Problem · Backward Diffusion · Modelling · Convex Energy · Gradient Descent · Contrast Enhancement · Image Processing

Mathematics Subject Classification (2010): 76R50 · 60J60 · 58J65 · 37N30 · 37N40 · 65D18 · 68U10

1 Introduction

Forward diffusion processes are well-suited to describe the smoothing of a given signal or image. This smoothing operation implies a loss of high frequencies or details in the original data. As a result, the inverse problem, backward diffusion, suffers from deficient information which is needed to uniquely reconstruct the original data. The introduction of noise due to measured data increases this difficulty even further. Consequently, a solution to the inverse problem – if it exists at all – is highly sensitive and heavily depends on the input data: Even the smallest flaw in the initial data can have a large impact on the evolution and may cause large deviations. Therefore, it becomes clear that backward diffusion processes necessitate further stabilisation.

Previous Work on Backward Diffusion. Already more than 60 years ago, John discussion the quality of a numerical solution to the inverse diffusion problem given that a solution exists, and that it is bounded and non-negative. Since then, a large number of different regularisation methods have evolved which achieve stability by
bounding the noise of the measured and the unperturbed data \[49\], by operator splitting \[27\], by Fourier regularisation \[13\], or by a modified Tikhonov regulariser \[54\]. Hào and Duc \[23\] suggest a mollification method where stability for the inverse diffusion problem follows from a convolution with the Dirichlet kernel. In \[24\] the same authors provide a regularisation method for backward parabolic equations with time-dependent coefficients. Ternat et al. \[51\] suggest low-pass filters and fourth-order regularisation terms for stabilisation.

Backward parabolic differential equations also enjoy high popularity in the image analysis community where they have e.g. been used for image restoration and image deblurring respectively. The first contribution to backward diffusion dates back to 1955 when Kovásznay and Joseph \[28\] proposed to use the scaled negative Laplacian for contour enhancement. Gabor \[14\] observed that the isotropy of the Laplacian operator leads to amplification of accidental noise at contour lines at the same time it enhances the contour lines. As a remedy, he proposed to restrict the contrast enhancement to the orthogonal contour direction and – in a second step – suggested additional smoothing in tangent direction. Lindenbaum et al. \[29\] make use of averaged derivatives in order to improve the directional sensitive filter by Gabor. However, the authors point out that smoothing in only one direction favours the emergence of artefacts in nearly isotropic image regions. They recommend to use the Perona-Malik filter \[40\] instead. Forces of Perona-Malik type are also used by Pollak et al. \[43\] who specify a family of evolution equations to sharpen edges and suppress noise in the context of image segmentation. In \[50\], ter Haar Romeny et al. stress the influence of higher order time derivatives on the Gaussian deblurred image. Referring to the heat equation, the authors express the time derivatives in the spatial domain and approximate them using Gaussian derivatives. Steiner et al. \[47\] highlight how backward diffusion can be used for feature enhancement in planar curves.

In the field of image processing, a frequently used stabilisation technique constrains the extrema in order to enforce a maximum-minimum principle. This is e.g. implemented in the inverse diffusion filter of Osher and Rudin \[38\]. It imposes zero diffusivities at extrema and applies backward diffusion everywhere else. The so-called forward-and-backward (FAB) diffusion of Gilboa et al. \[19\] follows a slightly different approach. Closely related to the Perona-Malik filter \[40\] it uses negative diffusivities for a specific range of gradient magnitudes. On the other hand, it imposes forward diffusion for values of low and zero gradient magnitude. By doing so, the filter prevents the output values from exploding at extrema. However, it is worth mentioning that – so far – all adequate implementations of inverse diffusion processes with forward or zero diffusivities at extrema require sophisticated numerical schemes \[38,53\].

Another, less popular stabilisation approach implies the application of a fidelity term and has been used to penalise deviations from the input image \[46,6\] or from the average grey value of the desired range \[45\]. Consequently, both the weights of the fidelity and the diffusion term control the range of the filtered image.

Further methods achieve stabilisation using a regularisation strategy built on FFT-based operators \[7,8,9\] and by the restriction to polynomials of fixed finite
degree [25]. Mair et al. [31] discuss the well-posedness of deblurring Gaussian blur in the discrete image domain based on analytic number theory.

In summary, the presented methods offer an insight into the challenge of handling backward diffusion in practice and underline the demand for careful stabilisation strategies and sophisticated numerical methods.

Diffusion and Energy Minimisation. In our paper we are going to present an alternative approach to deal with backward diffusion problems. To understand it better, it is useful to recapitulate some relations between diffusion and energy minimisation. For the sake of convenience we assume a one-dimensional evolution that smoothes an initial signal $f : [a, b] \to \mathbb{R}$. In this context, the original signal $f$ serves as initial state of the diffusion equation

$$\partial_t u = \partial_x (g(u_x^2) u_x)$$  \hspace{1cm} (1)

where $u = u(x, t)$ represents the filtered outcome with $u(x, 0) = f(x)$. Additionally, let $u_x = \partial_x u$ and assume reflecting boundary conditions at $x = a$ and $x = b$.

Given a nonnegative diffusivity function $g$, growing diffusion times $t$ lead to simpler representations of the input signal. From Perona and Malik’s work [40] we know that the smoothing effect at signal edges can be reduced if $g$ is a decreasing function of the contrast $u_x^2$. As long as the flux function $\Phi(u_x) := g(u_x^2) u_x$ is strictly increasing in $u_x$ the corresponding forward diffusion process $\partial_t u = \Phi(u_x) u_{xx}$ involves no edge sharpening. This diffusion can be regarded as the gradient descent evolution which minimises the energy

$$E[u] = \int_a^b \Psi(u_x^2) \, dx.$$  \hspace{1cm} (2)

The potential function $\tilde{\Psi}(u_x) = \Psi(u_x^2)$ is strictly convex in $u_x$, increasing in $u_x^2$, and fulfils $\tilde{\Psi}'(u_x^2) = g(u_x^2)$. Furthermore, the energy functional has a flat minimiser which is – due to the strict convexity of the energy functional – unique. The gradient descent / diffusion evolution is well-posed and converges towards this minimiser for $t \to \infty$. Due to this classical emergence of well-posed forward diffusion from strictly convex energies it seems natural to believe that backward diffusion processes are necessarily associated with non-convex energies. However, as we will see, this conjecture is wrong.

Our Contribution. In our article, we show that a specific class of backward diffusion processes are gradient descent evolutions of energies that have one unexpected property: They are convex! Our second innovation is the incorporation of a specific constraint: We impose reflecting boundary conditions in the diffusion co-domain. While such range constraints have shown their usefulness in some other context (see e.g. [33]), to our knowledge they have never been used for stabilising backward diffusions. For our novel backward diffusion models, we show also a surprising numerical fact: A simple explicit scheme turns out to be stable and convergent. Last but not least, we apply our models to the contrast enhancement of greyscale and colour images.

This article is a revised version of our conference contribution [4] which we extend in several aspects. First, we enhance our model for convex backward diffusion to
support not only a global and weighted setting but also a localised variant. We analyse this extended model in terms of well-posedness, stability, and convergence towards a unique minimiser. Furthermore, we formulate a simple explicit scheme for our newly proposed approach which shares all important properties with the time-continuous evolution. In this context, we provide a detailed discussion on the selection of suitable time step sizes. Additionally, we suggest two new applications: global contrast enhancement of digital colour images and local contrast enhancement of digital grey and colour images.

Structure of the Paper. In Section 2, we present our model for convex backward diffusion with range constraints. We describe a general approach which allows to formulate weighted local and global evolutions. Section 3 includes proofs for model properties such as range and rank-order preservation as well as convergence analysis and the derivation of explicit steady-state solutions. Section 4 provides a simple explicit scheme which can be used to solve the occurring initial value problem. In Section 5, we explain how to enhance the global and local contrast of digital images using the proposed model. Furthermore, we discuss the relation to existing literature on contrast enhancement. In Section 6, we draw conclusions from our findings and give an outlook on future research.

2 Model

Let us now explore the roots of our model and derive – in a second step – the particle evolution which forms the heart of our method and which is given by the gradient descent of a convex energy.

2.1 Motivation from Swarm Dynamics

The idea behind our model goes back to the scenario of describing a one-dimensional evolution of particles within a closed system. Recent literature on mathematical swarm models employs a pairwise potential $U : \mathbb{R}^n \to \mathbb{R}$ to characterise the behaviour of individual particles (see e.g. [12,11,17,15,10] and the references therein). The potential function allows to steer attractive and repulsive forces among swarm mates. Physically simplified models like [16] neglect inertia and describe the individual particle velocity $\partial_t v_i$ within a swarm of size $N$ directly as

$$\partial_t v_i = - \sum_{j=1}^{N} \nabla U(|v_i - v_j|), \quad i = 1, \ldots, N,$$

where $v_i$ and $v_j$ denote particle positions in $\mathbb{R}^n$. These models are also referred to as first order models. Often they are inspired by biology and describe long-range attractive and short-range repulsive behaviour between swarm members. The interplay of attractive and repulsive forces leads to flocking and allows to gain stability for the whole swarm. Inverting this behaviour – resulting in short-range attractive and long-range repulsive forces – leads to a highly unstable scenario in which the swarm splits up into small separating groups which might never reach a point where
they stand still. One would expect that a restriction to repulsive forces only will increase this instability even further. However, we will present a model which copes well with exactly this situation. In our set-up every particle moves within the open interval $(0, 1)$ and has an interaction radius of size 1. Keeping this in mind, let us briefly examine the two main assumptions of the evolution. First, there exist reflections for all particles at the left and right domain boundary. Secondly, the particles repel each other and – furthermore – get repelled by the reflections. However, due to the limited viewing range, only one of the two reflections of a certain particle is considered at any given time, namely the one which is closer to the reference particle (see Figure 1). A special case occurs if the reference particle is located at position 0.5: the repulsive forces of both of its own reflections equal out.

2.2 Discrete Variational Model

We propose a dynamical system which has its roots in a spatial discretisation of the energy functional (2). Furthermore, we make use of a decreasing energy function $\Psi : \mathbb{R}_0^+ \rightarrow \mathbb{R}$ and a global range constraint on $u$. The corresponding flux function $\Phi$ is defined as $\Phi(s) := \Psi'(s^2)s$.

Our goal is to describe the evolution of one-dimensional – not necessarily distinct – particle positions $v_i \in (0, 1)$, where $i = 1, \ldots, N$. Therefore, we extend the position vector $\mathbf{v} = (v_1, \ldots, v_N)^T$ with the additional coordinates $v_{N+1}, \ldots, v_{2N}$ defined as $v_{2N+1-i} := 2 - v_i \in (1, 2)$. This extended position vector $\mathbf{v} \in (0, 2)^{2N}$ allows to evaluate the energy function

$$E(\mathbf{v}, \mathbf{W}) = \frac{1}{4} \cdot \sum_{i=1}^{2N} \sum_{j=1}^{2N} w_{i,j} \cdot \Psi((v_j - v_i)^2),$$

which models the repulsion potential between all positions $v_i$ and $v_j$. The coefficient $w_{i,j}$ denotes entry $j$ in row $i$ of a constant non-negative weight matrix $\mathbf{W} = (w_{i,j}) \in (\mathbb{R}_0^+)^{2N \times 2N}$. It models the importance of the interaction between position $v_i$ and $v_j$. All diagonal elements of the weight matrix are positive, i.e. $w_{i,i} > 0, \forall i \in \{1, 2, \ldots, 2N\}$. In addition, we assume that the weights for all extended positions are the same as those for the original ones. Namely, we have

$$w_{i,j} = w_{i,2N+1-j} = w_{2N+1-i,j} = w_{2N+1-i,2N+1-j}$$

![Fig. 1. Four particles with positions in $(0, 1)$ and their reflections at the left and right domain boundary (labelled $l$ and $r$ accordingly). Particle 2, for example, gets repelled by the particles $1, 3, 4, 1; 2, 3, 4$.](image-url)
| $\Psi_{a,n}(s^2)$ | $\Psi'_{a,n}(s^2)$ | $\Psi_{a,n}(s)$ |
|-----------------|-----------------|-----------------|
| $\alpha \cdot (s-1)^{2n-1}$ | $\frac{\alpha}{n} \cdot (s-1)^{2n-2}$ | $\alpha \cdot (s-1)^{2n-1}$ |

Table 1. One exemplary class of penaliser functions $\Psi(s^2)$ for $s \in [0,1]$ with $n \in \mathbb{N}$, $\alpha > 0$ and corresponding diffusivity $\Psi'(s^2)$ and flux $\Phi(s)$ functions.

for $1 \leq i, j \leq N$.

For the penaliser function $\Psi$ we impose several restrictions which we discuss subsequently. Table 1 shows one reasonable class of functions $\Psi_{a,n}$ as well as the corresponding diffusivities $\Psi'_{a,n}$ and flux functions $\Phi_{a,n}$. In Figure 2 we provide an illustration of three functions using $a = 1$ and $n = 1, 2, 3$. The penaliser is constructed following a three step procedure. First, the function $\Psi(s^2)$ is defined as a continuously differentiable, decreasing, and strictly convex function for $s \in [0,1]$ with $\Psi(0) = 0$ and $\Phi_-(1) = 0$ (left-sided derivative). Next, $\Psi$ is extended to $[-1,1]$ by symmetry and to $\mathbb{R}$ by periodicity $\Psi((2+s)^2) = \Psi(s^2)$. This results in a penaliser $\Psi(s^2)$ which is continuously differentiable everywhere except at even integers, where it is still continuous. Note that $\Psi(s^2)$ is increasing on $[−1,0]$ and $[1,2]$. The flux $\Phi$ is continuous and increasing in $(0,2)$ with jump discontinuities at 0 and 2 (see Figure 2). Furthermore, we have that $\Phi(s) = -\Phi(−s)$ and $\Phi(2+s) = \Phi(s)$. Exploiting the properties of $\Psi$ allows us to rewrite (4) without the redundant entries $v_{N+1}, \ldots, v_{2N}$ as

$$ E(v, W) = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} w_{i,j} \cdot \left( \Psi((v_j - v_i)^2) + \Psi((v_j + v_i)^2) \right). \quad (6) $$

A gradient descent for (4) is given by

$$ \partial_t v_i = -\partial_{v_i} E(v, W) = \sum_{j \in J_i^1} w_{i,j} \cdot \Phi(v_j - v_i), \quad i = 1, \ldots, 2N, \quad (7) $$

where $v_j$ now are functions of the time $t$ and $J_i^1 := \{ j \in \{1,\ldots,2N\} | v_j \neq v_i \}$. Note that for $1 \leq i, j \leq N$, thus $|v_j - v_i| < 1$, the flux $\Phi(v_j - v_i)$ is negative for $v_j > v_i$ and positive otherwise, thus driving $v_i$ always away from $v_j$. This implies that we have negative diffusivities $\Psi'$ for all $|v_j - v_i| < 1$. Due to the convexity of $\Psi(s^2)$, the absolute values of the repulsive forces $\Phi$ are decreasing with the distance between $v_i$ and $v_j$. We remark that the jumps of $\Phi$ at 0 and 2 are not problematic here, as all positions $v_i$ and $v_j$ in the argument of $\Phi$ are distinct by the definition of $J_i^1$.

Let us discuss shortly how the interval constraint for the $v_i$, $i = 1, \ldots, N$, is enforced in (4) and (7). First, notice that $v_{2N+1-i}$ for $i = 1, \ldots, N$ is the reflection of $v_i$ on the right interval boundary 1. For $v_i$ and $v_{2N+1-j}$ with $1 \leq i, j \leq N$ and $v_{2N+1-j} - v_i < 1$ there is a repulsive force due to $\Phi(v_{2N+1-j} - v_i) < 0$ that drives $v_i$ and $v_{2N+1-j}$ away from the right interval boundary. The closer $v_i$ and $v_{2N+1-j}$ come to this boundary, the stronger is the repulsion. For $v_{2N+1-j} - v_i > 1$, we have $\Phi(v_{2N+1-j} - v_i) > 0$. By $\Phi(v_{2N+1-j} - v_i) = \Phi((2-v_j) - v_i) = \Phi((-v_j) - v_i)$, this can equally be interpreted as a repulsion between $v_i$ and $-v_j$ where $-v_j$ is the reflection of $v_j$ at the left interval boundary 0. In this case the interaction between $v_i$ and $v_{2N+1-j}$ drives $v_i$ and $-v_j$ away from the left interval boundary. Recapitulating both
Fig. 2. **Top:** Exemplary penaliser functions $\tilde{\Psi}_{1,1}$, $\tilde{\Psi}_{1,2}$, and $\tilde{\Psi}_{1,3}$ extended to the interval $[-1, 3]$ by imposing symmetry and periodicity with $\tilde{\Psi}_{a,n}(s) := \Psi_{a,n}(s^2)$. **Middle:** Corresponding diffusivities $\tilde{\Psi}'_{1,1}$, $\tilde{\Psi}'_{1,2}$, and $\tilde{\Psi}'_{1,3}$ with $\tilde{\Psi}'_{a,n}(s) := \Psi'_{a,n}(s^2)$. **Bottom:** Corresponding flux functions $\Phi_{1,1}$, $\Phi_{1,2}$, and $\Phi_{1,3}$ with $\Phi_{a,n}(s) = \Psi_{a,n}(s^2)s$. 
possible cases, it becomes clear that every \( v_i \) is either repelled from the reflection of \( v_j \) at the left or at the right interval boundary, but never from both at the same time. As \( \partial_t v_{2N+1-i} = -\partial_t v_i \) holds in (7), the symmetry of \( v \) is preserved. Dropping the redundant entries \( v_{N+1}, \ldots, v_{2N} \), Equation (7) can be rewritten as
\[
\partial_t v_i = \sum_{j \in J^i_2} w_{i,j} \cdot \Phi(v_j - v_i) - \sum_{j=1}^{N} w_{i,j} \cdot \Phi(v_j + v_i),
\] (8)
for \( i = 1, \ldots, N \), where the second sum emphasises the repulsions between original and reflected coordinates in a symmetric way. The set \( J^i_2 \) is defined as
\[
J^i_2 := \{ j \in \{1, 2, \ldots, N\} | v_j \neq v_i \}.
\]
Equation (8) denotes a formulation for pure repulsion amongst \( N \) different positions \( v_i \) with stabilisation being achieved through the consideration of their reflections at the domain boundary. It is worth mentioning that within (6) and (8) we only make use of the first \( N \times N \) entries of \( \tilde{W} \). In the following, we denote this submatrix by \( \tilde{W} \) and refer to its elements as \( \tilde{w}_{i,j} \).

Given an initial vector \( f \in (0,1)^N \) and initialising \( v_{i}(0) = f_i, v_{2N+1-i}(0) = 2 - f_i \) for \( i = 1, \ldots, N \), the gradient descent (7) and (8) evolves \( v \) towards a minimiser of \( E \).

3 Theory

Below we provide a detailed analysis of the evolution and discuss its main properties. For this purpose we consider the Hessian of (6) whose entries for \( 1 \leq i \leq N \) read
\[
\partial_{v_i v_j} E(v, \tilde{W}) = \sum_{j \in J^i_2} \tilde{w}_{i,j} \cdot \Phi'(v_j - v_i) + \sum_{j=1}^{N} \tilde{w}_{i,j} \cdot \Phi'(v_j + v_i),
\] (9)
\[
\partial_{v_i v_j} E(v, \tilde{W}) = \begin{cases} 
\tilde{w}_{i,j} \cdot (\Phi'(v_j + v_i) - \Phi'(v_j - v_i)), & \forall j \in J^i_2, \\
\tilde{w}_{i,j} \cdot \Phi'(v_j + v_i), & \forall j \in J^i_3,
\end{cases}
\] (10)
where \( J^i_3 := \{ j \in \{1, 2, \ldots, N\} | v_i = v_j \} \).

3.1 General Results

In a first step, let us investigate the well-posedness of the underlying initial value problem in the sense of Hadamard [22].

**Theorem 1 (Well-Posedness).** Let \( \Psi = \Psi_{a,n} \) as defined in Table 1. Then the initial value problem (8) is well-posed since

(a) it has a solution,
(b) the solution is unique, and
(c) it depends continuously on the initial conditions.

**Proof.** The initial value problem (8) can be written as
\[
\dot{v}(t) = f(v(t)) := -\nabla_v E(v(t), W) \quad (11)
\]
\[
v(0) = v_0 \quad (12)
\]
with \( \mathbf{v}(t) \in \mathbb{R}^{2N} \) and \( t \in \mathbb{R}_0^+ \) where we make use of the fact that \( \mathbf{W} \) is a constant weight matrix.

In case \( \mathbf{f}(\mathbf{v}(t)) \) is continuously differentiable and Lipschitz continuous all three conditions (a)–(c) hold. Existence and uniqueness directly follow from [39], chapter 3.1, Theorem 3. Continuous dependence on the initial conditions is guaranteed due to [39], chapter 2.3, Theorem 1 which is based on Gronwall’s Lemma [21]. Thus, let us now prove differentiability and Lipschitz continuity of \( \mathbf{f}(\mathbf{v}(t)) \).

**Differentiability:** Differentiability follows from the fact that all functions \( \Phi_{a,n} \) are continuously differentiable. As a consequence the partial derivatives of (8) w.r.t. \( v_i \) exist for \( i = 1, \ldots, 2N \).

**Lipschitz Continuity:** The Gershgorin circle theorem [18] allows to estimate a valid Lipschitz constant \( L \) as an upper bound of the spectral radius of the Jacobian of (8). For \( 1 \leq i \leq N \) the entries read

\[
\partial_{v_i}(\partial_t v_i) = -\sum_{j \in J_2} \tilde{w}_{i,j} \cdot (\Phi'(v_j - v_i) + \Phi'(v_j + v_i)) - 2 \cdot \sum_{j \in J_3} \tilde{w}_{i,j} \cdot \Phi'(v_j + v_i) \tag{13}
\]

\[
\partial_{v_j}(\partial_t v_i) = \begin{cases} 
\tilde{w}_{i,j} \cdot (\Phi'(v_j + v_i) - \Phi'(v_j - v_i)), & \forall j \in J_2, \\
-\tilde{w}_{i,j} \cdot \Phi'(v_j - v_i), & \forall j \in J_3.
\end{cases} \tag{14}
\]

The radii of the Gershgorin discs fulfil

\[
\begin{align*}
\lambda_i &= \sum_{j=1}^{N} |\partial_{v_j}(\partial_t v_i)| \\
&= \sum_{j \in J_2} \tilde{w}_{i,j} \cdot |\Phi'(v_j - v_i) - \Phi'(v_j + v_i)| + \sum_{j \in J_3} \tilde{w}_{i,j} \cdot |\Phi'(v_j + v_i)| \tag{15} \\
&< \sum_{j \in J_2} \tilde{w}_{i,j} \cdot |\Phi'(v_j - v_i) - \Phi'(v_j + v_i)| + \sum_{j \in J_3} \tilde{w}_{i,j} \cdot |\Phi'(v_j + v_i)| =: \tilde{r}_i
\end{align*}
\]

for \( i = 1, \ldots, N \). Then we have \( |\lambda_i - \partial_{v_i}(\partial_t v_i)| < \tilde{r}_i \) for \( 1 \leq i \leq N \) where \( \lambda_i \) denotes the \( i \)-th eigenvalue of the Jacobian of (8). This leads to the bounds

\[
\lambda_i < \sum_{j \in J_2} \tilde{w}_{i,j} \cdot (|\Phi'(v_j - v_i)| - (\Phi'(v_j - v_i) + \Phi'(v_j + v_i))) \\
+ \sum_{j \in J_3} \tilde{w}_{i,j} \cdot (|\Phi'(v_j + v_i)| - 2 \cdot \Phi'(v_j + v_i)) \\
\leq \sum_{j \in J_2} \tilde{w}_{i,j} \cdot (|\Phi'(v_j - v_i)| + |\Phi'(v_j + v_i)| + |\Phi'(v_j - v_i)| + |\Phi'(v_j + v_i)|) \\
+ \sum_{j \in J_3} \tilde{w}_{i,j} \cdot (|\Phi'(v_j + v_i)| + 2 \cdot |\Phi'(v_j + v_i)|)
\]
\[
\leq 4 \cdot L_\Phi \cdot \sum_{j \in J_2} \hat{w}_{i,j} + 3 \cdot L_\Phi \cdot \sum_{j \in J_3} \hat{w}_{i,j}
< 4 \cdot L_\Phi \cdot \sum_{j = 1}^N \hat{w}_{i,j}, \quad i = 1, \ldots, N,
\]

where \( L_\Phi \) represents the Lipschitz constant of the flux function \( \Phi \). Using the same reasoning one can show that

\[
\lambda_i > -4 \cdot L_\Phi \cdot \sum_{j = 1}^N \hat{w}_{i,j}, \quad i = 1, \ldots, N.
\]

Consequently, an upper bound for the spectral radius – and thus for the Lipschitz constant \( L \) of the gradient of (8) – reads

\[
L \leq \max_{1 \leq i \leq N} |\lambda_i| < 4 \cdot L_\Phi \cdot \max_{1 \leq i \leq N} \sum_{j = 1}^N \hat{w}_{i,j} =: L_{\text{max}},
\]

For our specific class of flux functions \( \Phi_{a,n} \) a valid Lipschitz constant \( L_\Phi \) is given by

\[
L_\Phi = a \cdot n \cdot (2n - 1) \cdot 2^{2n-2}
\]

such that we have

\[
L < 4 \cdot a \cdot n \cdot (2n - 1) \cdot 2^{2n-2} \cdot \max_{1 \leq i \leq N} \sum_{j = 1}^N \hat{w}_{i,j}.
\]

This concludes the proof. \( \square \)

Next, let us show that no position \( v_i \) can ever reach or cross the interval boundaries 0 and 1.

**Theorem 2 (Avoidance of Range Interval Boundaries).** For any weighting matrix \( \hat{W} \in (\mathbb{R}_0^+)^{N \times N} \) all \( N \) positions \( v_i \) which evolve according to (8) and have an arbitrary initial value in \((0, 1)\) do not reach the domain boundaries 0 and 1 for any time \( t \geq 0 \).

**Proof.** Equation (8) can be written as

\[
\partial_t v_i = \sum_{j \in J_2} \hat{w}_{i,j} \cdot (\Phi(v_j - v_i) - \Phi(v_j + v_i)) - \sum_{j \in J_3} \hat{w}_{i,j} \cdot \Phi(2v_i),
\]

where \( 1 \leq i \leq N \). Notice that for \( j \in J_2 \) we have

\[
\lim_{v_i \to 0^+} \Phi(v_j - v_i) - \Phi(v_j + v_i) = 0,
\]

\[
\lim_{v_i \to 1^-} \Phi(v_j - v_i) - \Phi(v_j + v_i) = 0,
\]

where the latter follows from the periodicity of \( \Phi \). Consequently, any position \( v_i \) which gets arbitrarily close to one of the domain boundaries 0 or 1 experiences no
impact by positions $v_j$ with $j \in J^i_2$, and the first sum in (21) gets zero. The definition of $\Psi(s^2)$ implies that
\begin{align}
\Psi'(s^2) < 0, & \quad \forall s \in (0, 1), \quad (24) \\
\Psi'(s^2) > 0, & \quad \forall s \in (1, 2), \quad (25)
\end{align}
from which it follows for $1 \leq i \leq N$ that
\begin{align}
-\Phi(2v_i) > 0, & \quad \forall v_i \in \left(0, \frac{1}{2}\right), \quad (26) \\
-\Phi(2v_i) < 0, & \quad \forall v_i \in \left(\frac{1}{2}, 1\right). \quad (27)
\end{align}
Now remember that $\tilde{W} \in (\mathbb{R}^+)^{N \times N}$ and $\tilde{w}_{i,i} > 0$. In combination with (26) and (27) we get
$$\lim_{v_i \to 0^+} \partial_t v_i > 0 \quad \text{and} \quad \lim_{v_i \to 1^-} \partial_t v_i < 0,$$
which concludes the proof. □

Let us for a moment assume that the penaliser function is given by $\Psi = \Psi_{a,n}$ from Table 1. Below, we prove that this implies convergence to the global minimum of the energy $E(v, \tilde{W})$.

**Theorem 3 (Convergence for $\Psi = \Psi_{a,n}$.)** For $t \to \infty$, given a penaliser $\Psi_{a,1}$ with arbitrary $a > 0$, any initial configuration $v \in (0,1)^N$ converges to a unique steady state $v^*$ which is the global minimiser of the energy given in (6).

**Proof.** As a sum of convex functions, (6) is convex. Therefore, the function $V(v, \tilde{W}) := E(v, \tilde{W}) - E(v^*, \tilde{W})$ (where $v^*$ is the equilibrium point) is a Lyapunov function with $V(v^*, \tilde{W}) = 0$ and $V(v, \tilde{W}) > 0$ for all $v \neq v^*$. Furthermore, we have
$$\partial_t V(v, \tilde{W}) = -\sum_{i=1}^N \left(\partial_{v_i} E(v, \tilde{W})\right)^2 \leq 0.$$  \hfill (29)
According to Gershgorin’s theorem [18], one can show that the Hessian matrix of (6) is positive definite for $\Psi = \Psi_{a,1}$ from which it follows that $E(v, \tilde{W})$ has a strict (global) minimum. This implies that the inequality in (29) becomes strict except in case of $v = v^*$, and guarantees asymptotic Lyapunov stability [30] of $v^*$. Thus, we have convergence to $v^*$ for $t \to \infty$. □

**Remark 1.** Theorem 3 can be extended to the case of $n = 2$ and – in a weaker formulation – to arbitrary $n \in \mathbb{N}$. The proofs for both cases are based on a straightforward application of the Gershgorin circle theorem. For details we refer to the supplementary material.

(a) Given that $\Psi = \Psi_{a,n=2}$, let us assume that one of the following two conditions
- $v_i \neq \frac{1}{2}$, or
- there exists $j \in J_i^2$ for which $v_j \neq 1 - v_i$ and $\tilde{w}_{i,j} > 0$,
is fulfilled for every \( i \in [1, N] \) and \( t \geq 0 \). Then the Hessian matrix of (6) is positive definite and convergence to the strict global minimum of \( E(\mathbf{v}, \mathbf{W}) \) follows.

(b) For all penaliser functions \( \Psi = \Psi_{a,n} \), one can show that the Hessian matrix of (6) is positive semi-definite. This means that our method converges to a global minimum of \( E(\mathbf{v}, \mathbf{W}) \). However, this minimum does not have to be unique.

In general, the steady-state solution of (8) depends on the definition of the penaliser function \( \Psi \). Based on (21), and assuming that \( \Psi = \Psi_{a,n} \), a minimiser of \( E(\mathbf{v}, \mathbf{W}) \) necessarily fulfils the equation

\[
0 = \sum_{j \in J_i} \tilde{w}_{i,j} \cdot \left( (v_j^* - v_i^* - 1)^{2n-1} - (v_j^* + v_i^* - 1)^{2n-1} \right) - \sum_{j \in J'_i} \tilde{w}_{i,j} \cdot (2v_i^* - 1)^{2n-1},
\]

where \( i = 1, \ldots, N \).

### 3.2 Global Model

If all positions \( v_i \) interact with each other during the evolution, i.e. \( \tilde{w}_{i,j} > 0 \) for \( 1 \leq i, j \leq N \), we speak of our model as acting globally. Below, we prove the existence of weight matrices \( \mathbf{W} \) for which distinct positions \( v_i \) and \( v_j \) (with \( i \neq j \)) can never become equal (assuming that the positions \( v_i, i = 1, \ldots, N \), are distinct for \( t = 0 \)). This implies that the initial rank-order of \( v_i \) is preserved throughout the evolution.

**Theorem 4 (Distinctness of \( v_i \) and \( v_j \)).** Among \( N \) initially distinct positions \( v_i \in (0, 1) \) evolving according to (8), no two ever become equal if \( \tilde{w}_{j,k} = \tilde{w}_{i,k} > 0 \) for \( 1 \leq i, j, k \leq N, i \neq j \).

**Proof.** Given \( N \) distinct positions \( v_i \in (0, 1) \), equation (8) can be written as

\[
\partial_t v_i = \sum_{k=1}^{N} \tilde{w}_{i,k} \cdot \Phi(v_k - v_i) - \sum_{k=1}^{N} \tilde{w}_{i,k} \cdot \Phi(v_k + v_i),
\]

for \( i = 1, \ldots, N \). We use this equation to derive the difference

\[
\partial_t (v_j - v_i) = (\tilde{w}_{j,i} + \tilde{w}_{i,j}) \cdot \Phi(v_i - v_j)
\]

\[
+ \sum_{k=1}^{N} \left( \tilde{w}_{j,k} \cdot \Phi(v_k - v_j) - \tilde{w}_{i,k} \cdot \Phi(v_k - v_i) \right)
\]

\[
- \sum_{k=1}^{N} \left( \tilde{w}_{j,k} \cdot \Phi(v_k + v_j) - \tilde{w}_{i,k} \cdot \Phi(v_k + v_i) \right),
\]

where \( 1 \leq i, j \leq N \). Assume w.l.o.g. that \( v_j > v_i \) and consider (32) in the limit \( v_j - v_i \to 0 \). Then we have

\[
\lim_{v_j-v_i\to0} (\tilde{w}_{j,i} + \tilde{w}_{i,j}) \cdot \Phi(v_i - v_j) > 0,
\]

(33)
we have

\[
\lim_{v_j - v_i \to 0} \sum_{k=1 \atop k \neq i,j}^N \left( \tilde{w}_{j,k} \cdot \Phi(v_k - v_j) - \tilde{w}_{i,k} \cdot \Phi(v_k - v_i) \right) - \sum_{k=1}^N \left( \tilde{w}_{j,k} \cdot \Phi(v_k + v_j) - \tilde{w}_{i,k} \cdot \Phi(v_k + v_i) \right)
\]

\[
= \sum_{k=1 \atop k \neq i,j}^N (\tilde{w}_{j,k} - \tilde{w}_{i,k}) \cdot \Phi(v_k - v_i) - \sum_{k=1}^N (\tilde{w}_{j,k} - \tilde{w}_{i,k}) \cdot \Phi(v_k + v_i)
\]

\[
= 0 \quad \text{if } \tilde{w}_{j,k} = \tilde{w}_{i,k} \quad \text{for } 1 \leq k \leq N.
\]

(34)

In conclusion, we can guarantee for global models with distinct particle positions that

\[
\lim_{v_j - v_i \to 0} \partial_t (v_j - v_i) > 0,
\]

if \( \tilde{w}_{j,k} = \tilde{w}_{i,k} \) where \( 1 \leq i, j, k \leq N \) and \( i \neq j \). According to (35), \( v_j \) will always start moving away from \( v_i \) (and vice versa) when the difference between both gets sufficiently small. Since the initial positions are distinct, it follows that \( v_i \neq v_j \) for \( i \neq j \) for all times \( t \). \( \Box \)

A special case occurs if all entries of the weight matrix \( \tilde{W} \) are set to 1 i.e. \( \tilde{W} = 11^T \) with \( 1 := (1, \ldots, 1)^T \in \mathbb{R}^N \). For this scenario, we obtain an analytic steady-state solution which is independent of the penaliser \( \Psi \):

**Theorem 5 (Analytic Steady-State Solution for \( \tilde{W} = 11^T \)).** Under the assumption that \( (v_i) \) is in increasing order, \( \tilde{W} = 11^T \), and that \( \Psi(s^2) \) is twice continuously differentiable in \((0, 2)\) the unique minimiser of (4) is given by \( v^* = (v_1^*, \ldots, v_{2N}^*)^T, v_i^* = (i - 1/2)/N, i = 1, \ldots, 2N \).

**Proof.** With \( \tilde{W} = 11^T \), Equation (4) can be rewritten without the redundant entries of \( v \) as

\[
E(v) = \sum_{i=1}^{N-1} \sum_{j=i+1}^N \Psi((v_j - v_i)^2) + \frac{1}{2} \sum_{i=1}^N \Psi(4v_i^2) + \sum_{i=1}^{N-1} \sum_{j=i+1}^N \Psi((v_j + v_i)^2)
\]

(36)

from which one can verify by straightforward, albeit lengthy calculations that \( \nabla E(v^*) = 0 \), and the Hessian of \( E \) at \( v^* \) is

\[
D^2E(v^*) = \sum_{k=1}^N A_k \Phi'(\frac{k}{N})
\]

(37)

with sparse symmetric \( N \times N \)-matrices

\[
A_k = 2I - T_k - T_{-k} + H_{k+1} + H_{2N-k+1},
\]

(38)

\[
A_N = I + H_{N+1},
\]

(39)
Fig. 3. Application of the global model to a system of 7 particles with weight matrix $\tilde{W} = 11^T$.

Fig. 4. Application of the global model to a system of 7 particles with $\tilde{w}_{i,k} = 1/k$ for $1 \leq i, k \leq N$.

with $k = 1, \ldots, N - 1$, where the unit matrix $I$, the single-diagonal Toeplitz matrices $T_k$, and the single-antidiagonal Hankel matrices $H_k$ are defined as

\[
I = (\delta_{i,j})_{i,j=1}^{N},
\]
\[
T_k = (\delta_{j-i,k})_{i,j=1}^{N},
\]
\[
H_k = (\delta_{i+j,k})_{i,j=1}^{N}.
\]

Here, $\delta_{i,j}$ denotes the Kronecker symbol, $\delta_{i,j} = 1$ if $i = j$, and $\delta_{i,j} = 0$ otherwise. All $A_k$, $k = 1, \ldots, N$ are weakly diagonally dominant with positive diagonal, thus positive semidefinite by Gershgorin’s Theorem. Moreover, the tridiagonal matrix $A_1$ is of full rank, thus even positive definite. By strict convexity of $\Psi(s^2)$, all $\Phi'(k/N)$ are positive, thus $D^2E(\psi^*)$ is positive definite.

As a consequence, the steady state of the gradient descent for any initial data $f$ (with arbitrary rank-order) can – under the condition that $\tilde{W} = 11^T$ – be computed directly by sorting the $f_i$: Let $\sigma$ be the permutation of $\{1, \ldots, N\}$ for which $(f_{\sigma^{-1}(i)})_{i=1,\ldots,N}$ is increasing (this is what a sorting algorithm computes), the steady state is given by $v_i^* = (\sigma(i) - 1/2)/N$ for $i = 1, \ldots, N$ (cf. Figure 3).

Additionally, we present an analytic expression for the steady state of the global model given a penaliser function $\Psi = \Psi_{a,n}$ (cf. Table 1) with $n = 1$.

**Theorem 6 (Analytic Steady-State Solution for $\Psi = \Psi_{a,n=1}$).** Given $N$ distinct positions $v_i$ in increasing order and a penaliser function $\Psi = \Psi_{a,n=1}$, the unique
minimiser of \( (4) \) is given by
\[
v_i^* = \frac{\sum_{j=1}^i \tilde{w}_{i,j} - \frac{1}{2} \tilde{w}_{i,i}}{\sum_{j=1}^N \tilde{w}_{i,j}}, \quad i = 1, \ldots, N.
\]

Proof. The presented minimiser follows directly from (30). Figure 4 provides an illustration of the steady state. \( \square \)

Finally, and in case all entries of the weight matrix \( \tilde{W} \) are set to 1, we show that the global model converges – independently of \( \Psi \) – to a unique steady state:

**Theorem 7 (Convergence for \( \tilde{W} = 11^T \)).** Given that \( \tilde{W} = 11^T \), any initial configuration \( v \in (0, 1)^N \) with distinct entries converges to a unique steady state \( v^* \) for \( t \to \infty \). This is the global minimiser of the energy given in (6).

Proof. Using the same reasoning as in the proof for Theorem 3 we know that inequality (29) holds. Due to the positive definiteness of (37) it follows that \( E(v, \tilde{W}) \) has a strict (global) minimum which implies that the inequality in (29) becomes strict except in case of \( v = v^* \). This guarantees asymptotic Lyapunov stability of \( v^* \) and thus convergence to \( v^* \) for \( t \to \infty \).

### 3.3 Relation to Variational Signal and Image Filtering

Let us now interpret \( v_1, \ldots, v_N \) as samples of a smooth 1D signal \( u : \Omega \to [0, 1] \) over an interval \( \Omega \) of the real axis, taken at sampling positions \( x_i = x_0 + ih \) with grid mesh size \( h > 0 \). We consider the model \( \tilde{W} \) with \( w_{i,j} := \gamma(|x_j - x_i|) \), where \( \gamma : \mathbb{R}_0^+ \to [0, 1] \) is a non-increasing weighting function with compact support \([0, \varrho])\).

**Theorem 8 (Space-Continuous Energy).** Equation (4) can be considered as a discretisation of
\[
E[u] = \frac{1}{2} \int_{\Omega} (W(u_x^2) + B(u)) \, dx
\]
with penaliser \( W(u_x^2) \approx C \Psi(u_x^2) \) and barrier function \( B(u) \approx D \Psi(4u^2) \), where \( C \) and \( D \) are positive constants.

Remark 2.

(a) The penaliser \( W \) is decreasing and convex in \( u_x \). The barrier function \( B \) is convex and it enforces the interval constraint on \( u \) by favouring values \( u \) away from the interval boundaries. The discrete penaliser \( \Psi \) generates both the penaliser \( W \) for derivatives and the barrier function \( B \).

(b) Note that by construction of \( W \) the diffusivity \( g(u_x^2) := W'(u_x^2) \sim \Psi'(u_x^2) \) has a singularity at 0 with \( -\infty \) as limit.

(c) The cut-off of \( \gamma \) at radius \( \varrho \) implies the locality of the functional (44) that can thereby be linked to a diffusion equation of type (1). Without a cut-off, a nonlocal diffusion equation would arise instead.

Proof (of Theorem 8). We notice first that \( v_j - v_i \) and \( v_i + v_j \) for \( 1 \leq i, j \leq N \) are first-order approximations of \( (j - i) h u_x(x_i) \) and \( 2u(x_i) \), respectively.
Derivation of the Penaliser \( W \). Assume first for simplicity that \( \Psi(s^2) = -\kappa s, \kappa > 0 \) is linear in \( s \) on \([0, 1]\) (thus not strictly convex). Then we have for a part of the inner sums of (4) corresponding to a fixed \( i \):

\[
\frac{1}{2} \left( \sum_{j=1}^{N} \gamma(|x_j - x_i|) \cdot \Psi\left( (v_j - v_i)^2 \right) + \sum_{j=N+1}^{2N} \gamma(|x_j - x_{2N+1-i}|) \cdot \Psi\left( (v_j - v_{2N+1-i})^2 \right) \right) \\
= \sum_{j=1}^{N} \gamma(|x_j - x_i|) \cdot \Psi\left( (v_j - v_i)^2 \right) \\
\approx -\kappa \ h \ u_x(x_i) \sum_{j=1}^{N} \gamma(|j - i| \ h) \cdot |j - i| \\
= \ h \Psi\left( u_x(x_i)^2 \right) \sum_{k=1-i}^{N-i} |k| \ \gamma(|k| \ h) \\
\approx \ h \Psi\left( u_x(x_i)^2 \right) \cdot \frac{2}{h^2} \int_0^\varrho z \gamma(z) dz \\
=: \ hC\Psi\left( u_x(x_i)^2 \right),
\]

where in the last step the sum over \( k = 1-i, \ldots, N-i \) has been replaced with a sum \( k = -\lfloor \varrho/h \rfloor, \ldots, \lfloor \varrho/h \rfloor \), thus introducing a cutoff error for those locations \( x_i \) that are within the distance \( \varrho \) from the interval ends. Summation over \( i = 1, \ldots, N \) approximates

\[
\int_\Omega C\Psi(u_x^2) dx
\]

from which we can read off \( W(u_x^2) \approx C\Psi(u_x^2) \).

For \( \Psi(s^2) \) that are nonlinear in \( s \), \( \Psi(u_x(x_i)^2) \) in (45) is changed into a weighted sum of \( \Psi((ku_x(x_i))^2) \) for \( k = 1, \ldots, N-1 \), which still amounts to a decreasing function \( W(u_x^2) \) that is convex in \( u_x \). Qualitatively, \( W' \) then behaves the same way as before.

Derivation of the Barrier Function \( B \). Collecting the summands of (4) that were not used in (45), we have, again for fixed \( i \),

\[
\frac{1}{2} \left( \sum_{j=N+1}^{2N} \gamma(|x_j - x_i|) \cdot \Psi\left( (v_j - v_i)^2 \right) + \sum_{j=1}^{N} \gamma(|x_j - x_{2N+1-i}|) \cdot \Psi\left( (v_j - v_{2N+1-i})^2 \right) \right) \\
= \sum_{j=1}^{N} \gamma(|x_j - x_i|) \cdot \Psi\left( (v_i + v_j)^2 \right) \\
\approx \left( \frac{2}{h} \int_0^\varrho \gamma(z) dz \cdot 1 \right) \cdot \Psi(4u(x_i)^2) \\
=: \ hD \cdot \Psi(4u(x_i)^2),
\]

and thus after summation over \( i \) analogous to the previous step \( \int_\Omega B(u) \, dx \) with \( B(u) \approx D\Psi(4u^2) \).

\[ \square \]
Similar derivations can be made for patches of 2D images. A point worth noticing is that the barrier function $B$ is bounded. This differs from usual continuous models where such barrier functions tend to infinity at the interval boundaries. However, for each given sampling grid and patch size the barrier function is just strong enough to prevent $W$ from pushing the values out of the interval.

4 Explicit Time Discretisation

Up to this point we have established a theory for the time-continuous evolution of particle positions. In order to be able to employ our model in simulations and applications we need to discretise (8) in time. Subsequently, we provide a simple yet powerful discretisation which preserves all important properties of the time-continuous model. An approximation of the time derivative in (8) by forward differences yields the explicit scheme

$$v_{i}^{k+1} = v_i^k + \tau \cdot \sum_{\ell \in J_i} \tilde{w}_{i,\ell} \cdot \Phi(v_{\ell}^k - v_i^k) - \tau \cdot \sum_{\ell = 1}^{N} \tilde{w}_{i,\ell} \cdot \Phi(v_{\ell}^k + v_i^k), \quad (47)$$

for $i = 1, \ldots, N$, where $\tau$ denotes the time step size and an upper index $k$ refers to the time $k \tau$. In the following, we derive necessary conditions for which the explicit scheme preserves the position range $(0, 1)$ and the position ordering. Furthermore, we show convergence of (47) in dependence of $\tau$.

**Theorem 9 (Avoidance of Range Interval Boundaries of the Explicit Scheme).**

Let $L_\Phi$ be the Lipschitz constant of $\Phi$ restricted to the interval $(0, 2)$. Moreover, let $0 < v_i^k < 1$, for every $1 \leq i \leq N$, and assume that the time step size $\tau$ of the explicit scheme (47) satisfies

$$0 < \tau < \frac{1}{2 \cdot L_\Phi \cdot \max_{1 \leq i \leq N} \sum_{\ell = 1}^{N} \tilde{w}_{i,\ell}}. \quad (48)$$

Then it follows that $0 < v_i^{k+1} < 1$ for every $1 \leq i \leq N$.

**Proof.** In accordance with (21) the explicit scheme (47) can be written as

$$v_{i}^{k+1} = v_i^k + \tau \cdot \sum_{\ell \in J_i} \tilde{w}_{i,\ell} \cdot \left( \Phi(v_{\ell}^k - v_i^k) - \Phi(v_{\ell}^k + v_i^k) \right) - \tau \cdot \sum_{\ell \in J_i} \tilde{w}_{i,\ell} \cdot \Phi(2v_i^k), \quad (49)$$

where $i = 1, \ldots, N$. Now assume that $0 < v_i^k, v_j^k < 1$ and let us examine the contribution of the two summation terms in (49). We need to distinguish the following five cases:

1. If $v_i^k = v_j^k \leq \frac{1}{2}$ then $2v_i^k \in (0, 1]$. Thus,

$$0 \leq -\Phi(2v_i^k). \quad (50)$$

2. If $\frac{1}{2} < v_i^k = v_j^k$ then $2v_i^k \in (1, 2)$. Thus, using $\Phi(1) = 0$,

$$|\Phi(2v_i^k)| = |\Phi(2v_i^k) - \Phi(1)| \leq |2v_i^k - 1| \cdot L_\Phi < 2v_i^k \cdot L_\Phi. \quad (51)$$
3. If \( v_i^k < v_j^k \) then \( v_j^k - v_i^k, v_j^k + v_i^k \in (0, 2) \). Thus,

\[
|\Phi(v_j^k + v_i^k) - \Phi(v_j^k - v_i^k)| \leq L_\Phi \cdot 2v_i^k.
\] (52)

4. If \( v_j^k < v_i^k \leq \frac{1}{2} \) then \( v_j^k - v_i^k \in (-1, 0) \) and \( v_j^k + v_i^k \in (0, 1) \). Thus,

\[
0 \leq \Phi(v_j^k - v_i^k) - \Phi(v_j^k + v_i^k),
\] (53)

\[
0 \leq -\Phi(2v_i^k).
\] (54)

5. Finally, if \( v_i^k < v_j^k \) and \( \frac{1}{2} < v_i^k \), using the periodicity of \( \Phi \) we get

\[
|\Phi(v_j^k - v_i^k) - \Phi(v_j^k + v_i^k)| = |\Phi(v_j^k + v_i^k) - \Phi(2 + v_j^k - v_i^k)| \leq 2v_i^k \cdot L_\Phi.
\] (55)

Combining (47) with (48) and (50)- (55) we obtain that

\[
v_i^{k+1} - v_i^k = -\tau \cdot \sum_{\ell \in J_1^k} \hat{w}_{i,\ell} \cdot \left(\Phi(v_{\ell}^k + v_{i}^k) - \Phi(v_{\ell}^k - v_{i}^k)\right) - \tau \cdot \sum_{\ell \in J_2^k} \hat{w}_{i,\ell} \cdot \Phi(2v_{i}^k)
\]

\[
\geq -\tau \cdot L_\Phi \cdot 2v_i^k \sum_{\ell=1}^N \hat{w}_{i,\ell}
\]

\[
> - v_i^k,
\] (56)

from which it directly follows that \( v_i^{k+1} > 0 \), as claimed.

The proof for \( v_i^{k+1} < 1 \) is straightforward. Assume w.l.o.g. that \( \tilde{v}_i^k := 1 - v_i^k \). For the reasons given above, we obtain \( \tilde{v}_i^{k+1} > 0 \). Consequently, \( 1 - v_i^{k+1} > 0 \) and \( v_i^{k+1} < 1 \) follows.

**Theorem 10 (Rank-Order Preservation of the Explicit Scheme).** Let \( L_\Phi \) be the Lipschitz constant of \( \Phi \) restricted to the interval (0, 2). Furthermore, let \( v_i^0 \), for \( i = 1, \ldots, N \), denote the initially distinct positions in (0, 1) and – in accordance with Theorem 4 – let the weight matrix \( \tilde{W} \) have constant columns, i.e. \( \tilde{w}_{i,\ell} = \tilde{w}_{i,\ell} \) for \( 1 \leq i, j, \ell \leq N \). Moreover, let \( 0 < v_i^k < v_j^k < 1 \) and assume that the time step size \( \tau \) used in the explicit scheme (47) satisfies

\[
0 < \tau < \frac{1}{2 \cdot L_\Phi \cdot \max_{1 \leq i \leq N} \sum_{\ell=1}^N \tilde{w}_{i,\ell}}.
\] (57)

Then we have \( v_i^{k+1} < v_j^{k+1} \).

**Proof.** For distinct positions, (47) reads

\[
v_i^{k+1} = v_i^k + \tau \cdot \sum_{\ell=1}^N \tilde{w}_{i,\ell} \cdot \Phi(v_{\ell}^k - v_{i}^k) - \tau \cdot \sum_{\ell \neq i} \tilde{w}_{i,\ell} \cdot \Phi(v_{\ell}^k + v_{i}^k)
\] (58)
for \( i = 1, \ldots, N \). Considering this explicit discretisation for \( \partial_t v_i \) and \( \partial_t v_j \) we obtain for \( i, j \in \{1, 2, \ldots, N\} \):

\[
v_{j}^{k+1} - v_{i}^{k+1} = v_{j}^{k} - v_{i}^{k} + \tau \cdot (\bar{w}_{j,i} + \bar{w}_{i,j}) \cdot \Phi(v_{i}^{k} - v_{j}^{k}) + \tau \cdot \sum_{\ell=1}^{N} \left( \bar{w}_{j,\ell} \cdot \Phi(v_{\ell}^{k} - v_{j}^{k}) - \bar{w}_{i,\ell} \cdot \Phi(v_{\ell}^{k} - v_{i}^{k}) \right)
- \tau \cdot \sum_{\ell=1}^{N} \left( \bar{w}_{j,\ell} \cdot \Phi(v_{\ell}^{k} + v_{j}^{k}) - \bar{w}_{i,\ell} \cdot \Phi(v_{\ell}^{k} + v_{i}^{k}) \right).
\]

(59)

Now remember that \( v_{i}^{k} < v_{j}^{k} \) by assumption and that – as a consequence –

\[
\tau \cdot (\bar{w}_{j,i} + \bar{w}_{i,j}) \cdot \Phi(v_{i}^{k} - v_{j}^{k}) > 0.
\]

(60)

Using the fact that \( \bar{w}_{j,k} = \bar{w}_{i,k} \) for \( 1 \leq i, j, k \leq N \) and that \( \Phi \) is Lipschitz in \((0, 2)\), we also know that

\[
T_1 := \tau \cdot \sum_{\ell=1}^{N} \left| \bar{w}_{j,\ell} \cdot \Phi(v_{\ell}^{k} - v_{j}^{k}) - \bar{w}_{i,\ell} \cdot \Phi(v_{\ell}^{k} - v_{i}^{k}) \right|
= \tau \cdot \sum_{\ell=1}^{N} \bar{w}_{j,\ell} \cdot \left| \Phi(v_{\ell}^{k} - v_{j}^{k}) - \Phi(v_{\ell}^{k} - v_{i}^{k}) \right|
\leq \tau \cdot L_{\Phi} \cdot |v_{i}^{k} - v_{j}^{k}| \cdot \sum_{\ell=1}^{N} \bar{w}_{j,\ell}.
\]

(61)

\[
T_2 := \tau \cdot \sum_{\ell=1}^{N} \left| \bar{w}_{j,\ell} \cdot \Phi(v_{\ell}^{k} + v_{j}^{k}) - \bar{w}_{i,\ell} \cdot \Phi(v_{\ell}^{k} + v_{i}^{k}) \right|
= \tau \cdot \sum_{\ell=1}^{N} \bar{w}_{j,\ell} \cdot \left| \Phi(v_{\ell}^{k} + v_{j}^{k}) - \Phi(v_{\ell}^{k} + v_{i}^{k}) \right|
\leq \tau \cdot L_{\Phi} \cdot |v_{i}^{k} - v_{j}^{k}| \cdot \sum_{\ell=1}^{N} \bar{w}_{j,\ell}.
\]

(62)

Let the time step size \( \tau \) fulfil (57). Then we can write

\[
T_1 + T_2 < 2 \cdot L_{\Phi} \cdot 2 \cdot |v_{j}^{k} - v_{i}^{k}| \cdot \sum_{\ell=1}^{N} \bar{w}_{j,\ell} < v_{j}^{k} - v_{i}^{k}.
\]

(63)

In combination with \( T_1, T_2 \geq 0 \), it follows that

\[
T_2 - T_1 \geq -T_2 - T_1 > -(v_{j}^{k} - v_{i}^{k}),
\]

(64)

and we immediately know that \( v_{j}^{k} - v_{i}^{k} - T_1 + T_2 > 0 \). Together with (59) and (60) we get \( 0 < v_{j}^{k+1} - v_{i}^{k+1} \), as claimed. \( \square \)
Theorem 11 (Convergence of the Explicit Scheme). Let \( E(v, \tilde{W}) \) be a twice continuously differentiable convex function. Then the explicit scheme (47) converges for time step sizes
\[
0 < \tau \leq \frac{1}{2 \cdot L_\Phi \cdot \max_{1 \leq i \leq N} \sum_{j=1}^{N} \tilde{w}_{i,j}} < \frac{2}{L},
\]
where \( L_\Phi \) denotes the Lipschitz constant of \( \Phi \) restricted to the interval \((0, 2)\) and \( L \) refers to the Lipschitz constant of the gradient of \((6)\).

Proof. Convergence of the gradient method to the global minimum of \( E(v, \tilde{W}) \) is well-known for continuously differentiable convex functions with Lipschitz continuous gradient and time step sizes \( 0 < \tau < 2/L \) (see e.g. [33, Theorem 2.1.14]). A valid Lipschitz constant is given by \( L_{\text{max}} \) as defined in (18). Consequently, the time step sizes \( \tau \) need to fulfill (65) in order to ensure convergence of (47). The smaller or equal relation results from the strictness in (18). \( \Box \)

Remark 3 (Optimal Time Step Size). The optimal time step size, i.e. the value of \( \tau \) which leads to most rapid descent, is given by \( \tau = 1/L \) (see e.g. [33, 2.1.5]). Thus, we suggest to use \( \tau = 1/L_{\text{max}} \).

5 Application to Image Enhancement

Now that we have presented a stable and convergent numerical scheme, we apply (47) to enhance the contrast of digital greyscale and colour images. Throughout all experiments we use \( \Psi = \Psi_{1,1} \) (cf. Table 1 and Figure 2).

5.1 Greyscale Images

The application of the proposed model to greyscale images follows the ideas presented in [5]. We define a digital greyscale image as a mapping \( f : \{1, \ldots, n_x\} \times \{1, \ldots, n_y\} \rightarrow [0, 1] \). Note that all grey values are mapped to the interval \((0, 1)\) to ensure the validity of our model before processing. The grid position of the \( i \)-th image pixel is given by the vector \( x_i \) whereas \( v_i \) denotes the corresponding grey value. Subsequently, we will see that a well-considered choice of the weighting matrix \( \tilde{W} \) allows either to enhance the global or the local contrast of a given image.

Global Contrast Enhancement. For global contrast enhancement we make use of the global model as discussed in Section 3.2. Only the \( N \) different occurring grey values \( v_i \) – and not their positions in the image – are considered. We let every entry \( \tilde{w}_{i,j} \) of the weighting matrix denote the frequency of grey value \( v_j \) in the image. Assuming an 8-bit greyscale image this leads to a weighting matrix of size \( 256 \times 256 \) which is independent of the image size. As illustrated in Figure 5, global contrast enhancement can be achieved in two ways: As a first option one can use the explicit scheme (47) to describe the evolution of all grey values \( v_i \) up to some time \( t \) (see column two of Figure 5). The amount of contrast enhancement grows with increasing values of \( t \). In our experiments an image size of \( 481 \times 321 \) pixels and the application
of the flux function $\Phi_{1,1}$ with $L_{\Phi} = 1$ imply an upper bound of $1/(2 \cdot 481 \cdot 321)$ for $\tau$. Thus, we can achieve the time $t = 2 \cdot 10^{-6}$ in Figure 5 in a single iteration. If one is only interested in an enhanced version of the original image with maximum global contrast there is an alternative, namely the derived steady state solution for linear flux functions (43). The results are shown in the last column of Figure 5. This figure also confirms that the solution of the explicit scheme (47) converges to the steady-state solution (43) for $t \to \infty$. From (43) it is clear that this steady state is equivalent to histogram equalisation. In summary, this means that the application of our global model to greyscale images offers an evolution equation histogram equalisation which allows to control the amount of contrast enhancement in an intuitive way through the time parameter $t$.

**Local Contrast Enhancement.** In order to achieve local contrast enhancement we use our model to describe the evolution of grey values $v_i$ at all $n_x \cdot n_y$ image grid positions. The change of every grey value $v_i$ depends on all grey values within a disk-shaped neighbourhood of radius $\varrho$ around its grid position $x_i$. We assume that

$$\tilde{w}_{i,j} := \gamma(|x_j - x_i|), \quad \forall i, j \in \{1, 2, \ldots, N\},$$

(66)
where we weight the spatial distance $|x_j - x_i|$ by a function $\gamma : \mathbb{R}_0^+ \rightarrow [0, 1]$ with compact support $[0, \varrho)$ which fulfils

$$\gamma(x) \in (0, 1], \quad \text{if } x < \varrho,$$

$$\gamma(x) = 0, \quad \text{if } x \geq \varrho.$$

The choice of $\gamma$ is application dependent. However, it usually makes sense to define $\gamma(x)$ as a non-increasing function in $x$. Possible choices are e.g.

$$\gamma_1(x) = \begin{cases} 1, & \text{if } x < \varrho, \\ 0, & \text{else,} \end{cases} \quad (68)$$

$$\gamma_2(x) = \begin{cases} 1 - 6x^2 + 6x^3, & \text{if } 0 \leq x < \frac{\varrho}{2}, \\ 2 \cdot (1 - x)^3, & \text{if } \frac{\varrho}{2} \leq x < \varrho, \\ 0, & \text{else,} \end{cases} \quad (69)$$

which are both sketched in Figure 6. When applying this local model to images we make use of mirroring boundary conditions in order to avoid artefacts at the image boundaries. Figure 7 provides an example for local contrast enhancement of digital greyscale images. Again, we describe the grey value evolution with the explicit scheme (47). Furthermore, we use $\gamma_1$ to model the influence of neighbouring grey values. As is evident from Figure 7, increasing the values for $t$ goes along with enhanced local contrast.

5.2 Colour Images

Based on the assumption that our input data is given in sRGB colour space [48] (in the following denoted by RGB) we represent a digital colour image by the mapping $f : \{1, \ldots, n_x\} \times \{1, \ldots, n_y\} \rightarrow [0, 1]^3$. Subsequently, our aim is the contrast enhancement of digital colour images without distorting the colour information. This means that we only want to adapt the luminance but not the chromaticity of a given image. For this purpose, we convert the given image data to YCbCr colour space [44, Section 3.5] since this representation provides a separate luminance channel. Next, we perform contrast enhancement on the luminance channel only. Just as for greyscale images we map all Y-values to the interval $(0, 1)$ to fulfil our model.
requirements. After enhancing the contrast, we transform the colour information of the image back to RGB colour space.

At this point it is important to mention that the colour gamut of the RGB colour space is a subset of the YCbCr colour gamut and during the conversion process of colour coordinates from YCbCr to RGB colour space the so-called colour gamut problem may occur: Colours from the YCbCr colour gamut may lie outside the RGB colour gamut and thus cannot be represented in RGB colour coordinates. Naik and Murthy [32] state that a simple clipping of the values to the bounds creates undesired shift of hue and may lead to colour artefacts. In order to avoid the colour gamut problem we adapt the ideas presented by Nikolova and Steidl [34] which are based on the intensity representation of the HSI colour space [20, Section 6.2.3]. Using the original and enhanced intensities, they define an affine colour mapping and transform the original RGB values. This preserves the hue and results in an enhanced RGB image. It is straightforward to show that their algorithms are valid for any intensity $\hat{f}$ of type

$$\hat{f} = c_r \cdot r + c_g \cdot g + c_b \cdot b,$$

with $c_r + c_g + c_b = 1$ and $c_r, c_g, c_b \in [0, 1]$, where $r$, $g$, and $b$ denote RGB colour coordinates. Thus, they are applicable to the luminance representation of the YCbCr colour space, too, i.e. $c_r = 0.299$, $c_g = 0.587$, $c_b = 0.114$. Tian and Cohen make use of the same idea in [52]. As in [33], our result image is a convex combination of the

---

**Fig. 7.** Local contrast enhancement using $\Phi = \Phi_{1,1}$, $\gamma = \gamma_1$, $\varrho = 60$, and greyscale versions of images from the BSDS500 [2].
outcomes of a multiplicative and an additive algorithm (see [34, Algorithm 4 and 5]) with coefficients $\lambda$ and $1 - \lambda$ for $\lambda \in [0, 1]$. During our experiments we use a fixed value of $\lambda = 0.5$ (for details on how to choose $\lambda$ we refer to [34]). An overview of our strategy for contrast enhancement of digital colour value images is given in Figure 8.

![Diagram](image)

**Fig. 8.** Procedure of contrast enhancement for digital colour images following [34].

**Global Contrast Enhancement.** Again, we apply the global model from Section 3.2 in order to achieve global contrast enhancement. As mentioned before, we consider the $N$ different occurring Y-values of the YCbCr representation of the input image and denote them by $v_i$ (similar to Section 5.1 we neglect their positions in the image). Every entry of the weighting matrix $\hat{w}_{i,j}$ contains the number of occurrences of the value $v_j$ in the Y-channel of the image. It becomes clear that the application of our model – in this setting – basically comes down to histogram equalisation of the Y-channel. Figure 9 shows the resulting RGB images after global contrast enhancement. Similar to the greyscale scenario, we can either apply the explicit scheme (47) or – for $\Phi = \Phi_{a,1}$ – estimate the steady state solution following (43). For the first case the amount of contrast enhancement grows with the positive time parameter $t$. The second column of Figure 9 shows the results for $\Phi = \Phi_{1,1}$ given time $t$. The corresponding steady state solutions are illustrated in the last column of Figure 9.

**Local Contrast Enhancement.** In a similar manner – and adapting the ideas from Subsection 5.1 – we achieve local contrast enhancement in colour images. For this purpose we describe the evolution of Y-values $v_i$ at all $n_x \cdot n_y$ image grid positions using a disk-shaped neighbourhood of radius $\rho$ around the corresponding grid positions $x_i$. The entries of the weighting matrix $\hat{W}$ follow (66). In combination with mirrored boundary conditions the explicit scheme (47) allows to increase the
Fig. 9. Global contrast enhancement using $\Phi = \Phi_{1,1}$, $\lambda = 0.5$, and images from [1].

Fig. 10. Local contrast enhancement using $\Phi = \Phi_{1,1}$, $\gamma = \gamma_1$, $\varrho = 60$, $\lambda = 0.5$, and images from [1].
local contrast of an image with growing $t$. Figure 10 shows exemplary results for $\Phi = \Phi_{1,1}$ and $\gamma = \gamma_1$ (cf. (68)). Note, how well – in comparison to the global set-up in Figure 9 – the structure of the door gets enhanced while the details of the door knob are preserved. The differences are even larger in the second image: For both the couple in the foreground and the background scenery, contrast increases which implies visibility also for larger times $t$.

5.3 Parameters

In total, our model has up to six parameters: $\Phi$, $\tau$, $t$, $\lambda$, $\varrho$, and $\gamma$. During our experiments we have fixed $\Phi(s)$ to the linear flux function $\Phi_{1,1}(s)$ and $\lambda$ to 0.5. Valid bounds for the time step size $\tau$ are given in Theorems 3–11. From the theory in Section 3 and the subsequent experiments on greyscale and colour images it becomes clear that the amount of contrast enhancement grows with the diffusion time. Thus, it remains to discuss the influence of the parameters $\varrho$ and $\gamma$. We found out that the neighbourhood radius $\varrho$ affects the diffusion time and controls the amount of perceived local contrast enhancement, i.e. it steers the localisation of the contrast enhancement process. Whereas small radii lead to high contrast in already small image areas, the size of image sections with high contrast increases with $\varrho$. For sufficiently large values of $\varrho$ global histogram equalisation is approximated. Another interesting point is the choice of the weighting function $\gamma$. Overall, choosing $\gamma = \gamma_1$ leads to more homogeneous contrast enhancement resulting in a smoother perception. For $\gamma = \gamma_2$ the focus always lies on the neighbourhood centre which implies even more enhancement of local structures than in the preceding case. In summary, $\gamma_2$ leads to more enhancement which, however, also creates undesired effects in smooth or noisy regions. Thus, we prefer $\gamma_1$ over $\gamma_2$. Further experiments which visualise the effect of the parameters can be found in the supplementary material.

5.4 Related Work from an Application Perspective

Now that we have demonstrated the applicability of our model to digital images we want to discuss briefly its relation to other existing theories in the context of image processing.

As mentioned in Section 5.1 applying the global model – with $\tilde{W} = 11^T$ – is identical to histogram equalisation (a common formulation can e.g. be found in [20]). Furthermore, there exist other closely related histogram specification techniques – such as [45, 35, 36] – which can have the same steady state. If we compare our evolution with the histogram modification flow introduced by Sapiro and Caselles [45], we see that their flow can also be translated into a combination of repulsion among grey-values and a barrier function. However, in [45] the repulsive force is constant, and the barrier function quadratic. Thus, they cannot be derived from the same kind of interaction between the $v_i$ and their reflected counterparts as in our paper.

Referring to Section 5.1 there also exist well-known approaches which aim to enhance the local image contrast such as adaptive histogram equalisation – see [42] and the references therein – or contrast limited adaptive histogram equalisation [55]. The
latter technique tries to overcome the over-amplification of noise in mostly homogeneous image regions when using adaptive histogram equalisation. Both approaches share the basic idea with our approach in Section 5.1 and perform histogram equalisation for each pixel, i.e. the mapping function for every pixel is determined using a neighbourhood of predefined size and its corresponding histogram.

Another related research topic is the rich field of colour image enhancement which we broach in Section 5.2. A short review of existing methods – as well as two new ideas – is presented in [3]. Therein, Bassiou and Kotropoulus also mention the colour gamut problem for methods which perform contrast enhancement in a different colour space and transform colour coordinates to RGB afterwards. Of particular interest are the publications by Naik and Murthy [32] and Nikolova and Steidl [34] whose ideas are used in Section 5.2. Both of them suggest – based on an affine colour transform – strategies to overcome the colour gamut problem while avoiding colour artefacts in the resulting image. A recent approach which also makes use of these ideas is presented by Tian and Cohen [52]. Ojo et al. [37] make use of the HSV colour space to avoid the colour gamut problem when enhancing the contrast of colour images. A variational approach for contrast enhancement which tries to approximate the hue of the input image was recently published by Pierre et al. [41].

6 Conclusions and Outlook

In our paper we have presented a mathematical model which describes pure backward diffusion as gradient descent of strictly convex energies. The underlying evolution makes use of ideas from the area of collective behaviour and – in terms of the latter – our model can be understood as a fully repulsive discrete first order swarm model. Not only it is surprising that our model allows backward diffusion to be formulated as a convex optimisation problem but also that it is sufficient to impose reflecting boundary conditions in the diffusion co-domain in order to guarantee stability. This strategy is contrary to existing approaches which either assume forward or zero diffusion at extrema or add classical fidelity terms to avoid instabilities. Furthermore, discretisation of our model does not require sophisticated numerics. We have proven that a straightforward explicit scheme is sufficient to preserve the stability of the time-continuous evolution. In our experiments, we show that our model can directly be applied to contrast enhancement of digital greyscale and colour images.

We see our contribution mainly as an example of stable modelling of backward parabolic evolutions that create neither theoretical nor numerical problems. We are convinced that this concept has far more widespread applications in inverse problems, image processing, and computer vision. Exploring them will be part of our future research.

Acknowledgements

Our research on intrinsic stability properties of specific backward diffusion processes goes back to 2005, when Joachim Weickert was visiting Mila Nikolova in Paris. Unfortunately the diffusion equation considered at that time did not have the expected properties, such that it took one more decade to come up with more appropriate
models. Leif Bergerhoff wishes to thank Antoine Gautier and Peter Ochs for fruitful discussions around the topic of the paper. This project has received funding from the DFG Cluster of Excellence *Multimodal Computing and Interaction* and from the European Union’s Horizon 2020 research and innovation programme (grant agreement No 741215, ERC Advanced Grant *INCOVID*). This is gratefully acknowledged.

References

1. Kodak Lossless True Color Image Suite. [http://www.r0k.us/graphics/kodak/](http://www.r0k.us/graphics/kodak/), last visited August 31, 2018
2. Arbelaez, P., Maire, M., Fowlkes, C., Malik, J.: Contour detection and hierarchical image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence 33(5), 898–916 (2011)
3. Bassiou, N., Kotropoulos, C.: Color image histogram equalization by absolute discounting back-off. Computer Vision and Image Understanding 107(1), 108–122 (Jul 2007)
4. Bergerhoff, L., Cárdenas, M., Weickert, J., Welk, M.: Modelling stable backward diffusion and repulsive swarms with convex energies and range constraints. In: Pelillo, M., Hancock, E. (eds.) Energy Minimization Methods in Computer Vision and Pattern Recognition, 11th International Conference, EMMCVPR 2017, Venice, Italy. Lecture Notes in Computer Science, vol. 10746, pp. 409–423. Springer, Cham (Mar 2018)
5. Bergerhoff, L., Weickert, J.: Modelling image processing with discrete first-order swarms. In: Pillay, N., Engelbrecht, P.A., Abraham, A., du Plessis, C.M., Snášel, V., Muda, K.A. (eds.) Advances in Nature and Biologically Inspired Computing, vol. 419, pp. 261–270. Springer, Cham (2016)
6. Carasso, A., Sanderson, J., Hyman, J.: Digital removal of random media image degradations by solving the diffusion equation backwards in time. SIAM Journal on Numerical Analysis 15(2), 344–367 (Apr 1978)
7. Carasso, A.S.: Compensating operators and stable backward in time marching in nonlinear parabolic equations. GEM - International Journal on Geomathematics 5(1), 1–16 (Apr 2014)
8. Carasso, A.S.: Stable explicit time marching in well-posed or ill-posed nonlinear parabolic equations. Inverse Problems in Science and Engineering 24(8), 1364–1384 (2016)
9. Carasso, A.S.: Stabilized Richardson leapfrog scheme in explicit stepwise computation of forward or backward nonlinear parabolic equations. Inverse Problems in Science and Engineering 25(12), 1719–1742 (2017)
10. Carrillo, J.A., Fornasier, M., Toscani, G., Vecil, F.: Particle, kinetic, and hydrodynamic models of swarming. In: Naldi, G., Pareschi, L., Toscani, G. (eds.) Mathematical Modeling of Collective Behavior in Socio-Economic and Life Sciences, pp. 297–336. Birkhäuser, Boston (2010)
11. Chuang, Y.L., Huang, Y.R., D’Orsogna, M.R., Bertozzi, A.L.: Multi-vehicle flocking: Scalability of cooperative control algorithms using pairwise potentials. In: 2007 IEEE International Conference on Robotics and Automation, ICRA 2007, 10-14 April 2007, Roma, Italy. pp. 2292–2299 (2007)
12. D’Orsogna, M.R., Chuang, Y.L., Bertozzi, A.L., Chayes, L.S.: Self-propelled particles with soft-core interactions: Patterns, stability, and collapse. Physical Review Letters 96, 104302 (Mar 2006)
13. Fu, C.L., Xiong, X.T., Qian, Z.: Fourier regularization for a backward heat equation. Journal of Mathematical Analysis and Applications 331(1), 472–480 (2007)
14. Gabor, D.: Information theory in electron microscopy. Laboratory Investigation 14, 801–807 (Jun 1965)
15. Gazi, V., Fidan, B.: Coordination and control of multi-agent dynamic systems: Models and approaches. In: Sahin, E., Spears, W.M., Winfield, A.F.T. (eds.) Swarm Robotics, Lecture Notes in Computer Science, vol. 4433, pp. 71–102. Springer, Berlin (2007)
16. Gazi, V., Passino, K.M.: Stability analysis of swarms. IEEE Transactions on Automatic Control 48(4), 692–697 (Apr 2003)
17. Gazi, V., Passino, K.M.: Stability analysis of social foraging swarms. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics 34(1), 539–557 (Feb 2004)
18. Gerschgorin, S.: Über die Abgrenzung der Eigenwerte einer Matrix. Bulletin de l’Académie des Sciences de l’URSS. Classe des sciences mathématiques et naturelles pp. 749–754 (1931)
19. 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(7), 689–703 (2002)
20. Gonzalez, R., Woods, R.: Digital Image Processing. Prentice-Hall, Upper Saddle River, NJ, USA, third edn. (2008)
21. Gronwall, T.H.: Note on the derivatives with respect to a parameter of the solutions of a system of differential equations. Annals of Mathematics 20(4), 292–296 (Jul 1919)
22. Hadamard, J.: Sur les problèmes aux dériviées partielles et leur signification physique. Princeton University Bulletin 13, 49-52 (1902)
23. Hào, D.N., Duc, N.V.: Stability results for the heat equation backward in time. Journal of Mathematical Analysis and Applications 353(2), 627–641 (May 2009)
24. Hào, D.N., Duc, N.V.: Stability results for backward parabolic equations with time-dependent coefficients. Inverse Problems 27(2), 025063 (Jan 2011)
25. Hummel, R.A., Kimia, B.B., Zucker, S.W.: Deblurring Gaussian blur. Computer Vision, Graphics, and Image Processing 38(1), 66–80 (Apr 1987)
26. John, F.: Numerical solution of the equation of heat conduction for preceding times. Annali di Matematica Pura ed Applicata 40(1), 129–142 (Dec 1955)
27. Kirkup, S., Wadsworth, M.: Solution of inverse diffusion problems by operator-splitting methods. Applied Mathematical Modelling 26(10), 1003–1018 (2002)
28. Kovásznay, L.S.G., Joseph, H.M.: Image processing. Proceedings of the IRE 43(5), 560–570 (May 1955)
29. Lindenbaum, M., Fischer, M., Bruckstein, A.M.: On Gabor’s contribution to image enhancement. Pattern Recognition 27(1), 1–8 (Jan 1994)
30. Lyapunov, A.M.: The general problem of the stability of motion. International Journal of Control 55(3), 531–534 (1992)
31. Mair, B.A., Wilson, D.C., Reti, Z.: Deblurring the discrete Gaussian blur. In: Proceedings of the Workshop on Mathematical Methods in Biomedical Image Analysis. pp. 273–277. IEEE, Los Alamitos (Jun 1996)
32. Naik, S.K., Murthy, C.A.: Hue-preserving color image enhancement without gamut problem. IEEE Transactions on Image Processing 12(12), 1591–1598 (Dec 2003)
33. Nesterov, Y.: Introductory Lectures On Convex Optimization. Springer, New York (2004)
34. Nikolova, M., Steidl, G.: Fast hue and range preserving histogram specification: Theory and new algorithms for color image enhancement. IEEE Transactions on Image Processing 23(9), 4087–4100 (Sep 2014)
35. Nikolova, M., Steidl, G.: Fast ordering algorithm for exact histogram specification. IEEE Transactions on Image Processing 23(12), 5274–5283 (Dec 2014)
36. Nikolova, M., Wen, Y.W., Chan, R.: Exact histogram specification for digital images using a variational approach. Journal of Mathematical Imaging and Vision 46(3), 309–325 (Jul 2013)
37. Ojo, J.A., Solomon, I.D., Adeniran, S.A.: Colour-preserving contrast enhancement algorithm for images. In: Chen, L., Kapoor, S., Bhattacharya, I. (eds.) Emerging Trends and Advanced Technologies for Computational Intelligence: Extended and Selected Results from the Science and Information Conference 2015. pp. 207–222. Springer, Cham (2016)
38. Osher, S., Rudin, L.: Shocks and other nonlinear filtering applied to image processing. In: Tescher, A.G. (ed.) Applications of Digital Image Processing XIV, Proceedings of SPIE, vol. 1567, pp. 414–431. SPIE Press, Bellingham (1991)
39. Perko, L.: Differential Equations and Dynamical Systems. No. 7 in Texts in Applied Mathematics, Springer, third edn. (2001)
40. Perona, P., Malik, J.: Scale space and edge detection using anisotropic diffusion. IEEE Transactions on Pattern Analysis and Machine Intelligence 12, 629–639 (1990)
41. Pierre, F., Aujol, J.F., Bugeau, A., Steidl, G., Ta, V.T.: Variational contrast enhancement of gray-scale and RGB images. Journal of Mathematical Imaging and Vision 57(1), 99–116 (Jan 2017)
42. Puers, S.M., Ambrun, E.P., Austin, J.D., Cromartie, R., Geselowitz, A., Greer, T., ter Haar Romeny, B.M., Zimmerman, J.B., Zuiderveld, K.: Adaptive histogram equalization and its variations. Computer Vision, Graphics, and Image Processing 39(3), 355–368 (1987)
43. Pollak, I., Willsky, D.S., Krim, H.: Image segmentation and edge enhancement with stabilized inverse diffusion equations. IEEE Transactions on Image Processing 9(2), 256–266 (Feb 2000)
44. Pratt, W.K.: Digital Image Processing. Wiley & Sons, New York, third edn. (2001)
45. Sapiro, G., Caselles, V.: Histogram modification via differential equations. Journal of Differential Equations 135, 238–268 (1997)
46. Sochen, N.A., Zeevi, Y.Y.: Resolution enhancement of colored images by inverse diffusion processes. In: Proc. 1998 IEEE International Conference on Acoustics, Speech and Signal Processing. pp. 2853–2856. Seattle, WA (May 1998)
47. Steiner, A., Kimmel, R., Bruckstein, A.M.: Planar shape enhancement and exaggeration. Graphical Models and Image Processing 60(2), 112–124 (Mar 1998)
48. Stokes, M., Anderson, M., Chandrasekar, S., Motta, R.: A standard default color space for the internet - sRGB (version 1.10). https://www.w3.org/Graphics/Color/sRGB (Nov 1996), last visited August 31, 2018
49. Tautenhahn, U., Schröter, T.: On optimal regularization methods for the backward heat equation. Zeitschrift für Analysis und ihre Anwendungen 15(2), 475–493 (1996)
50. ter Haar Romeny, B.M., Florack, L.M.J., de Swart, M., Wilting, J., Viergever, M.A.: Deblurring Gaussian blur. In: Bookstein, F.L., Duncan, J.S., Lange, N., Wilson, D.C. (eds.) Mathematical Methods in Medical Imaging III, Proceedings of SPIE, vol. 2299, pp. 139–148. SPIE Press, Bellingham (Jul 1994)

51. Ternat, F., Orellana, O., Daripa, P.: Two stable methods with numerical experiments for solving the backward heat equation. Applied Numerical Mathematics 61(2), 266–284 (Feb 2011)

52. Tian, Q.C., Cohen, L.D.: Color consistency for photo collections without gamut problems. In: Ansøleg, L., Guðmundsson, G., Gurrin, C., Jónsson, B., Satoh, S. (eds.) MultiMedia Modeling: 23rd International Conference, MMM 2017, Reykjavik, Iceland, January 4-6, 2017, Proceedings, Part I. pp. 90–101. Springer, Cham (Jan 2017)

53. Welk, M., Weickert, J.: An efficient and stable two-pixel scheme for 2D forward-and-backward diffusion. In: Lauze, F., Dong, Y., Dahl, A. (eds.) Scale Space and Variational Methods in Computer Vision, Lecture Notes in Computer Science, vol. 10302, pp. 94–106. Springer, Cham (2017)

54. Zhao, Z., Meng, Z.: A modified Tikhonov regularization method for a backward heat equation. Inverse Problems in Science and Engineering 19(8), 1175–1182 (2011)

55. Zuiderveld, K.: Contrast Limited Adaptive Histogram Equalization. In: Heckbert, P.S. (ed.) Graphics Gems IV, pp. 474–485. Academic Press Professional, Inc., San Diego, CA, USA (1994)