


Cross-Diffusion Systems for Image Processing: II. The Nonlinear Case

A. Araújo¹ · S. Barbeiro¹ · E. Cuesta² · A. Durán² 

Abstract In this paper we study the application of 2×2 nonlinear cross-diffusion systems as mathematical models of image filtering. These are systems of two nonlinear, coupled partial differential equations of parabolic type. The nonlinearity and cross-diffusion character are provided by a nondiagonal matrix of diffusion coefficients that depends on the variables of the system. We prove the well-posedness of an initial-boundary-value problem with Neumann boundary conditions and uniformly positive definite cross-diffusion matrix. Under additional hypotheses on the coefficients, the models are shown to satisfy the scale-space properties of shift, contrast, average grey and translational invariances. The existence of Lyapunov functions and the asymptotic behaviour of the solutions are also studied. According to the choice of the cross-diffusion matrix (on the basis of the results on filtering with linear cross-diffusion, discussed by the authors in a companion paper and the use of edge stopping functions) the performance of the models is compared by computational means in a filtering problem. The numerical results reveal differences in the evolution of the filtering as well as in the quality of edge detection given by one of

the components of the system, in terms of the cross-diffusion matrix.

Keywords Cross-diffusion · Complex diffusion · Image denoising

1 Introduction

This paper is concerned with the use of nonlinear cross-diffusion systems for the mathematical modelling of image filtering. In this approach, a grey-scale image is represented by a vector field $\mathbf{u} = (u, v)^T$ of two real-valued functions u, v defined on some domain in \mathbb{R}^2 . Additionally, an image restoration problem is modelled by an evolutionary process such that, from an initial distribution of a noisy image and with the time as a scale parameter, the restored image at any time satisfies an initial-boundary-value problem (IBVP) of a nonlinear system of partial differential equations (PDE) of cross-diffusion type, where the coupled evolution of the two components of the image and the nonlinearity are determined by a cross-diffusion coefficient matrix.

The use of cross-diffusion systems for modelling, especially in population dynamics, is well known, see e. g. Galiano et al. [11,12] and Ni [22] (along with references therein). To our knowledge, in the case of image processing, two previous proposals are related. The first one concerns the use of complex diffusion (Gilboa et al. [15]), where the image is represented by a complex function and the filtering process is governed by a nonlinear PDE of diffusion type with a complex-valued diffusion coefficient. This equation can be written as a cross-diffusion system for the real and imaginary parts of the image. The application of complex diffusion to image filtering and edge-enhancing problems brings advantages based on the role of the imaginary part

✉ A. Durán
angel@mac.uva.es

A. Araújo
alma@mat.uc.pt

S. Barbeiro
silvia@mat.uc.pt

E. Cuesta
eduardo@mat.uva.es

¹ CMUC, Department of Mathematics, University of Coimbra, Coimbra, Portugal

² Department of Applied Mathematics, University of Valladolid, Valladolid, Spain

as edge detector in the linear case (the so-called small theta approximation) and its use, in the nonlinear case, instead of the size of the gradient of the image as the main variable to control the diffusion coefficient, Gilboa et al. [13–15].

A second reference on nonlinear cross-diffusion is the unpublished manuscript by Lorenz et al. [20], where the authors prove the existence of a global solution of a cross-diffusion problem, related to the complex diffusion approach proposed by Gilboa and collaborators. This already represents an advance with respect to the ill-posed Perona–Malik formulation, Perona and Malik [25] and Kinchenassamy [17]. Additionally, a better behaviour of cross-diffusion models with respect to the textures of the image is numerically suggested.

Concerning the use of coupled diffusion problems with vector-valued images, several references can be emphasized. In the monograph ter Haar Romeny (Ed.) [16], Whitaker and Gerig [28] develop a framework for variable conductance diffusion, which is applied to vector-valued images consisting of geometric features of a single scalar image. They use two strategies: the first one, in order to study high-order geometric features, is the geometry-limited diffusion, which makes use of partial derivatives as input to a vector-valued diffusion problem. The second strategy, or spectra-limited diffusion, is applied in texture problems and uses a local frequency decomposition as input to a vector-valued diffusion process. The idea of describing an image by using vector-valued functions is also present in Proesmans et al. [26] (also in ter Haar Romeny [16]), where several systems of coupled, nonlinear diffusion problems for such maps are introduced. The behaviour of the properties described by the components of the vector image is investigated in edge-preserving and filtering problems arising in optical flow, stereo-matching and multi-spectral analysis. Additionally, the use of coupled systems of diffusion-reaction equations in optical flow is made in Álvarez et al. [4], by introducing an improved formulation of the classical model by Nagel and Enkelmann [21]. Finally, Peter et al. [24] (see also references therein) derive and justify tensor-driven diffusion models by using statistics of natural images, discussing the corresponding performance in nonlinear filtering.

The present paper is a continuation of a companion work by the same authors devoted to the application of linear cross-diffusion processes to image filtering (Araújo et al. [2]). The linear cross-diffusion is analysed as a scale-space representation, and an axiomatic, based on scale invariance, is built. Then those convolution kernels satisfying shift, rotational and scale invariance as well as recursivity (semigroup property) are characterized. The resulting filters are determined by a positive definite matrix, directing the diffusion and a positive parameter which, as in the scalar case, Pauwels et al. [23], delimits the locality property. Furthermore, since complex diffusion can be seen as a particular case of cross-

diffusion, some properties of the former are generalized in the latter. More precisely, the use of one of the components of the cross-diffusion system as edge detector is investigated, extending the property of small theta approximation.

The general purpose of the present paper is to continue the research on cross-diffusion models for image processing, by incorporating nonlinearity. The contributions of the paper are the following:

- We formalize nonlinear cross-diffusion IBVP as mathematical models for image processing, by proving the following theoretical results:
 1. Well-posedness. By assuming that the coefficient matrix is uniformly positive definite and has globally Lipschitz and bounded entries, the IBVP of a nonlinear cross-diffusion system of PDE with Neumann boundary conditions is studied. The existence of a unique weak solution, continuous dependence on the initial data and the existence of an extremum principle are proved. Some of the arguments of Lorenz et al. [20] for the system under study will be used and generalized here. Some extensions, not treated here, are the use of nonlocal operators and different types of boundary conditions in the PDE formulation.
 2. The previous IBVP is also studied from the scale-space representation viewpoint, see e.g. Álvarez et al. [3] and Lindeberg [19]. Specifically, grey-level shift invariance, reverse contrast invariance and translational invariance are proved under additional assumptions on the diffusion coefficients.
 3. The theoretical results are completed by analysing the existence of Lyapunov functionals associated to the cross-diffusion problem, Weickert [27]. The first result here is the decreasing of the energy (defined as the Euclidean norm of the solution) by cross-diffusion. The existence of Lyapunov functionals different from this energy depends on the relation between the cross-diffusion coefficient matrix and the function defining the functional. Finally, the solution is proved to evolve asymptotically to a constant image consisting of the average values of the components of the initial distribution.
- A numerical comparison of the performance of the models is made. The computational study is carried out on the basis of the results about the linear models, presented in Araújo et al. [2] and the numerical treatment of complex diffusion in Gilboa et al. [15]. More precisely, the performance of the experiments is based on the choices of the cross-diffusion coefficient matrix and the scheme of approximation to the continuous problem. As far as the coefficients are concerned, we select a matrix which combines linear cross-diffusion, including a constant positive definite matrix, with the use of standard

edge detection functions, depending on the component of the image that plays the role of edge detector from the generalized small theta approximation. The resulting form of the diffusion matrix generalizes the complex diffusion approach, Gilboa et al. [15]. Two strategies for the treatment of the edge detection functions are also implemented.

On the other hand, an adaptation to cross-diffusion systems of an explicit numerical method, considered and analysed in Araújo et al. [5] and Bernardes et al. [7], for complex diffusion problems was used to perform the numerical experiments in filtering problems. The numerical results reveal differences in the behaviour of the models, according to the choice of the positive definite matrix and the edge stopping function. They are mainly concerned with a delay of the blurring effect (already observed in the linear case) and the influence of the generalized small theta approximation in the detection of the edges during the filtering problem.

The paper is structured according to these highlights. In Sect. 2, the IBVP of a cross-diffusion PDE with Neumann boundary conditions is introduced and the theoretical results of well-posedness, scale-space properties, Lyapunov functions and long time behaviour are proved. Section 3 is devoted to the computational study of the performance of the models. The main conclusions and future research are outlined in Section 4.

The following notation will be used throughout the paper. A bounded (typically rectangular) domain in \mathbb{R}^2 will be denoted by Ω , with boundary $\partial\Omega$ and where $\bar{\Omega} := \Omega \cup \partial\Omega$. By \mathbf{n} we denote the outward normal vector to $\partial\Omega$. For p positive integer, $L^p(\Omega)$ denotes the normed space of L^p -functions on Ω with $\|\cdot\|_{L^p}$ as the associated norm. From the Sobolev space $H^k(\Omega)$ on Ω (k is a nonnegative integer), where $H^0(\Omega) = L^2(\Omega)$ we define $X_k := H^k(\Omega) \times H^k(\Omega)$ with norm denoted by

$$\|\mathbf{u}\|_{X_k} = \left(\|u\|_k^2 + \|v\|_k^2 \right)^{1/2}, \quad \mathbf{u} = (u, v)^T,$$

where $\|\cdot\|_k$ is the norm in $H^k(\Omega)$. On the other hand, the dual space of $H^k(\Omega)$ will be denoted by $(H^k(\Omega))'$; this is characterized as the completion of $L^2(\Omega)$ with respect to the norm, [1],

$$\|v\|_{-k,2} = \sup_{u \in H^k(\Omega), \|u\|_k=1} |\langle u, v \rangle|, \quad \langle u, v \rangle = \int_{\Omega} uv \, d\Omega.$$

Additionally, $(X_k)'$ will stand for $(H^k(\Omega))' \times (H^k(\Omega))'$.

For $T > 0$, $Q_T = \Omega \times (0, T]$ will denote the set of points (\mathbf{x}, t) with $\mathbf{x} \in \Omega$, $0 < t \leq T$ and $\bar{Q}_T := \bar{\Omega} \times [0, T]$. The space of infinitely continuously differentiable real-valued functions in $\bar{\Omega} \times (0, T]$ will be denoted by $C^\infty(\bar{\Omega} \times (0, T])$

as well as the space of m -th order continuously differentiable functions $\mathbf{u} : (0, T] \rightarrow X_k$ by $C^m(0, T, X_k)$, m, k nonnegative integers. Additionally, $L^2(0, T, H^k)$ will stand for the normed space of functions $u : (0, T] \rightarrow H^k(\Omega)$ with associated norm

$$\|u\|_{L^2(0,T,H^k)} = \left(\int_0^T \|u(t)\|_k^2 dt \right)^{1/2}.$$

We also denote by $L^\infty(0, T, H^k)$ the normed space of functions $u : (0, T] \rightarrow H^k(\Omega)$ with norm

$$\|u\|_{L^\infty(0,T,H^k)} = \text{ess sup}_{t \in (0,T)} \|u(t)\|_k,$$

with ess sup as the essential supremum. (The essential infimum will be denoted as ess inf .)

In Sect. 2, we will make use of the convolution operator

$$(g * f)(\mathbf{x}) = \int_{\mathbb{R}^2} f(\mathbf{x} - \mathbf{y})g(\mathbf{y})d\mathbf{y}, \quad (1.1)$$

for $g \in L^1(\mathbb{R}^2)$, $f \in L^2(\mathbb{R}^2)$ and the Fourier transform

$$\widehat{f}(\xi) = \int_{\mathbb{R}^2} f(\mathbf{x})e^{-i\xi \cdot \mathbf{x}}d\mathbf{x}, \quad f \in L^2(\mathbb{R}^2), \quad \xi \in \mathbb{R}^2, \quad (1.2)$$

where \cdot denotes the Euclidean inner product in \mathbb{R}^2 with the norm represented by $|\cdot|$. In order to define (1.1), (1.2) when $f \in L^2(\Omega)$, a continuous extension of f in \mathbb{R}^2 will be considered and denoted by \tilde{f} .

Finally, div , ∇ will stand, respectively, for the divergence and gradient operators. Concerning the gradient, if $\mathbf{u} = (u, v)^T$ then $J\mathbf{u}$ stands for the Jacobian matrix of \mathbf{u} , $\mathbf{u}_x = (u_x, v_x)^T$, $\mathbf{u}_y = (u_y, v_y)^T$ and

$$\|J\mathbf{u}\|_{X_0} := \left(\|\mathbf{u}_x\|_{X_0}^2 + \|\mathbf{u}_y\|_{X_0}^2 \right)^{1/2}.$$

Additional notation for the numerical experiments will be specified in Sect. 3.

2 Nonlinear Cross-Diffusion Model

We consider the following IBVP of cross-diffusion for $\mathbf{u} = (u, v)^T$,

$$\begin{aligned} \frac{\partial u}{\partial t}(\mathbf{x}, t) &= \text{div} (D_{11}(\mathbf{u}(\mathbf{x}, t))\nabla u(\mathbf{x}, t) \\ &\quad + D_{12}(\mathbf{u}(\mathbf{x}, t))\nabla v(\mathbf{x}, t)), \\ \frac{\partial v}{\partial t}(\mathbf{x}, t) &= \text{div} (D_{21}(\mathbf{u}(\mathbf{x}, t))\nabla u(\mathbf{x}, t) \\ &\quad + D_{22}(\mathbf{u}(\mathbf{x}, t))\nabla v(\mathbf{x}, t)), \quad (\mathbf{x}, t) \in Q_T, \end{aligned} \quad (2.1)$$

with the initial data given by

$$u(\mathbf{x}, 0) = u_0(\mathbf{x}), \quad v(\mathbf{x}, 0) = v_0(\mathbf{x}), \quad \mathbf{x} \in \Omega, \quad (2.2)$$

and Neumann boundary conditions in $\partial\Omega \times [0, T]$,

$$\begin{aligned} \langle D_{11}(\mathbf{u})\nabla u + D_{12}(\mathbf{u})\nabla v, n \rangle &= 0, \\ \langle D_{21}(\mathbf{u})\nabla u + D_{22}(\mathbf{u})\nabla v, n \rangle &= 0. \end{aligned} \quad (2.3)$$

In (2.1), (2.3), the scalar functions D_{ij} , $i, j = 1, 2$, are the entries of a cross-diffusion 2×2 matrix operator

$$\mathbf{u} \mapsto D(\mathbf{u}) : \overline{Q_T} \rightarrow M_{2 \times 2}(\mathbb{R}),$$

with, for $(\mathbf{x}, t) \in \overline{Q_T}$,

$$D(\mathbf{u})(\mathbf{x}, t) = D(\mathbf{u}(\mathbf{x}, t)) = \begin{pmatrix} D_{11}(\mathbf{u}(\mathbf{x}, t)) & D_{12}(\mathbf{u}(\mathbf{x}, t)) \\ D_{21}(\mathbf{u}(\mathbf{x}, t)) & D_{22}(\mathbf{u}(\mathbf{x}, t)) \end{pmatrix},$$

and which satisfies the following hypotheses:

(H1) There exists $\alpha > 0$ such that for each $\mathbf{u} : \overline{Q_T} \rightarrow \mathbb{R}^2$

$$\xi^T D(\mathbf{u}(\mathbf{x}, t)) \xi \geq \alpha |\xi|^2, \quad \xi \in \mathbb{R}^2, (\mathbf{x}, t) \in \overline{Q_T}. \quad (2.4)$$

(H2) There exists $L > 0$ such that for $\mathbf{u}, \mathbf{v} : \overline{Q_T} \rightarrow \mathbb{R}^2$, $(\mathbf{x}, t) \in \overline{Q_T}$, $i, j = 1, 2$,

$$|D_{ij}(\mathbf{v}(\mathbf{x}, t)) - D_{ij}(\mathbf{u}(\mathbf{x}, t))| \leq L |\mathbf{v}(\mathbf{x}, t) - \mathbf{u}(\mathbf{x}, t)|.$$

(H3) There exists $M > 0$ such that for each $\mathbf{u} : \overline{Q_T} \rightarrow \mathbb{R}^2$

$$|D_{ij}(\mathbf{u}(\mathbf{x}, t))| \leq M, \quad (\mathbf{x}, t) \in \overline{Q_T}, i, j = 1, 2.$$

Conditions (H1)–(H3) will also be complemented with other assumptions, required by scale-space properties, see Sect. 2.2.

In what follows the weak formulation of (2.1)–(2.3) will be considered. This consists of finding $\mathbf{u} = (u, v)^T : (0, T] \rightarrow X_1$ satisfying, for any $t \in (0, T]$

$$\begin{aligned} \int_{\Omega} ((\partial_t u)w_1 + (\partial_t v)w_2) \, d\Omega \\ + \int_{\Omega} \text{tr} \left((J\mathbf{w})^T D(\mathbf{u})(J\mathbf{u}) \right) \, d\Omega = 0, \end{aligned} \quad (2.5)$$

for all $\mathbf{w} = (w_1, w_2)^T \in X_1$ and where tr denotes the trace of the matrix.

2.1 Well-Posedness

This section is devoted to the study of well-posedness of (2.1)–(2.3). More precisely, we prove the existence of a unique solution of (2.5), regularity, continuous dependence on the initial data and finally an extremum principle. The

proofs follow standard arguments, see Catté et al. [9], Weickert [27] (see also Galiano et al. [11] and references therein). We first consider a related linear problem and prove a maximum–minimum principle as well as estimates of the solution in different norms. These results are crucial to prove the existence of the solution for the nonlinear case by using the Schauder fixed-point theorem, Brezis [8]. The same arguments as in Catté et al. [9] and Weickert [27] apply to prove the uniqueness, as well as regularity and continuous dependence on the initial data. Finally, the proof of the extremum principle for the linear problem can be adapted to obtain the corresponding result for (2.1)–(2.3).

Theorem 1 *Let us assume that (H1)–(H3) hold and let $\mathbf{u}_0 = (u_0, v_0)^T \in X_1$. Then (2.5) admits a unique solution $\mathbf{u} \in C(0, T, X_0) \cap L^2(0, T, X_1)$ that depends continuously on the initial data. Furthermore, if D is in $C^\infty(\mathbb{R}^2, M_{2 \times 2}(\mathbb{R}))$ then \mathbf{u} is a strong solution of (2.1)–(2.3) with $\mathbf{u} \in C^\infty(\overline{\Omega} \times (0, T])$.*

Proof We first define

$$W(0, T) = \left\{ w \in L^2(0, T, H^1(\Omega)) : \frac{dw}{dt} \in L^2(0, T, (H^1(\Omega))) \right\},$$

with the graph norm.

Existence

In order to study the existence of solution of (2.5) we first consider, for $\mathbf{U} = (U, V)^T$, with

$$U, V \in W(0, T) \cap L^\infty(0, T, L^2(\Omega)),$$

the following linear IBVP in Q_T :

$$\begin{aligned} \frac{\partial u}{\partial t}(\mathbf{x}, t) &= \text{div} (D_{11}(\mathbf{U}(\mathbf{x}, t))\nabla u(\mathbf{x}, t) \\ &\quad + D_{12}(\mathbf{U}(\mathbf{x}, t))\nabla v(\mathbf{x}, t)), \\ \frac{\partial v}{\partial t}(\mathbf{x}, t) &= \text{div} (D_{21}(\mathbf{U}(\mathbf{x}, t))\nabla u(\mathbf{x}, t) \\ &\quad + D_{22}(\mathbf{U}(\mathbf{x}, t))\nabla v(\mathbf{x}, t)), \\ u(\mathbf{x}, 0) &= u_0(\mathbf{x}), \quad v(\mathbf{x}, 0) = v_0(\mathbf{x}), \quad \mathbf{x} \in \Omega, \end{aligned} \quad (2.6)$$

with Neumann boundary conditions in $\partial\Omega \times [0, T]$

$$\begin{aligned} \langle D_{11}(\mathbf{U})\nabla u + D_{12}(\mathbf{U})\nabla v, n \rangle &= 0, \\ \langle D_{21}(\mathbf{U})\nabla u + D_{22}(\mathbf{U})\nabla v, n \rangle &= 0. \end{aligned} \quad (2.7)$$

Since $D(\mathbf{U}) = D(U, V)$ is uniformly positive definite (hypothesis (H1)), then, e.g. Ladyzenskaya et al. [18], there is a unique weak solution of (2.6), (2.7), $\mathbf{u}(U, V) = (U_1(U, V), U_2(U, V))$, with

$$U_1, U_2 \in W(0, T) \cap L^\infty(0, T, L^2(\Omega)).$$

We now establish some estimates for this solution in different norms, Lorenz et al. [20]. Consider first the weak formulation of (2.6): find $\mathbf{u}(U, V) = (U_1(U, V), U_2(U, V))$ in $L^2(0, T, X_1)$ satisfying

$$\int_{\Omega} ((\partial_t U_1)v_1 + (\partial_t U_2)v_2) \, d\Omega + \int_{\Omega} \operatorname{tr} \left((J\mathbf{v})^T D(U, V)(J\mathbf{u}) \right) \, d\Omega = 0, \quad (2.8)$$

for every $\mathbf{v} = (v_1, v_2) \in X_1$ and all $0 \leq t \leq T$. We take the test functions $v_1 = (U_1 - b_1)_+$, $v_2 = (U_2 - b_2)_+$ for some $b_1, b_2 > 0$ that will be specified later and where $f_+ = \max\{f, 0\}$ (Lorenz et al. [20], Weickert [27]). Then (2.8) becomes

$$\frac{1}{2} \int_{\Omega} \left(\partial_t (U_1 - b_1)_+^2 + \partial_t (U_2 - b_2)_+^2 \right) \, d\Omega + \int_{U_1 > b_1, U_2 > b_2} \operatorname{tr} \left((J\mathbf{u})^T D(U, V)(J\mathbf{u}) \right) \, d\Omega = 0.$$

Then (H1) implies that

$$\frac{d}{dt} \int_{\Omega} \left((U_1 - b_1)_+^2 + (U_2 - b_2)_+^2 \right) \, d\Omega \leq 0.$$

Thus integrating between 0 and t , for any $0 \leq t \leq T$, we have

$$\int_{\Omega} \left((U_1(t) - b_1)_+^2 + (U_2(t) - b_2)_+^2 \right) \, d\Omega \leq \int_{\Omega} \left((U_1(0) - b_1)_+^2 + (U_2(0) - b_2)_+^2 \right) \, d\Omega. \quad (2.9)$$

Now we take b_1, b_2 such that the integral on the right-hand side of (2.9) becomes zero. If we assume that $U_1(0), U_2(0) \in L^\infty(\Omega)$ and define

$$b_1 = \|U_1(0)\|_{L^\infty(\Omega)}, \quad b_2 = \|U_2(0)\|_{L^\infty(\Omega)},$$

then (2.9) implies

$$\int_{\Omega} \left((U_1(t) - b_1)_+^2 + (U_2(t) - b_2)_+^2 \right) \, d\Omega \leq 0$$

and consequently $(U_1(t) - b_1)_+ = (U_2(t) - b_2)_+ = 0$ for $0 \leq t \leq T$, that is

$$\begin{aligned} U_1(x, t) &\leq b_1 = \|U_1(0)\|_{L^\infty(\Omega)}, \\ U_2(x, t) &\leq b_2 = \|U_2(0)\|_{L^\infty(\Omega)}. \end{aligned} \quad (2.10)$$

Similarly, taking $v_1 = (U_1 - a_1)_-$, $v_2 = (U_2 - a_2)_-$ for some $a_1, a_2 > 0$ and where $f_- = \min\{f, 0\}$, the same argument leads to

$$\begin{aligned} &\int_{\Omega} \left((U_1(t) - a_1)_-^2 + (U_2(t) - a_2)_-^2 \right) \, d\Omega \\ &\leq \int_{\Omega} \left((U_1(0) - a_1)_-^2 + (U_2(0) - a_2)_-^2 \right) \, d\Omega. \end{aligned}$$

If we now define

$$a_1 = \operatorname{ess\,inf} U_1(0), \quad a_2 = \operatorname{ess\,inf} U_2(0),$$

then

$$\int_{\Omega} \left((U_1(t) - a_1)_-^2 + (U_2(t) - a_2)_-^2 \right) \, d\Omega \leq 0$$

and therefore $(U_1(t) - a_1)_- = (U_2(t) - a_2)_- = 0$ for $0 \leq t \leq T$, that is

$$U_1(x, t) \geq \operatorname{ess\,inf} U_1(0), \quad U_2(x, t) \geq \operatorname{ess\,inf} U_2(0). \quad (2.11)$$

In particular, if $U_1(0), U_2(0) \geq 0$ then $U_1(x, t), U_2(x, t) \geq 0$ for all $(x, t) \in Q_T$.

A second estimate for the solution of the linear problem (2.6) is now obtained from the functional of energy

$$E_L(t) = \frac{1}{2} \int_{\Omega} \operatorname{tr} \left((J\mathbf{u})^T D(U, V)(J\mathbf{u}) \right) \, d\Omega.$$

Note that if in the weak formulation (2.8) we take $\mathbf{v} = (U_1, U_2)^T$ then

$$\frac{d}{dt} E_L(t) + \int_{\Omega} (\nabla U_1 \nabla U_2) D(U, V) (\nabla U_1 \nabla U_2)^T \, d\Omega = 0,$$

which implies

$$\frac{d}{dt} E_L(t) \leq 0,$$

that is $E_L(t)$ decreases. This leads to the L^∞ estimates

$$\begin{aligned} \|U_1\|_{L^\infty(0, T, L^2(\Omega))} &\leq \|U_1(0)\|_{L^2(\Omega)}, \\ \|U_2\|_{L^\infty(0, T, L^2(\Omega))} &\leq \|U_2(0)\|_{L^2(\Omega)}. \end{aligned} \quad (2.12)$$

We now search for estimates of $U_1(t), U_2(t)$ as functions in $H^1(\Omega)$ (and also of $\frac{d}{dt} U_1(t), \frac{d}{dt} U_2(t)$ as functions in $(H^1(\Omega))'$). Note first that from the previous argument we have, for $t \in [0, T]$,

$$\begin{aligned} &\int_{\Omega} \left(U_1(\mathbf{x}, t)^2 + U_2(\mathbf{x}, t)^2 \right) \, d\Omega \\ &\leq \int_{\Omega} \left(U_1(\mathbf{x}, 0)^2 + U_2(\mathbf{x}, 0)^2 \right) \, d\Omega, \end{aligned}$$

and also

$$\begin{aligned} & \frac{d}{dt} \frac{1}{2} \int_{\Omega} (U_1(\mathbf{x}, t)^2 + U_2(\mathbf{x}, t)^2) d\Omega \\ & + \alpha \int_{\Omega} (|\nabla U_1(\mathbf{x}, t)|^2 + |\nabla U_2(\mathbf{x}, t)|^2) d\Omega \leq 0. \end{aligned} \quad (2.13)$$

Then (2.13) implies that for any $t \in [0, T]$

$$\begin{aligned} & \frac{1}{2} \int_{\Omega} (U_1(\mathbf{x}, t)^2 + U_2(\mathbf{x}, t)^2) d\Omega \\ & + \alpha \int_0^t \int_{\Omega} (|\nabla U_1(\mathbf{x}, s)|^2 + |\nabla U_2(\mathbf{x}, s)|^2) d\Omega ds \\ & \leq \frac{1}{2} \int_{\Omega} (U_1(\mathbf{x}, 0)^2 + U_2(\mathbf{x}, 0)^2) d\Omega. \end{aligned}$$

Therefore,

$$\begin{aligned} & \int_0^T \frac{1}{2} \int_{\Omega} (U_1(\mathbf{x}, t)^2 + U_2(\mathbf{x}, t)^2) d\Omega dt \\ & + \alpha \int_0^T \int_{\Omega} (|\nabla U_1(\mathbf{x}, t)|^2 + |\nabla U_2(\mathbf{x}, t)|^2) d\Omega dt \\ & = \int_0^T E_L(t) dt \\ & + \alpha \int_0^T \int_{\Omega} (|\nabla U_1(\mathbf{x}, t)|^2 + |\nabla U_2(\mathbf{x}, t)|^2) d\Omega dt \\ & = \int_0^T E_L(t) dt - E_L(T) + E_L(0) \\ & + \alpha \int_0^T \int_{\Omega} (|\nabla U_1(\mathbf{x}, t)|^2 + |\nabla U_2(\mathbf{x}, t)|^2) d\Omega dt \\ & \leq \int_0^T E_L(t) dt - E_L(T) + E_L(0) \leq (T+1)E_L(0). \end{aligned}$$

Thus, if $\mathbf{U}_0 = (U_1(0), U_2(0))^T$ then there exists a constant $C_1 = C_1(\alpha, \mathbf{U}_0, T)$ such that

$$\begin{aligned} \|U_1\|_{L^2(0,T,H^1(\Omega))} & \leq C_1, \\ \|U_2\|_{L^2(0,T,H^1(\Omega))} & \leq C_1. \end{aligned} \quad (2.14)$$

On the other hand, if $\|\mathbf{v}\|_{L^2(0,T,X_1)} = 1$, the weak formulation (2.8), assumption (H3) and Cauchy–Schwarz inequality imply that

$$\begin{aligned} & \left| \int_0^T \int_{\Omega} ((\partial_t U_1)v_1 + (\partial_t U_2)v_2) d\Omega dt \right| \\ & = \left| \int_0^T \left(\int_{\Omega} \text{tr} \left((J\mathbf{v})^T D(U, V)(J\mathbf{u}) \right) d\Omega \right) dt \right| \\ & \leq \int_0^T M \|\nabla \mathbf{v}(t)\|_{X_0} \|\nabla \mathbf{u}(t)\|_{X_0} dt \\ & \leq \int_0^T M \|\mathbf{v}(t)\|_{X_1} \|\mathbf{u}(t)\|_{X_1} dt \\ & \leq M \|\mathbf{v}\|_{L^2(0,T,X_1)} \|\mathbf{u}\|_{L^2(0,T,X_1)} = M \|\mathbf{u}\|_{L^2(0,T,X_1)}. \end{aligned}$$

Therefore, this and (2.14) lead to

$$\left\| \frac{d}{dt} \mathbf{u} \right\|_{L^2(0,T,(X_1)^2)} \leq M \|\mathbf{u}\|_{L^2(0,T,X_1)} \leq MC_1. \quad (2.15)$$

The existence of a solution of (2.5) is now derived, making use of the estimates (2.12), (2.14) and (2.15) and by using the Schauder fixed-point theorem, Brezis [8]. (Analogous arguments were used in Catté et al. [9], see also Weickert [27].) We first assume that $\mathbf{u}_0 = (u_0, v_0)^T \in X_0$ in (2.2). Consider the following subset of $W(0, T)^2 := W(0, T) \times W(0, T)$:

$$K = \{\mathbf{w} = (w_1, w_2)^T \in W(0, T)^2 : \mathbf{w} \text{ satisfies (2.12), (2.14) and (2.15) with } \mathbf{w}(0) = \mathbf{u}_0\},$$

and the mapping $T : K \rightarrow W(0, T)^2$ such that $T(\mathbf{w}) := \mathbf{u}(\mathbf{w})$ is the (weak) solution of (2.6) with $(U, V)^T = \mathbf{w}$.

It is not hard to see that K is a nonempty, convex subset of $W(0, T)^2$. Our goal is to apply the Schauder fixed-point theorem to the operator T in the weak topology. To this end, we need to prove that:

- (1) $T(K) \subset K$.
- (2) K is a weakly compact subset of $W(0, T)^2$.
- (3) T is weakly continuous.

Observe that by construction (1) is satisfied. In order to prove (2), consider a sequence $\{\mathbf{w}_n\}_n \subset K$ and $t \in [0, T]$. Since K is a bounded set, then

$$\{\mathbf{w}_n(t)\}_n, \left\{ \frac{d}{dt} \mathbf{w}_n(t) \right\}_n$$

are uniformly bounded in X_1 which implies the existence of a subsequence (denoted again by $\{\mathbf{w}_n(t)\}_n, \{\frac{d}{dt} \mathbf{w}_n(t)\}_n$) and $\varphi(t), \psi(t) \in X_1$ such that

$$\mathbf{w}_n(t) \rightarrow \varphi(t), \quad \frac{d}{dt} \mathbf{w}_n(t) \rightarrow \psi(t),$$

weakly in X_1 and for $0 \leq t \leq T$. On the other hand, since $W(0, T) \subset L^2(0, T, L^2(\Omega))$ and the embedding is compact, Catté et al. [9], there exists $\mathbf{w} \in L^2(0, T, X_0)$ such that $\|\mathbf{w}_n - \mathbf{w}\|_{L^2(0,T,X_0)} \rightarrow 0$ for some subsequence $\{\mathbf{w}_n\}_n$. Consequently, $\mathbf{w} = \varphi \in L^2(0, T, X_1)$. Actually, $\psi = \frac{d}{dt} \varphi$ and then K is weakly compact in $W(0, T)^2$.

Finally, consider a sequence $\{\mathbf{w}_n\}_n \subset K$ which converges weakly to some $\mathbf{w} \in K$. Let $\mathbf{u}_n = T(\mathbf{w}_n)$. In order to prove property (3), we have to see that \mathbf{u}_n converges weakly to $\mathbf{u} = T(\mathbf{w})$. Here the proof is similar to that of Catté et al. [9]. Previous arguments applied to \mathbf{u}_n and property (2) establish the existence of a subsequence $\{\mathbf{u}_n\}_n$ and $\phi \in L^2(0, T, X_1)$ satisfying

- (i) $\mathbf{u}_n \rightarrow \phi$ weakly in $L^2(0, T, X_1)$;
- (ii) $\frac{d}{dt}\mathbf{u}_n \rightarrow \frac{d}{dt}\phi$ weakly in $L^2(0, T, (X_1)')$;
- (iii) $\mathbf{u}_n \rightarrow \phi$ in $L^2(0, T, X_0)$ and almost everywhere on $\Omega \times [0, T]$, (e. g. Brezis [8], Theorem 4.9);
- (iv) $\mathbf{w}_n \rightarrow \mathbf{w}$ in $L^2(0, T, X_0)$ and almost everywhere on $\Omega \times [0, T]$.

These convergence properties imply two additional ones:

- (v) $\mathbf{u}_n(0) \rightarrow \phi(0)$ in $(X_1)'$;
- (vi) $\nabla \mathbf{u}_n \rightarrow \nabla \phi$ weakly in $L^2(0, T, X_0)$.

Now, note that due to (H2) and property (v) we have

$$D(\mathbf{w}_n) \rightarrow D(\mathbf{w})$$

in $L^2(0, T, X_0)$. Then if we take limit in (2.8) we have $\phi = T(\mathbf{w})$. Finally, since the whole sequence $\{\mathbf{u}_n\}_n$ is bounded in K which is weakly compact, then it converges weakly in $W(0, T)$. By uniqueness of solution of (2.8), the whole sequence $\mathbf{u}_n = T(\mathbf{w}_n)$ must converge weakly to $\phi = T(\mathbf{w})$; therefore, T is weakly continuous and (3) holds.

Thus, Schauder fixed-point theorem proves the existence of a solution \mathbf{u} of (2.5). The solution \mathbf{u} is in K and therefore $\mathbf{u} \in L^2(0, T, X_1)$, $\frac{d\mathbf{u}}{dt} \in L^2(0, T, (X_1)')$, and it satisfies (2.12), (2.14) and (2.15). Furthermore, due to the conditions (H1)–(H3) on D , at least $\mathbf{u} \in C(0, T, X_0)$.

Regularity of Solution

The same bootstrap argument as in Catté et al. [9] and Weickert [27] applies to obtain that \mathbf{u} is a strong solution and $\mathbf{u} \in C^\infty(\overline{\Omega} \times (0, T])$ if (H2) is substituted by the hypothesis that D is in $C^\infty(\mathbb{R}^2, M_{2 \times 2}(\mathbb{R}))$.

Uniqueness of Solution

Consider $\mathbf{u}^{(1)} = (u^{(1)}, v^{(1)})^T$, $\mathbf{u}^{(2)} = (u^{(2)}, v^{(2)})^T$ solutions of (2.5) with the same initial condition. Then for all $\mathbf{w} = (w_1, w_2)^T \in X_1$

$$\begin{aligned} & \int_{\Omega} \left((\partial_t(u^{(1)} - u^{(2)}))w_1 + (\partial_t(v^{(1)} - v^{(2)}))w_2 \right) d\Omega \\ & + \int_{\Omega} \text{tr} \left((J\mathbf{w})^T D(\mathbf{u}^{(1)})(J\mathbf{u}^{(1)}) \right) d\Omega \\ & - \int_{\Omega} \text{tr} \left((J\mathbf{w})^T D(\mathbf{u}^{(2)})(J\mathbf{u}^{(2)}) \right) d\Omega = 0, \end{aligned}$$

which can be written as

$$\begin{aligned} & \int_{\Omega} \left((\partial_t(u^{(1)} - u^{(2)}))w_1 + (\partial_t(v^{(1)} - v^{(2)}))w_2 \right) d\Omega \\ & + \int_{\Omega} \text{tr} \left((J\mathbf{w})^T D(\mathbf{u}^{(1)})(J(\mathbf{u}^{(1)} - \mathbf{u}^{(2)})) \right) d\Omega \\ & + \int_{\Omega} \text{tr} \left((J\mathbf{w})^T \left(D(\mathbf{u}^{(1)}) - D(\mathbf{u}^{(2)}) \right) (J\mathbf{u}^{(2)}) \right) d\Omega = 0. \end{aligned}$$

Now we take $\mathbf{w} = \mathbf{u}^{(1)} - \mathbf{u}^{(2)}$ and use (H1), (H2) to write

$$\begin{aligned} & \frac{1}{2} \frac{d}{dt} \|\mathbf{u}^{(1)}(t) - \mathbf{u}^{(2)}(t)\|_{X_0}^2 + \alpha \left\| J \left(\mathbf{u}^{(1)}(t) - \mathbf{u}^{(2)}(t) \right) \right\|_{X_0}^2 \\ & \leq L \|\mathbf{u}^{(1)}(t) - \mathbf{u}^{(2)}(t)\|_{X_0} \|J\mathbf{u}^{(2)}(t)\|_{X_0} \\ & \quad \left\| J \left(\mathbf{u}^{(1)}(t) - \mathbf{u}^{(2)}(t) \right) \right\|_{X_0} \\ & \leq \frac{1}{\alpha} L^2 \|\mathbf{u}^{(1)}(t) - \mathbf{u}^{(2)}(t)\|_{X_0}^2 \|J\mathbf{u}^{(2)}(t)\|_{X_0}^2 \\ & \quad + \frac{\alpha}{4} \left\| J \left(\mathbf{u}^{(1)}(t) - \mathbf{u}^{(2)}(t) \right) \right\|_{X_0}^2. \end{aligned}$$

(In the last step the inequality $ab \leq a^2/4\epsilon^2 + \epsilon^2 b^2$ has been used, with $\epsilon^2 = \alpha/4$.) Therefore,

$$\begin{aligned} & \frac{d}{dt} \|\mathbf{u}^{(1)}(t) - \mathbf{u}^{(2)}(t)\|_{X_0}^2 \\ & \leq \frac{2}{\alpha} L^2 \|\mathbf{u}^{(1)}(t) - \mathbf{u}^{(2)}(t)\|_{X_0}^2 \|J\mathbf{u}^{(2)}(t)\|_{X_0}^2. \end{aligned}$$

Finally, Gronwall's lemma leads to

$$\begin{aligned} & \|\mathbf{u}^{(1)}(t) - \mathbf{u}^{(2)}(t)\|_{X_0}^2 \\ & \leq \|\mathbf{u}^{(1)}(0) - \mathbf{u}^{(2)}(0)\|_{X_0}^2 \exp \left(C \int_0^t \|J\mathbf{u}^{(2)}(s)\|_{X_0}^2 ds \right), \end{aligned} \quad (2.16)$$

with $C = \frac{2}{\alpha} L^2$ and since $\mathbf{u}^{(1)}(0) = \mathbf{u}^{(2)}(0)$ then uniqueness is proved.

Continuous Dependence on Initial Data

Since \mathbf{u} is bounded on $\overline{Q_T}$, then $J\mathbf{u}$ is bounded and hypothesis (H1) on D implies

$$\begin{aligned} & \int_0^t \|J\mathbf{u}(\cdot, s)\|_{X_0}^2 ds \\ & \leq \int_0^T \|J\mathbf{u}(\cdot, s)\|_{X_0}^2 ds = \frac{1}{\alpha} \int_0^T \alpha \|J\mathbf{u}(\cdot, s)\|_{X_0}^2 ds \\ & \leq \frac{1}{\alpha} \left| \int_0^T \int_{\Omega} \nabla \mathbf{u}(\mathbf{x}, t) D(\mathbf{u}(\mathbf{x}, t)) \nabla \mathbf{u}(\mathbf{x}, t)^T d\Omega \right| ds \\ & = \frac{1}{\alpha} \left| \int_0^T \int_{\Omega} \mathbf{u}(\mathbf{x}, t)^T \mathbf{u}_t(\mathbf{x}, t) d\Omega \right| ds \\ & \leq \frac{1}{\alpha} \int_0^T \|\mathbf{u}(\cdot, s)\|_{X_0} \|\mathbf{u}_t(\cdot, s)\|_{X_0} ds \\ & \leq \frac{1}{\alpha} \|\mathbf{u}\|_{L^2(0, T, X_1)} \|\mathbf{u}_t\|_{L^2(0, T, (X_1)')}. \end{aligned}$$

Now, let $\epsilon > 0$ and take

$$\delta := \epsilon \exp \left(-\frac{C}{\alpha} \|\mathbf{u}\|_{L^2(0, T, X_1)} \|\mathbf{u}_t\|_{L^2(0, T, (X_1)')} \right).$$

If $\|\mathbf{u}^{(1)}(0) - \mathbf{u}^{(2)}(0)\|_{X_0} < \delta$ and using (2.16) then

$$\|\mathbf{u}^{(1)}(t) - \mathbf{u}^{(2)}(t)\|_{X_0} < \epsilon,$$

for all $t \in [0, T]$. This proves the continuous dependence on the initial data. \square

Extremum Principle

Well-posedness results are finished off with the following extremum principle.

Theorem 2 *Let us assume that in (2.2) $\mathbf{u}_0 = (u_0, v_0)^T \in L^\infty(\Omega) \times L^\infty(\Omega)$ and define:*

$$a_1 = \text{ess inf } u_0, \quad a_2 = \text{ess inf } v_0, \\ b_1 = \|u_0\|_{L^\infty(\Omega)}, \quad b_2 = \|v_0\|_{L^\infty(\Omega)}.$$

Let $\mathbf{u} = (u, v)^T$ be the weak solution of (2.1)–(2.3). Then for all $(\mathbf{x}, t) \in Q_T$

$$a_1 \leq u(\mathbf{x}, t) \leq b_1, \quad a_2 \leq v(\mathbf{x}, t) \leq b_2.$$

Proof Note that the same argument as that of the linear problem (2.6) can be adapted to this nonlinear case straightforwardly, by taking, in the case of the maximum principle, $w_1 = (u - b_1)_+$, $w_2 = (v - b_2)_+$ in the weak formulation (2.5) and, in the case of the minimum principle, $w_1 = (u - a_1)_-$, $w_2 = (v - a_2)_-$. \square

Remark 1 In Gilboa et al. [15], a nonlinear complex diffusion problem with diffusion coefficient of the form

$$c = c(v) = \frac{e^{i\theta}}{1 + \left(\frac{v}{\kappa\theta}\right)^2}, \quad (2.17)$$

is considered. In (2.17) the image is represented by a complex function $u + iv$, κ is a threshold parameter and θ is a phase angle parameter. In the corresponding cross-diffusion formulation (2.1) for $\mathbf{u} = (u, v)^T$, the coefficient matrix is

$$D(u, v) = g(v) \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}, \\ g(v) = \frac{1}{1 + \left(\frac{v}{\kappa\theta}\right)^2}. \quad (2.18)$$

Thus, for $\xi \in \mathbb{R}^2$,

$$\xi^T D(u, v) \xi = (g(v) \cos \theta) |\xi|^2.$$

The function g in (2.18) is decreasing for $v \geq 0$ and satisfies $g(0) = 1$, $\lim_{s \rightarrow +\infty} g(s) = 0$. Consequently, D in (2.18) would not satisfy (H1) for $v \geq 0$. In addition to assuming $\theta \in (0, \pi)$ (in order to have $\cos \theta > 0$), two strategies to overcome this drawback are suggested.

- The first one is to replace $g(v)$ in (2.18) by $g(M(v))$, where $M(\cdot)$ is a cut-off operator

$$M(v)(\mathbf{x}, t) = \min_{(\mathbf{x}, t) \in Q_T} \{v(\mathbf{x}, t), M\}, \quad (2.19)$$

with M a sufficiently large constant. The same approach can be generalized for the cross-diffusion matrix operator D .

- A second strategy is to replace $g(v)$ in (2.18) by $g(|w_\sigma|)$ where w_σ is the second component of the matrix convolution $\mathbf{v}_\sigma = K_\sigma * \mathbf{u}$, $K_\sigma := K(\cdot, \sigma) = (k_{ij}(\cdot, \sigma))_{i,j=1,2}$ is the matrix such that (Araújo et al. [2])

$$\widehat{K}_\sigma(\xi) = (\widehat{k}_{ij}(\cdot, \sigma))_{i,j=1,2} = e^{-|\xi|^2 \sigma d}, \quad \xi \in \mathbb{R}^2,$$

where

$$d = d_\theta = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}.$$

(The matrix convolution $K_\sigma * \mathbf{u}$ is defined as the vector

$$\begin{pmatrix} k_{11} * \tilde{u} + k_{12} * \tilde{v} \\ k_{21} * \tilde{u} + k_{22} * \tilde{v} \end{pmatrix}$$

where $*$ denotes the usual convolution operator in \mathbb{R}^2 and $\tilde{\mathbf{u}} = (\tilde{u}, \tilde{v})^T$ is a continuous extension of \mathbf{u} in \mathbb{R}^2 .)

We observe that the weak formulation (2.5) with the corresponding modified matrix $D(u, v) = g(|w_\sigma|)d_\theta$ satisfies the conclusions of Theorem 1 by adapting the proof as follows (see Catté et al. [9]): Let $U, V \in W(0, T) \cap L^\infty(0, T, H^0(\Omega))$ such that

$$\|U\|_{L^\infty(0, T, H^0(\Omega))} \leq \|u_0\|_0, \\ \|V\|_{L^\infty(0, T, H^0(\Omega))} \leq \|v_0\|_0.$$

Since $U, V \in L^\infty(0, T, H^0(\Omega))$ and g as well as each entry of K_σ are C^∞ then $g(|w_\sigma|) \in L^\infty(0, T, C^\infty(\Omega))$. Thus, since g is decreasing, there is $C > 0$, which only depends on g , K_σ and $\|u_0\|_0, \|v_0\|_0$ such that

$$g(|w_\sigma|) \leq C$$

almost everywhere in Q_T . Thus, the corresponding matrix $D(u, v) = g(|w_\sigma|)d_\theta$ satisfies (H1) for almost any $(\mathbf{x}, t) \in Q_T$. With this modification, the rest of Theorem 1 is proved in the same way.

The previous argument can be generalized to a general cross-diffusion problem (2.1) with cross-diffusion matrices of the form

$$D(\mathbf{u}) = D(u, v) = g(|w_\sigma|)d, \quad (2.20)$$

where

- (i) w_σ is the second component of $\mathbf{v}_\sigma = K_\sigma * \mathbf{u}$ with K_σ satisfying

$$\widehat{K}_\sigma(\xi) = e^{-|\xi|^2 \sigma d}, \quad \xi \in \mathbb{R}^2,$$

for some positive definite matrix d and,

- (ii) $g : [0, +\infty) \rightarrow (0, +\infty)$ is a smooth, decreasing function with $g(0) = 1$ and $\lim_{s \rightarrow +\infty} g(s) = 0$.

2.2 Scale-Space Properties

For $t \geq 0$ let us define the scale-space operator

$$T_t : \mathbf{u}_0 \mapsto T_t(\mathbf{u}_0) := \mathbf{u}(t) = \mathbf{u}(\cdot, t), \quad (2.21)$$

such that $\mathbf{u}(t)$ is the unique weak solution at time t of (2.1)–(2.3) with initial data (2.2) given by \mathbf{u}_0 . Some properties of (2.21) will be here analysed. More precisely additional hypotheses on D in (2.1) allow (2.21) to satisfy grey-level shift, reverse contrast, average grey and translational invariances. In what follows we assume that (H1)–(H3) hold.

2.2.1 Grey-Level Shift Invariance

Lemma 1 *Let us assume that D in (2.1) additionally satisfies*

$$D(\mathbf{u}(\mathbf{x}, t) + \mathbf{C}) = D(\mathbf{u}(\mathbf{x}, t)), \quad (2.22)$$

for all $(\mathbf{x}, t) \in Q_T$, $\mathbf{u}(\cdot, t) \in X_1$ and $\mathbf{C} = (C_1, C_2)^T \in \mathbb{R}^2$. Then

$$T_t(\mathbf{0}) = \mathbf{0}, \quad T_t(\mathbf{u}_0 + \mathbf{C}) = T_t(\mathbf{u}_0) + \mathbf{C}, \quad t \geq 0. \quad (2.23)$$

Proof The main argument for the proof is the uniqueness of solution of (2.1)–(2.3). Note first that $\mathbf{u} = \mathbf{0}$ is a solution with $\mathbf{u}_0 = \mathbf{0}$ and consequently it is clear that $T_t(\mathbf{0}) = \mathbf{0}$. On the other hand, because of (2.22) we have that

$$\mathbf{w}(t) = T_t(\mathbf{u}_0) + \mathbf{C}, \quad t \geq 0,$$

satisfies (2.1) with initial condition $\mathbf{u}_0 + \mathbf{C}$ and therefore, by uniqueness, it must coincide with $T_t(\mathbf{u}_0 + \mathbf{C})$. \square

Remark 2 From Araújo et al. [2], we know that the kernel matrices K_σ satisfying (2.21) are mass preserving, that is $K_\sigma * \mathbf{C} = \mathbf{C}$, $\mathbf{C} \in \mathbb{R}^2$ and therefore

$$K_\sigma * (\mathbf{u} + \mathbf{C}) = (K_\sigma * \mathbf{u}) + \mathbf{C}.$$

This implies that cross-diffusion coefficient matrices (2.20) satisfy (2.23) but only for constants $\mathbf{C} = (C_1, 0)^T$, $C_1 \in \mathbb{R}$. If the first component of $\mathbf{u}(t)$ represents the grey-level values of the filtered image at time t then this weaker version of (2.23) can be interpreted as shift invariance of the grey values.

2.2.2 Reverse Contrast Invariance

Lemma 2 *Let us assume that D in (2.1) additionally satisfies*

$$D(-\mathbf{u}(\mathbf{x}, t)) = D(\mathbf{u}(\mathbf{x}, t)), \quad (2.24)$$

for all $(\mathbf{x}, t) \in Q_T$, $\mathbf{u}(\cdot, t) \in X_1$. Then

$$T_t(-\mathbf{u}_0) = -T_t(\mathbf{u}_0), \quad t \geq 0. \quad (2.25)$$

Proof By (2.24) the functions $\mathbf{w}_1(t) = T_t(-\mathbf{u}_0)$ and $\mathbf{w}_2(t) = -T_t(-\mathbf{u}_0)$ satisfy (2.5) with the same initial data $-\mathbf{u}_0$. Therefore, by uniqueness, $\mathbf{w}_1 = \mathbf{w}_2$ and (2.25) holds. \square

Remark 3 Matrices D of the form (2.20) satisfy (2.24), and therefore the corresponding operator (2.21) satisfies (2.25).

2.2.3 Average Grey Invariance

For $f \in L^2(\Omega)$ we define

$$m(f) = \frac{1}{A(\Omega)} \int_{\Omega} f(\mathbf{x}) d\Omega,$$

where $A(\Omega)$ stands for the area of Ω .

Lemma 3 *For $\mathbf{u} = (u, v)^T$ let $\mathbf{M}(\mathbf{u}) = (m(u), m(v))^T$. Then*

$$\mathbf{M}(T_t(\mathbf{u}_0)) = \mathbf{M}(\mathbf{u}_0), \quad t \geq 0. \quad (2.26)$$

Proof We consider the vector function

$$\begin{aligned} \mathbf{G}(t) &= (G_1(t), G_2(t))^T, \\ G_1(t) &= \int_{\Omega} u(\mathbf{x}, t) d\Omega, \quad G_2(t) = \int_{\Omega} v(\mathbf{x}, t) d\Omega, \quad t \geq 0, \end{aligned}$$

where $\mathbf{u} = (u, v)^T = T_t(\mathbf{u}_0)$. As in Weickert [27], we have, for $i = 1, 2$,

$$|G_i(t) - G_i(0)| \leq A(\Omega)^{1/2} \|\mathbf{u}(t) - \mathbf{u}(0)\|_{X_0}.$$

Since at least $\mathbf{u} \in C(0, T, X_0)$ then \mathbf{G} is continuous at $t = 0$. On the other hand, divergence theorem and the boundary conditions (2.3) imply that, for $i = 1, 2$

$$\begin{aligned} \frac{d}{dt} G_i(t) &= \int_{\Omega} \operatorname{div}(D_{i1}(\mathbf{u}) \nabla u + D_{i2}(\mathbf{u}) \nabla v) d\Omega \\ &= \int_{\partial\Omega} \langle D_{i1}(\mathbf{u}) \nabla u + D_{i2}(\mathbf{u}) \nabla v, \mathbf{n} \rangle d\Gamma = 0. \end{aligned}$$

Then $G_i(t)$ is constant for all $t \geq 0$. Thus, the quantity $\mathbf{M}(\mathbf{u}_0) = (m(u_0), m(v_0))^T$ is preserved by cross-diffusion. \square

Remark 4 Actually, each component $G_i(t)$, $i = 1, 2$ is preserved. This may be used to establish a suitable definition of average grey level in this formulation, using these two quantities, and its preservation by cross-diffusion; we refer Araújo et al. [2] for a discussion about this question.

2.2.4 Translational Invariance

Let us define the translational operator $\tau_{\mathbf{h}}$ as

$$\begin{aligned}\tau_{\mathbf{h}}\mathbf{u}(\mathbf{x}) &= (u(\mathbf{x} + \mathbf{h}), v(\mathbf{x} + \mathbf{h}))^T, \\ \mathbf{u} &= (u, v)^T \in X_0, \quad \mathbf{x}, \mathbf{h} \in \mathbb{R}^2.\end{aligned}$$

Lemma 4 We assume that D in (2.1) additionally satisfies

$$D(\tau_{\mathbf{h}}\mathbf{u}(\mathbf{x}, t)) = D(\mathbf{u}(\mathbf{x}, t)), \quad (2.27)$$

for all $(\mathbf{x}, t) \in Q_T$, $\mathbf{u}(\cdot, t) \in X_1$. Then

$$T_t(\tau_{\mathbf{h}}\mathbf{u}_0) = \tau_{\mathbf{h}}(T_t(\mathbf{u}_0)), \quad t \geq 0. \quad (2.28)$$

Proof Due to (2.27), the functions $\mathbf{w}_1(t) = T_t(\tau_{\mathbf{h}}\mathbf{u}_0)$ and $\mathbf{w}_2(t) = \tau_{\mathbf{h}}T_t(\mathbf{u}_0)$ are solutions of (2.5) with the same initial data $\tau_{\mathbf{h}}\mathbf{u}_0$. Therefore, by uniqueness, (2.28) holds. \square

Remark 5 Matrices D of the form (2.20) satisfy (2.27), and therefore the corresponding operator (2.21) satisfies (2.28).

2.3 Lyapunov Functions and Behaviour at Infinity

The previous study is finished off by analysing the existence of Lyapunov functionals and the behaviour of the solution when $t \rightarrow +\infty$. As far as the first question is concerned, we have the following result.

Lemma 5 Let $\mathbf{u} = (u, v)^T$ be the unique weak solution of (2.1)–(2.3) and let us consider the functional

$$V(t) = \Phi(\mathbf{u}(t)) := \frac{1}{2} \int_{\Omega} (u(\mathbf{x}, t)^2 + v(\mathbf{x}, t)^2) d\Omega. \quad (2.29)$$

Then V defines a Lyapunov function for (2.1)–(2.3).

Proof Note first that from the weak formulation (2.5) with $\mathbf{w} = \mathbf{u}$ we obtain

$$\frac{d}{dt} V(t) + \int_{\Omega} \operatorname{tr} \left((J\mathbf{u})^T D(\mathbf{u}(\mathbf{x}, t)) (J\mathbf{u}) \right) d\Omega = 0,$$

which, due to (H1) and (2.5), implies

$$\frac{d}{dt} V(t) \leq 0, \quad t \geq 0.$$

Note also that since $\tilde{r}(z) = z^2/2$ is convex and (2.26) holds, then Jensen inequality implies that

$$\begin{aligned}\Phi(M\mathbf{u}_0) &= \int_{\Omega} \frac{(m(u_0))^2 + (m(v_0))^2}{2} d\Omega \\ &= \int_{\Omega} \tilde{r}(m(u_0)) + \tilde{r}(m(v_0)) d\Omega \\ &= \int_{\Omega} \tilde{r}(m(u(t))) + \tilde{r}(m(v(t))) d\Omega \\ &\leq \int_{\Omega} \left(\frac{1}{A(\Omega)} \int_{\Omega} \tilde{r}(u(\mathbf{x}, t)) d\Omega \right. \\ &\quad \left. + \frac{1}{A(\Omega)} \int_{\Omega} \tilde{r}(v(\mathbf{x}, t)) d\Omega \right) d\Omega \\ &= \int_{\Omega} (\tilde{r}(u(\mathbf{x}, t)) + \tilde{r}(v(\mathbf{x}, t))) d\Omega = \Phi(\mathbf{u}(t)).\end{aligned}$$

Therefore, (2.29) is a Lyapunov functional. \square

The search for more Lyapunov functions will make use of convex functions.

Lemma 6 Let a_j, b_j , $j = 1, 2$ be defined in Theorem 2, $I = I(\mathbf{a}, \mathbf{b}) = (a_1, b_1) \times (a_2, b_2)$ and let us assume that r is a $C^2(I)$ convex function satisfying

$$\nabla^2 r(\mathbf{u}(\mathbf{x})) D(\mathbf{u}(\mathbf{x})) = D(\mathbf{u}(\mathbf{x})) \nabla^2 r(\mathbf{u}(\mathbf{x})), \quad (2.30)$$

where \mathbf{u} is the weak solution of (2.1)–(2.3) with $u_0, v_0 \geq 0$ and $\nabla^2 r(u, v)$ stands for the Hessian of r . Then

$$V_r(t) = \Phi_r(\mathbf{u}(t)) = \int_{\Omega} r(u(t), v(t)) d\Omega, \quad (2.31)$$

is a Lyapunov function for (2.1)–(2.3).

Proof Observe that using divergence theorem, boundary conditions (2.3) and after some computations we have

$$\begin{aligned}V_r'(t) &= \int_{\Omega} (r_u(u, v)u_t + r_v(u, v)v_t) d\Omega \\ &= \int_{\Omega} (r_u(u, v) \operatorname{div} (D_{11}(\mathbf{u})\nabla u + D_{12}(\mathbf{u})\nabla v) \\ &\quad + r_v(u, v) \operatorname{div} (D_{21}(\mathbf{u})\nabla u + D_{22}(\mathbf{u})\nabla v)) d\Omega \\ &= - \int_{\Omega} \left(\left\langle \nabla^2 r(u, v) \begin{pmatrix} u_x \\ v_x \end{pmatrix}, D(\mathbf{u}) \begin{pmatrix} u_x \\ v_x \end{pmatrix} \right\rangle \right. \\ &\quad \left. + \left\langle \nabla^2 r(u, v) \begin{pmatrix} u_y \\ v_y \end{pmatrix}, D(\mathbf{u}) \begin{pmatrix} u_y \\ v_y \end{pmatrix} \right\rangle \right) d\Omega.\end{aligned}$$

Since r is convex, then $\nabla^2 r(u, v)$ is positive semi-definite. Thus, due to (2.30), $\nabla^2 r(u, v) D(\mathbf{u})$ is positive semi-definite and therefore $V_r'(t) \leq 0$, $t \geq 0$. Similarly, the application of a generalized version of Jensen inequality, Zabandan and Kiliman [29], and the convexity of r imply

$$r(\mathbf{M}(\mathbf{u})) \leq m(r(\mathbf{u})).$$

This and (2.26) lead to

$$\begin{aligned} \Phi_r(\mathbf{M}(\mathbf{u}_0)) &= \int_{\Omega} r(\mathbf{M}(\mathbf{u}_0)) d\Omega = \int_{\Omega} r(\mathbf{M}(\mathbf{u}(t))) d\Omega \\ &\leq \int_{\Omega} m(r(\mathbf{u}(t))) d\Omega \\ &= \int_{\Omega} \frac{1}{A(\Omega)} \int_{\Omega} r(\mathbf{u}(t)) d\mathbf{x} d\Omega = \int_{\Omega} r(\mathbf{u}(t)) d\mathbf{x} \\ &= \Phi_r(\mathbf{u}(t)), \end{aligned}$$

and (2.31) is a Lyapunov functional. \square

Remark 6 The choice $r(x, y) = \frac{x^2+y^2}{2}$ leads to the Lyapunov functional (2.29). More generally, if $p \geq 2$ then taking

$$r(x, y) = |x|^p + |y|^p,$$

implies that the $L^p \times L^p$ norm

$$\|\mathbf{u}\|_{L^p \times L^p} = (\|u\|_{L^p}^p + \|v\|_{L^p}^p)^{1/p}, \quad \mathbf{u} = (u, v)^T,$$

is a Lyapunov functional, see Weickert [27].

As far as the behaviour at infinity of the solution of (2.1)–(2.3) is concerned, the arguments in Weickert [27] can also be adapted here.

Lemma 7 *Let $\mathbf{u}(t)$, $t \geq 0$ be the weak solution of (2.1)–(2.3) and let us consider $\mathbf{w} = \mathbf{u} - \mathbf{M}(\mathbf{u}_0)$, where \mathbf{M} is given by Lemma 3. If (2.23) holds then*

$$\lim_{t \rightarrow \infty} \|\mathbf{w}(t)\|_{X_0} = 0. \quad (2.32)$$

Proof Since (2.23) holds then \mathbf{w} satisfies the diffusion equation of (2.1). By using the weak formulation (2.5), divergence theorem and the boundary conditions (2.3), we have

$$\frac{1}{2} \frac{d}{dt} \int_{\Omega} (w_1^2 + w_2^2) d\Omega = - \int_{\Omega} \text{tr} \left((J\mathbf{w})^T D(J\mathbf{w}) \right) d\Omega.$$

Now, (H1) and (2.5) imply

$$\text{tr} \left((J\mathbf{w})^T D(J\mathbf{w}) \right) \geq \alpha \|J\mathbf{w}\|_{X_0}^2.$$

Therefore,

$$\frac{d}{dt} \|\mathbf{w}\|_{X_0}^2 \leq -2\alpha \|J\mathbf{w}\|_{X_0}^2.$$

Note now that if we apply the Poincaré inequality to each w_i , $i = 1, 2$, then there is $C_0 > 0$ such that

$$\|\mathbf{w}\|_{X_0}^2 \leq C_0 \|J\mathbf{w}\|_{X_0}^2.$$

This implies that

$$\frac{d}{dt} \|\mathbf{w}\|_{X_0}^2 \leq -2\alpha C_0 \|\mathbf{w}\|_{X_0}^2.$$

By Gronwall's lemma

$$\|\mathbf{w}(t)\|_{X_0}^2 \leq e^{-2\alpha C_0 t} \|\mathbf{w}(0)\|_{X_0}^2$$

and (2.32) holds. \square

Remark 7 Since we are assuming strong ellipticity (H1) in the model, the asymptotic behaviour (2.32) is expected. In particular, we conclude that the model does not preserve the discontinuities of the initial conditions, which is a serious limitation in the computer vision context. Assumption (H1) could be relaxed by considering degenerate elliptic cross-diffusion operators D , that is, substituting (2.4) by

$$\xi^T D(\mathbf{u}(\mathbf{x}, t)) \xi \geq 0, \quad (\mathbf{x}, t) \in \overline{Q_T}.$$

The analysis of such models is beyond the scope of this work.

3 Numerical Experiments

The performance of (2.1)–(2.3) in filtering problems is numerically illustrated in this section.

3.1 The Numerical Procedure

In order to implement (2.1)–(2.3), some details are described. The first point concerns the choice of the cross-diffusion matrix D . We have considered to this end the results on linear cross-diffusion shown in the companion paper Araújo et al. [2] and the complex diffusion approach, Gilboa et al. [15], see Remark 1. According to them, matrices of the form

$$D(u, v) = g(|w|)d, \quad d = \begin{pmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{pmatrix}, \quad (3.1)$$

were used for the experiments, with

$$g(v) = \frac{1}{1 + \left(\frac{v}{\kappa}\right)^2}, \quad (3.2)$$

with κ a threshold parameter, see Gilboa et al. [15] and d a positive definite matrix. Both possibilities $w = M(v)$ in (2.19) and $w = w_{\sigma}$ in (2.20) have been implemented. The form (3.1), (3.2) takes into account (2.17) by using the extended version of the small theta approximation, see Araújo et al. [2] [which justifies the presence of d_{12} in (3.2)] as well as the classical nonlinear diffusion approach with the form of g , see Aubert and Kornprobst, [6], Catté et al. [9], Perona and Malik [25].

The guidance about the choice of the matrix d was also based on linear cross-diffusion. Thus if $s = (d_{22} - d_{11})^2 + 4d_{12}d_{21}$, three types of matrices d (for which $s > 0$, $s < 0$ and $s = 0$) have been considered. (The parameter s determines if the eigenvalues of d are real or complex, see Araújo et al. [2].) The specific examples of d for the experiments are given in Sect. 3.2.

A second question on the implementation concerns the choice of a numerical scheme to approximate (2.1)–(2.3). Thus, the explicit numerical method introduced and analysed in Araújo et al. [5] for the complex diffusion case has been adapted here. The method is now briefly described. By using the notation of Sect. 2, \overline{Q}_T is first discretized as follows. We define a uniform grid on $\overline{\Omega} = [l_1, r_1] \times [l_2, r_2]$ with mesh step size $h > 0$ as

$$\begin{aligned} \overline{\Omega}_h = \{\mathbf{x}_{ij} = (x_i, y_j) \in \overline{\Omega} : x_i = l_1 + ih, y_j = l_2 + jh, \\ i = 0, \dots, N_1 - 1, j = 0, \dots, N_2 - 1\}, \end{aligned} \quad (3.3)$$

for integers $N_1, N_2 > 1$ such that $hN_i = r_i - l_i, i = 1, 2$. As far as the time discretization is concerned, fixed $T > 0$, for an integer $M \geq 1$ and $\Delta t > 0$ such that $M\Delta t = T$, the interval $[0, T]$ is partitioned in

$$0 = t^0 < t^1 < \dots < t^{M-1} < t^M = T, \quad (3.4)$$

with $t^{m+1} = t^m + \Delta t, m = 0, \dots, M - 1$. The resulting discretization of \overline{Q}_T with (3.3) and (3.4) is denoted by $\overline{Q}_h^{\Delta t} = Q_h^{\Delta t} \cup \Gamma_h^{\Delta t}$, where $Q_h^{\Delta t}, \Gamma_h^{\Delta t}$ stand for the interior and boundary meshes, respectively.

If $M_{N_1 \times N_2}(\mathbb{R})$ denotes the space of $N_1 \times N_2$ real matrices then let us consider some initial distribution $U^0, V^0 : \overline{\Omega}_h \rightarrow M_{N_1 \times N_2}(\mathbb{R})$ with $U^0 = (U_{ij}^0)_{i=1, j=1}^{N_1, N_2}, V^0 = (V_{ij}^0)_{i=1, j=1}^{N_1, N_2}$. From U^0, V^0 and for $m = 0, \dots, M - 1$ the approximate image at time t^{m+1} is defined as $(U^{m+1}, V^{m+1})^T$, where $U^{m+1} = (U_{ij}^{m+1})_{i=1, j=1}^{N_1, N_2}, V^{m+1} = (V_{ij}^{m+1})_{i=1, j=1}^{N_1, N_2} : \overline{\Omega}_h \rightarrow M_{N_1 \times N_2}(\mathbb{R})$ satisfy the system (in vector form)

$$\begin{aligned} \frac{U^{m+1} - U^m}{\Delta t} &= \nabla_h \cdot (g(V^m)(d_{11} \nabla_h U^m + d_{12} \nabla_h V^m)), \\ \frac{V^{m+1} - V^m}{\Delta t} &= \nabla_h \cdot (g(V^m)(d_{21} \nabla_h U^m + d_{22} \nabla_h V^m)), \end{aligned} \quad (3.5)$$

where g and d are given by (3.1), (3.2). In (3.5), ∇_h is the discrete operator such that if $W = (W_{ij})_{i=1, j=1}^{N_1, N_2}$ then

$$(\nabla_h W)_{ij} = \left(\frac{W_{i+1, j} - W_{i-1, j}}{2h}, \frac{W_{i, j+1} - W_{i, j-1}}{2h} \right), \\ i = 1, \dots, N_1 - 1, j = 1, \dots, N_2 - 1.$$

The scheme (3.5) is completed with the discretization of the Neumann boundary conditions (2.2) by using ∇_h .

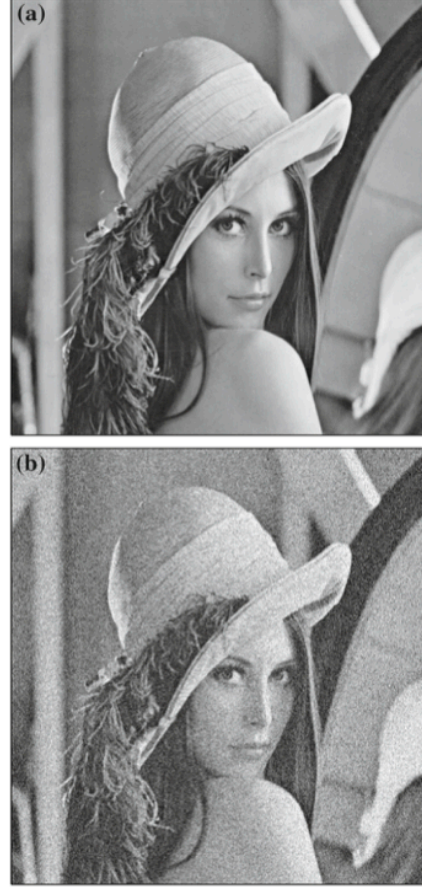


Fig. 1 **a** Original image S of Lena and **b** noisy image of Lena with Gaussian noise of $\sigma' = 30$

We note that for the complex diffusion case, a stability condition for (3.5) with diffusion coefficient (2.17) was derived in Araújo et al. [5] and Bernardes et al. [7],

$$\Delta t := \max_{0 \leq m \leq M-1} \Delta t^m \leq \frac{\cos \theta}{4} \left(1 + \frac{\min_m (V^m)^2}{\kappa^2 \theta^2} \right). \quad (3.6)$$

Condition (3.6) was taken into account in the numerical experiments below, where $h = 1$ and $\Delta t = 0.05$ (with $\kappa = 10$) were used.

3.2 Numerical Results

In this section, several numerical experiments have been performed according to the following steps. Assume that the discrete values $S_{ij} = S(\mathbf{x}_{ij}), \mathbf{x}_{ij} \in \overline{\Omega}_h$ of some real-valued function $S : \overline{\Omega} \rightarrow \mathbb{R}$ represent an original image on $\overline{\Omega}_h$. From S , some noise of Gaussian type with zero mean and standard deviation σ' at pixel \mathbf{x}_{ij} is added to S_{ij} . This is rep-

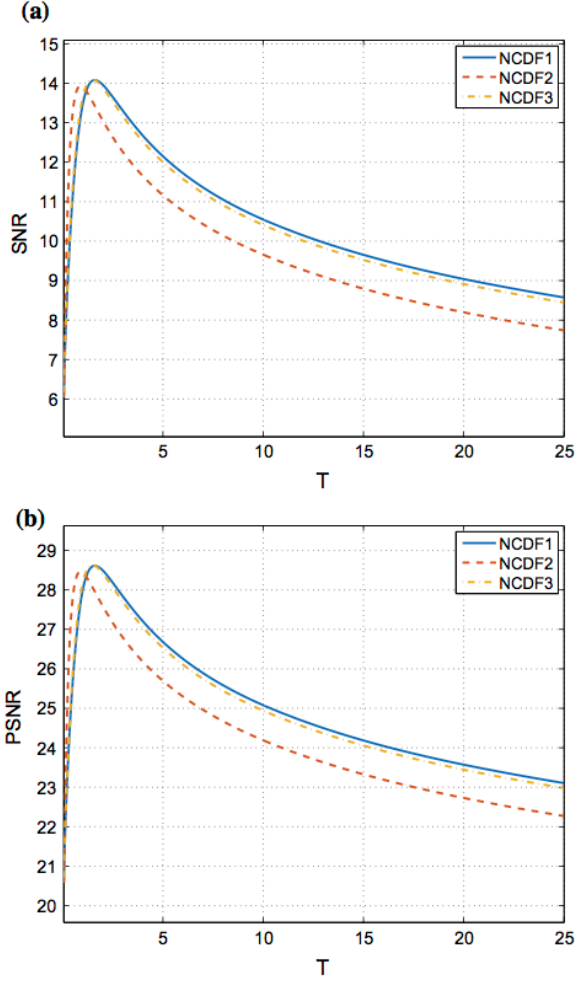


Fig. 2 SNR **a** and **b** PSNR versus time: NCDF1 (solid line), NCDF2 (dashed line) and NCDF3 (dashed-dotted line)

resented by a matrix $N(\sigma') = (N_{ij}(\sigma'))_{ij}$ and generates the initial noisy image values

$$u^0 = (u_{ij}^0)_{ij}, \quad u_{ij}^0 = S_{ij} + N_{ij}(\sigma'), \\ i = 1, \dots, N_1, j = 1, \dots, N_2.$$

Then the explicit method (3.5) with initial distribution $U_{ij}^0 = u_{ij}^0$, $V_{ij}^0 = 0$, $i = 1, \dots, N_1$, $j = 1, \dots, N_2$ and D from (3.1), (3.2) is run; the corresponding numerical solutions U^m, V^m , $m = 1, \dots, M$ are monitored in such a way that U^m approximates the original signal S at time t^m . In order to measure the quality of restoration, three metrics are used:

- Signal-to-Noise-Ratio (SNR):

$$SNR(S, U^m) = 10 \log_{10} \left(\frac{\text{var}(S)}{\text{var}(U^m - S)} \right), \quad (3.7)$$

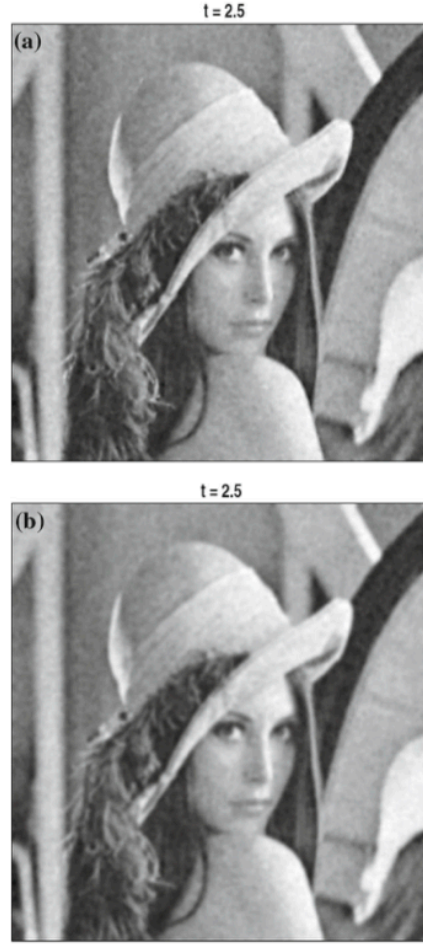


Fig. 3 First component of solution of (3.5) at time $t = 2.5$ with **a** NCDF1 and **b** NCDF2

where the variance (var) of an image U is defined by

$$\text{var}(U) = \frac{1}{N_1 N_2} \|U - \bar{U}\|_F^2,$$

$\|\cdot\|_F$ stands for the Frobenius norm, and \bar{U} is a uniform image with intensities equal to the mean value of the intensities of U .

- Peak Signal-to-Noise-Ratio (PSNR):

$$PSNR(S, U^m) = 20 \log_{10} \left(\frac{255}{RMSE(S, U^m)} \right), \quad (3.8)$$

where the root-mean-square-error (RMSE) is defined as

$$RMSE(S, U^m) = \frac{1}{\sqrt{N_1 N_2}} \|S - U^m\|_F^2;$$

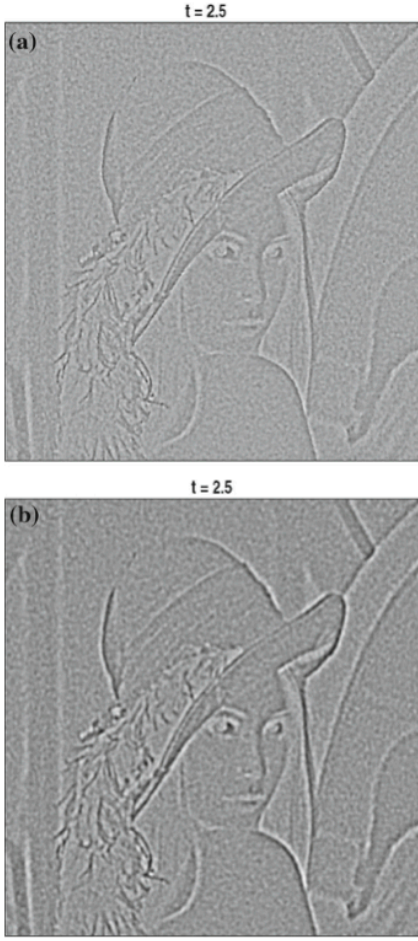


Fig. 4 Second component of solution of (3.5) at time $t = 2.5$ with a NCDF1 and b NCDF2

- The no-reference perceptual blur metric (NPB) proposed by Cr  t  -Roffet et al. [10]. This is based on evaluating the blur annoyance of the image by comparing the variations between neighbouring pixels before and after the application of a low-pass filter. The estimation ranges from 0 (the best quality blur perception) to 1 (the worst one).

The following numerical results illustrate the behaviour of (2.1)–(2.3) according to the choice of the matrix d in (3.1) and the implementation of (3.2). The experiments are concerned with the filtering of a noisy image of Lena (Fig. 1) and a first group makes use of the matrices

- (i) NCDF1 ($s > 0$): $d_{11} = 1, d_{12} = 0.025, d_{21} = 1, d_{22} = 1$.
- (ii) NCDF2 ($s < 0$): $d_{11} = 1, d_{12} = -0.025, d_{21} = 0.025, d_{22} = 1$.

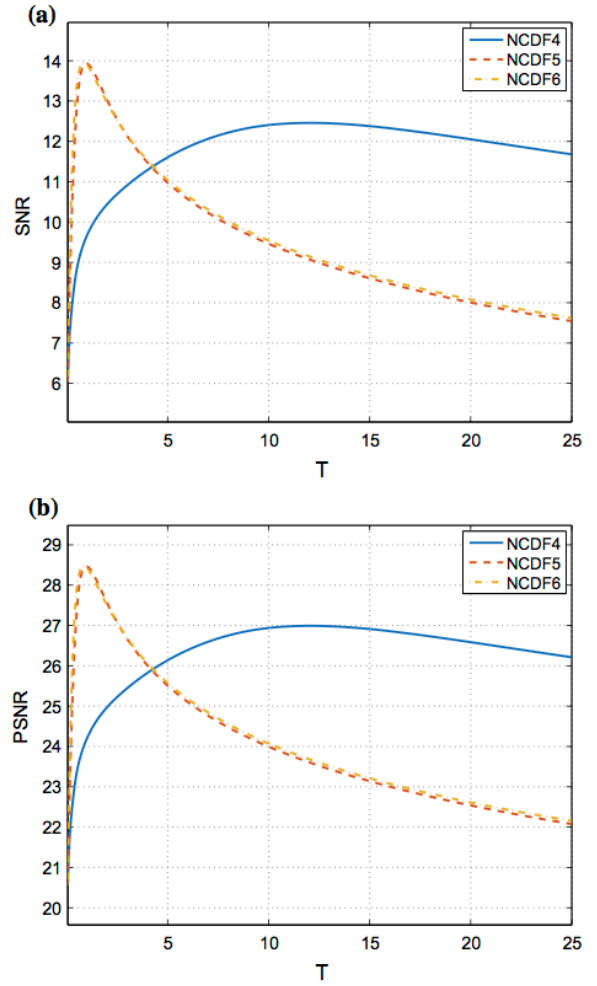


Fig. 5 SNR a and b PSNR versus time: NCDF4 (solid line), NCDF5 (dashed line) and NCDF6 (dashed-dotted line)

- (iii) NCDF3 ($s = 0$): $d_{11} = 1, d_{12} = -0.025, d_{21} = 1, d_{22} = 1.1$.

These models were taken to study three points of the filtering: the restoration process from the first component of the numerical solution of (3.5), the behaviour of the edges from the second component and the quality of filtering from the computation of the evolution of the three metrics.

The numerical experiments in Fig. 2 show the time evolution of the SNR and PSNR parameters given by the models NCDF1-3. For the three models, the metrics attain a maximum value from which the quality of restoration is decreasing. The main difference appears in the time at which the maximum holds, being longer in the case of NCDF1 (corresponding to $s > 0$) and NCDF3 (for which $s = 0$) then in the model NCDF2 (where $s < 0$: this



Fig. 6 First component of the solution of (3.5) at times $t = 2.5, 15, 25$ with NCDF4

would illustrate the complex diffusion case, see Araújo et al. [2]). Note also from Fig. 2 that NCDF1 and NCDF3 will provide a better evolution of the two metrics: they

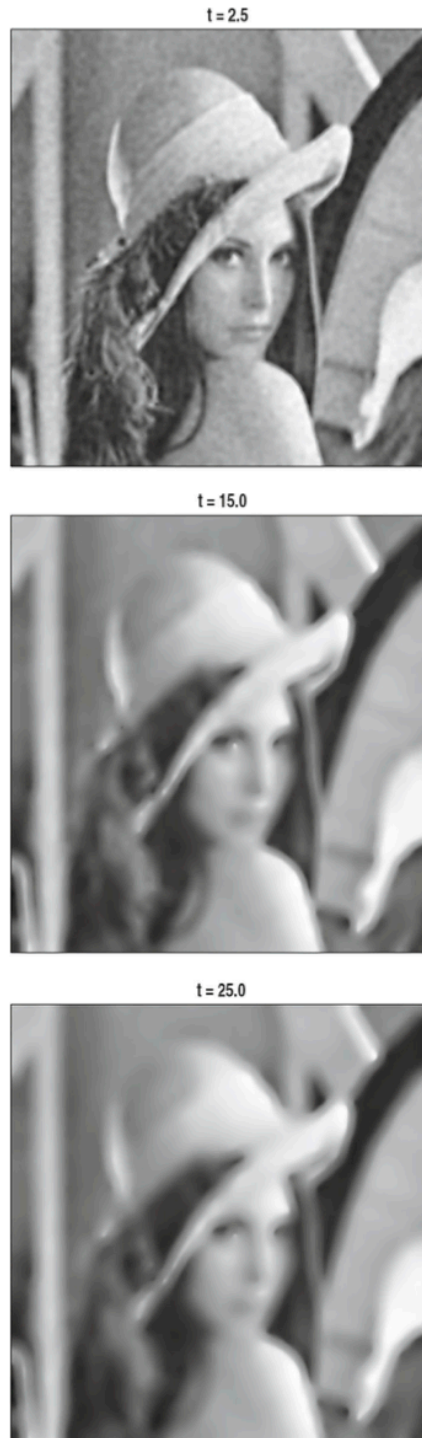


Fig. 7 First component of the solution of (3.5) at times $t = 2.5, 15, 25$ with NCDF5

will be more suitable than NCDF2 for long time restoration processes, while NCDF2 performs better in short computations.



Fig. 8 Second component of solution of (3.5) at times $t = 2.5, 15, 25$ with NCDF4

Since the longer the evolution the more noise is removed, models NCDF1 and NCDF3 suggest a better control of the diffusion to improve the quality of the restored images. This



Fig. 9 Second component of solution of (3.5) at times $t = 2.5, 15, 25$ with NCDF5

is observed in Figs. 3 and 4, which show the two components of the solution of (3.5) at time $t = 2.5$ given by NCDF1 and NCDF2 (the images corresponding to NCDF3 are similar to

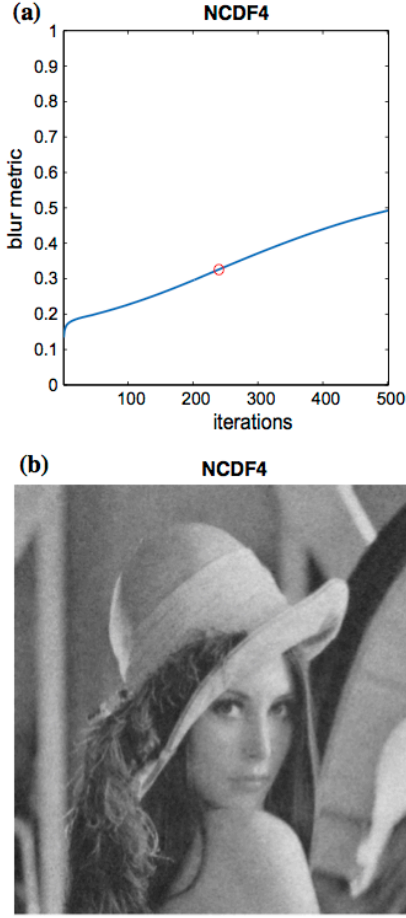


Fig. 10 NCDF4: **a** NPB values versus time and **b** first component of the solution of (3.5) at the time marked by the *small circle*

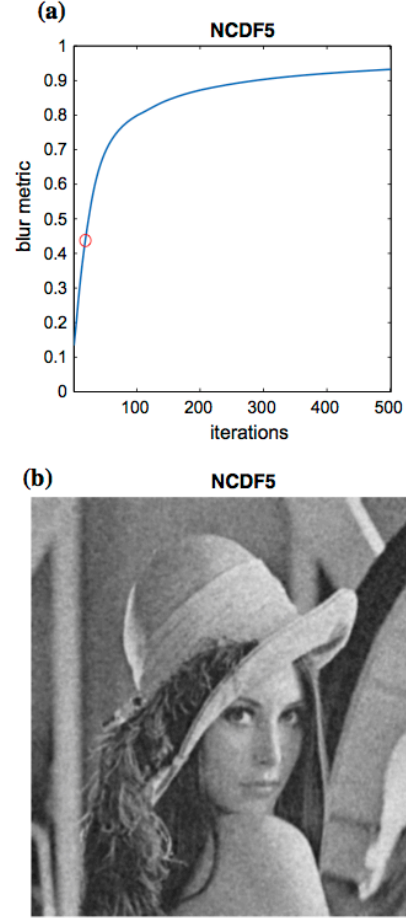


Fig. 11 NCDF5: **a** NPB values versus time and **b** first component of the solution of (3.5) at the time marked by the *small circle*

those of NCDF1 and will not be shown here). Observe that the second component has the role of edge detector and it is less affected by noise and over diffusion in the case of NCDF1.

For large values in magnitude of the entries of the matrix d in (3.2), the differences in the models are more significant. This is illustrated by a second group of experiments, for which the matrices are

- (i) NCDF4 ($s > 0$): $d_{11} = 1, d_{12} = 0.9, d_{21} = 1, d_{22} = 1$.
- (ii) NCDF5 ($s < 0$): $d_{11} = 1, d_{12} = -0.9, d_{21} = 0.9, d_{22} = 1$.
- (iii) NCDF6 ($s = 0$): $d_{11} = 1, d_{12} = -0.9, d_{21} = 0.225, d_{22} = 1.9$,

and the rest of the implementation data is the same as that of the previous experiments. The evolution of the SNR and PSNR values is now shown in Fig. 5. Note that the behaviour of NCDF5 and NCDF6 is very similar and their quality met-

rics, compared to those of NCDF4, are more suitable up to a time of filtering close to $t = 5$. From this time, NCDF4 behaves better and becomes a better choice to filter for a longer time.

The comparison between the solutions of (3.5) with the three models reveals these differences in a significant way, see Figs. 6, 7, 8 and 9, where the images corresponding to NCDF4 and NCDF5 at several times are displayed. (The results with NCDF6 are very similar to those of NCDF5.) In the case of the first component (Figs. 6, 7), the performance of the models by $t = 2.5$ is similar, but at longer times NCDF4 delays the blurring and leads to a restored image with better quality. This control of the diffusion is confirmed in Figs. 10, 11 and 12, which show, for the three models, the time evolution of the NPB metric (right) and the corresponding first component of the solution of (3.5) at the time for which the SNR value is maximum (left). (In each case, this time corresponds to the iteration of the numerical scheme associated to the small circle in the figure on the right.) The reduction

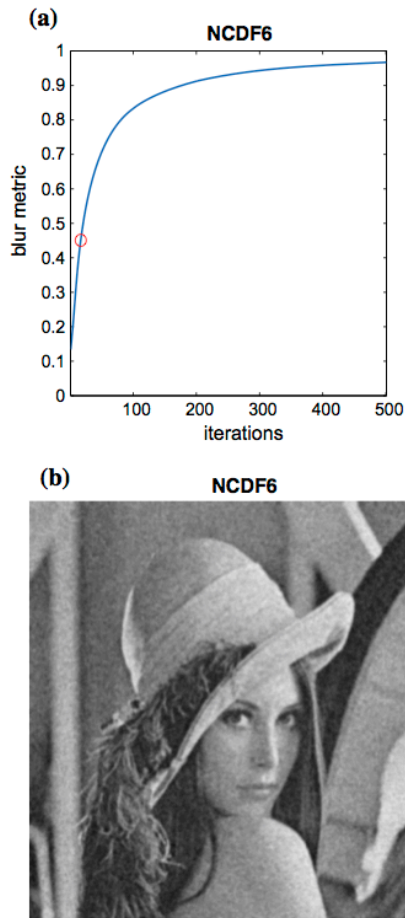


Fig. 12 NCDF6: **a** NPB values versus time and **b** first component of the solution of (3.5) at the time marked by the *small circle*

in the edge spreading is also observed in the detection of the edges by using the second components, see Figs. 8 and 9. The evolution of the NPB curve for NCDF4 implies the best quality in terms of blur perception, among the three models.

4 Concluding Remarks

In the present paper, nonlinear cross-diffusion systems as mathematical models for image filtering are studied. This is a continuation of the companion paper, Araújo et al. [2], devoted to the linear case. Here the nonlinearity is introduced through 2×2 , uniformly positive definite cross-diffusion coefficient matrices with bounded, globally Lipschitz entries. In the first part of the paper, well-posedness of the corresponding IBVP with Neumann boundary conditions is proved, as well as several scale-space properties and the limiting behaviour to the constant average grey value of the image at infinity. The second part is devoted to some numer-

ical comparisons on the performance of the filtering process from some noisy images using three models distinguished by different choices of the cross-diffusion matrix. The computational part does not intent to be exhaustive and instead aims to suggest and anticipate some preliminary conclusions that may motivate further research. As in the linear case, the systems incorporate some degrees of freedom. This diversity is mainly represented by the choice of the cross-diffusion matrix. The numerical study performed here makes use of cross-diffusion matrices whose derivation was based on the choices made in Gilboa et al. [15] for the complex diffusion case, combined with the results on linear cross-diffusion. The numerical results reveal that the structure of the diffusion coefficients affects the evolution of the filtering process and the quality in the detection of the edges through one of the components of the system.

Additional lines of future research concern the extension of the cross-formulation to study edge-enhancing problems as well as the introduction and analysis of discrete cross-diffusion systems, as discrete models for image filtering and as schemes of approximation to the continuous problem.

Acknowledgements This work was supported by Spanish Ministerio de Economía y Competitividad under the Research Grant MTM2014-54710-P. A. Araújo and S. Barbeiro were also supported by the Centre for Mathematics of the University of Coimbra—UID/MAT/00324/2013, funded by the Portuguese Government through FCT/MCTES and co-funded by the European Regional Development Fund through the Partnership Agreement PT2020.

References

1. Adams, R.A.: Sobolev Spaces. Academic Press, New York (1975)
2. Araújo, A., Barbeiro, S., Cuesta, E., Durán, A.: Cross-diffusion systems for image processing: I. The linear case. Preprint available at <http://arxiv.org/abs/1605.02923>
3. Álvarez, L., Guichard, F., Lions, P.-L., Morel, J.-M.: Axioms and fundamental equations for image processing. Arch. Ration. Mech. Anal. **123**, 199–257 (1993)
4. Álvarez, L., Weickert, J., Sánchez, J.: Reliable estimation of dense optical flow fields with large displacements. Int. J. Comput. Vis. **39**(1), 41–56 (2000)
5. Araújo, A., Barbeiro, S., Serranho, P.: Stability of finite difference schemes for complex diffusion processes. SIAM J. Numer. Anal. **50**, 1284–1296 (2012)
6. Aubert, J., Kornprobst, P.: Mathematical Problems in Image Processing. Springer, Berlin (2001)
7. Bernardes, R., Maduro, C., Serranho, P., Araújo, A., Barbeiro, S., Cunha-Vaz, J.: Improved adaptive complex diffusion despeckling filter. Opt. Express. **18**(23), 24048–24059 (2010)
8. Brezis, H.: Functional Analysis. Sobolev Spaces and Partial Differential Equations. Springer, New York (2011)
9. Catté, F., Lions, P.-L., Morel, J.-M., Coll, B.: Image selective smoothing and edge detection by nonlinear diffusion. SIAM J. Numer. Anal. **29**(1), 182–193 (1992)
10. Créte-Roffet, F., Dolmiere, T., Ladret, P., Nicolas, M.: The blur effect: perception and estimation with a new no-reference perceptual blur metric. In: Proc. SPIE on Human Vision and Electronic Imaging XII, San José, California, USA, 11 pages (2007)

11. Galiano, G., Garzón, M.J., Jüngel, A.: Analysis and numerical solution of a nonlinear cross-diffusion system arising in population dynamics. *Rev. R. Acad. Cie. Ser. A Mat.* **95**(2), 281–295 (2001)
12. Galiano, G., Garzón, M.J., Jüngel, A.: Semi-discretization in time and numerical convergence of solutions of a nonlinear cross-diffusion population model. *Numer. Math.* **93**, 655–673 (2003)
13. Gilboa, G., Zeevi, Y. Y., Sochen, N. A.: Complex diffusion processes for image filtering. In: *Scale-Space and Morphology in Computer Vision*, pp. 299–307, Springer, Berlin (2001)
14. Gilboa, G., Sochen, N. A., Zeevi, Y. Y.: Generalized shock filters and complex diffusion. In: *Computer Vision-ECCV*, pp. 399–413, Springer, Berlin (2002)
15. Gilboa, G., Sochen, N.A., Zeevi, Y.Y.: Image enhancement and denoising by complex diffusion processes. *IEEE Trans. Pattern Anal. Mach. Intell.* **26**(8), 1020–1036 (2004)
16. ter Haar Romeny, B. (Ed.): *Geometry-Driven Diffusion in Computer Vision*, Springer, New York (1994)
17. Kinchenassamy, S.: The Perona–Malik paradox. *SIAM J. Appl. Math.* **57**, 1328–1342 (1997)
18. Ladyzenskaya, O.A., Solonnikov, V.A., Ural'ceva, N.N.: *Linear and Quasilinear Equations of Parabolic Type*. Amer. Math. Soc, Providence (1968)
19. Wah, B. (ed.): *Scale space*. Encyclopedia of Computer Science and Engineering, vol. IV, pp. 2495–2509. Wiley, Hoboken (2009)
20. Lorenz, D.A., Bredies, K., Zeevi, Y.Y.: *Nonlinear complex and cross diffusion*. Unpublished report, University of Bremen (2006). Freely available in: <https://www.researchgate.net>
21. Nagel, H.-H., Enkelmann, W.: An investigation of smoothness constraints for the estimation of displacement vector fields from images sequences. *IEEE Trans. Pattern Anal. Mach. Intell.* **8**, 565–593 (1986)
22. Ni, W.-M.: Diffusion, cross-diffusion and their spike-layer steady states. *Notices AMS.* **45**(1), 9–18 (1998)
23. Pauwels, E.J., Van Gool, L.J., Fiddelaers, P., Moons, T.: An extended class of scale-invariant and recursive scale space filters. *Pattern Anal. Mach. Intell.* **17**(7), 691–701 (1995)
24. Peter, P., Weickert, J., Munk, A., Krivobokova, T., Li, H.: Justifying tensor-driven diffusion from structure-adaptive statistics of natural images. In: *Energy Minimization Methods in Computer Vision and Pattern Recognition*, Volume 8932 of the series Lecture Notes in Computer Science, pp. 263–277 (2015)
25. Perona, P., Malik, J.: Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. Pattern Anal. Mach. Intell.* **12**, 629–639 (1990)
26. Proesmans, M., Pauwels, E., Van Gool, L.: Coupled geometry-driven diffusion equations for low-level vision. In: ter Haar-Romeny, B. (ed.) *Geometry-Driven Diffusion in Computer Vision*, pp. 191–228. Springer, New York (1994)
27. Weickert, J.: *Anisotropic Diffusion in Image Processing*. B.G. Teubner, Stuttgart (1998)
28. Whitaker, R., Gerig, G.: Vector-valued diffusion. In: ter Haar-Romeny, B. (ed.) *Geometry-Driven Diffusion in Computer Vision*, pp. 93–124. Springer, New York (1994)
29. Zabandan, G., Kiliman, A.: A new version of Jensen's inequality and related results. *J. Inequal. Appl.* **238** (2012)



Adérito Araújo is an Associate Professor at the University in Coimbra, Portugal. He received his Ph.D. in 1998 in applied mathematics from the University of Coimbra on the topic of Additive Symplectic Integrators. Since that date, he has been an active member of the Numerical Analysis and Optimization group of the Centre of Mathematics of the University of Coimbra maintaining active collaboration with engineers, doctors, physicists and chemists in several interdisciplinary projects.

His research interests include the development of mathematical and numerical models for biomedical applications, in particular, for image processing and light propagation in human retina. He is in the Management Committee of the COST Action TD1409 Mathematics for Industry network (MI-NET).



Sílvia Barbeiro is an Assistant Professor at the Department of Mathematics of the University of Coimbra. She studied Mathematics at the University of Coimbra and at Technical University of Berlin. She received her Ph.D. degree in Applied Mathematics in 2005 from the University of Coimbra. She was a visiting research fellow at the University of Texas at Austin during the Spring and Fall semesters of 2008, at the Acoustics Research Institute of

the Austrian Academy of Sciences during the academic year of 2015/2016 and at the University of Oxford during the trinity term of 2016. Her research interests are in numerical analysis and computational mathematics, including modelling and the development and analysis of numerical methods for partial differential equations and integro-differential equations. The main applications driving her recent research are in the fields of electromagnetic wave propagation and image processing.



Eduardo Cuesta is Senior Lecturer in the University of Valladolid (Spain) since 2010. He obtained the five-years B.Sc. in Mathematics in 1991 and his Ph.D. in Numerical Analysis in 2001. He was a Postdoctoral Fellow in the Institute of Applied & Computational Mathematics (FORTH-IACM) in Crete (Greece) 2003. After several publications, attending to congress, and invited seminar and conferences, his current interests include evolutionary mathematical

models and numerical methods for image processing, and nonlocal problems and its applications.



Ángel Durán received the Ph.D. degree in Mathematics from the University of Valladolid, Spain, in 1998. He is currently a senior lecturer in Applied Mathematics in the Higher Technical School of Telecommunications Engineering of the University of Valladolid. His research interests are the theoretical and numerical analysis of nonlinear waves, in particular on the dynamics of solitary waves, and, more recently, the mathematical formulation of models for image

processing.