Pré-Publicações do Departamento de Matemática
Universidade de Coimbra
Preprint Number 16–23

CROSS-DIFFUSION SYSTEMS FOR IMAGE PROCESSING: II. THE NONLINEAR CASE

ADÉRITO ARAÚJO, SÍLVIA BARBEIRO, EDUARDO CUESTA AND ÁNGEL DURÁN

Abstract: In this paper the use of nonlinear cross-diffusion systems to model image restoration is investigated, theoretically and numerically. In the first case, well-posedness, scale-space properties and long time behaviour are analyzed. From a numerical point of view, a computational study of the performance of the models is carried out, suggesting their diversity and potentialities to treat image filtering problems. The present paper is a continuation of a previous work of the same authors, devoted to linear cross-diffusion models.

1. Introduction

This paper is concerned with the use of nonlinear cross-diffusion systems for the mathematical modeling of image filtering. Here the grey-scale image is represented by a vector field of two components $\mathbf{u} = (u, v)^T$. From an initial noisy image the filtering process for $\mathbf{u}$ will be governed by an evolution problem, [6, 16] of nonlinear cross-diffusion type. The nonlinearity is identified by some cross-diffusion coefficient matrix satisfying certain properties which will be analyzed along the text.

The paper is a continuation of a previous work devoted to the linear case, [2]. There, those systems satisfying relevant scale-space properties were identified. The performance of the models was analyzed by numerical means in terms of some features: the way how the information about the initial noisy image is distributed between the components of the vector field, the role of each of the components to control diffusion and the choice of some entries of the convolution kernel.

The goal is to continue into this approach by incorporating nonlinearity into the model. As in the linear case, a foregoing related proposal is the use of complex diffusion problems, developed by Gilboa and collaborators, [14], where the image is represented by a complex function which evolves according to some complex diffusion process. The application of nonlinear complex diffusion to image filtering and edge enhancing is discussed in [12, 13, 14].

Received May 20, 2016.

The interpretation of complex diffusion as a cross diffusion system motivates the introduction of a more general set of equations and its study as a mathematical model in image processing. As in the complex diffusion case, the idea of dividing the information of the image in two components evolving in a cross way is behind the approach. This may provide a relevant diversity of the resulting models that can be used to adapt their performance to the problem under study. This was shown for the complex diffusion, both in the linear case when studying the role of the imaginary part as smoothed Laplacian of the initial image (the so-called small theta approximation) and in the nonlinear case when some complex shock filter models are proposed, [14]. The present paper is concerned with the study of the nonlinear cross-diffusion model for image restoration, following the same approach of the previous work for the linear case, [2], leaving the study for edge-detection for a future research.

The application of cross-diffusion systems for modeling, especially in population dynamics, is well known (see e. g. [10, 11, 21] and references therein). To our knowledge, the only reference concerning image processing with cross-diffusion models is the unpublished manuscript [20], where the authors prove the existence of 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, [17, 22]. The improvement is also confirmed by a numerical comparison in a restoration problem.

The main contribution of this paper is the formalization of nonlinear cross-diffusion problems as mathematical models for image processing. Under some hypothesis on the cross-diffusion matrix, well-posedness of the nonlinear cross-diffusion initial boundary-value problem with Neumann boundary conditions is stablished, along with the proof of some scale-space properties, [3, 19], and the study of the behaviour of the solution at infinity. Some of the arguments of [20] for the system under study will be used and generalized here.

Additionally, the performance of the models are estimated by computational means in several examples. Specifically, three particular cases (one of them corresponding to the complex diffusion case) are applied in a filtering problem and the quality of restoration is measured and compared by using some standard performance metrics. The numerical experiments presented here seek to be illustrative of the potential of these models to image filtering. Rather than being exhaustive, this numerical study is a proof-of-concept which aims to motivate future research in this direction.

The paper is structured as follows: In Section 2 an initial boundary-value problem of cross-diffusion type with Neumann boundary conditions is introduced. Using standard results, existence and uniqueness of solution of the weak formulation is proved, the regularity is studied and a maximum principle is established. These make up the main body of well-posedness results. The satisfaction of some scale-space properties, a discussion on the existence of Lyapunov functions and the behaviour at infinity complete the theoretical analysis of the model. In Section 3 some different versions of the system, according to some choice of the cross-diffusion matrix, are compared, by numerical means, in a standard problem of image restoration. The comparison is focused on the computation of some quality indexes and the evolution of the filtering process. Finally the main conclusions and some future research are outlined in Section 4.

2. Nonlinear cross-diffusion model

The following initial boundary-value problem of cross-diffusion for $\mathbf{u} = (u, v)^T$

$$\begin{aligned}
\frac{\partial u}{\partial t}(\mathbf{x},t) &= \operatorname{div}\left(D_{11}(\mathbf{u}(\mathbf{x},t))\nabla u(\mathbf{x},t)\right. \\
&\qquad \left.+D_{12}(\mathbf{u}(\mathbf{x},t))\nabla v(\mathbf{x},t)\right), \\
\frac{\partial v}{\partial t}(\mathbf{x},t) &= \operatorname{div}\left(D_{21}(\mathbf{u}(\mathbf{x},t))\nabla u(\mathbf{x},t)\right. \\
&\qquad \left.+D_{22}(\mathbf{u}(\mathbf{x},t))\nabla v(\mathbf{x},t)\right), \\
u(\mathbf{x},0) &= u_0(\mathbf{x}), v(\mathbf{x},0) = v_0(\mathbf{x}), \quad \mathbf{x} \in \Omega,
\end{aligned} \tag{2.1}$$

for $(\mathbf{x},t) \in Q_T := \Omega \times (0,T]$, with Neumann boundary conditions in $\partial\Omega \times [0,T]$

$$\begin{aligned}
\langle D_{11}(\mathbf{u})\nabla u + D_{12}(\mathbf{u})\nabla v, n\rangle = 0, \\
\langle D_{21}(\mathbf{u})\nabla u + D_{22}(\mathbf{u})\nabla v, n\rangle = 0,
\end{aligned} \tag{2.2}$$

is considered. In (2.1), (2.2), Ω is a (squared) domain in $\mathbb{R}^2$ with boundary $\partial\Omega$, n stands for the exterior normal vector to $\partial\Omega$. From the Sobolev spaces on Ω $H^k(\Omega), k \geq 0$ (with $H^0(\Omega) = L^2(\Omega)$) we define $X_k := H^k(\Omega) \times H^k(\Omega)$

with norm, [1]

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

Finally, D stands for a cross-diffusion matrix operator

$$\mathbf{u} \in X_1 \mapsto D(\mathbf{u}) : \overline{Q_T} \to \mathbb{R}^4,$$

with, for $(\mathbf{x}, t) \in \overline{Q_T} := \overline{\Omega} \times [0, T]$

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

and satisfying the following hypotheses:

(H1) There exists $\alpha > 0$ such that for each $\mathbf{u} : \overline{Q_T} \to \mathbb{R}^2$ with $\mathbf{u}(\cdot, t) \in X_1, t \in [0, T]$

$$\xi^T D(\mathbf{u}(\mathbf{x}, t))\xi \geq \alpha|\xi|^2, \quad \xi \in \mathbb{R}^2, (\mathbf{x}, t) \in \overline{Q_T}. \tag{2.3}$$

(H2) There exists $L > 0$ such that for $\mathbf{u}, \mathbf{v} : \overline{Q_T} \to \mathbb{R}^2$ with $\mathbf{u}(\cdot, t), \mathbf{v}(\cdot, t) \in X_1, (\mathbf{x}, t) \in \overline{Q_T}, i, j = 1, 2$

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

(H3) There exists $M > 0$ such that for each $\mathbf{u} : \overline{Q_T} \to \mathbb{R}^2$ with $\mathbf{u}(\cdot, t) \in X_1, t \in [0, T]$

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

Conditions (H1)-(H3) can also be complemented by some other assumptions required by well-posedness or some scale-space properties (see Section 2.2).

As an example we have the case of complex diffusion, [14]. This can be formulated as a cross-diffusion problem (2.2) for the real and imaginary parts of the (complex) image function $u+iv$. In [14] the following diffusion coefficient is used:

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

where κ is a threshold parameter and θ is a phase angle parameter. In the cross-diffusion formulation, this corresponds to the matrix

$$D(u, v) = g(v) \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \; g(v) = \frac{1}{1 + \left(\frac{v}{\kappa\theta}\right)^2}. \tag{2.4}$$

In what follows the weak formulation of (2.1), (2.2) will be considered. This consists of finding $\mathbf{u} = (u, v)^T : [0, T] \longrightarrow X_1$ satisfying, for any $t \in [0, T]$

$$\begin{aligned}
&\int_\Omega ((\partial_t u)w_1 + (\partial_t v)w_2)\, d\Omega \\
&+ \int_\Omega (\nabla w_1 \nabla w_2) D(\mathbf{u})(\nabla u \nabla v)^T d\Omega = 0,
\end{aligned} \tag{2.5}$$

for all $\mathbf{w} = (w_1, w_2)^T \in X_1$. In (2.5), the integrand of the second term is defined as

$$\begin{aligned}
(\nabla w_1 \nabla w_2) D (\nabla v_1 \nabla v_2)^T &= \nabla w_1^T D_{11} \nabla v_1 + \nabla w_1^T D_{12} \nabla v_2 \\
&\quad + \nabla w_2^T D_{21} \nabla v_1 \\
&\quad + \nabla w_2^T D_{22} \nabla v_2,
\end{aligned}$$

where $D_{ij} = D_{ij}(\mathbf{u}), i, j = 1, 2, D = D(\mathbf{u})$ and

$$\nabla f D_{ij} \nabla g = f_x D_{ij} g_x + f_y D_{ij} g_y, \quad i, j = 1, 2.$$

This implies that

$$\begin{aligned}
(\nabla w_1 \nabla w_2) D (\nabla v_1 \nabla v_2)^T &= (w_{1x}, w_{2x}) D \begin{pmatrix} v_{1x} \\ v_{2x} \end{pmatrix} \\
&\quad + (w_{1y}, w_{2y}) D \begin{pmatrix} v_{1y} \\ v_{2y} \end{pmatrix} \\
&= \mathbf{w}_x^T D \mathbf{v}_x + \mathbf{w}_y^T D \mathbf{v}_y.
\end{aligned} \tag{2.6}$$

Problem (2.1), (2.2) will be now studied when understood as an evolution system for image processing. The analysis is concerned with well-posedness, scale-space properties and long time behaviour.

2.1. Well-posedness. The goal of this section is to prove the existence, uniqueness, regularity and continuous dependence on the initial data of the solution for (2.5), which corresponds to the weak solution of the nonlinear cross-diffusion initial boundary-value problem (2.1), (2.2). Some notation is first introduced. We define

$$\begin{aligned}
W(0, T) &= \{w \in L^2(0, T, H^1(\Omega)) : \\
&\qquad \frac{dw}{dt} \in L^2(0, T, (H^1(\Omega))')\},
\end{aligned}$$

where $(H^1(\Omega))'$ stands for the dual space of $H^1(\Omega)$, characterized as the completion of $L^2(\Omega)$ with respect to the norm, [1],

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

In $W(0,T)$ we consider the graph norm and we will also make use of the L^∞ norm

$$||v||_{L^\infty(0,T,L^2(\Omega))} = \operatorname*{ess\,sup}_{t\in[0,T]} ||u(t)||_{L^2(\Omega)}.$$

Finally $(X_1)'$ will stand for $(H^1(\Omega))' \times (H^1(\Omega))'$.

To prove the well-posedness of the problem, we follow the standard arguments used in [9, 23] (see also [10] and references therein). We first consider a related linear problem and we establish a maximum-minimum principle and the 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, [8]. The same arguments as in [9, 23] apply to prove the uniqueness of solution, as well as its 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), (2.2), finishing off the study of well-posedness.

Theorem 1. *Under hypotheses (H1)-(H3), (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 and which is a strong solution of (2.1), (2.2) for* $\mathbf{u}_0 \in X_1$ *when* D *is smooth with* $\mathbf{u} \in C^\infty((0,T] \times \overline{\Omega})$*. Furthermore, if*

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

and $\mathbf{u} = (u,v)^T$ *then for all* $(\mathbf{x},t) \in Q_T$

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

2.1.1. *Existence.* In order to study the existence of solution of (2.5) the following linear initial-boundary-value problem is considered. Let $\mathbf{U} = (U,V)^T$, with

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

be given. In order to study the existence of solution of (2.5) the following linear initial-boundary-value problem n Q_T is considered.

$$\begin{aligned}
\frac{\partial u}{\partial t}(\mathbf{x},t) &= \operatorname{div}\left(D_{11}(\mathbf{U}(\mathbf{x},t))\nabla u(\mathbf{x},t)\right.\\
&\quad \left.+D_{12}(\mathbf{U}(\mathbf{x},t))\nabla v(\mathbf{x},t)\right),\\
\frac{\partial v}{\partial t}(\mathbf{x},t) &= \operatorname{div}\left(D_{21}(\mathbf{U}(\mathbf{x},t))\nabla u(\mathbf{x},t)\right.\\
&\quad \left.+D_{22}(\mathbf{U}(\mathbf{x},t))\nabla v(\mathbf{x},t)\right),\\
u(\mathbf{x},0) &= u_0(\mathbf{x}), v(\mathbf{x},0) = v_0(\mathbf{x}), \quad \mathbf{x}\in\Omega,
\end{aligned} \tag{2.7}$$

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

$$\begin{aligned}
\langle D_{11}(\mathbf{u})\nabla u + D_{12}(\mathbf{u})\nabla v, n\rangle = 0,\\
\langle D_{21}(\mathbf{u})\nabla u + D_{22}(\mathbf{u})\nabla v, n\rangle = 0.
\end{aligned} \tag{2.8}$$

Since $D(\mathbf{U}) = D(U,V)$ is uniformly positive definite (hypothesis (H1)), then, [18], there is a unique weak solution of (2.7), (2.8), $\mathbf{u}(U,V) = (U_1(U,V),U_2(U,V))$, with

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

We now establish some estimates of this solution in different norms, [20]. Consider first the weak formulation of (2.7): find $\mathbf{u}(U,V) = (U_1(U,V),U_2(U,V)) \in L^2(0,T,X_1)$ satisfying

$$\begin{aligned}
&\int_\Omega \left((\partial_t U_1)v_1 + (\partial_t U_2)v_2\right) d\Omega\\
&+\int_\Omega (\nabla v_1 \nabla v_2) D(U,V)(\nabla U_1\nabla U_2)^T d\Omega = 0,
\end{aligned} \tag{2.9}$$

for every $\mathbf{v} = (v_1,v_2)\in X_1$ and all $0\le t\le T$. 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 $f_+ = \max\{f,0\}$, [20, 23]. Then (2.9) becomes

$$\begin{aligned}
&\frac{1}{2}\int_\Omega \left(\partial_t(U_1-b_1)_+^2 + \partial_t(U_2-b_2)_+^2\right) d\Omega\\
&+\int_{U_1>b_1,U_2>b_2} (\nabla U_1\nabla U_2)D(U,V)(\nabla U_1\nabla U_2)^T d\Omega = 0.
\end{aligned}$$

By using (H1)

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

Integrating we have, for any $0\leq t\leq T$,

$$\begin{aligned}&\int_{\Omega}\left((U_1(t)-b_1)_+^2+(U_2(t)-b_2)_+^2\right)d\Omega\\ &\leq\int_{\Omega}\left((U_1(0)-b_1)_+^2+(U_2(0)-b_2)_+^2\right)d\Omega.\end{aligned}\tag{2.10}$$

Now we can take b_1,b_2 such that the integral on the right hand side of (2.10) becomes zero. For instance, if we assume $U_1(0),U_2(0)\in L^{\infty}(\Omega)$ and

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

then (2.10) 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 all $0\leq t\leq T$, that is

$$\begin{aligned}U_1(x,t)&\leq b_1=||U_1(0)||_{L^{\infty}(\Omega)},\\ U_2(x,t)&\leq b_2=||U_2(0)||_{L^{\infty}(\Omega)}.\end{aligned}\tag{2.11}$$

Formulas (2.11) can be understood as a maximum principle and will be adapted to the nonlinear case below. On the other hand, 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 consider

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

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 all $0 \leq t \leq T$, that is

$$U_1(x,t) \geq \operatorname{ess\,inf} U_1(0),\; U_2(x,t) \geq \operatorname{ess\,inf} U_2(0). \tag{2.12}$$

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 on the solution of the linear problem (2.7) can be obtained from the 'energy'

$$E_L(t) = \frac{1}{2}\int_\Omega \left(U_1(\mathbf{x},t)^2 + U_2(\mathbf{x},t)^2\right) d\Omega.$$

Note that if in the weak formulation (2.9) 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

$$\begin{aligned} ||U_1||_{L^\infty(0,T,L^2(\Omega))} &\leq ||U_1(0)||_{L^2(\Omega)}, \\ ||U_2||_{L^\infty(0,T,L^2(\Omega))} &\leq ||U_2(0)||_{L^2(\Omega)}. \end{aligned} \tag{2.13}$$

We now look 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

$$\begin{aligned} &\frac{d}{dt}\frac{1}{2}\int_\Omega \left(U_1(\mathbf{x},t)^2 + U_2(\mathbf{x},t)^2\right) d\Omega \\ &+\alpha \int_\Omega \left(|\nabla U_1(\mathbf{x},t)|^2 + |\nabla U_2(\mathbf{x},t)|^2\right) d\Omega \leq 0. \end{aligned} \tag{2.14}$$

Then (2.14) implies that for any $t \in [0,T]$

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

Therefore

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

Thus, if $U_0 = (U_1(0), U_2(0))$ there exists a constant $C_1 = C_1(\alpha, U_0, T)$ such that

$$\begin{aligned}
&||U_1||_{L^2(0,T,H^1(\Omega))} \leq C_1, \\
&||U_2||_{L^2(0,T,H^1(\Omega))} \leq C_1. \qquad (2.15)
\end{aligned}$$

On the other hand, if $||v||_{L^2(0,T,H^1(\Omega))} = 1$, the weak formulation (2.9), assumption (H3) and Cauchy-Schwarz inequality imply that

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

Therefore, this and (2.15) lead to

$$
||\frac{d}{dt}u||_{L^2(0,T,(X_1)')} \leq M||u||_{L^2(0,T,X_1)} \leq MC_1. \tag{2.16}
$$

2.1.2. *Schauder fixed point theorem. Existence of solution.* Estimates (2.13), (2.15) and (2.16) will be used below to study the existence of solution of (2.5) by using the Schauder fixed-point theorem, [8]. (Analogous arguments were used in [9, 23].) We first assume $u_0 \in L^2(\Omega) \times L^2(\Omega)$ in (2.1) and consider the subset of
$W(0,T)^2 := W(0,T) \times W(0,T)$:

$$
\begin{aligned}
K \;=\; &\{\mathbf{w} = (w_1, w_2) \in W(0,T)^2 : w \text{ satisfies} \\
&(2.13),\ (2.15) \text{ and } (2.16) \text{ with } w(0) = u_0\},
\end{aligned}
$$

and the mapping $T : K \longrightarrow W(0,T)^2$ such that $T(w) := u(w)$ is the (weak) solution of (2.7) with $(U,V) = w$.

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 $\{w_n\}_n \subset K$ and $t \in [0,T]$. Since K is a bounded set, then

$$\{w_n(t)\}_n, \{\frac{d}{dt}w_n(t)\}_n$$

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

$$w_n(t) \to \varphi(t), \quad \frac{d}{dt}w_n(t) \to \psi(t),$$

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

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

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

These convergence properties imply two additional ones:

(v) $u_n(0) \to \phi(0)$ in $(X_1)'$.
(vi) $\nabla u_n \to \nabla\phi$ weakly in $L^2(0,T,X_0)$.

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

$$D(w_n) \to D(w)$$

in $L^2(0,T,X_0)$. Then if we take limit in (2.9) we have $\phi = T(w)$. Finally, since the whole sequence $\{u_n\}_n$ is bounded in K which is weakly compact, then it converges weakly in $W(0,T)$. By uniqueness of solution of (2.9) the whole sequence $u_n = T(w_n)$ must converge weakly to $\phi = T(w)$ and therefore T is weakly continuous and (3) holds.

Thus, Schauder fixed point theorem proves the existence of solution $\mathbf{u}$ of (2.5). The solution $\mathbf{u}$ is in K and therefore $\mathbf{u} \in L^2(0,T,X_1), \frac{d\mathbf{u}}{dt} \in$

$L^2(0,T,(X_1)')$, it satisfies (2.13), (2.15) and (2.16). Furthermore, due to the conditions (H1)-(H3) on D, at least $\mathbf{u} \in C(0,T,X_0)$.

2.1.3. *Regularity of solution.* The same bootstrap argument as in [9, 23] applies to obtain that u is a strong solution and $\mathbf{u} \in C^\infty((0,T] \times \overline{\Omega})$ if (H2) is substituted by the hypothesis that D is smooth enough.

2.1.4. *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 $\mathbf{w} = (w_1, w_2) \in X_1$

$$\begin{aligned}&\int_\Omega \Big((\partial_t(u^{(1)} - u^{(2)}))w_1 + (\partial_t(v^{(1)} - v^{(2)}))w_2\Big)\, d\Omega\\ &+ \int_\Omega (\nabla w_1 \nabla w_2) D(\mathbf{u}^{(1)})(\nabla u^{(1)} \nabla v^{(1)})^T d\Omega\\ &- \int_\Omega (\nabla w_1 \nabla w_2) D(\mathbf{u}^{(2)})(\nabla u^{(2)} \nabla v^{(2)})^T d\Omega = 0,\end{aligned}$$

which can be written as

$$\begin{aligned}&\int_\Omega \Big((\partial_t(u^{(1)} - u^{(2)}))w_1 + (\partial_t(v^{(1)} - v^{(2)}))w_2\Big)\, d\Omega\\ &+ \int_\Omega (\nabla w_1 \nabla w_2) D(\mathbf{u}^{(1)}) \begin{pmatrix} \nabla(u^{(1)} - u^{(2)}) \\ \nabla(v^{(1)} - v^{(2)}) \end{pmatrix} d\Omega\\ &+ \int_\Omega (\nabla w_1 \nabla w_2) \Big(D(\mathbf{u}^{(1)}) - D(\mathbf{u}^{(2)})\Big) \begin{pmatrix} \nabla u^{(2)} \\ \nabla v^{(2)} \end{pmatrix} d\Omega = 0.\end{aligned}$$

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

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

(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

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

Now, Gronwall's lemma leads to

$$\begin{aligned}&||\mathbf{u}^{(1)}(t) - \mathbf{u}^{(2)}(t)||_{X_0}^2\\ &\leq ||\mathbf{u}^{(1)}(0) - \mathbf{u}^{(2)}(0)||_{X_0}^2 exp\left(C\int_0^t ||\nabla\mathbf{u}^{(2)}(s)||_{X_0}^2 ds\right),\end{aligned} \tag{2.17}$$

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

2.1.5. *Extremum principle.* Note that the same argument as that of the linear problem (2.7) can be adapted to the nonlinear case straightforwardly to prove an extremum principle (2.11), (2.12) for (2.1).

2.1.6. *Continuous dependence on initial data.* Since u is bounded on $\overline{Q_T}$, then $\nabla\mathbf{u}$ is bounded and hypothesis (H1) on D implies

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

Now, let $\epsilon > 0$ and take

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

If $||\mathbf{u}^{(1)}(0) - \mathbf{u}^{(2)}(0)||_{X_0} < \delta$ and using (2.17) 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, cf. [23].

2.2. Scale-space properties. We now deal with scale-space properties for (2.1). Let $\mathbf{u}(\mathbf{x}, t)$ be the unique solution of (2.1) and consider the scale-space operator

$$T_t : \mathbf{u}_0 \longmapsto T_t(\mathbf{u}_0) := \mathbf{u}(\cdot, t) = \mathbf{u}(t), \quad t \geq 0.$$

2.2.1. *Grey level shift invariance.* It looks clear that $T_t(0) = 0$. If the function

$$\mathbf{w}(t) = T_t(\mathbf{u}_0) + C$$

$\mathbf{C} = (C_1, C_2)^T \in \mathbb{R}^2$, satisfies (2.1) with initial condition $\mathbf{u}_0 + C$, by uniqueness

$$T_t(\mathbf{u}_0 + C) = T_t(\mathbf{u}_0) + \mathbf{C}, \quad t \geq 0. \tag{2.18}$$

If we assume that the diffusion tensor is only a function of $J(u)$,

$$J(\mathbf{u}) = \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix},$$

then the property (2.18) is verified.

Remark 1. *If grey level range is governed by one of the components, say u, then (2.18) holds if $D = D(v)$ and for any $\mathbf{C} = (C_1, 0)^T$. This is related to the role of small theta approximation in complex diffusion,* [14].

2.2.2. *Reverse contrast invariance.* Property

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

holds if $D(-\mathbf{u}) = D(\mathbf{u})$. (For instance, if D depends only on $|\mathbf{u}|$ or $|\nabla u|, |\nabla v|$.)

2.2.3. *Average grey invariance.* We consider the vector function

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

As in [23], we have, for $i = 1, 2$

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

where $A(\Omega)$ stands for the area of Ω. Since at least $\mathbf{u} \in C(0, T, L^2(\Omega) \times L^2(\Omega))$ then $\mathbf{G}$ is continuous at $t = 0$. On the other hand, divergence theorem and the boundary conditions imply that, for $i = 1, 2$

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

Then $G_i(t)$ is constant for all $t \geq 0$. Thus the quantity

$$M\mathbf{u}_0 = (m(u_0), m(v_0))^T = \begin{pmatrix} \frac{1}{A(\Omega)} \int_\Omega u_0(\mathbf{x}) d\Omega \\ \frac{1}{A(\Omega)} \int_\Omega v_0(\mathbf{x}) d\Omega \end{pmatrix}$$

is preserved by cross-diffusion, that is

$$\frac{1}{A(\Omega)} \int_\Omega T_t(\mathbf{u}_0)(\mathbf{x}) d\Omega = M\mathbf{u}_0, \quad t \geq 0. \tag{2.19}$$

Remark 2. *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 [2] for a discussion about this question.*

2.2.4. *Translational invariance.* Consider a translational operator

$$\tau_{\mathbf{h}} f(\mathbf{x}) = f(\mathbf{x} + \mathbf{h}), \quad \mathbf{x}, \mathbf{h} \in \mathbb{R}^2.$$

Then it is clear that if D does not depend explicitly on $\mathbf{x}$ then

$$T_t(\tau_{\mathbf{h}} u_0) = \tau_{\mathbf{h}}(T_t(u_0)), \quad t \geq 0,$$

since both functions satisfy (2.1) with the same initial condition $\tau_{\mathbf{h}} u_0$.

2.3. Lyapunov functions and behaviour at infinity. The previous study can be finished off by analyzing the existence of Lyapunov functionals and the behaviour of the solution when $t \to \infty$, [23]. We already have a Lyapunov functional given by

$$V(t) = \Phi(\mathbf{u}(t)) := \frac{1}{2} \int_\Omega \left(u(\mathbf{x},t)^2 + v(\mathbf{x},t)^2 \right) d\Omega, \tag{2.20}$$

where $\mathbf{u} = (u,v)^T$ is the solution of (2.1). In order to prove this, from the weak formulation with $\mathbf{w} = \mathbf{u}$ we obtain

$$\frac{d}{dt} V(t) + \int_\Omega (\nabla u \nabla v) D(\mathbf{u}) (\nabla u \nabla v)^T d\Omega = 0,$$

which, due to (H1) and (2.6), implies

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

Note also that since $\widetilde{r}(z) = \frac{z^2}{2}$ is convex and (2.19) holds, then using Jensen inequality implies that

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

Therefore, (2.20) is a Lyapunov functional. These arguments can be generalized by considering functionals of the form

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

where r is a C^2 strongly convex function of parameter $p \geq 0$, that is

$$\langle \nabla r(\mathbf{x}) - \nabla r(\mathbf{y}), \mathbf{x} - \mathbf{y} \rangle \leq p||\mathbf{x} - \mathbf{y}||^2, \quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^2,$$

(which implies that r is strictly convex). For instance, (2.20) corresponds to taking

$$r(x, y) = \frac{x^2 + y^2}{2}.$$

Then, divergence theorem, boundary conditions (2.2) and (2.6) imply

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

where $\nabla^2 r(u, v)$ stands for the Hessian of r. Since r is strongly convex of parameter $p \geq 0$ then $\nabla^2 r(u, v)$ is positive semi-definite. Thus if we assume that $\nabla^2 r(u, v)$ and $D(\mathbf{u})$ commute, then $\nabla^2 r(u, v)D(\mathbf{u})$ is positive semi-definite and therefore $V'(t) \leq 0, t \geq 0$. Similarly, the application of a generalized version of Jensen inequality, [24], and convexity of r imply

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

This and (2.19) lead to

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

and V_r is a Lyapunov functional.

As far as the behaviour at infinity is concerned, the arguments in [23] can also be adapted here. We define $\mathbf{w} = \mathbf{u} - M\mathbf{u}_0$, where $\mathbf{u}$ is the solution of (2.1). If we assume If we assume grey level shift invariance then by the grey level shift invariance (2.18) $\mathbf{w}$ satisfies the diffusion equation of (2.1). By using the weak formulation (2.3), divergence theorem and the boundary conditions, we have

$$\frac{1}{2}\frac{d}{dt}\int_{\Omega}(w_1^2+w_2^2)d\Omega \;=\; -\int_{\Omega}[\nabla w_1 \nabla w_2]D\begin{bmatrix}\nabla w_1\\ \nabla w_2\end{bmatrix}d\Omega.$$

Now, (H1) and (2.6) imply

$$[\nabla w_1 \nabla w_2]D\begin{bmatrix}\nabla w_1\\ \nabla w_2\end{bmatrix} \geq \alpha\left(||\nabla w_1||_{L^2}^2 + ||\nabla w_2||_{L^2}^2\right).$$

Therefore

$$\frac{d}{dt}||\mathbf{w}||_{L^2\times L^2}^2 \leq -2\alpha||\nabla \mathbf{w}||_{L^2\times L^2}^2,$$

where

$$\begin{aligned}||\mathbf{w}||_{L^2\times L^2}^2 &= ||w_1||_{L^2}^2 + ||w_2||_{L^2}^2,\\ ||\nabla\mathbf{w}||_{L^2\times L^2}^2 &= ||\nabla w_1||_{L^2}^2 + ||\nabla w_2||_{L^2}^2.\end{aligned}$$

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

$$||\mathbf{w}||_{L^2\times L^2}^2 \leq C_0||\nabla \mathbf{w}||_{L^2\times L^2}^2.$$

This implies

$$\frac{d}{dt}||\mathbf{w}||_{L^2\times L^2}^2 \leq -2\alpha C_0||\mathbf{w}||_{L^2\times L^2}^2,$$

and by Gronwall's lemma

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

Thus we obtain the asymptotic behaviour

$$\lim_{t\to\infty}||\mathbf{u}(t) - M\mathbf{u}_0||_{X_0} = 0. \tag{2.21}$$

3. Numerical experiments

The performance of (2.1), (2.2) for image restoration problems is numerically illustrated in this section.

3.1. The numerical procedure. Some details of the implementation are first given. Three models of the form (2.1) will be considered, with a cross-diffusion matrix

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

where

$$g(v) = \frac{\cos\theta}{1 + \left(\frac{v}{k\theta}\right)^2},$$

(see (2.4)) with k, θ as parameters, [14], and d a positive definite matrix which will take one of the three forms:

$$d = \begin{pmatrix} \nu & -\mu \\ \mu & \nu \end{pmatrix}; \quad d = \begin{pmatrix} \nu & -\mu \\ \lambda & \nu \end{pmatrix}; \quad d = \begin{pmatrix} \nu & \mu \\ \lambda & \nu \end{pmatrix}, \tag{3.2}$$

for $\lambda, \mu, \nu \in \mathbb{R}, \lambda \neq \mu$. The first one in (3.2) corresponds to complex diffusion and will lead to a nonlinear complex diffusion filter (3.1), denoted by (NCDF1), in a similar form considered in (2.4). The other two choices of d in (3.2) will generate two different cross-diffusion matrices (3.1). The corresponding models will be denoted by (NCDF2) and (NCDF3) respectively. In order to simplify some of the parameters involved, the values $\lambda = \nu = 1$ are fixed and we take $\mu = \tan(\theta)$. It is not hard to see that for (3.1) conditions (H1)-(H3) are satisfied.

The semi-implicit numerical method introduced and analyzed in [5] for the complex diffusion case has been adapted here to approximate (2.1). (The convergence of this method for cross-diffusion models, as well as that of the implicit one also considered in [5] will be studied as future task elsewhere.) We make a brief description. By using the notation introduced in Section 2, we consider $\Omega = (0, N_1 - 1) \times (0, N_2 - 1)$, where N_1, N_2 are positive integers, and $Q_T = \Omega \times (0, T]$, with $T > 0$. Let us construct an equidistant rectangular grid on $\overline{Q_T}$. We define the space grid with mesh size $h = 1$ by

$$\begin{aligned} \overline{\Omega}_h \;=\; & \{x_{ij} \in \overline{\Omega} : x_{ij} = (i,j),\ i = 0, ..., N_1 - 1, \\ & j = 0, ..., N_2 - 1\}. \end{aligned}$$

For the temporal interval we consider the mesh

$$0 = t^0 < t^1 < \cdots < t^{M-1} < t^M = T,$$

where $M \geq 1$ is an integer and $\Delta t^m = t^{m+1} - t^m$, $m = 0, \ldots, M-1$. We denote by $\overline{Q}_h^{\Delta t}$ the mesh in $\overline{Q}$ defined by the cartesian product of the space grid $\overline{\Omega}_h$ and a grid in the temporal domain. Let $Q_h^{\Delta t} = \overline{Q}_h^{\Delta t} \cap Q$ and $\Gamma_h^{\Delta t} = \overline{Q}_h^{\Delta t} \cap \partial\Omega \times [0, T]$. For a real-valued function u^0 defined on $\overline{\Omega}_h$, representing the initial image, we define $U^0 = u^0(x_{ij})$, $x_{ij} \in \overline{\Omega}_h$, and V^0 as the null matrix of the same dimension as U^0. The restored image U^{m+1}, $m = 0, ..., M-1$, is obtained by the numerical method in $Q_h^{\Delta t}$

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

where ∇_h is the standard second order gradient approximation, completed with discrete Neumann boundary conditions. We note that, for the complex diffusion case, the stability condition is given by (see [4], [7])

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

This condition implies that $\Delta t \leq 0.2486$ for $\theta = \frac{\pi}{30}$ and $\Delta t \leq 0.1858$ for $\theta = \frac{7\pi}{30}$. This was taken into account here and $\Delta t = 0.05$ was used in the experiments below (with $k = 10$).

3.2. Numerical results. In order to study the quality of restoration of the three models (NCDF1), (NCDF2) and (NCDF3), we have performed numerical experiments by monitoring the corresponding approximate evolution of (2.1), given by the semi-implicit method, from several initial, noisy images and with different types of noise. The results shown here is a summary of the computations. Following [15], we consider the input U^0 as composed of the original image S and a Gaussian noise with zero mean and standard deviation σ, given by the Matlab function `randn`:

$$U^0 = S + \sigma * \texttt{randn}(\texttt{size}(U^0)). \tag{3.3}$$

With $V^0 = 0$ our aim is then to find a restored image $(U^m, V^m)^T$, $m = 1, 2, ..., M$, such that U^m approximates the original signal S at time level t^m. Three quality indexes are used at that time:

- Signal-to-Noise-Ratio (SNR):
$$SNR(S,U^m) = 10\log_{10}\left(\frac{\mathrm{var}(S)}{\mathrm{var}(U^m - S)}\right), \tag{3.4}$$
where the variance (var) of an image U is defined by
$$\mathrm{var}(U) = \frac{1}{N_1 N_2}\|U - \bar{U}\|_F^2,$$
$\|\cdot\|_F$ stands for the Frobenius norm, $\bar{U}$ is an uniform image with intensities equal to the mean value of the intensities of U, and $N_1 \times N_2$ is the dimension of U.
- Peak Signal-to-Noise-Ratio (PSNR):
$$PSNR(S,U^m) = 20\log_{10}\left(\frac{255}{RMSE(S,U^m)}\right), \tag{3.5}$$
where the Root-Mean-Square-Error ($RMSE$) is defined as
$$RMSE(S,U^m) = \frac{1}{\sqrt{N_1 N_2}}\|S - U^m\|_F^2;$$
- The correlation coefficient (CC) between the original and restored images:
$$CC(S,U^m) = \frac{\sum_{i,j}(S_{ij} - \bar{S})(U^m_{ij} - \bar{U}^m)}{\|S - \bar{S}\|_F \|U^m - \bar{U}^m\|_F}.$$

The experiment is concerned with an image S of Lena with $\sigma = 30$ in (3.3). In Table 1 the three quality indexes are computed at time $T = 2.5$ for three values of θ (corresponding to, say, small, medium and large values of μ). The results observed for the three indexes show a better performance of (NCDF1) and (NCDF2) against (NCDF3). If we compare these first two models, (NCDF1) looks more competitive for the smallest value of θ ($\mu \approx 1.051042 \times 10^{-1}$); this illustrates the influence of the small theta approximation, explained in [14] for complex diffusion and generalized for linear cross-diffusion in [2]. The results suggest that for (NCDF2) being competitive at that time, smallest values of the non-diagonal entries of the associated matrix d in (3.2) might be required. The performance of (NCDF2) improves in a relevant way when $\theta = 7\pi/30$ ($\mu \approx 9.004040 \times 10^{-1}$) and is better than that of (NCDF1) for $\theta = 10\pi/30$ ($\mu \approx 1.732051$). This confirms the loss of efficiency of complex diffusion out of the small theta approximation, [14], and for which models like (NCDF2) may appear as alternative.

The evolution of the filtering process is now analyzed. Figures 1-3 show the evolution of the two components of the cross-diffusion equations for the three models and from the initial noisy image, along with the corresponding histograms of the intensities for both components, when $\theta = \pi/30$. The results suggest some delay of (NCDF2) and (NCDF3) in the filtering with respect to (NCDF1). (Compare, for example, the first component of the image and the histogram corresponding to (NCDF1) at $T = 2.5$ with those of the other two at $T = 25$.) On the other hand, this has the advantage for the last two models of a delay in the blurring effect in both components. In the case of (NCDF2), this delay is less remarkable when θ increases, while (NCDF3) shows a more accentuated attenuation of blurring, see Figures 4-6. Finally, the images corresponding to $\theta = 10\pi/30$ (Figures 7-9) confirm the good behaviour of (NCDF2) suggested by Table 1 as well as the loss of performance of (NCDF3).

The delay mentioned above motivates Table 2, where the maximum values of the indexes and the time where they occur are displayed, for an evolution of the models up to $T = 25$. The effect of this delay looks remarkable in the case of $\theta = \pi/30$, where (NCDF2) and (NCDF3) improve their SNR value up to be comparable to that of (NCDF1) and overcome the PSNR value of this last one. The results confirm the best behaviour of (NCDF2) for $\theta = 10\pi/30$ and a relevant improvement of the performance of (NCDF3) for $\theta = 7\pi/30$ is also observed. Figures 10-12 show, for the three models, the time evolution up to $T = 25$ of the first two quality indexes with $\theta = \pi/30, 7\pi/30, 10\pi/30$.

4. Concluding remarks

In the present paper nonlinear cross-diffusion models for image filtering are studied. This is a continuation of a previous work, [2], devoted to the linear case. Here the nonlinear character is introduced by a cross-diffusion matrix satisfying some hypotheses. In the first part of the paper well-posedness 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 numerical 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 numerical study does not intent to be exhaustive and instead aims to suggest and anticipate some preliminary conclusions that may motivate further research about cross-diffusion as mathematical models for image processing. As in the linear case, the systems incorporate some degrees of freedom providing diversity and adaptability in the search for the best one for the restoration problem under study. This diversity is mainly represented by the choice of the cross-diffusion matrix and the decomposition of the initial noisy image into two components. While the numerical study shown here is focused on a specific type of cross-diffusion (with the goal of taking the known case of complex diffusion as a guide for comparison) this choice is the first point to be analyzed in a more detailed way in the future. Additionally, the decomposition of initial noisy image into two components shows, according to the theoretical properties of the models and the results concerning the linear case, a second point of analysis. Thirdly, although the numerical experiments performed to elaborate this paper implemented restoration problems with different types of noise and they did not show different information, some degree of adaptability of the models according to the noise involved cannot be dismissed and should be a thorough aspect of research. Finally, comparisons with existing models, especially of Perona-Malik type, are mandatory, [20].

On the other hand, the experiments evaluate in some sense the general philosophy of the decomposition of the information of the image in two components, one of the motivations to introduce these models. This was here explored for a restoration problem and is expected to be studied in edge-detection problems in a future research, with the presentation of cross-diffusion shock filters. A final task to be made in the future concerns the corresponding semi-discrete and fully discrete models that may be formulated. They are expected to be studied in similar terms, including additionally the convergence analysis when considered as schemes of approximation to a continuous problem.

5. Acknowledgments

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. Rational Mech. Anal. 123, 199–257 (1993).

[4] Araújo, A., Barbeiro, S., Serranho, P.: Stability of finite difference schemes for complex diffusion processes. SIAM J. Numer. Anal. 50, 1284–1296 (2012).

[5] Araújo, A., Barbeiro, S., Serranho, P.: Convergence of finite difference schemes for nonlinear complex reaction-diffusion processes. SIAM J. Numer. Anal. 53(1), 228–250 (2015).

[6] Aubert, J., Kornprobst, P.: Mathematical Problems in Image Processing, Springer-Verlag, 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] 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).

[11] 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).

[12] 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).

[13] Gilboa, G., Sochen, N. A., Zeevi, Y. Y.: Generalized shock filters and complex diffusion. In: Computer Vision-ECCV, pp. 399-413, Springer, Berlin (2002).

[14] Gilboa, G., Sochen, N. A., Zeevi, Y. Y.: Image enhancement and denoising by complex diffusion processes, IIEEE Trans. Pattern Anal. Mach. Intell.. 26(8), 1020–1036 (2004).

[15] Gilboa, G., Sochen, N. A., Zeevi, Y. Y.: Estimation of optimal PDE-based denoising in the SNR sense. IEEE Trans. Image Processing. 15(8), 2269–2280 (2006).

[16] Guichard, F., Morel, J. M.: Image Analysis and P.D.E.'s. IPAM GBM Tutorials, 2001. http://www.ipam.ucla.edu/publications/gbm2001/gbmtut jmorel.pdf.

[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] Lindeberg, T.: Scale Space. In: Encyclopedia of Computer Science and Engineering (B. Wah ed.), Vol IV, pp. 2495–2509, John Willey and Sons, Hoboken New Jersey (2009).

[20] Lorenz, D. A., Bredies, K., Zeevi, Y. Y.: Nonlinear complex and cross diffusion. Unpublished report, University of Bremen, 2006.

[21] Ni, W.-M.: Diffusion, cross-diffusion and their spike-layer steady states. Notices AMS. 45(1) 9-18 (1998).

[22] Perona, P., Malik, J.: Scale–space and edge detection using anisotropic diffusion. IEEE Trans. on Pattern Anal. and Mach. Intell. 12 629–639 (1990).

[23] Weickert, J.: Anisotropic Diffusion in Image Processing. B.G. Teubner, Stuttgart (1998).

[24] Zabandan, G., Kiliman, A.: A new version of Jensen's inequality and related results. J. Inequalities and Applications. 238 (2012).

Table 1. Lena: Signal-to-Noise Ratio, Peek-Signal-to-Noise Ratio and the Correlation Coefficient at $T = 2.5$.

	$NCDF1$			$NCDF2$			$NCDF3$		
	SNR	PSNR	CC	SNR	PSNR	CC	SNR	PSNR	CC
$\theta = \pi/30$	13.58	28.11	0.9782	8.96	23.49	0.9397	8.74	23.27	0.9368
$\theta = 7\pi/30$	13.47	28.00	0.9775	13.51	28.04	0.9776	10.08	24.61	0.9514
$\theta = 10\pi/30$	13.03	27.57	0.9749	13.23	27.77	0.9761	3.96	18.50	0.8381

Table 2. Lena: Maximum values of Signal-to-Noise Ratio, Peek-Signal-to-Noise Ratio and the Correlation Coefficient.

		$NCDF1$	
	SNR	PSNR	CC
$\theta = \pi/30$	14.08 ($t = 1.5$)	28.62 ($t = 1.5$)	0.9804 ($t = 1.55$)
$\theta = 7\pi/30$	13.98 ($t = 1.5$)	28.51 ($t = 1.5$)	0.9798 ($t = 1.5$)
$\theta = 10\pi/30$	13.20 ($t = 1.9$)	27.73 ($t = 1.9$)	0.9758 ($t = 1.95$)
		$NCDF2$	
	SNR	PSNR	CC
$\theta = \pi/30$	13.55 ($t = 16.4$)	28.08 ($t = 16.4$)	0.9778 ($t = 16.65$)
$\theta = 7\pi/30$	13.96 ($t = 1.55$)	28.49 ($t = 1.55$)	0.9797 ($t = 1.6$)
$\theta = 10\pi/30$	13.61 ($t = 1.6$)	28.14 ($t = 1.6$)	0.9780 ($t = 1.6$)
		$NCDF3$	
	SNR	PSNR	CC
$\theta = \pi/30$	13.52 ($t = 19.7$)	28.05 ($t = 19.7$)	0.9777 ($t = 20$)
$\theta = 7\pi/30$	12.47 ($t = 20.9$)	27.01 ($t = 20.9$)	0.9726 ($t = 23.6$)
$\theta = 10\pi/30$	6.65 ($t = 0.5$)	21.18 ($t = 0.5$)	0.9040 ($t = 0.45$)

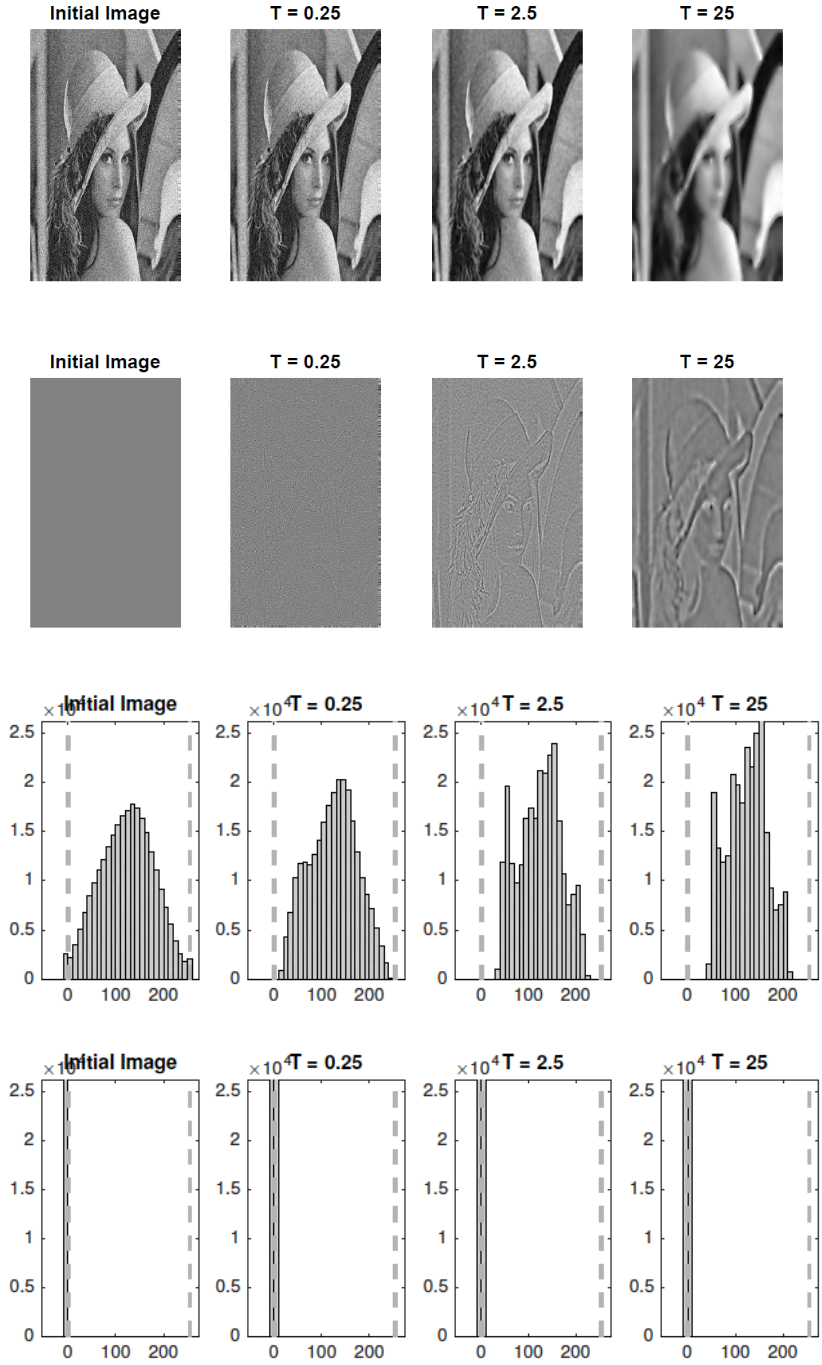

FIGURE 1. Original and restored images for (NCDF1) with $\theta = \pi/30$ (top) and their histograms of intensities (bottom).

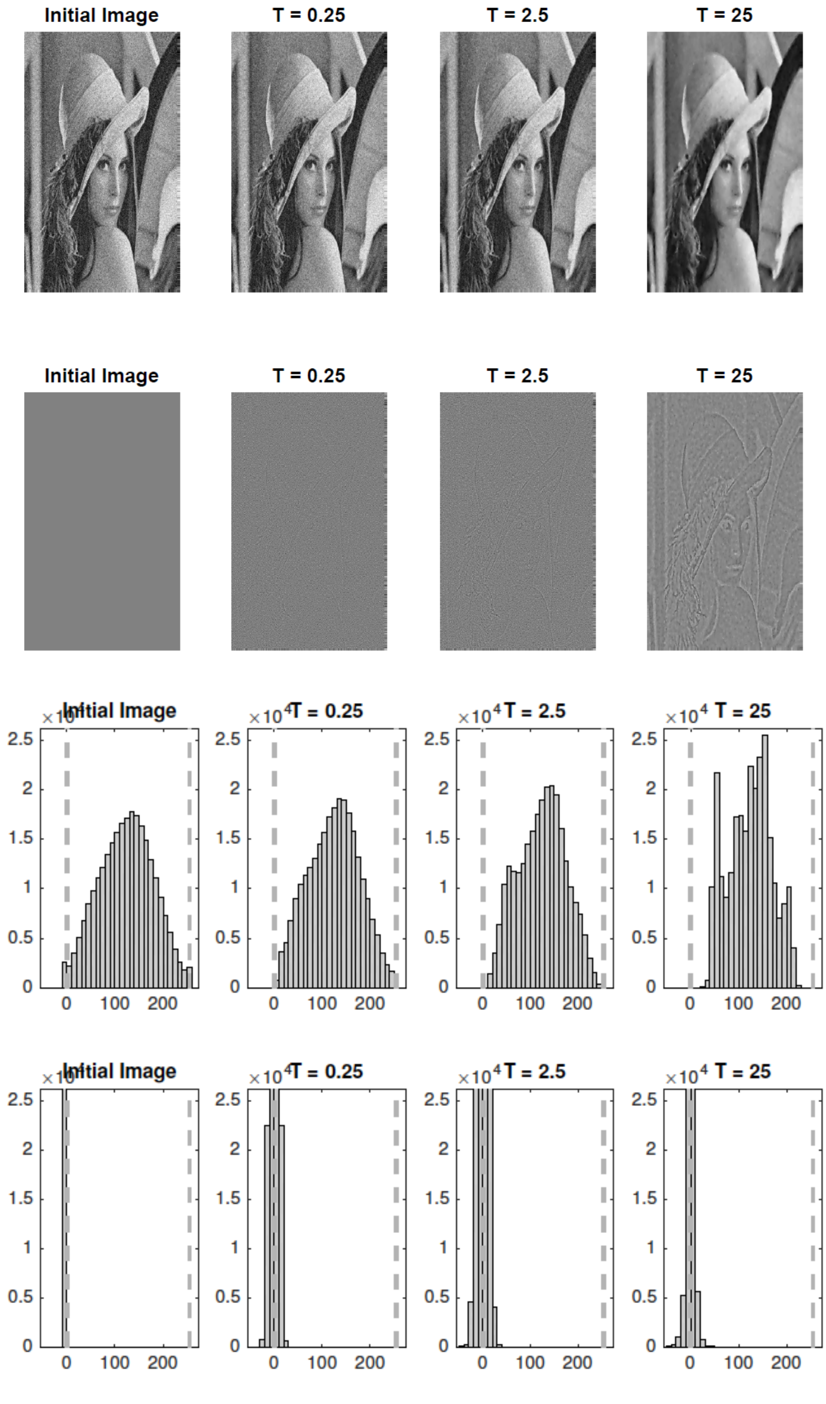

FIGURE 2. Original and restored images for (NCDF2) with $\theta = \pi/30$ (top) and their histograms of intensities (bottom).

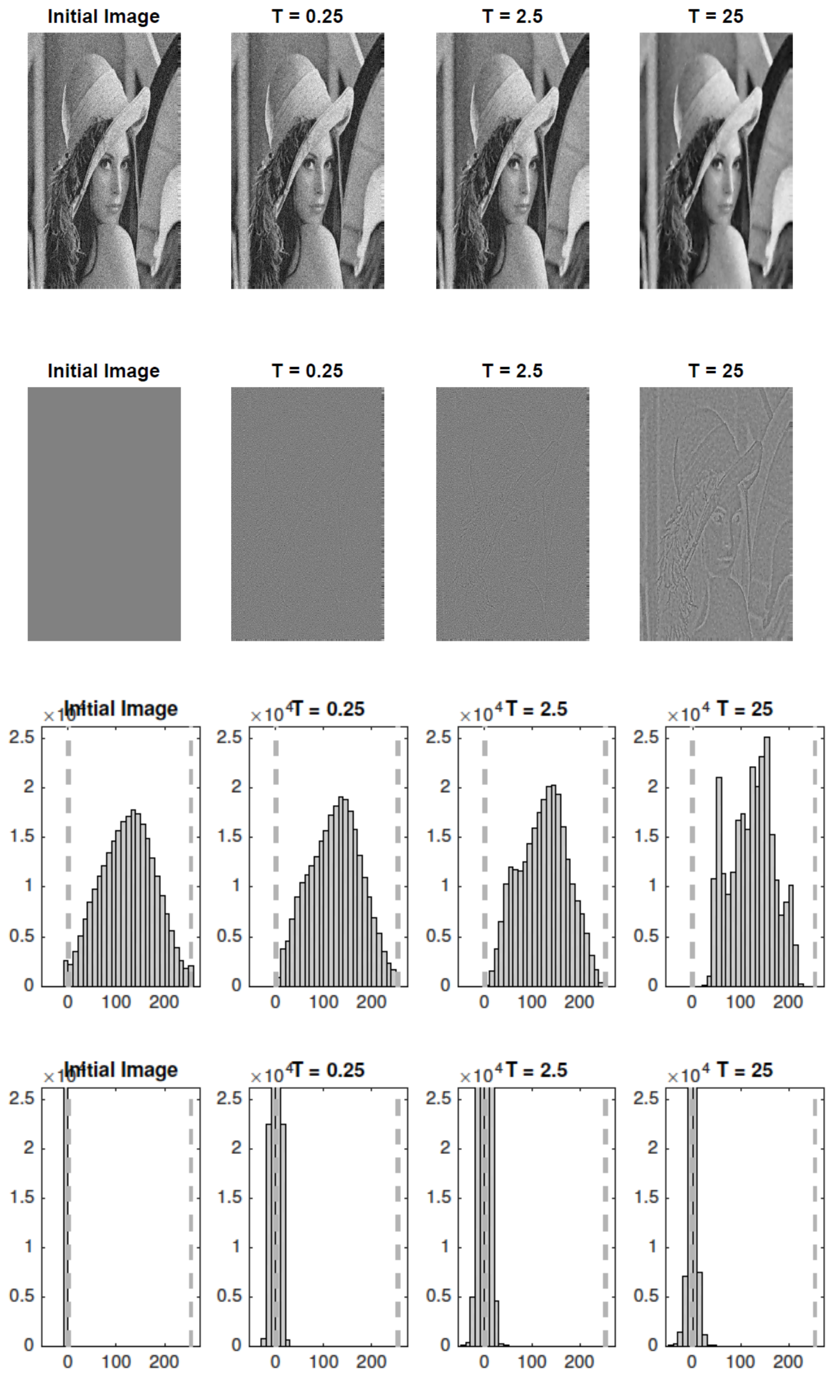

FIGURE 3. Original and restored images for (NCDF3) with $\theta = \pi/30$ (top) and their histograms of intensities (bottom).

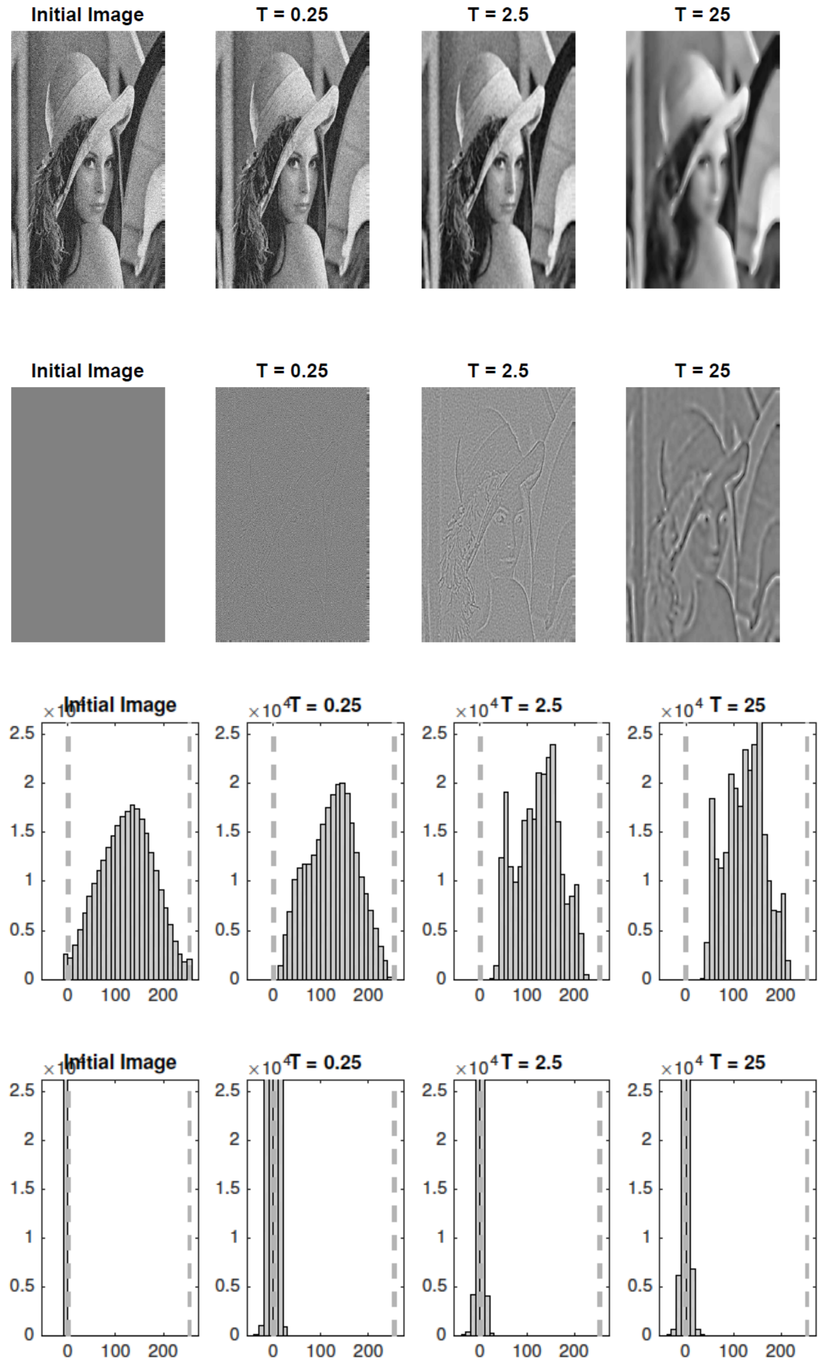

FIGURE 4. Original and restored images for (NCDF1) with $\theta = 7\pi/30$ (top) and their histograms of intensities (bottom).

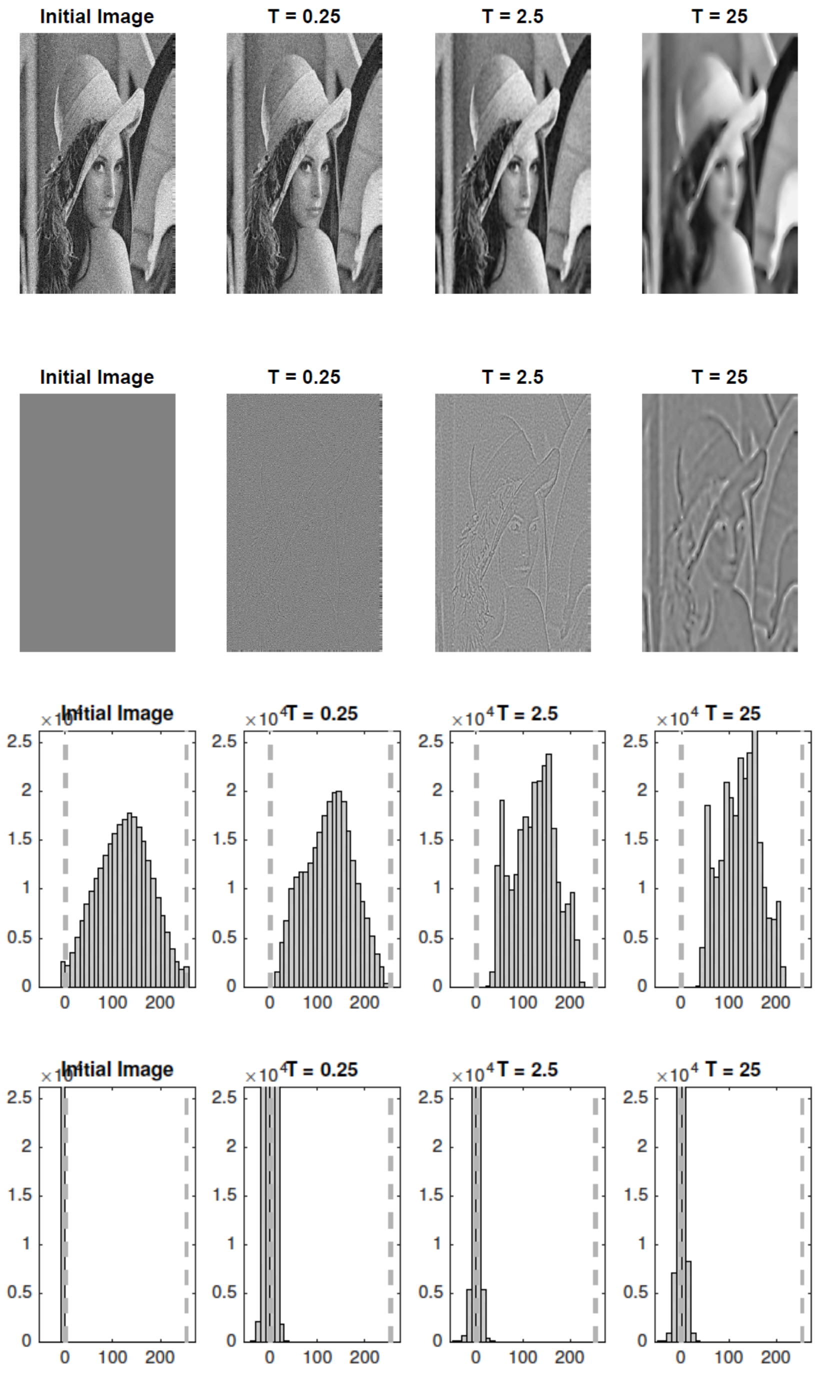

FIGURE 5. Original and restored images for (NCDF2) with $\theta = 7\pi/30$ (top) and their histograms of intensities (bottom).

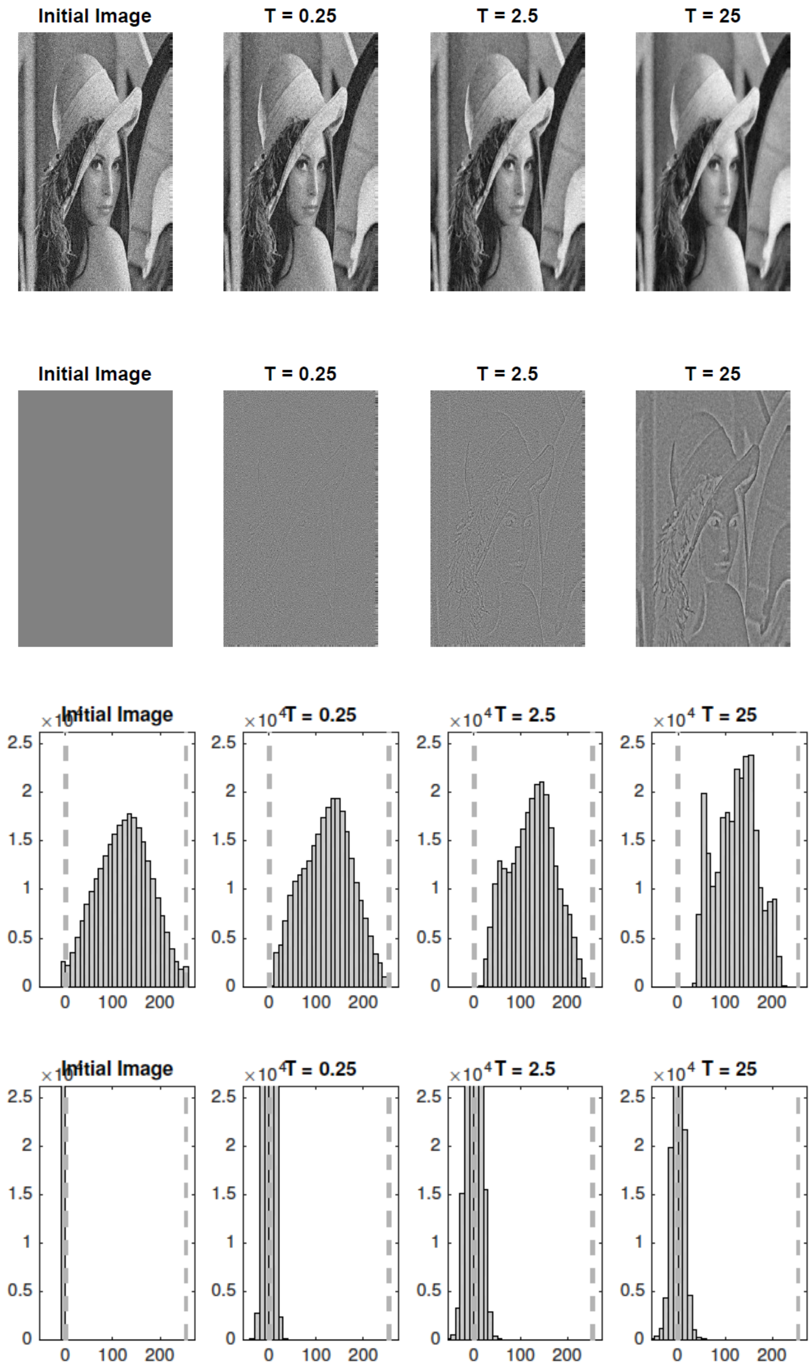

FIGURE 6. Original and restored images for (NCDF3) with $\theta = 7\pi/30$ (top) and their histograms of intensities (bottom).

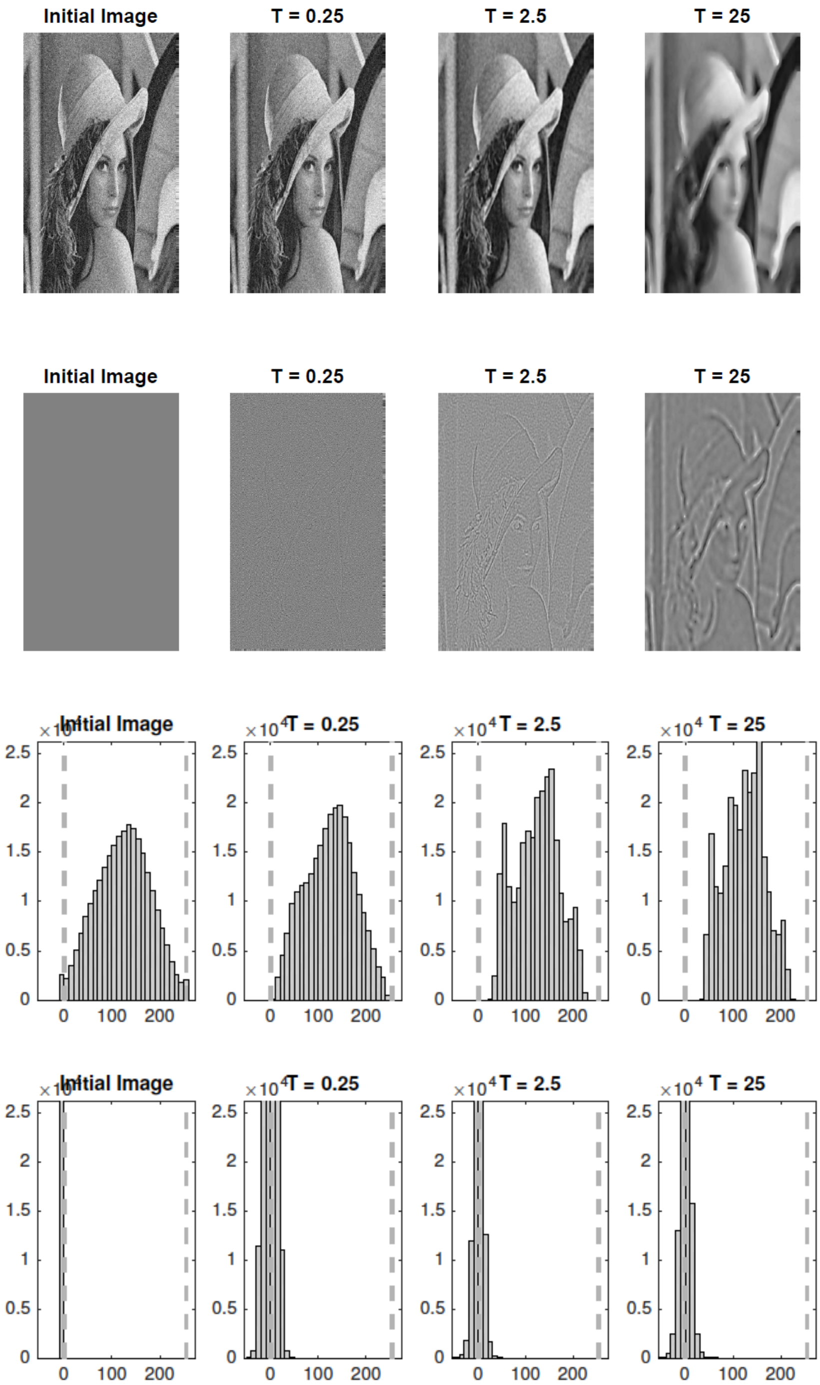

FIGURE 7. Original and restored images for (NCDF1) with $\theta = 10\pi/30$ (top) and their histograms of intensities (bottom).

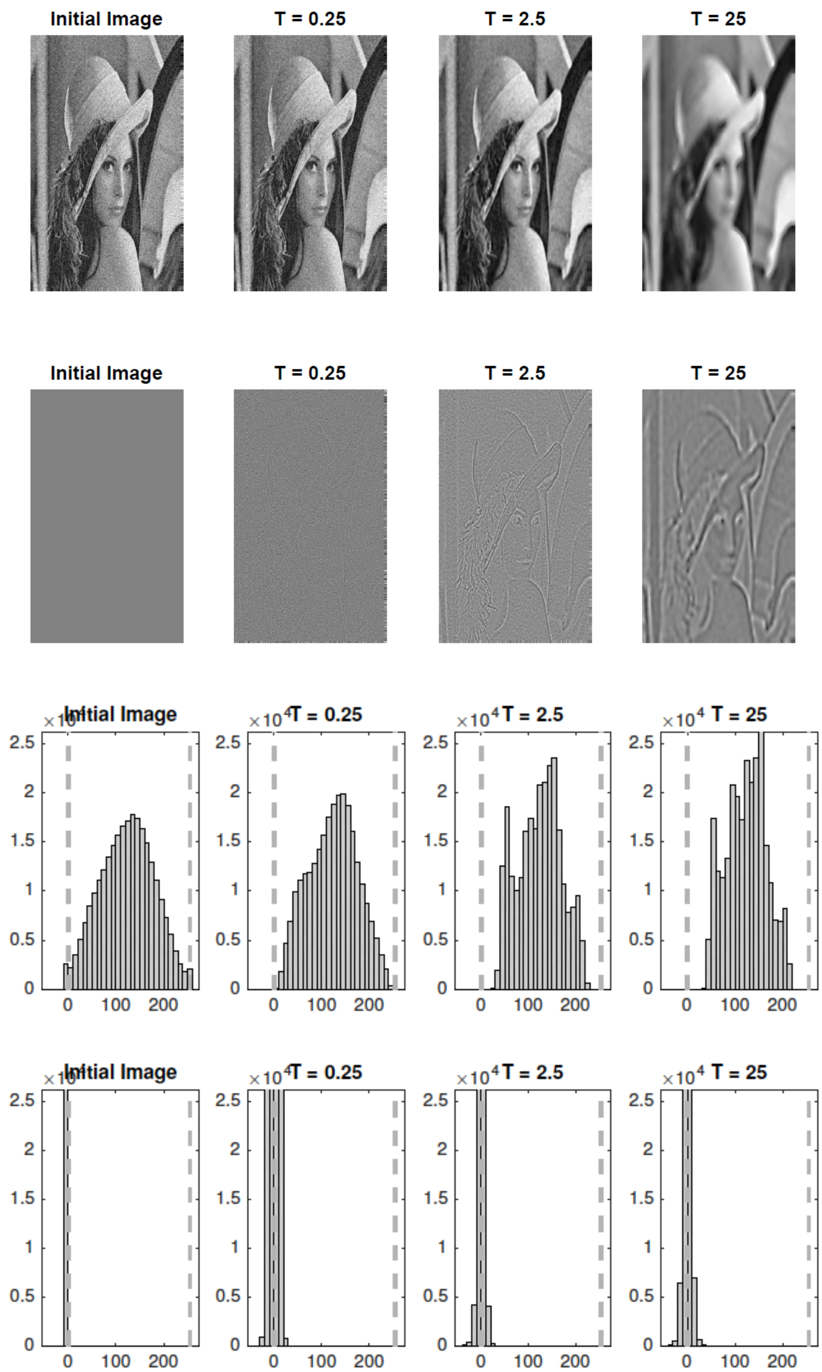

FIGURE 8. Original and restored images for (NCDF2) with $\theta = 10\pi/30$ (top) and their histograms of intensities (bottom).

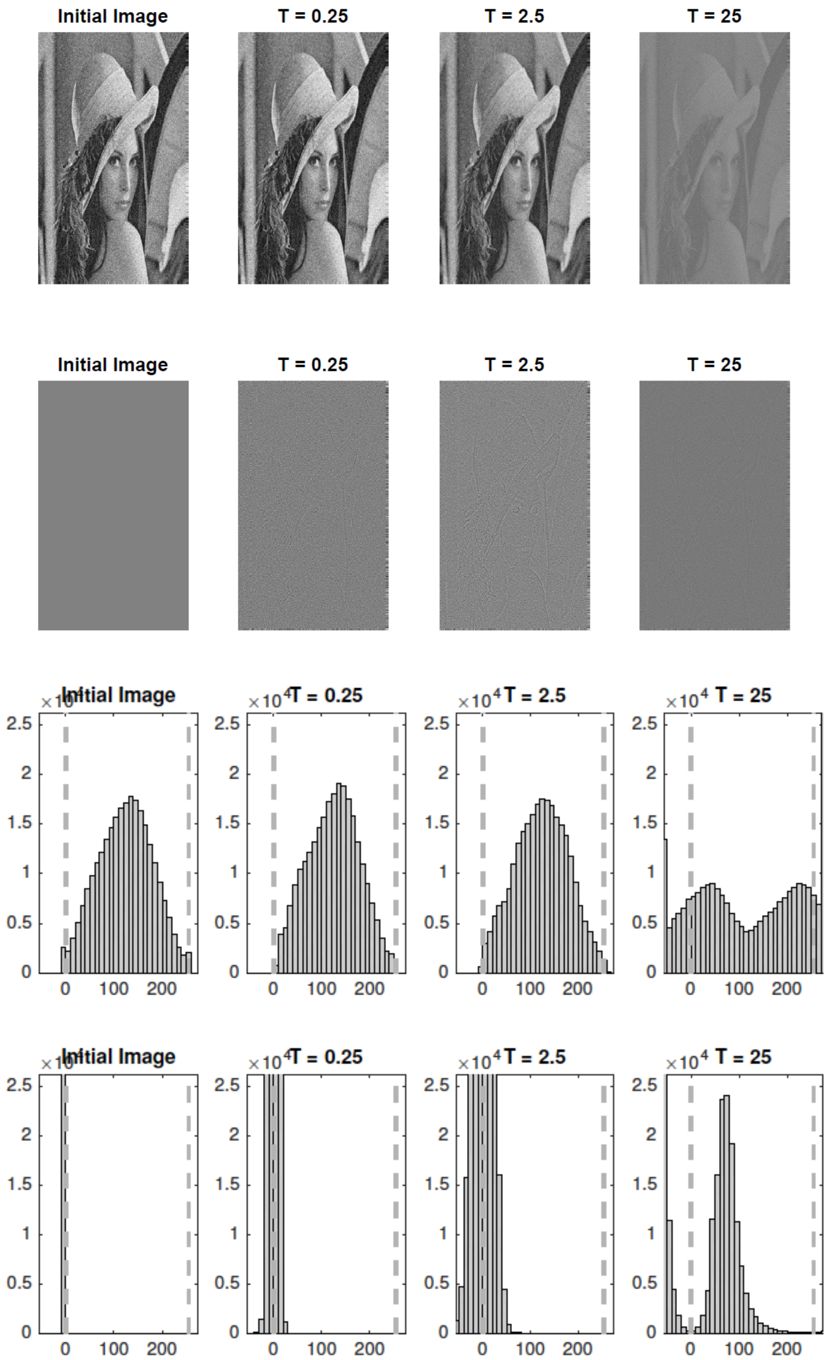

FIGURE 9. Original and restored images for (NCDF3) with $\theta = 10\pi/30$ (top) and their histograms of intensities (bottom).

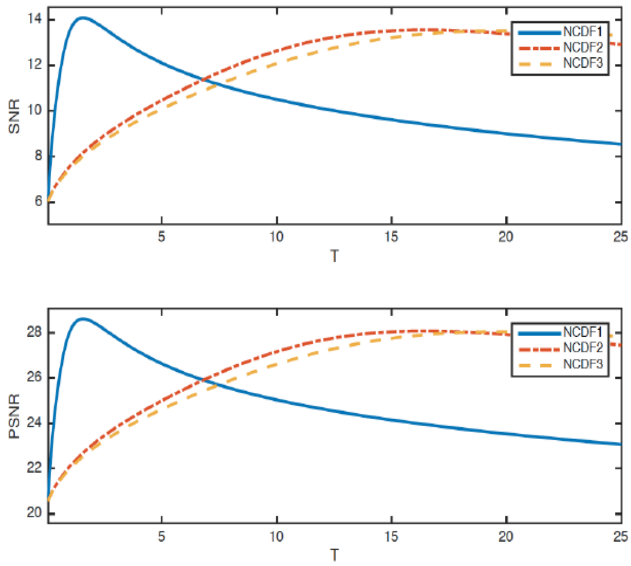

FIGURE 10. Time evolution of quality indexes for the restored images with the three models for $\theta = \pi/30$.

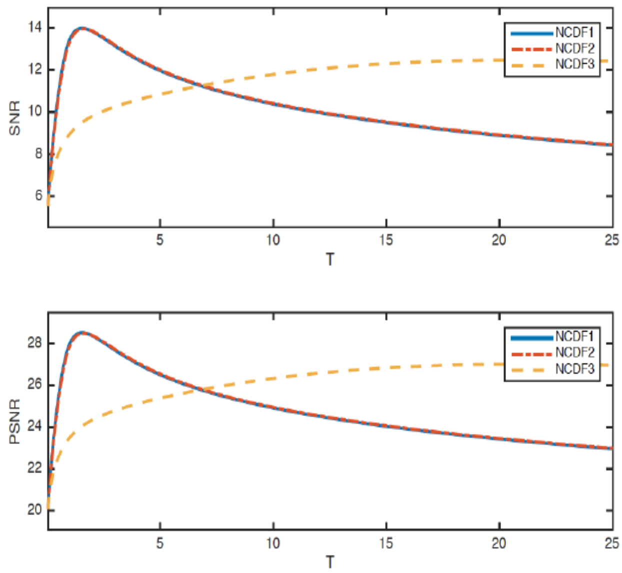

FIGURE 11. Time evolution of quality indexes for the restored images with the three models for $\theta = 7\pi/30$.

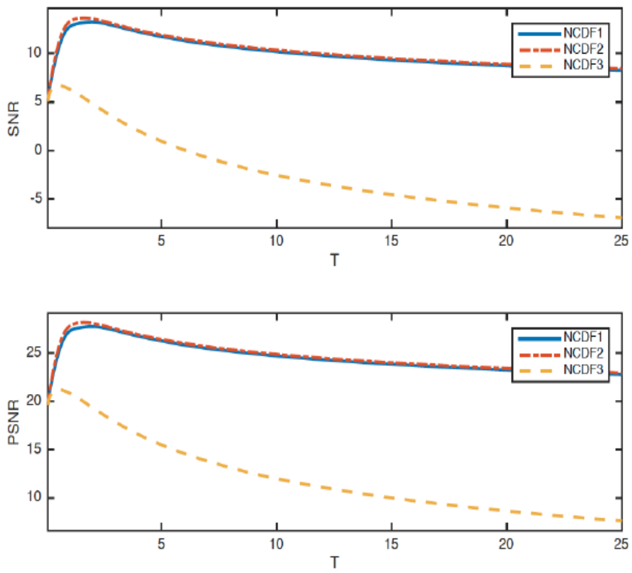

FIGURE 12. Time evolution of quality indexes for the restored images with the three models for $\theta = 10\pi/30$.

ADÉRITO ARAÚJO
CMUC, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF COIMBRA, PORTUGAL
E-mail address: alma@mat.uc.pt

SÍLVIA BARBEIRO
CMUC, DEPARTMENT OF MATHEMATICS, UNIVERSITY OF COIMBRA, PORTUGAL
E-mail address: silvia@mat.uc.pt

EDUARDO CUESTA
DEPARTMENT OF APPLIED MATHEMATICS, UNIVERSITY OF VALLADOLID, SPAIN
E-mail address: eduardo@mat.uva.es

ÁNGEL DURÁN
DEPARTMENT OF APPLIED MATHEMATICS, UNIVERSITY OF VALLADOLID, SPAIN
E-mail address: angel@mac.uva.es