Math. Model. Nat. Phenom. 13 (2018) 34
https://doi.org/10.1051/mmnp/2018014

Mathematical Modelling of Natural Phenomena
www.mmnp-journal.org

SHEARLET-BASED REGULARIZED RECONSTRUCTION IN REGION-OF-INTEREST COMPUTED TOMOGRAPHY

T.A. Bubba[1,*], D. Labate[2], G. Zanghirati[3] and S. Bonettini[4]

Abstract. Region of interest (ROI) tomography has gained increasing attention in recent years due to its potential to reducing radiation exposure and shortening the scanning time. However, tomographic reconstruction from ROI-focused illumination involves truncated projection data and typically results in higher numerical instability even when the reconstruction problem has unique solution. To address this problem, both *ad hoc* analytic formulas and iterative numerical schemes have been proposed in the literature. In this paper, we introduce a novel approach for ROI tomographic reconstruction, formulated as a convex optimization problem with a regularized term based on shearlets. Our numerical implementation consists of an iterative scheme based on the scaled gradient projection method and it is tested in the context of fan-beam CT. Our results show that our approach is essentially insensitive to the location of the ROI and remains very stable also when the ROI size is rather small.

Mathematics Subject Classification. 44A12, 68T60, 65K10, 65F22, 68U10, 92C55

Received November 30, 2017. Accepted January 17, 2018.

1. Introduction

Computed tomography (CT) is a noninvasive imaging technique designed to visualize the internal structure of a body or an object without surgical intervention or destructive material testing. The impact of CT has been enormous in different areas, including industrial nondestructive testing, security screening and medical diagnostics.

Region of interest (ROI) tomography aims at the reconstruction of a particular area of interest within a field of view while the rest of the image is neglected. Since this operation would require to irradiate only a region within the field of view, it is possible in principle to speed up data collection by shortening scanning time, and to reduce potentially harmful X-ray radiation exposure. As a result ROI tomography has received increasing attention in recent years in applications including biomedical imaging, *e.g.*, contrast-enhanced cardiac imaging and positioning of intracranial stents, and industrial testing.

Keywords and phrases: Computed tomography, region-of-interest reconstruction, shearlets, wavelets, gradient projection methods.

[1] Department of Mathematics and Statistics, University of Helsinki, Gustaf Hällströmin katu 2B, Helsinki 00014, Finland.
[2] Department of Mathematics, University of Houston, 651 Phillip G. Hoffman, Houston, TX 77204-3008, USA.
[3] Deptartment of Mathematics and Computer Science, University of Ferrara, and INdAM-GNCS, via G. Saragat 1, Ferrara 44122, Italy.
[4] Department of Physics, Computer Science and Mathematics, University of Modena and Reggio Emilia, and INdAM-GNCS, via G. Campi 213/B, Modena 41125, Italy.

* Corresponding author: tatiana.bubba@helsinki.fi

CT reconstruction consists in recovering the unknown density function of an object by collecting *projections* from multiple views, *i.e.*, by measuring the attenuation of X-rays propagated through the object. This problem is mathematically ill-posed in general, and the ill-posedness is more severe when projections are truncated [31], as in the case of ROI-focused irradiation. It is known that the direct application of classical reconstruction algorithms such as the Filtered Back-Projection (FBP) or the FDK algorithms [32] with ROI-truncated projections and the missing projection data set to zero can produce serious visual artifacts and numerical instability. These undesirable effects become worse as the size of the ROI decreases.

During the last decade, a number of analytic and algebraic methods have been proposed to address the problem of ROI reconstruction from truncated projections [8]. While this problem is not guaranteed to have unique solution in general [31], it was shown that it is possible to derive analytic ROI reconstruction formulas from truncated data under restrictive assumptions on the location of the ROI or *a priori* information [9, 34, 43]. For example, the Differentiated Back-Projection (DBP) method [10, 35, 41] can be applied if there exists at least one projection view in which the complete (*i.e.*, non-truncated) projection data are available. Algebraic or iterative methods, on the other hand, are generally more flexible, since they can be applied to essentially any type of acquisition mode. They are usually computationally more demanding, even if advances in high-performance computing make algebraic methods more and more competitive [3, 22]. Some common iterative algorithms, which have been adapted to ROI tomography, are the simultaneous iterative reconstruction technique (SIRT) [26], the maximum likelihood expectation-maximization algorithm (MLEM) [36] and the least-squares Conjugate Gradient (LSCG) method [27]. The performance of such reconstruction methods is usually rather sensitive to the ROI size and the presence of noise. Other contributions includes a Bayesian multiresolution method for local tomography reconstruction in dental X-ray imaging [33], where the authors use a wavelet basis for the representation of the dental structures, with high resolution inside the ROI and coarser resolution outside the ROI. This approach is closely related to the one more recently proposed in [28]. A wavelet-based regularization algorithm based on iterative reconstruction-reprojection is proposed in [1], where a smoothing convolution operator for the re-projecting phase is included. In all these approaches some form of prior knowledge on the object attenuation function, or a regularization term, is used to ensure a stable ROI reconstruction. Other regularization-based approaches, which do not rely on multiscale methods, can be found in [24, 40, 42].

In this paper, we address the ROI CT reconstruction problem using an iterative algorithm formulated as a convex optimization problem with a regularized functional based on shearlets, a multiscale methods that is especially designed to sparsely approximate images with edges [13] and, thus, is expected to provide an efficient representation for the images found in common CT applications. We recall that shearlet-based approaches for CT reconstruction were recently proposed in [16, 37], for the non-truncated case.

The regularized approach we consider in this paper is inspired by the reconstruction recently proposed in [17] that combines a data fidelity constraint to ensure that reconstructed projections match the observed data inside the ROI and a data consistency condition to enforce that solutions are consistent with the tomographic reconstruction problem. We adopt the same setting and propose two new formulations of the regularized reconstruction problem that we call implicit and explicit formulations. In addition to using a shearlet-based regularization, another novelty of this paper is that our numerical implementation is based on the scaled gradient projection method (SGP) [4] – an accelerated first-order descent method for convex and non-convex objective functions. This method is particularly effective when the feasible region is given by "simple" constraints, as it is in our case where the projection onto a feasible region has a simple implementation. The numerical tests reported in this paper show that our algorithm produces accurate ROI reconstructions for any ROI location and also for ROI sizes that are small with respect to the field of view.

The paper is organized as follows: in Section 2, we introduce the problem formulation in both the continuous and the discrete settings, and derive two possible formulations of the objective function. We also recall some background about shearlets and the SGP algorithm. In Section 3 we recall the distance-driven method [12] that we use to simulate acquisition. Numerical experiments and results are given in Section 4. Concluding remarks are in Section 5. A preliminary version of our results was recently presented in a conference paper [5].

2. A SGP-shearlet approach for ROI tomography

The framework we introduce below applies to different projection geometries, including parallel and fan-beam, and extends to higher dimensions. However, for brevity, we only consider here the 2D case. The higher dimensional case is similar.

2.1. The Radon transform and the ROI solution problem

The CT reconstruction problem consists in reconstructing a density function from a set of projections, obtained by measuring the attenuation of X-rays over straight lines. Mathematically, it can be formulated as a "line-integral model" through the Radon transform. Given a function $f \in L^1(\mathbb{R}^2)$, the *Radon transform* of f at (θ, τ) is the line integral of f over the lines (or rays) $\ell(\theta,\tau)$ perpendicular to $\mathbf{e}_\theta = \big(\cos(\theta), \sin(\theta)\big)^T \in \mathbb{S}^1$ with (signed) distance $\tau \in \mathbb{R}$ from the origin:

$$\mathcal{R}f(\theta,\tau) = \int_{\ell(\theta,\tau)} f(\mathbf{x})\,\mathrm{d}\mathbf{x} = \int_{\mathbb{R}^2} \delta(\tau - \mathbf{x}\cdot \mathrm{e}_\theta)\, f(\mathbf{x})\,\mathrm{d}\mathbf{x}, \tag{2.1}$$

where $\ell(\theta,\tau) = \{\mathbf{x} \in \mathbb{R}^2 : \mathbf{x}\cdot\mathbf{e}_\theta = \tau\}$. Therefore, the Radon transform maps f into the set of its linear projections defined on the tangent space $\mathcal{T} = \{(\theta,\tau) : \theta \in [0, 2\pi),\ \tau \in \mathbb{R}\}$. In what follows, we will refer to the Radon projections as the *full sinogram* and we shall denote it by:

$$y(\theta,\tau) = \mathcal{R}f(\theta,\tau), \quad \theta \in [0,2\pi),\ \tau \in \mathbb{R}. \tag{2.2}$$

We will refer to the domain of (2.2) as the *projection domain* and to the domain of the density function as the *image domain*.

In the ROI tomographic problem, *projections are only collected for those rays meeting a region-of-interest* inside the field of view. The goal is to recover the density function inside the ROI, while the rest of the function is essentially ignored. Denoting the ROI as $S \subset \mathbb{R}^2$, the set of ROI-truncated projections is identified as the set

$$\mathcal{P}(S) = \{(\theta,\tau) \in \mathcal{T} : \ell(\theta,\tau) \cap S \neq \emptyset\} \subset \mathcal{T}. \tag{2.3}$$

Thus, the ROI reconstruction problem can be formulated as the problem of reconstructing the function f restricted to the ROI S from the truncated Radon projections:

$$y_0(\theta,\tau) = M(\theta,\tau)\,\mathcal{R}f(\theta,\tau), \quad M(\theta,\tau) = 1_{\mathcal{P}(S)}(\theta,\tau), \tag{2.4}$$

where M is a *mask function* selecting only the rays passing through the ROI, and 1_A is the indicator function of the set A. We will refer to y_0 as the *truncated sinogram*. In the following, we will assume that the ROI is a disk $S \subset \mathbb{R}^2$ with center $\mathbf{p}_{\mathrm{ROI}} \in \mathbb{R}^2$ and radius $R_{\mathrm{ROI}} \in \mathbb{R}$. Thus, $\mathcal{P}(S) = |\tau - \mathbf{p}_{\mathrm{ROI}}\cdot\mathbf{e}_\theta| < R_{\mathrm{ROI}}$. This is not a restrictive assumption as more general ROIs can be handled by finding the minimal enclosing disk for this ROIs.

A standard approach for obtaining a reconstruction of f from equation (2.4) is by computing the least squares solution $\widehat{f}$

$$\widehat{f} = \arg\min_f \|M\mathcal{R}f - y_0\|_2^2. \tag{2.5}$$

However, the minimizer of this problem is not unique in general. In fact, it is known that the solution of the ROI problem is not guaranteed to be unique in general [31]. Even when the uniqueness is ensured, the inversion of the Radon transform is an ill-posed problem, and the ill-posedness may be more severe when the projections are truncated, as in the case of the ROI CT problem. A classical approach to achieve uniqueness is by using Tikhonov

regularization, which imposes an additional norm condition by searching for the minimum-norm solution. This norm condition can be applied in the image or in the projection domain, leading to possibly different solutions.

To derive a stable formulation of the ROI reconstruction problem we combine a Tikhonov-like regularization with the following observation from [17]. Notice that

$$y = y_0 + (1 - M)\, y, \tag{2.6}$$

showing that the ROI reconstruction problem can be viewed as an extrapolation problem on y_0. Clearly, we cannot choose *any* extrapolated function y outside $\mathcal{P}(S)$ but a global constraint is needed to ensure that y belongs to the range of the Radon transform of the density function $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$:

$$Rf = y. \tag{2.7}$$

By applying M and $1 - M$ to the left-hand and right-hand sides of (2.7), respectively, we hence obtain (cf. [17]):

$$M\mathcal{R}f = M\, y = y_0, \quad \text{(data fidelity)} \tag{2.8}$$
$$(1 - M)\mathcal{R}f = (1 - M)\, y. \quad \text{(data consistency)} \tag{2.9}$$

Equation (2.8) defines a constraint inside the ROI, forcing *fidelity* to the truncated sinogram, while equation (2.9) imposes a constraint on the extrapolation scheme, asking for *consistency* with the (hypothetical) full sinogram. In Section 2.2 we show how to combine these equations with regularization to obtain a computationally suitable algorithm for the ROI CT problem.

2.2. Explicit and implicit discrete formulations of the objective function

To derive a discrete formulation for the ROI CT reconstruction problem, we start by discretizing equations (2.8) and (2.9). We shall denote by K the number of projection angles and by P the number of detector cells (*i.e.*, the samples along the detector array); N is both the width and the height in pixels of the object to reconstruct. Equation (2.2) can be understood as a Fredholm integral equation of the first kind, where the kernel $\delta(\tau - \mathbf{x} \cdot \mathbf{e}_\theta)$ describes how to map the image domain into the projection domain (namely, the projection geometry) can be discretized by means of a matrix $\mathbf{W}$ which has dimensions $KP \times N^2$. The mask identifying the ROI is the diagonal matrix $\mathbf{M}$ of dimensions $KP \times KP$ whose entries are either 0 or 1. The unknown discrete density function f to be reconstructed is represented as a vector $\mathbf{f}$ of length N^2 in which the entries are stacked column by column. Similarly, the full sinogram y and the truncated sinogram y_0 are represented as vectors $\mathbf{y}$ and $\mathbf{y}_0$, respectively, of length KP, obtained by stacking the entries column-wise. We recall that $\mathbf{y}$ and $\mathbf{y}_0$ are related to each other by the data fidelity equation

$$\mathbf{y}_0 = \mathbf{M}\mathbf{y} = \mathbf{M}\mathbf{W}\mathbf{f}, \tag{2.10}$$

and the data consistency equation, that sets the extrapolation scheme outside the ROI, reads as

$$(\mathbf{I}_{KP} - \mathbf{M})\mathbf{W}\mathbf{f} = (\mathbf{I}_{KP} - \mathbf{M})\mathbf{y}, \tag{2.11}$$

where $\mathbf{I}_{KP}$ is the $KP \times KP$ identity matrix. As indicated above, data fidelity and data consistency equations need to be coupled with a regularization term to yield a unique solution.

Rather than requiring the exact equalities stated by the equations (2.10) and (2.11), we shall minimize the L^2-norm error associated to them, according to a maximum likelihood approach [39]. By exploiting this idea, we can state two different optimization problems and objective functions. On the one hand, we can consider the L^2-norm error of the data fidelity equation only and incorporate the data consistency information in the

regularization term. In this case, the only variable we are minimizing on is the image $\mathbf{f}$ to be reconstructed. This approach yields:

$$\widehat{\mathbf{f}} = \arg\min_{\mathbf{f}\in\Omega_{\mathbf{f}}} \mathbf{\Psi}(\mathbf{f}),$$

where

$$\mathbf{\Psi}(\mathbf{f}) = \frac{1}{2}\|\mathbf{MWf} - \mathbf{y}_0\|_2^2 + \lambda\,\|\mathbf{\Phi}((\mathbf{I}_{KP} - \mathbf{M})\mathbf{Wf} + \mathbf{y}_0)\|_2^2, \tag{2.12}$$

where $\mathbf{\Phi}$ is a discrete filter which will be discussed in more details in Section 2.3. We will refer to (2.12) as the *implicit formulation*, since the full sinogram $\mathbf{y}$ does not appear explicitly.

On the other hand, we can consider both the data fidelity and the data consistency L^2-norm errors into the objective function to assess if, by explicitly incorporating the extrapolation scheme in the minimization problem, we achieve a more accurate reconstruction. In this case, the image $\mathbf{f}$ to be reconstructed and the full sinogram $\mathbf{y}$ are both unknown. This second approach yields:

$$\left(\widehat{\mathbf{f}}, \widehat{\mathbf{y}}\right) = \arg\min_{\substack{\mathbf{f}\in\Omega_{\mathbf{f}} \\ \mathbf{y}\geq\mathbf{0}}} \mathbf{\Psi}(\mathbf{f}, \mathbf{y}), \tag{2.13}$$

where

$$\mathbf{\Psi}(\mathbf{f}, \mathbf{y}) = \frac{1}{2}\|\mathbf{MWf} - \mathbf{y}_0\|_2^2 + \frac{1}{2}\|(\mathbf{I}_{KP} - \mathbf{M})(\mathbf{Wf} - \mathbf{y})\|_2^2 + \lambda\,\|\mathbf{\Phi}((\mathbf{I}_{KP} - \mathbf{M})\mathbf{y} + \mathbf{y}_0)\|_2^2. \tag{2.14}$$

In contrast with the above formulation, we will refer to (2.14) as the *explicit formulation*. Notice that each term of the objective function is convex with respect to the unknowns. In both cases, λ denotes the regularization parameter and the feasible region $\Omega_{\mathbf{f}}$ is either defined as $\mathbf{f} \geq 0$ (nonnegativity constraint) or – when the maximum pixel value L of the image happens to be known (for instance, in the case of synthetic data) – as $0 \leq \mathbf{f} \leq L$ (box constraint). In both cases the inequalities are meant component-wise. Clearly, we can always assume to be in the case of box constraints, and the limit case $L = +\infty$ recovers the nonnegativity constraint case. The case $L < +\infty$ has mostly a numerical impact, not as much for the computational cost but rather on the sequence generated by the algorithm since we are "narrowing" the feasible region.

We will also include a smoothed total variation (sTV) term [38] in the objective function to control oscillatory artifacts in the numerical solution. In conclusion, our objective functions are

$$\mathbf{\Psi}(\mathbf{f}) = \frac{1}{2}\|\mathbf{MWf} - \mathbf{y}_0\|_2^2 + \lambda\,\|\mathbf{\Phi}((\mathbf{I}_{KP} - \mathbf{M})\mathbf{Wf} + \mathbf{y}_0)\|_2^2 + \rho\,\mathrm{TV}_\delta(\mathbf{f}), \tag{2.15}$$

for the implicit formulation and

$$\mathbf{\Psi}(\mathbf{f}, \mathbf{y}) = \frac{1}{2}\|\mathbf{MWf} - \mathbf{y}_0\|_2^2 + \frac{1}{2}\|(\mathbf{I}_{KP} - \mathbf{M})(\mathbf{Wf} - \mathbf{y})\|_2^2 + \lambda\,\|\mathbf{\Phi}((\mathbf{I}_{KP} - \mathbf{M})\mathbf{y} + \mathbf{y}_0)\|_2^2 + \rho\,\mathrm{TV}_\delta(\mathbf{f}), \tag{2.16}$$

for the explicit one. Here, ρ is a regularization parameter and δ is the TV smoothing parameter. sTV was chosen since it is widely used in medical imaging. In the following, we will investigate if using the Tikhonov-like regularization term is superior to the sTV minimization and/or to the combined minimizations (2.15) and (2.16).

We remark that, in the absence of noise, the solution $\widehat{\mathbf{f}}$ is the same for both the implicit and the explicit cases (up to a vector in the nullspace of $\mathbf{W}$). However, in the presence of noise, the explicit formulation should benefit from the presence of the norm $\|\mathbf{Wf} - \mathbf{y}\|_2^2$ that is in general non-zero, forcing a more accurate reconstruction

inside the ROI. Thus, in the following, we will compare the implicit and explicit formulations, to asses which yields more accurate reconstructions.

2.3. Regularization term: discrete shearlet transform

In the Section above, the term $\mathbf{\Phi}$ was intentionally left unspecified. The filter $\mathbf{\Phi}$ is obtained from a suitable discrete shearlet transform by getting rid of (or *thresholding*) the 90% of its smallest coefficients. To better clarify, we briefly recall below the main ideas about shearlets, a multiscale method which refines the conventional wavelet framework by combining multiscale analysis and directional sensitivity [19, 29].

Let $\phi \in L^2(\mathbb{R}^2)$. A 2D affine family generated by ϕ is a collection of functions of the form:

$$\Big\{\phi_{M,t}(x) = |\det(M)|^{-\frac{1}{2}}\phi\Big(M^{-1}(x-t)\Big) \ : \ M \in G, t \in \mathbb{R}^2\Big\}, \tag{2.17}$$

where G is a subset of the group $GL_2(\mathbb{R})$ of invertible 2×2 matrices. The function ϕ is called a 2D continuous wavelet if

$$g(x) = \int_G \int_{\mathbb{R}^2} \langle g, \, \phi_{M,t}\rangle \, \phi_{M,t}(x) \, \mathrm{d}t \, d\lambda(M), \quad \forall \, g \in L^2(\mathbb{R}^2), \tag{2.18}$$

where $\lambda(M)$ is an suitable measure on G. The corresponding 2D continuous wavelet transform of g is the mapping

$$g \mapsto \mathcal{W}g(M,t) = \langle g, \, \phi_{M,t}\rangle, \quad M \in G, t \in \mathbb{R}^2. \tag{2.19}$$

Discrete wavelet transforms are obtained by discretizing $\mathcal{W}g(M,t)$ on an appropriate grid. Usually, $M \in G$ and $t \in \mathbb{R}^2$ are replaced by A^j, $j \in \mathbb{Z}$, and $k \in \mathbb{Z}^2$, respectively, for appropriate choices of the matrix A. Similarly to wavelets, *continuous shearlets* are affine systems of the form:

$$\Big\{\phi_{a,s,t}(x) = |\det(M_{a,s})|^{-\frac{1}{2}}\phi\Big(M_{a,s}^{-1}(x-t)\Big) \ : \ M \in \Gamma, t \in \mathbb{R}^2\Big\}, \tag{2.20}$$

where $\phi \in L^2(\mathbb{R}^2)$. To construct shearlets, $\Gamma = \left\{M_{a,s} = \begin{pmatrix} a & s\sqrt{a} \\ 0 & \sqrt{a} \end{pmatrix} \ : \ a \in \mathbb{R}^+, s \in \mathbb{R}\right\}$ is the subset of $GL_2(\mathbb{R})$ comprising the matrices $M_{a,s}$ obtained by multiplying an *anisotropic dilation* matrix A_a with a *shear* matrix S_s:

$$M_{a,s} = \begin{pmatrix} a & s\sqrt{a} \\ 0 & \sqrt{a} \end{pmatrix} = \begin{pmatrix} a & 0 \\ 0 & \sqrt{a} \end{pmatrix}\begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix} =: A_a S_s. \tag{2.21}$$

Here the variables $a \in \mathbb{R}^+$, $s \in \mathbb{R}$ and $t \in \mathbb{R}^2$ denote the scale, orientation and the spatial location, respectively. Thus, continuous shearlets are formed by (anisotropically) dilating, shearing and translating an appropriate mother shearlet function $\phi \in L^2(\mathbb{R}^2)$ [21], while continuous wavelets are obtained by (isotropically) dilating and translating. Roughly speaking, shearlets are well localized waveforms whose orientation is controlled by the shear parameter s and that become increasingly elongated at fine scales (as $a \to 0$). The *continuous shearlet transform* of g is the map

$$g \mapsto \mathscr{SH}g(a,s,t) = \langle g, \, \phi_{a,s,t}\rangle, \quad a \in \mathbb{R}^+, s \in \mathbb{R}, t \in \mathbb{R}^2. \tag{2.22}$$

Discrete shearlets are formally defined by sampling continuous shearlets on an appropriate discrete set. In particular, we choose

$$\left\{\phi_{j,k,m}(x) = 2^{\frac{3}{4}j}\phi\Big(S_k A_{2^j} x - m\Big) \;:\; j,k \in \mathbb{Z}, m \in \mathbb{Z}^2\right\} \quad \text{with } \phi \text{ s. t.} \quad \mathcal{F}(\phi)(\omega) = \mathcal{F}(\phi_1)(\omega_1)\,\mathcal{F}(\phi_2)\left(\frac{\omega_2}{\omega_1}\right), \tag{2.23}$$

where $\omega = (\omega_1, \omega_2)$, $\mathcal{F}(\phi_1)$ is the Fourier transform of a wavelet function with compact support away from the origin and $\mathcal{F}(\phi_2)$ is a compactly supported bump function with $\mathrm{supp}(\mathcal{F}(\phi_2)) \subset [-1, 1]$. Such mother shearlet ϕ is called *classical mother shearlet*. Notice that there exist several choices for ϕ_1 and ϕ_2 satisfying the classical mother shearlet definition. One possible choice is to set ϕ_1 to be a Lemariè–Meyer wavelet and ϕ_2 to be a spline bump function. Starting from a classical mother shearlet, and under other suitable assumptions (see [20]), one can obtain a tight frame for $L^2(\mathbb{R}^2)$, and hence an invertible decomposition.

Summarizing, shearlets exhibit very appealing mathematical properties: they are well localized, namely they are compactly supported in the frequency domain and have fast decay in the spatial domain; the parabolic scaling facilitates the handling of anisotropic structures; highly directional sensitivity is provided by the shearing parameter. Thanks to these properties, shearlets provide optimally sparse approximations with images containing C^2-edges, outperforming conventional wavelets [19]. This is potentially relevant in CT-like applications, since point-like structures in the image domain map onto sine-shaped curvilinear structures in the projection domain.

In the following, after constructing the band-limited shearlet system above, we will threshold the 90% of the smallest (in absolute value) shearlet coefficients $\{\langle f,\, \phi_{j,k,m}\rangle\}_{j,k,m}$. The ratio for this is that the smallest coefficients are usually connected to noise, while the largest are associated to edges. By getting rid of the smallest ones, we are able to separate the structural components of the object from the noise. Hence, $\mathbf{\Phi}$ is the filter resulting from the combination of the thresholding operation and the shearlet transform.

2.4. Scaled gradient projection method

In this section, we describe our algorithm for the solution of the optimization problems (2.15)–(2.16). A step-by-step description is given in Algorithm 2.1, for the case of the implicit formulation. A similar step-by-step description can be derived for the explicit case, as discussed at the end of this section.

The proposed method, called *scaled gradient projection* (SGP) method [4], is an iterative approach that belongs to a family of first-order descent methods and is defined for convex (and non-convex), differentiable and constrained problems with "simple" feasible regions. This is indeed our case since the feasible region $\Omega_{\mathbf{f}}$ is usually either a box or a non-negativity constraint and, as indicated above, our functionals are convex and differentiable.

When the objective function reads as in (2.15), the $(k+1)$th iteration is given by

$$\mathbf{f}^{(k+1)} = (1-\lambda_k)\,\mathbf{f}^{(k)} + \lambda_k\,\mathcal{P}_{\Omega_{\mathbf{f}}}\Big(\mathbf{f}^{(k)} - \alpha_k D_k \nabla\mathbf{\Psi}(\mathbf{f}^{(k)})\Big), \quad k = 0, 1, 2, \ldots, \tag{2.24}$$

where λ_k, α_k are suitable step lengths, D_k is an appropriate scaling matrix, and $\mathcal{P}_{\Omega_{\mathbf{f}}}$ is the projection onto the feasible region. The main feature of this method consists in the combination of non-expensive diagonally scaled gradient directions with step length selection rules specially designed for these directions. In details, it is allowed any choice of the step length α_k in a closed interval $[\alpha_{\min}, \alpha_{\max}] \subset \mathbb{R}^+$ and of the scaling matrix D_k among the set of the symmetric and positive definite scaling matrix such that all its eigenvalues are contained in an interval $\left[\frac{1}{L_k}, L_k\right]$ where $L_k > 1$ and $\sum_{k=0}^{\infty}(L_k^2 - 1) < \infty$. Such a degree of freedom is very important from a practical point of view because it makes the updating rules problem-related and performance-aware. Additionally, global convergence properties are ensured by exploiting a nonmonotone line search strategy along the feasible direction [11, 18]. Indeed, the nonmonotone line search strategy, as implemented in step 5 of Algoritm 2.1, ensures that $\mathbf{\Psi}(\mathbf{f}^{(k+1)})$ ($\mathbf{\Psi}(\mathbf{f}^{(k+1)}, \mathbf{y}^{(k+1)})$, respectively) is lower than the maximum of the objective

function on the last μ iterations by successive reductions of λ_k; of course, if $\mu = 1$ then the strategy reduces to the standard monotone Armijo rule.

Algorithm 2.1. Scaled gradient projection method.

Choose the starting point $\mathbf{f}^{(0)} \in \Omega_{\mathbf{f}}$, set the parameters $\beta, \tau \in (0,1)$ and $0 < \alpha_{\min} < \alpha_{\max}$.
Fix a positive integer μ.
for $k = 0, 1, 2, \ldots$ **do**
  Step 1. Choose the parameter $\alpha_k \in [\alpha_{\min}, \alpha_{\max}]$ and the scaling matrix $D_k \in \mathcal{D}_{L_k}$
  Step 2. Projection: $\mathbf{z}^{(k)} = \mathcal{P}_{\Omega_f}\left(\mathbf{f}^{(k)} - \alpha_k D_k \nabla \boldsymbol{\Psi}(\mathbf{f}^{(k)})\right)$
    if $\left(\mathbf{z}^{(k)} = \mathbf{f}^{(k)}\right)$ **then** stop: $\mathbf{f}^{(k)}$ is a stationary point; **end if**
  Step 3. Descent direction: $\mathbf{d}^{(k)} = \mathbf{z}^{(k)} - \mathbf{f}^{(k)}$
  Step 4. Set $\lambda_k = 1$ and $\boldsymbol{\Psi}_{\max} = \max_{0 \le j \le \min(k,\mu-1)} \boldsymbol{\Psi}(\mathbf{f}^{(k-j)})$
  Step 5. Backtracking loop:
    while $\left(\boldsymbol{\Psi}(\mathbf{f}^{(k)} + \lambda_k \mathbf{d}^{(k)}) > \boldsymbol{\Psi}_{\max} + \beta \lambda_k \nabla \boldsymbol{\Psi}(\mathbf{f}^{(k)})^T \mathbf{d}^{(k)}\right)$ **do**
      $\lambda_k \leftarrow \tau \lambda_k$
    end while
  Step 6. Set $\mathbf{f}^{(k+1)} = \mathbf{f}^{(k)} + \lambda_k \mathbf{d}^{(k)}$
end for

Our implementation of SGP is equipped with an adaptive step length selection based on the Barzilai-Borwein (BB) updating rules [2, 14]. In practice, by means of a variable threshold, one of the two different selection strategies

$$\alpha_k^{\mathrm{BB1}} = \arg\min_{\alpha_k \in \mathbb{R}} \left\| B(\alpha_k)\, \mathbf{s}^{(k-1)} - \zeta^{(k-1)} \right\|,$$
$$\alpha_k^{\mathrm{BB2}} = \arg\min_{\alpha_k \in \mathbb{R}} \left\| \mathbf{s}^{(k-1)} - B(\alpha_k)^{-1} \zeta^{(k-1)} \right\|,$$

is selected [15], where $B(\alpha_k) = (\alpha_k D_k)^{-1}$ approximates the Hessian matrix $\nabla^2 \boldsymbol{\Psi}(\mathbf{f}^{(k)})$, $\mathbf{s}^{(k-1)} = \mathbf{f}^{(k)} - \mathbf{f}^{(k-1)}$ and $\zeta^{(k-1)} = \nabla \boldsymbol{\Psi}(\mathbf{f}^{(k)}) - \nabla \boldsymbol{\Psi}(\mathbf{f}^{(k-1)})$. As far as the scaling matrix concerns, D_k is chosen according to a split-gradient idea [30] based on the decomposition of the gradient of the objective function into a positive and a negative part. The updating rule for each entry $d_i^{(k)}$ is:

$$d_i^{(k)} = \min\left\{ L_k, \max\left\{ \frac{1}{L_k}, \frac{\mathbf{f}_i^{(k)}}{V_i(\mathbf{f}_i^{(k)})} \right\} \right\}, \quad i = 1, \ldots, KP \tag{2.25}$$

where V_i is the ith component of the positive part V of the gradient decomposition. The role of D_k is to improve the convergence rate without introducing significant computational costs. We can setup the SGP method for the case of two unknowns, as in equation (2.16), by applying the SGP iteration to the "enlarged" variable $(\mathbf{f}, \mathbf{y})$. This is accomplished by tacking a block-diagonal scaling matrix: the block corresponding to the image $\mathbf{f}$ reads as described above, while for the block corresponding to the full sinogram $\mathbf{y}$ we take $D_k = \mathbf{I}_{KP}$ for each $k = 0, 1, 2, \ldots$. That is, the $(k+1)$th iteration, with $k = 0, 1, 2, \ldots$, is

$$\begin{pmatrix} \mathbf{f}^{(k+1)} \\ \mathbf{y}^{(k+1)} \end{pmatrix} = (1 - \lambda_k) \begin{pmatrix} \mathbf{f}^{(k)} \\ \mathbf{y}^{(k)} \end{pmatrix} + \lambda_k \begin{pmatrix} \mathcal{P}_{\Omega_{\mathbf{f}}}\left(\mathbf{f}^{(k)} - \alpha_k D_k \nabla_{\mathbf{f}} \boldsymbol{\Psi}(\mathbf{f}^{(k)}, \mathbf{y}^{(k)})\right) \\ \mathcal{P}_{\mathbf{y} \ge 0}\left(\mathbf{y}^{(k)} - \alpha_k \nabla_{\mathbf{y}} \boldsymbol{\Psi}(\mathbf{f}^{(k)}, \mathbf{y}^{(k)})\right) \end{pmatrix}. \tag{2.26}$$

The step length selection is still performed as for the implicit case, but $\mathbf{s}^{(k-1)}$ and $\zeta^{(k-1)}$ are now the following "enlarged" variables:

$$\mathbf{s}^{(k-1)} = \begin{pmatrix} \mathbf{f}^{(k)} - \mathbf{f}^{(k-1)} \\ \mathbf{y}^{(k)} - \mathbf{y}^{(k-1)} \end{pmatrix}, \quad \zeta^{(k-1)} = \begin{pmatrix} \nabla_{\mathbf{f}} \boldsymbol{\Psi}(\mathbf{f}^{(k)}, \mathbf{y}^{(k)}) - \nabla_{\mathbf{f}} \boldsymbol{\Psi}(\mathbf{f}^{(k-1)}, \mathbf{y}^{(k-1)}) \\ \nabla_{\mathbf{y}} \boldsymbol{\Psi}(\mathbf{f}^{(k)}, \mathbf{y}^{(k)}) - \nabla_{\mathbf{y}} \boldsymbol{\Psi}(\mathbf{f}^{(k-1)}, \mathbf{y}^{(k-1)}) \end{pmatrix}. \tag{2.27}$$

Finally, we notice that the SGP algorithm can be also used as an iterative regularization method applied to the un-regularized functional by means of an early stopping technique.

3. Data simulation

To demonstrate and validate our approach, we use both simulated and real data. While for the real data, consisting of a 2D cross-section of a real 3D walnut (see Sect. 4), the measured data available at www.fips.fi/dataset [23], includes also the system matrix $\mathbf{W}$, for the simulated data set, consisting of the Shepp-Logan phantom, we construct our own system matrix $\mathbf{W}$ by simulating the geometry of a μCT scanner used for real measurements. In particular, we simulate a 2D fan-beam data from 182 uniformly spaced angles over 2π. The detector consists of 130 elements with a pixel pitch of 0.8 mm. The distance between the X-ray tube and the detector is set to 291.20 mm and the radius of rotation is 115.84 mm. The detector is offset by 1.5 pixels. Such 2D fan-beam geometry is implemented by using a state-of-the-art technique called distance-driven method [12]. For completeness we report in Section 3.1 a brief description of this method.

3.1. Distance-driven method

We recall that $\mathbf{W}$ is the $KP \times N^2$ matrix mapping the image domain onto the projection domain. Every row of $\mathbf{W}$ contains weights that relate the pixel intensities in the image domain to the corresponding sample in the projection domain. Obviously, the value of each weight depends on the chosen interpolation scheme. The distance-driven method combines a highly sequential memory access pattern with relatively low arithmetic complexity, without introducing interpolation artifacts in the image or projection domains. Essentially, this approach is based on converting the projection problem into a 1D re-sampling problem. There are two main ingredients in the distance-driven method. The first one is the *kernel operation*:

$$b_n = \sum_j w_j c_j \quad \text{with} w_j = \frac{\left[\min(\xi_{m+1}, \upsilon_{n+1}) - \max(\xi_m, \upsilon_n)\right]_+}{\upsilon_{n+1} - \upsilon_n}, \quad [x]_+ = \max(x, 0), \tag{3.1}$$

which allows to compute the destination signal values $\{b_j\}_j$ from the sample values $\{c_i\}_i$ of a source signal, the sample source locations $\{\xi_i\}_i$ and the sample destination locations $\{\upsilon_j\}_j$. The second element of the method is that there is a (possibly zero) length of overlap between each image pixel and each detector cell due to the bijection between the position on the detector and the position within an image row (or column). Thus, every point within an image row is uniquely mapped onto a point on the detector, and vice versa. In practice, to compute the overlap length, all pixel boundaries in an image row and all detector cell boundaries are mapped onto a common line, *e.g.*, a line parallel to a coordinate axis. That is, each length of overlap is the interval length between two adjacent intersections, obtained by connecting the boundaries midpoints of all detector cells and pixels in a image row to the X-ray source and by computing the intercepts of these lines with the common axis. The final weights are achieved by normalizing the overlap length by the detector cell width. This corresponds exactly to apply the kernel operation from equation (3.1). In our case, b_n is the theoretical (unblurred and noiseless) value measured at the nth detector cell and c_j is the estimate of the jth pixel attenuation function. For example, if the nth detector cell is "shadowed" by only two pixels in a row, equation (3.1) reads as:

$$b_n = \frac{\xi_{m+1} - \upsilon_n}{\upsilon_{n+1} - \upsilon_n} c_m + \frac{\upsilon_{n+1} - \xi_{m+1}}{\upsilon_{n+1} - \upsilon_n} c_{m+1}. \tag{3.2}$$

TABLE 1. Optimal results for the Shepp-Logan phantom with the *implicit* formulation of the objective function. The corresponding reconstructed images are reported in Figure 1. SH = Shearlets; sTV = smoothed Total Variation.

		ROI Rel. Err.	ROI PSNR	Iter	Sec	Params
	SH+sTV	0.16	29.57	93	9.02	$\lambda = 10^{-4}$, $\rho = 10^{-1}$
$\gamma = 0.5\,\mathrm{N}$	SH	0.17	28.93	101	9.29	$\lambda = 10^{-3}$
	sTV	0.16	29.60	92	5.36	$\rho = 10^{-1}$
	SH+sTV	0.13	32.44	107	9.11	$\lambda = 10^{-3}$, $\rho = 1$
$\gamma = 0.25\,\mathrm{N}$	SH	0.21	27.94	109	9.40	$\lambda = 10^{-4}$
	sTV	0.14	31.41	103	6.39	$\rho = 1$
	SH+sTV	0.18	33.37	118	10.04	$\lambda = 10^{-4}$, $\rho = 1$
$\gamma = 0.15\,\mathrm{N}$	SH	0.48	24.69	156	13.65	$\lambda = 10^{-4}$
	sTV	0.19	33.05	90	5.22	$\rho = 10^{-1}$

An efficient implementation of the distance-driven method can be achieved by using a vector-oriented approach instead of processing one pixel at a time. Indeed, one can update an entire row of the matrix (*i.e.*, selecting one detector cell at a fixed angular view) in one shot by identifying *all* those pixels that shadow that detector cell and their corresponding length of overlap. Moreover, the detector can be chosen as the common line hence avoiding the need to calculate the projections of the detector cell boundaries. Finally, when the common line is not the detector, an effective choice is to adaptively select the common line by mapping onto the x-axis or y-axis, depending on the angular view.

4. Numerical experiments

In this section, we present numerical demonstrations of our ROI CT reconstruction algorithm in the setting of 2D fan-beam geometry. All the algorithms were implemented in Matlab 8.1.0 and all experiments performed on a dual CPU server, equipped with two 6-cores Intel Xeon X5690 at 3.46 GHz, 188 GB DDR3 central RAM memory and up to 12 TB of disk storage.

For our first set of tests we considered a modified Shepp-Logan phantom sized $N \times N$ pixels with $N = 128$. We assumed it to be placed in the first quadrant of the Cartesian coordinate system with the image lower left corner at the origin. All projections are corrupted by a white Gaussian process with zero mean and 0.5% variance, and the truncated projections are computed by discarding the samples outside the ROI projection $\mathcal{P}(S)$. We recall that this corresponds to a ROI disk in the image domain (see Sect. 2.1). We considered concentric ROI disks with a decreasing *radius* γ, placed off-center with respect to the field of view. Our results include ROIs that are fully inside the object being imaged (corresponding to the interior tomography problem). We also considered a ROI with a larger *radius* exceeding the object along one coordinate axis for comparison. The matrix $\mathbf{\Phi}$ was generated by the *classical shearlet* decomposition (Sect. 2.3, Eq. (2.23)), by using the shearlet software available at www.math.uh.edu/~dlabate/software. The number of scales for the shearlet transform is 3, with 2^3 directions and 16×16 for the support size of the shearing filters.

We performed numerical tests to assess the best approach with respect to both the objective function and the regularization parameters. Tables 1 and 2 show results for the implicit and the explicit formulations, respectively, where perfomance is measured using the peak-signal-to-noise ratio (PSNR) and the relative error. We recall

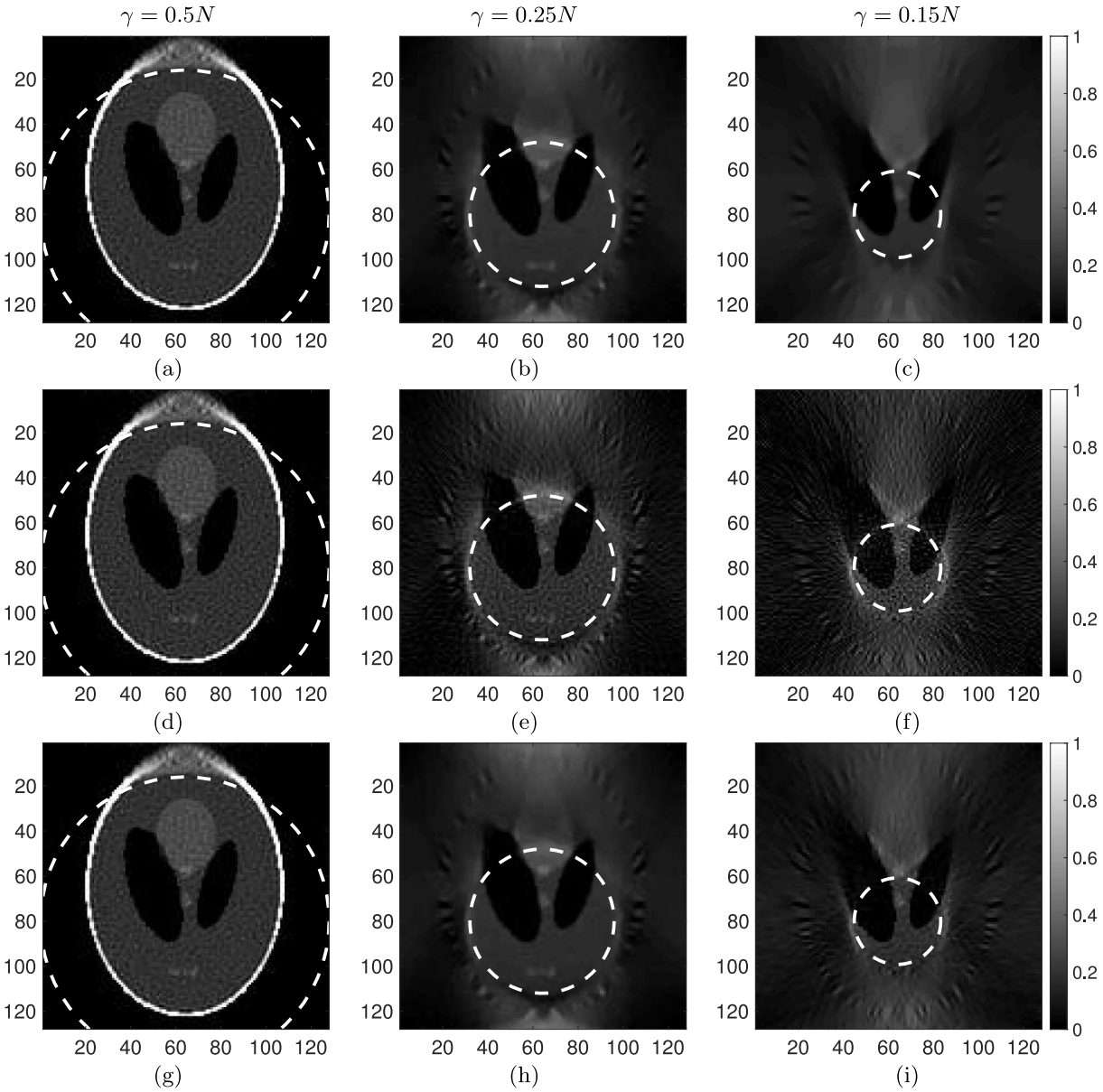

FIGURE 1. Optimal reconstructions of the Shepp-Logan phantom with implicit formulation, for decreasing *radii*: $\gamma = 0.5\,\mathrm{N}$ for (a), (d) and (g), $\gamma = 0.25\,\mathrm{N}$ for (b), (e) and (h), $\gamma = 0.15\,\mathrm{N}$ for (c), (f) and (i). *First row:* shealerts and smothed TV. *Second row:* shearlets only. *Third row:* smothed TV only.

that the PSNR, measured in dB, is defined as follows:

$$\mathrm{PSNR} = 10\log_{10}\left(\frac{\mathrm{MPV}^2}{\mathbf{e}_{\mathrm{MSE}}}\right), \tag{4.1}$$

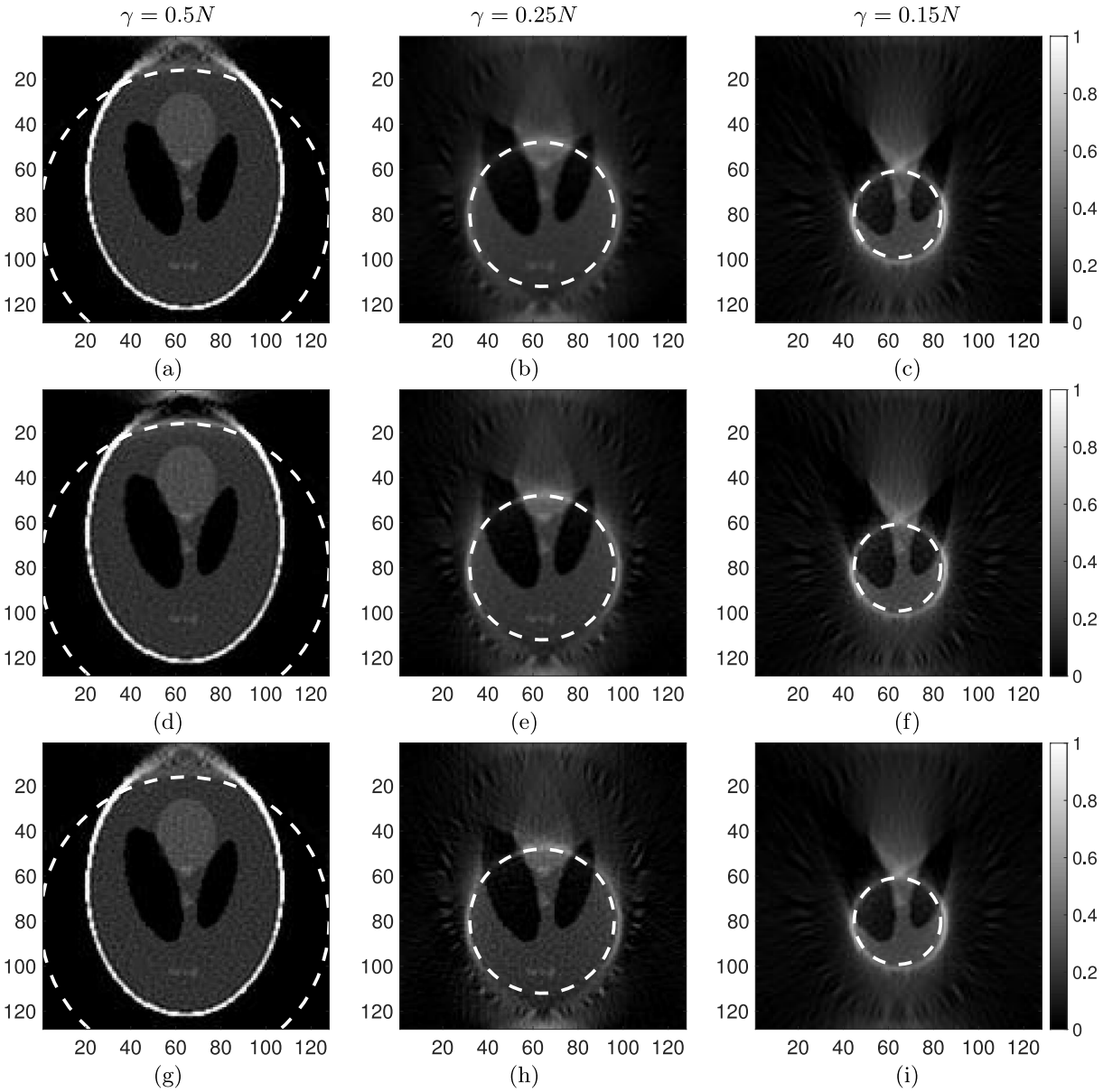

FIGURE 2. Optimal reconstructions of the Shepp-Logan phantom with explicit formulation, for decreasing *radii*: $\gamma = 0.5\,\mathrm{N}$ for (a), (d) and (g), $\gamma = 0.25\,\mathrm{N}$ for (b), (e) and (h), $\gamma = 0.15\,\mathrm{N}$ for (c), (f) and (i). *First row:* shearlets and smothed TV. *Second row:* shearlets only. *Third row:* smothed TV only.

where MPV is the maximum pixel value and $\mathbf{e}_{\mathrm{MSE}}$ is the mean squared error. We stress that both PSNR and relative error are evaluated inside the ROI S only. This is consistent with the motivation of ROI CT that aims at recovering the image inside the ROI only. Results in Tables 1 and 2 correspond to the reconstructions with the "best" ROI relative error and ROI PSNR, where "best" means the smallest ROI relative error and the largest ROI PSNR, compared against the tested range. The corresponding reconstructed images are reported in Figures 1 and 2.

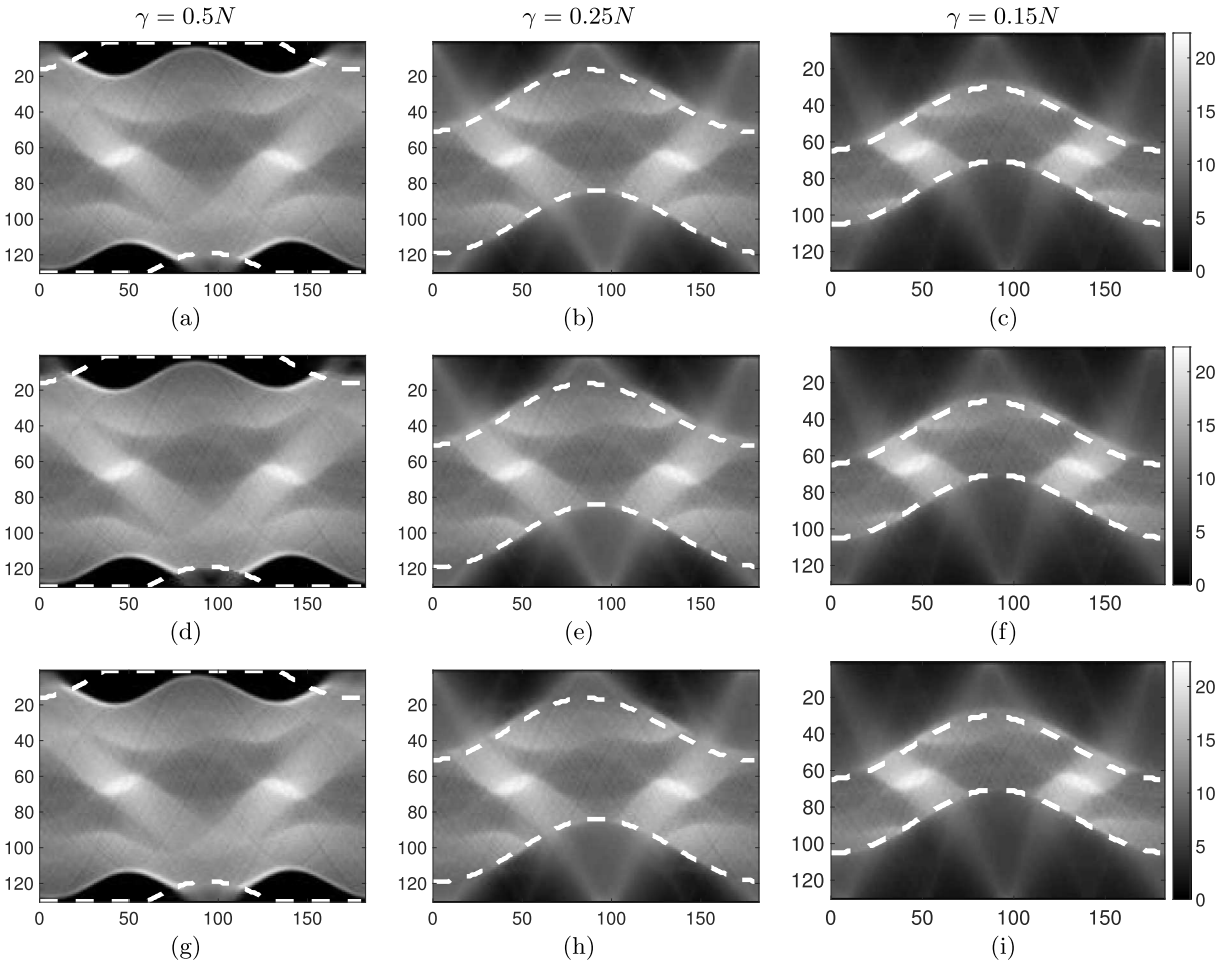

FIGURE 3. Optimal reconstructions of the sinogram of the Shepp-Logan phantom with the explicit formulation, for decreasing *radii*: $\gamma = 0.5\,\mathrm{N}$ for (a), (d) and (g), $\gamma = 0.25\,\mathrm{N}$ for (b), (e) and (h), $\gamma = 0.15\,\mathrm{N}$ for (c), (f) and (i). *First row:* shearlets and smoothed TV. *Second row:* shearlets only. *Third row:* smoothed TV only.

Concerning the setting of SGP parameters, the step length α_k was constrained in the interval $[\alpha_{\min}, \alpha_{\max}]$ with $\alpha_{\min} = 10^{-5}$ and $\alpha_{\max} = 10^{5}$, and the initial threshold L_0 for the scaling matrix D_k was set equal to 10^{10}. For the line search parameters we used $\tau = 0.4$, $\beta = 10^{-4}$, and $\mu = 1$.

We investigated the performance of the algorithm with respect to the regularization parameter $\lambda = 10^{\ell}$, $\ell = -4, -3, \ldots, 1$. For the sTV parameter ρ, we sampled the values 10^{-2}, 10^{-1} and 1. We considered both the explicit and the implicit formulations of the objective function by designing different versions of the SGP method. For the Tikhonov-like regularization term, we also considered a two-level undecimated Daubechies 4 wavelet transform in place of classical shearlets. In all the experiments we found the latter to outperform the former on a visual basis, even if the figures of merit are comparable. As an example, we compare in Figure 4 a shearlet-regularized implicit-based reconstruction versus a wavelet-regularized one. Especially for $\gamma = 0.15\,\mathrm{N}$, it is evident that shearlets outperform wavelets in the ROI reconstruction: notice, for instance, that the cupping artifact is more noticeable in the wavelet-based reconstruction (panel (f)).

Figures 1 and 2 show the reconstructed images obtained for the implicit and the explicit formulations, respectively, and for decreasing *radii*. In both figures, every row contains the reconstruction obtained with a

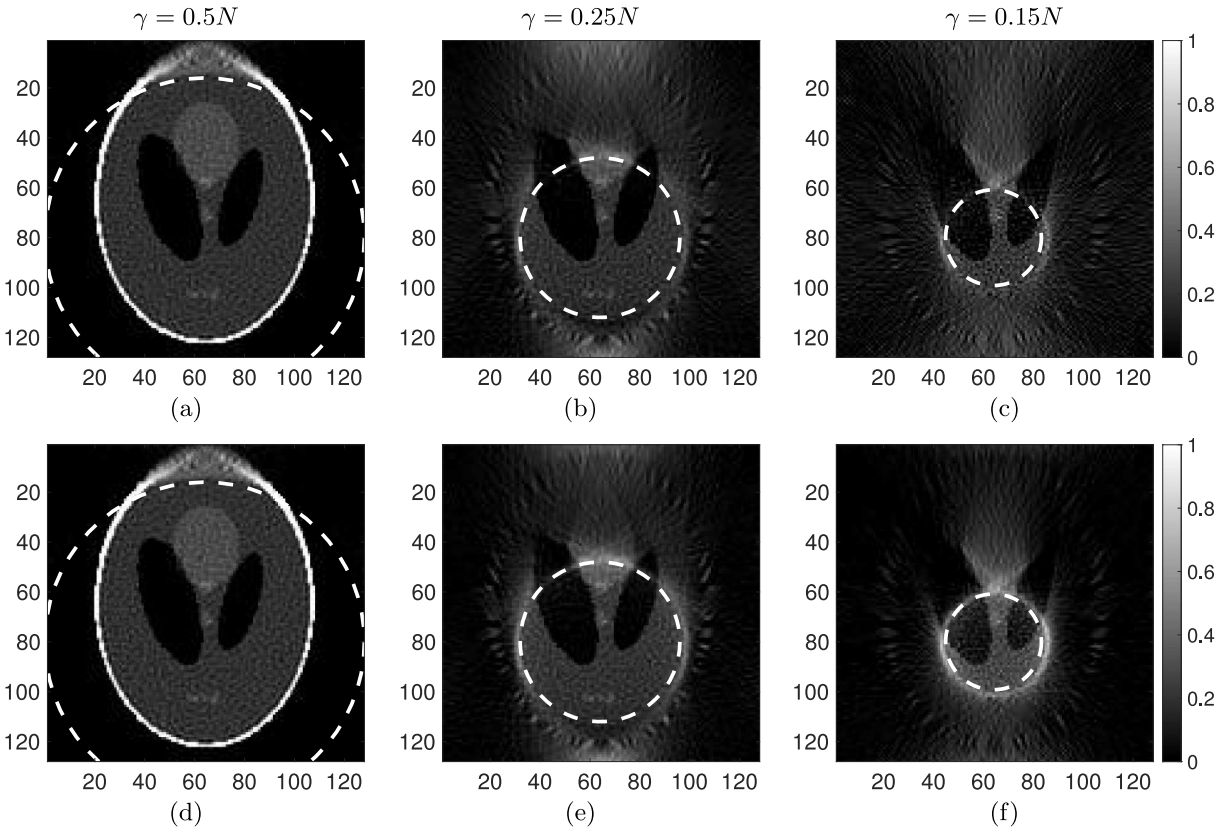

FIGURE 4. Comparison between wavelets only and shearlets only implicit-based reconstructions, with $\lambda = 10^{-3}$ (for $\gamma = 0.5\,\mathrm{N}$) and $\lambda = 10^{-4}$ (for $\gamma = 0.25N, 0.15N$), for decreasing *radii*: $\gamma = 0.5\,\mathrm{N}$ for (a) and (d), $\gamma = 0.25\,\mathrm{N}$ for (b) and (e), $\gamma = 0.15\,\mathrm{N}$ for (c) and (f). *First row:* shearlets (second row of Fig. 1). *Second row:* two-level undecimated Daubechies 4 wavelet.

different approach. In the first and second rows of Figure 1, the objective function reads as in (2.15) and (2.12), respectively. In the third row, the objective function is obtained by dropping the Tikhonov-like regularization term of (2.15). The same holds true for Figure 2, where the referring equation for the first two rows are (2.16) and (2.14), respectively. Finally, in both figures, every column contains the reconstructions computed for different *radii*: $\gamma = 0.5\,\mathrm{N}$ in the first column, $\gamma = 0.25\,\mathrm{N}$ in the second one and $\gamma = 0.15\,\mathrm{N}$ in the final one. In all pictures, the ROI is identified with a dashed white circle.

Figure 3 shows the reconstructed sinograms computed for the explicit formulation corresponding to the reconstructions in Figure 2. The circular ROI in the image domain corresponds to two sinusoidal dashed white lines in the projection domain.

In Figure 1, we observe that, when the *radius* is as large as half the number of pixels of the image edge (panel (a), (d) and (g)), all the reconstructions are good enough to recover even the finest details, but a mild checkerboard effect is still noticeable. The best approach for the objective function is the one based solely on sTV (panel (g)) obtained by dropping the shearlet-based term from formulation (2.15), even if the error rate has the same magnitude for all possible approaches and the reconstructions are almost identical. Notice that, for all possible formulations, the values for the parameters λ and ρ are quite small: they are always among the smallest ones of the investigated ranges. For smaller *radii* (panels (b), (e) and (h) with $\gamma = 0.25\,\mathrm{N}$), we notice that the fundamental structures are detected by all the approaches, even if the sTV-free approaches exhibits some checkerboard effects. The best reconstruction, with respect to both the figures of merit, is obtained by combining

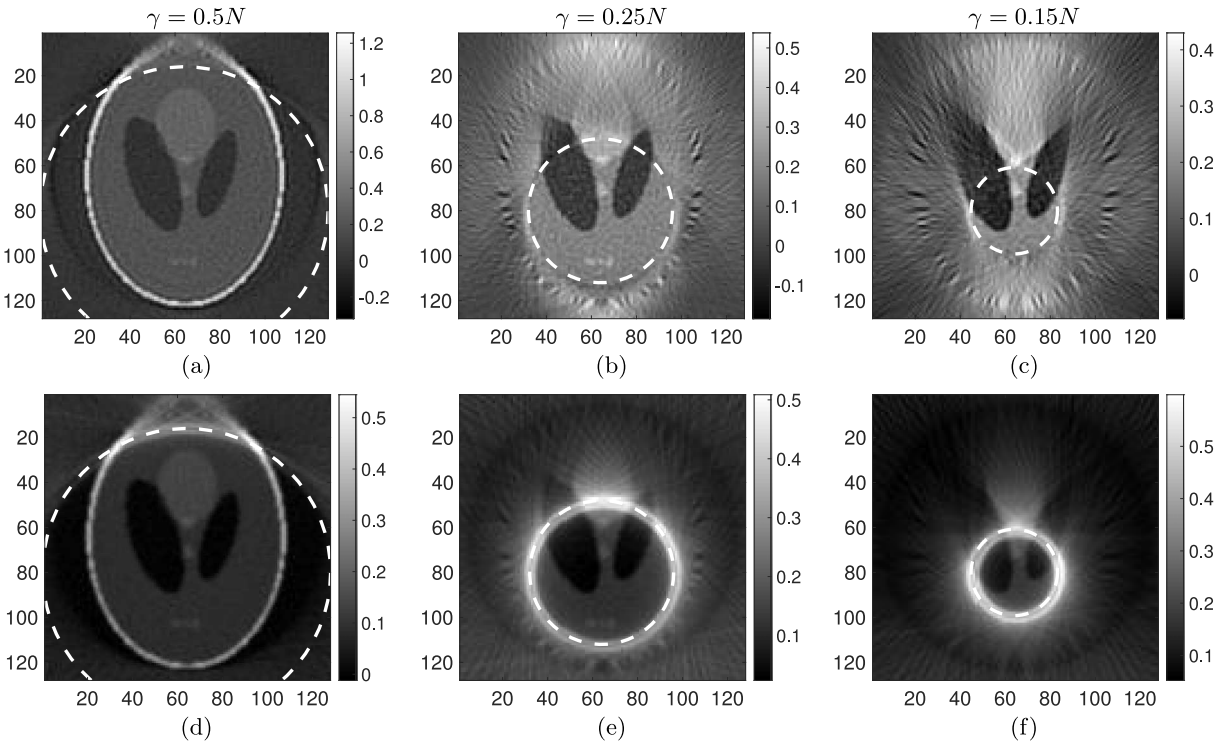

FIGURE 5. Reconstructions of the Shepp-Logan phantom with state-of-the-art methods, for decreasing *radii*: $\gamma = 0.5\,\mathrm{N}$ for (a) and (d), $\gamma = 0.25\,\mathrm{N}$ for (b) and (e), $\gamma = 0.15\,\mathrm{N}$ for (c) and (f). *First row:* least-squares Conjugated Gradient (LSCG). *Second row:* filtered back-projection (FBP).

TABLE 2. Optimal results for the Shepp-Logan phantom with the *explicit* formulation of the objective function. The corresponding reconstructed images are reported in Figure 2. SH = Shearlets; sTV = smoothed Total Variation.

		ROI Rel. Err.	ROI PSNR	Iter	Sec	Params
$\gamma = 0.5\,\mathrm{N}$	SH+sTV	0.19	28.03	131	12.11	$\lambda = 10^{-2}$, $\rho = 10^{-1}$
	SH	0.22	26.62	63	6.60	$\lambda = 10^{-2}$
	sTV	0.20	27.60	88	5.62	$\rho = 10^{-2}$
$\gamma = 0.25\,\mathrm{N}$	SH+sTV	0.25	26.74	51	4.70	$\lambda = 10^{-3}$, $\rho = 10^{-1}$
	SH	0.23	27.13	73	6.30	$\lambda = 10^{-4}$
	sTV	0.24	26.93	103	7.47	$\rho = 10^{-2}$
$\gamma = 0.15\,\mathrm{N}$	SH+sTV	0.70	21.40	54	5.09	$\lambda = 10^{-3}$, $\rho = 10^{-2}$
	SH	0.68	21.67	87	8.40	$\lambda = 10^{-4}$
	sTV	0.67	21.71	66	4.02	$\rho = 10^{-2}$

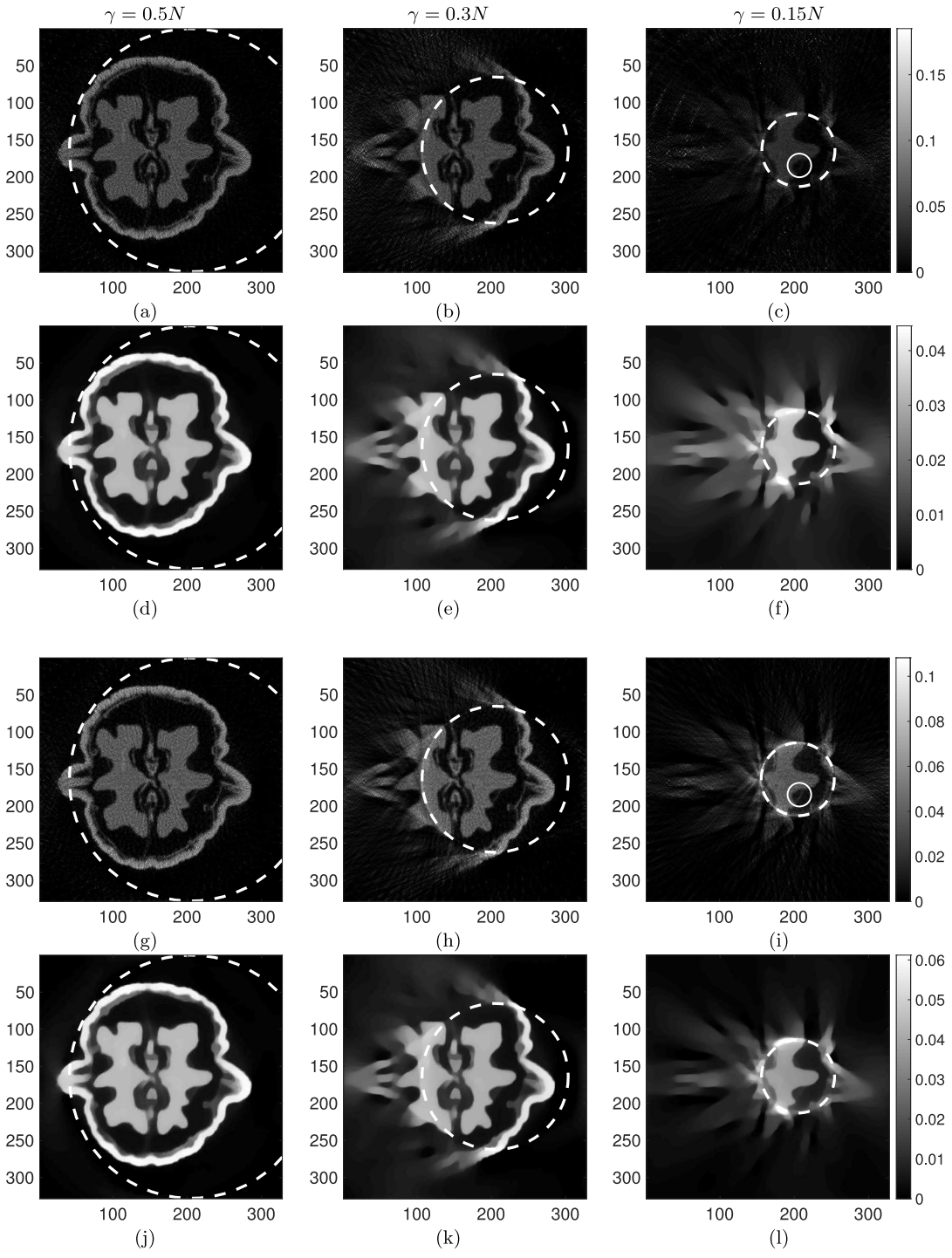

FIGURE 6. Reconstructions of a walnut (real data) with implicit (first and second rows) and explicit (third and fourth rows) formulation, for decreasing *radii*: $\gamma = 0.5\,\mathrm{N}$ for (a), (d), (g) and (j), $\gamma = 0.3\,\mathrm{N}$ for (b), (e), (h) and (k), $\gamma = 0.15\,\mathrm{N}$ for (c), (f), (i) and (l). *First and third rows:* shearlets only. *Second and fourth rows:* smoothed TV only.

shearlets and sTV (formulation (2.15)) and by setting to the largest value (in our investigated range) the sTV regularization parameter ρ. Notice that the pure sTV-based approaches (panels (b) and (h)) outperform the shearlet-based one (panel (e)) by almost one order of magnitude for both figures of merit. Similar considerations hold for the last case. Indeed, when the *radius* decreases (panels (c), (f) and (i) with $\gamma = 0.15\,\mathrm{N}$), the fundamental structures are still detected rather well in all reported images, but also the checkerboard effect remains noticeable in the sTV-free reconstructions. For this smaller *radius*, the optimal reconstruction is achieved for the purely sTV formulation of the objective function and for similar small value of the regularization parameter as the larger *radius* $\gamma = 0.25\,\mathrm{N}$. The relative error has the same magnitude as in the case $\gamma = 0.25\,\mathrm{N}$, outperforming the purely shearlet-based method by almost one order of magnitude.

In Figure 2, we observe that the fundamental structures are consistently well detected in all reconstructions. The checkerboard effect persists in the sTV-free reconstructions (second row), and for all formulations of the objective function in the case of the largest *radius* (panels (a), (d) and (g)). The corresponding sinogram reconstructions reported in Figure 3 are accurate inside the ROI and therein they show no artifacts. This is consistent with the accurate reconstruction of the corresponding images. The best approach for all *radii*, with respect to the figures of merit, is obtained by exploiting the sTV-based approaches. However, for all ROI *radii* the error rate (the PSNR, respectively) has the same magnitude for all possible approaches, namely the purely shearlets approach performs only slightly worse than the sTV-based ones. Notice that, even if panels (c), (f) and (i) exhibit a rather large relative error (70%), on a visual basis the reconstruction is acceptable or even good. Such a large relative error is probably due to the checkerboard effect (sub- or under-estimation of the pixel intensity).

Finally, for baseline comparison, we considered two state-of-the-art methods: an iterative approach, namely the least-squares Conjugate Gradient (LSCG) method, and an analytic approach, *i.e.*, the Filtered Back-Projection (FBP). We run the LSCG experiments using the Matlab default `pcg` function with different threshold levels (10^{-4}, 10^{-5} and 10^{-7}) of the residual norm for stopping the iteration, and with both 20 and 7000 as maximum number of iterations. The best results are obtained with 20 iterations and 10^{-4} as threshold level for the residual norm, yielding PSNR $= 25.05$, 29.27, 31.51, and rel. err. $= 0.26$, 0.18, 0.20, for $\gamma = 0.5N$, $0.25\,\mathrm{N}$ and $0.15\,\mathrm{N}$ respectively. For the FBP algorithm, we used the Matlab `fbp` function available in the package *AIR tools* [25], yielding PSNR $= 17.17$, 18.76, 15.45, and rel. err. $= 0.65$, 0.59, 1.29, respectively. The corresponding reconstructed images are shown in Figure 5. Here, we explicitly report the color bar to show that, even if at least LSCG seems to perform quite well with respect to the figures of merit, it is clear from the image reconstruction that the range for the image pixel intensity is not accurately preserved, since some "negative intensities" are found which is clearly not physical. Concerning FBP reconstructions, the results are considerably worse than SGP as evident from both the figure of merit as well as the reconstructed image quality.

In the numerical tests shown so far, the sTV-based regularization performs better than the shearlets-based regularization. This performance result is expected due to the fact that the Shepp-Logan phantom we considered in our first set of tests is a piecewise constant image.

To illustrate the performance of our algorithm under more realistic conditions, we considered a second set of experiments using a more realistic setting. Namely, we considered the tomographic X-ray *real* (and hence *noisy*) data of a walnut [23], available at www.fips.fi/dataset, of size 328×328 pixels, for the different ROI *radii* $\gamma = 0.5\,\mathrm{N}$, $0.3\,\mathrm{N}$, $0.15\,\mathrm{N}$. In Figure 6 we illustrate the ROI reconstruction images obtained using the implicit (first and second rows) and the explicit (third and fourth rows) formulations, with a non-negativity constraint. The first and third rows of the image show our shearlet-based reconstructions, for decreasing *radii*. The comparison with the result based on sTV regularization shows that the shearlet-based approach is very effective at recovering the finer structures of the image while the sTV-based approach (second and fourth rows) results in a over-smoothed reconstruction. We attribute this improved reconstruction result to the localization and directional properties of shearlets. In addition, we remark that the just-shearlet explicit approach (third row) is able to recover a more accurate reconstruction inside the ROI with respect to the implicit one (first row), as conjectured at the end of Section 2.2. This observation is more evident for $\gamma = 0.15\,\mathrm{N}$, as one can verify from the smallest white circles (continuous line) in the first and third rows of Figure 6 (panels (c) and (i)). All our reconstructions were obtained with only 150 iterations and correspond to $\lambda = 10^{-4}$ (first and third row)

and $\rho = 10^{-1}$ (second and fourth row). The shearlets-plus-sTV reconstruction are similar, on a visual basis, to the purely sTV reconstructions.

Concerning the computing time, Tables 1 and 2 show that, as expected, a higher number of iterations results in a more demanding running time for the sTV-based approach. However, codes were not optimized for computational efficiency as this would go beyond the scope of this paper.

The focus here is the assessment of a suitable objective function and an appropriate algorithm to face the ROI CT problem, without any assumption on the location or size of the ROI.

5. Conclusions

In this paper, we presented a numerical assessment of the ROI tomographic reconstruction in the 2D setting using a first-order iterative algorithm based on convex optimization. Our method makes no restricting assumptions on the location or size of the ROI. We considered two different objective functions, each with a variable level of regularization aiming at ensuring stable reconstruction from truncated data. Numerical experiments have shown that our reconstruction algorithms are very satisfactory for all versions of the objective function as the main structures and the intensity levels are recovered very accurately and with minimal ring artifacts. Using a piecewise constant phantom, the purely sTV-based approach performs slightly better than the shearlets-plus-TV approach, at least for smaller ROI *radii* (*i.e.*, when the ROI is fully inside the field of view). By using a more realistic (and noisy) test data, we verified that our shearlet-based regularized reconstruction outperforms the one based solely on the sTV term.

An important feature of our approach is that the performance of our ROI reconstructions algorithm is not sensitive to the location and size of the ROI and performs consistently well also for rather small ROI sizes, using both formulations of the objective function.

The extension of our method to the three-dimensional setting can be carried out following essentially the same ideas presented in this paper. In future work, we plan to consider more realistic phantoms, with different levels and types of noise, and examine a different formulation of the optimization problem that exploits sparsity in the regularization term. Some results are available in [6, 7].

Acknowledgements. T.A.B. acknowledges support by the Young Researchers Fellowship 2014 of the University of Ferrara, and the Academy of Finland project 1310822. T.A.B., S.B. and G.Z. acknowledge support by the Italian national research project FIRB2012, grant no. RBFR12M3AC, and by the local research project FAR2014 of the University of Ferrara. Also INdAM-GNCS is acknowledged. D.L. acknowledges support by NSF DMS 1720452, 1720487, GEAR 113491 and by a grant from the Simon Foundation (422488).

References

[1] R. Azencott, B.G. Bodmann, T. Chowdhury, D. Labate, A. Sen and D. Vera, Region-of-interest reconstruction from truncated cone-beam projections. *Inverse Probl. Imaging* **12** (2018) 29–57.
[2] J. Barzilai and J.M. Borwein, Two point step size gradient methods. *IMA J. Numer. Anal.* **8** (1988) 141–148.
[3] M. Beister, D. Kolditz, and W. Kalender. Iterative reconstruction methods in X-ray CT. *Phys. Med.* **28** (2012) 94–108.
[4] S. Bonettini, R. Zanella and L. Zanni, A scaled gradient projection method for constrained image deblurring. *Inverse Probl.* **25** (2009) 015002.
[5] T.A. Bubba, D. Labate, G. Zanghirati, S. Bonettini and B. Goossens, Shearlet-based regularized ROI reconstruction in fan beam computed tomography, in *SPIE Optics & Photonics, Wavelets And Applications XVI, San Diego, CA, USA* (2015) Vol. 9597, p. 95970K.
[6] T.A. Bubba, F. Porta, G. Zanghirati and S. Bonettini, The ROI CT problem: a shearlet-based regularization approach. *J. Phys.: Conf. Ser.* **756** (2016) 012009.
[7] T.A. Bubba, F. Porta, G. Zanghirati and S. Bonettini, A nonsmooth regularization approach based on shearlets for poisson noise removal in ROI tomography. *Appl. Math. Comput.* **318** (2018) 131–152.
[8] R. Clackdoyle and M. Defrise, Tomographic reconstruction in the 21st century, region-of-interest reconstruction from incomplete data. *IEEE Signal Process.* **60** (2010) 60–80.
[9] R. Clackdoyle and F. Noo. A large class of inversion formulae for the 2D Radon transform of functions of compact support. *Inverse Probl.* **20** (2004) 1281–1291.
[10] R. Clackdoyle, F. Noo, J. Guo and J. Roberts, Quantitative reconstruction from truncated projections in classical tomography. *IEEE Trans. Nucl. Sci.* **51** (2004) 2570–2578.

[11] Y.H. Dai and R. Fletcher, On the asymptotic behaviour of some new gradient methods. *Math. Program.* **103** (2005) 541–559.
[12] B. De Man and S. Basu, Distance-driven projection and backprojection in three dimensions. *Phys. Med. Biol.* **7** (2004) 2463–2475.
[13] G.R. Easley, D. Labate and W.Q. Lim, Sparse directional image representations using the discrete shearlet transform. *Appl. Comput. Harmon. Anal.* **25** (2008) 25–46.
[14] R. Fletcher, On the Barzilai-Borwein method, in Optimization and Control with Applications, edited by L. Qi, K. Teo and X. Yang. Springer, US, Boston, MA (2005) 235–256.
[15] G. Frassoldati, G. Zanghirati, and L. Zanni. New adaptive stepsize selections in gradient methods. *J. Ind. Manag. Optim.* **4** (2008) 299–312.
[16] E. Garduño and G.T. Herman, Computerized tomography with total variation and with shearlets. *Inverse Probl.* **33** (2017) 044011.
[17] B. Goossens, D. Labate and B. Bodmann, Region-of-interest computed tomography by regularity-inducing convex optimization. Submitted in: *Adv. Comput. Math.* (2017).
[18] L. Grippo, F. Lampariello and S. Lucidi, A nonmonotone line-search technique for newtons method. *SIAM J. Numer. Anal.* 23 (1986) 707–716.
[19] K. Guo and D. Labate, Optimally sparse multidimensional representation using shearlets. *SIAM J. Math. Anal.* **39** (2007) 298–318.
[20] K. Guo, D. Labate, W.-Q. Lim, G. Weiss and E. Wilson, Wavelets with composite dilations and their MRA properties. *Appl. Comput. Harmon. Anal.* **20** (2006) 202–236.
[21] K. Guo, D. Labate and W.-Q. Lim, Edge analysis and identification using the continuous shearlet transform. *Appl. Comput. Harmon. Anal.* **27** (2009) 24–46.
[22] Y. Guorui, J. Tian, Z. Shouping, C. Qin, D. Yakang, Y. Fei, D. Di and W. Ping, Fast Katsevich algorithm based on GPU for helical cone-beam computed tomography. *IEEE Trans. Inf. Tech. Biomed.* **14** (2010) 1053–1061
[23] K. Hämäläinen, L. Harhanen, A. Kallonen, A. Kujanpää, E. Niemi and S. Siltanen, Tomographic X-ray data of a walnut. Preprint arXiv:1502.04064 (2015).
[24] B. Hamelin, Y. Goussard, J.P. Dussault, G.G. Soulez, Design of iterative ROI transmission tomography reconstruction procedures and image quality analysis. *Med. Phys.* **37** (2010) 4577–4589.
[25] P.C. Hansen and M. Saxild-Hansen, AIR tools – a MATLAB package of algebraic iterative reconstruction methods. *J. Comput. Appl. Math.* **236** (2012) 2167–2178.
[26] G.T. Herman and A. Lent, Iterative reconstruction algorithms. *Comput. Biol. Med.* **6** (1976) 273–294.
[27] M.R. Hestenes and E. Stiefel, Methods of conjugate gradients for solving linear systems. *J. Res. Nat. Bur. Stand.* **6** (1952) 409–436.
[28] E. Klann, E.T. Quinto and R. Ramlau, Wavelet methods for a weighted sparsity penalty for region of interest tomography. *Inverse Probl.* **31** (2015) 025001.
[29] G. Kutyniok and D. Labate, Shearlets. Multiscale Analysis for Multivariate Data. Birkhäuser, Boston, MA, USA (2012).
[30] H. Lantéri, M. Roche, O. Cuevas and C. Aime, A general method to devise maximum-likelihood signal restoration multiplicative algorithms with non-negativity constraints. *Signal Process.* **81** (2001) 945–74.
[31] F. Natterer, The Mathematics of Computerized Tomography. SIAM: Society for Industrial and Applied Mathematics. (2001).
[32] F. Natterer and F. Wübbeling, Mathematical Methods in Image Reconstruction. SIAM: Society for Industrial and Applied Mathematics, Philadelphia, PA, USA (2001).
[33] K. Niinimaki, S. Siltanen and V. Kolehmainen, Bayesian multiresolution method for local tomography in dental X-ray imaging. *Phys. Med. Biol.* **52** (2007) 6663–6678.
[34] F. Noo, M. Defrise, R. Clackdoyle and H. Kudo, Image reconstruction from fan-beam projections on less than a short scan. *Phys. Med. Biol.* **47** (2002) 2525–2546.
[35] F. Noo, R. Clackdoyle and J. Pack, A two-step Hilbert transform method for 2D image reconstruction. *Phys. Med. Biol.* **49** (2004) 3903–3923.
[36] L.A. Shepp and Y. Vardi, Maximum likelihood reconstruction for emission tomography. *IEEE Trans. Med. Imaging* **1** (1982) 113–122.
[37] B. Vandeghinste, B. Goossens, R. Van Holen, C. Vanhove, A. Pižurica, S. Vandenberghe and S. Staelens, Iterative CT reconstruction using shearlet-based regularization. *IEEE Trans. Nucl. Sci.* **5** (2013) 3305–3317.
[38] C.R. Vogel, Computational Methods for inverse problems. SIAM, Philadelphia (2002).
[39] S. Wright, R.D. Nowak and M. Figueiredo, Sparse reconstruction by separable approximation. *IEEE Trans. Signal Proc.* **57** (2009) 2479–2493
[40] J. Yang, H. Yu, M. Jiang and G. Wang, High-order total variation minimization for interior tomography. *Inverse Probl.* **26** (2010) 035013.
[41] B. Zhang and G.L. Zeng, Two dimensional iterative region of interest reconstruction from truncated projection data. *Med. Phys.* **34** (2007) 935–944.
[42] A. Ziegler, T. Nielsen and M. Grass, Iterative reconstruction of a region of interest for transmission tomography. *Med. Phys.* **35** (2008) 1317–1327.
[43] Y. Zou, X. Pan and E. Sidky, Image reconstruction in regions-of-interest from truncated projections in a reduced fan-beam scan. *Phys. Med. Biol.* 50 (2005) 13–28.