Sparse regularization in limited angle tomography

Jürgen Frikel [a,b,*]

[a] *Institute of Biomathematics and Biometry, Helmholtz Zentrum München, German Research Center for Environmental Health, Ingolstädter Landstraße 1, D-85764, Germany*
[b] *Zentrum Mathematik, M6, Technische Universität München, Germany*

ARTICLE INFO

ABSTRACT

We investigate the reconstruction problem of limited angle tomography. Such problems arise naturally in applications like digital breast tomosynthesis, dental tomography, electron microscopy, etc. Since the acquired tomographic data is highly incomplete, the reconstruction problem is severely ill-posed and the traditional reconstruction methods, e.g. filtered backprojection (FBP), do not perform well in such situations. To stabilize the reconstruction procedure additional prior knowledge about the unknown object has to be integrated into the reconstruction process. In this work, we propose the use of the sparse regularization technique in combination with curvelets. We argue that this technique gives rise to an edge-preserving reconstruction. Moreover, we show that the dimension of the problem can be significantly reduced in the curvelet domain. To this end, we give a characterization of the kernel of the limited angle Radon transform in terms of curvelets and derive a characterization of solutions obtained through curvelet sparse regularization. In numerical experiments, we will show that the theoretical results directly translate into practice and that the proposed method outperforms classical reconstructions.

© 2012 Elsevier Inc. All rights reserved.

1. Introduction

Limited angle tomography problems arise naturally in many practical applications, such as digital breast tomosynthesis, dental tomography, etc. The underlying principle of these imaging techniques consists of two steps: first, the data acquisition step, where a few X-ray projections of an object are taken from different view angles (within a limited angular range). Second, the reconstruction step, where the attenuation coefficient $f : \mathbb{R}^2 \to \mathbb{R}$ of the object is approximately reconstructed from the given projection data. In this paper we are concerned with the second step, i.e., with the development of an appropriate (adapted) reconstruction technique which takes into account the special structure of the limited angle tomography.

To this end, we consider the *Radon transform* as mathematical model for the acquisition process which is defined by

$$\mathcal{R}f(\theta, s) = \int_{L(\theta, s)} f(x) \, dS(x), \tag{1}$$

where $L(\theta, s) = \{x \in \mathbb{R}^2 : x_1 \cos \theta + x_2 \sin \theta = s\}$ denotes a line with the normal direction $(\cos \theta, \sin \theta)^T$ and the signed distance from the origin $s \in \mathbb{R}$. Furthermore, we assume f to lie in the natural domain of the Radon transform, i.e., f is such that (1) exists for all (θ, s). Whenever we write $\mathcal{R}_\theta f(s)$ instead of $\mathcal{R}f(\theta, s)$, we consider $\mathcal{R}f(\theta, s)$ as a univariate function

---

* Correspondence to: Institute of Biomathematics and Biometry, Helmholtz Zentrum München, German Research Center for Environmental Health, Ingolstädter Landstraße 1, D-85764, Germany.
 E-mail address: juergen.frikel@helmholtz-muenchen.de.

of the second argument s with a fixed angular parameter θ. In this case, we will call the function $\mathcal{R}_\theta f$ a projection of f at angle θ.

In contrast to the classical computed tomography, in limited angle tomography the data $\mathcal{R}f(\theta, s)$ is known only within a limited angular range, that is, for $\theta \in [-\Phi, \Phi]$ with $\Phi < \pi/2$. To emphasize that the Radon transform $\mathcal{R}f$ of a function is defined only on a limited angle domain $[-\Phi, \Phi] \times \mathbb{R}$, we will write $\mathcal{R}_\Phi f$ and call it the *limited angle Radon transform*. As a consequence of the limited angular range, the reconstruction problem $y = \mathcal{R}_\Phi f$ becomes severely ill-posed [28,14]. Thus, small measurement errors can cause huge reconstruction errors. This is a serious drawback since, in practice, the acquired data is (to some extent) always corrupted by noise. The practical reconstruction problem therefore reads

$$y^\delta = \mathcal{R}_\Phi f + \eta, \tag{2}$$

where η denotes the noise, $\delta > 0$ is the noise level, i.e., $\|\eta\| < \delta$. The goal is to find an approximation to f from the noisy measurements y^δ.

It is well known that classical reconstruction methods, such as filtered backprojection (FBP), do not perform well in such situations [28]. These algorithms are sensitive to noise. To stabilize the inversion, additional prior knowledge about the solution has to be integrated into the reconstruction procedure [16]. Usually, variational methods are used to obtain a regularized solution f_α that is given as a minimizer of the so-called Tikhonov type functional

$$T_\alpha(f) = \left\| \mathcal{R}_\Phi f - y^\delta \right\|_2^2 + \alpha \Lambda(f), \tag{3}$$

where $\alpha > 0$ denotes a regularization parameter and $\Lambda : \mathrm{dom}(\Lambda) \to [0, \infty]$ is a convex and proper penalty functional [35]. The first term in (3) (data fidelity term) controls the data error, whereas the second term encodes the prior information about the object.

The choice among the various prior terms and, thus, regularization techniques, depends on the specific object (which is imaged) and, to some extent, on the desire to preserve or emphasize particular features of the unknown object. A possible choice for Λ may be any kind of a smoothness (semi-) norm [35]. For instance, the Besov norm allows to adjust the smoothness of the solution at a very fine scale [23,25,32]. Another prominent example in image reconstruction is the total variation (TV) norm. This norm is a smoothness norm that is particularly used for edge-preserving reconstructions [17,21]. However, as it was pointed out in [22], TV reconstruction may not be an appropriate choice for medical imaging purposes. One reason for this is that TV regularization favors piecewise constant functions and, hence, produces staircase effects (cf. [10,33]) which may destroy relevant information. To overcome this problem, higher order total variation priors were considered by some authors, see for example [1].

An alternative approach was introduced in [5,6], where the authors used curvelets for edge-preserving tomographic reconstruction at a full angular range. In this work, we also use curvelets for reconstructions at a limited angular range. To this end, we propose sparse regularization of curvelet coefficients as a reconstruction method. We will show that this method is stable and edge-preserving, which stems from the fact that curvelets provide sparse representation of functions with an optimal encoding of edges [7]. The main part of this work is devoted to a detailed analysis of this method. In particular, we will show that, by formulating the reconstruction problem in the curvelet domain, it is possible to predict the locations of relevant (visible) curvelet coefficients of the sought function. This a priori information depends only on the angular range and, hence, is available prior to the reconstruction. By exploiting this knowledge, we perform a significant dimensionality reduction in the curvelet domain and derive a version of the curvelet sparse regularization technique which is adapted to the limited angular range. The adapted method yields a significant speedup of the reconstruction algorithm while preserving the reconstruction quality. We will show in numerical experiments that these theoretical results directly translate into practice. Moreover, we will present a comprehensive performance analysis of curvelet sparse regularization and its adapted version. We would like to note here that first thoughts on this topic have been formulated in an extended abstract and published in Proceedings in Applied Mathematics and Mechanics [19].

The optimality of curvelets for the limited angle tomography is expressed by their ability for providing edge-preserving reconstructions as well as their ability to adapt to the limited angle geometry. As mentioned above, curvelets were already used for edge-preserving reconstruction in full angle tomography (cf. [5,6]). However, the method that we present in this work, is to our knowledge the first one that is adapted to the limited angle geometry. This adaption entails a significant dimensionality reduction in the curvelet domain and yields an additional stabilization of the reconstruction procedure. In [18], similar results were presented by using the characterization of visible singularities of E.T. Quinto [29], and "the resolution of wavefront set property" of the continuous curvelet transform [8]. However, the results therein were stated without proofs.

1.1. Organization of this paper

In Section 2 we will give a brief introduction to curvelets and the technique of sparse regularization. We will discuss that the combination of these mathematical tools yields a stable and edge-preserving reconstruction method which we call curvelet sparse regularization. In Section 3 we will discuss the relation between curvelet sparse regularization and the "curvelet thresholding" which was proposed in [6]. Our main results will be presented in Section 4. There, we will prove a characterization of the kernel of the limited angle Radon transform in terms of curvelets and derive a characterization of

limited angle reconstructions that are computed via curvelet sparse regularizations. These results will be applied to a finite dimensional reconstruction problem in Section 5. In particular, an adapted curvelet sparse regularization approach will be introduced. In Section 6 we will discuss some of our results. The implementation of our reconstruction method is addressed in Section 7. In the last section of this paper, we will present various numerical experiments that illustrate the theoretical results of this paper. In particular, a detailed performance analysis of curvelet sparse regularization and its adapted version will be presented.

1.2. Notation

The inner product of $x, y \in \mathbb{R}^n$ will be denoted as $x \cdot y$ or simply xy. When not otherwise stated, inner product in a function space X will be denoted by $\langle f, g \rangle_X$. The norm of a vector $x \in \mathbb{R}^n$ will be denoted by $|x|$ whereas the norm in a function space X will be denoted by $\|f\|_X$. We will be using some classical function spaces, such as the space of Schwartz functions $\mathcal{S}(\mathbb{R}^n)$ and the spaces of measurable functions $L^p(\Omega)$, without reference since they can be found in every book on functional analysis. Same holds for the classical sequence spaces ℓ^p. Probably the most important notation for this work is that of the Fourier transform $\hat{f}$ of a function $f \in \mathcal{S}(\mathbb{R}^n)$ which is defined as

$$\hat{f}(\xi) = (2\pi)^{-n/2} \int\limits_{\mathbb{R}^n} f(x) e^{-ix\xi} \, \mathrm{d}x.$$

The inverse Fourier transform is given by $\check{f}(x) = \hat{f}(-x)$. Basic properties of the Fourier transform will be used without proof. For details about the Fourier transform we refer to [36]. For $\eta \in [0, 2\pi]$, we define ρ_η to be the rotation operator $\rho_\eta f(x) = f(R_\eta x)$, where the rotation matrix R_η is defined by

$$R_\eta = \begin{pmatrix} \cos\eta & \sin\eta \\ -\sin\eta & \cos\eta \end{pmatrix}.$$

Eventually, we refer to [28] for notations and some basic facts about the Radon transform.

2. Stabilization of limited angle reconstructions by sparsity in the curvelet domain

Our aim is to solve the limited angle reconstruction problem (2) in a stable way such that the edges of the unknown function are well preserved. To stabilize the inversion, some a priori knowledge about the (unknown) object need to be integrated into the reconstruction procedure. In order to preserve edges, we will assume that the sought functions are C^2 except from discontinuities along C^2 curves. To translate this qualitative information into a mathematical language, we will use the fact that such functions are optimally sparse with respect to the curvelet frame [7]. Hence, determining solutions of (2) which are sparse or compressible with respect to the curvelet frame will stabilize the reconstruction procedure. To this end, the technique of sparse regularization of curvelet coefficients is an appropriate tool.

2.1. The curvelet dictionary

We briefly recall the definition of the curvelet frame [9,26]. At scale 2^{-j}, $j \in \mathbb{N}_0$, we first define the generating curvelets $\psi_{j,0,0}$ in the frequency domain by using polar coordinates (r, ω) as follows:

$$\hat{\psi}_{j,0,0}(r, \omega) = 2^{-3j/4} \cdot W\left(2^{-j} \cdot r\right) \cdot V\left(\frac{2^{\lceil j/2 \rceil + 1}}{\pi} \cdot \omega\right), \tag{4}$$

where $W(r)$ is a radial window and $V(\omega)$ is an angular window, respectively. Moreover, we assume the windows W and V to be real and smooth ($W, V \in C^\infty$) such that $\operatorname{supp} W \subset (1/2, 2)$, $\operatorname{supp} V \subset (-1, 1)$ and to satisfy the following admissibility conditions,

$$\sum_{j=-\infty}^{\infty} W^2\left(2^j r\right) = 1, \quad r \in (3/4, 3/2),$$

$$\sum_{l=-\infty}^{\infty} V^2(\omega - l) = 1, \quad \omega \in (-1/2, 1/2).$$

The family of curvelets $\{\psi_{j,l,k}\}_{j,l,k}$ is now constructed by translation and rotation of the generating curvelets $\psi_{j,0,0}$. That is, at scale 2^{-j}, the curvelet $\psi_{j,l,k}$ is defined via

$$\psi_{j,l,k}(x) = \psi_{j,0,0}\left(R_{\theta_{j,l}}\left(x - b_k^{j,l}\right)\right), \tag{5}$$
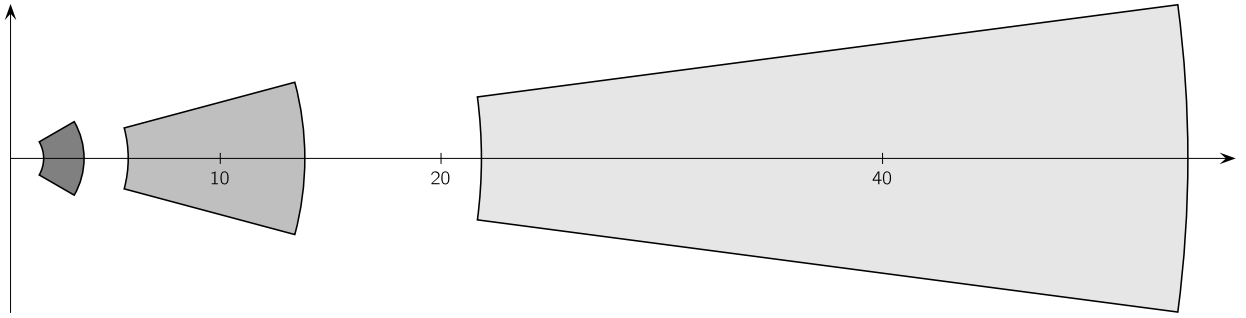
Fig. 1. Support of curvelets in the Fourier domain for $j = 1$ (dark gray), $j = 3$ (gray) and $j = 5$ (light gray). A similar image can be found in [26].
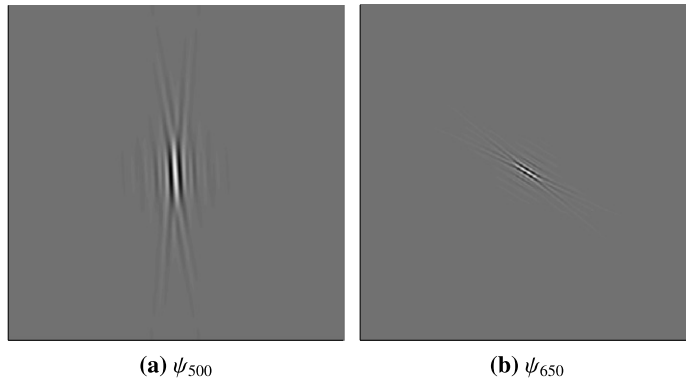

(a) ψ_{500} **(b)** ψ_{650}

Fig. 2. Curvelets at different scales and different orientations. Left image shows a curvelet with orientation $\theta_{5,0} = 0°$ whereas the right image shows a curvelet with orientation $\theta_{6,5} = 56.25°$.

where $R_{\theta_{j,l}}$ denotes the rotation matrix (cf. notations in Section 1) with respect to the scale-dependent rotation angles $\theta_{j,l}$ and the scale-dependent locations $b_k^{j,l}$ which are defined by

$$\theta_{j,l} = l \cdot \pi \cdot 2^{-\lceil j/2 \rceil - 1}, \quad -2^{\lceil j/2 \rceil + 1} \leqslant l < 2^{\lceil j/2 \rceil + 1},$$

$$b_k^{j,l} = R_{\theta_{j,l}}^{-1}\left(\frac{k_1}{2^j}, \frac{k_2}{2^{j/2}}\right), \quad k = (k_1, k_2) \in \mathbb{Z}^2.$$

Since the window functions W and V are compactly supported, and in particular, since the support of $W(2^j \cdot)$ is contained in $(1/2, \infty)$, it follows from (4) and (5) that each curvelet is supported on a polar wedge in the Fourier domain which has a positive distance to the origin, see Fig. 1. We have $\hat{\psi}_{j,l,k}(\xi) = 0$ for all $|\xi| < 1/2$ and all admissible indices (j, l, k). Thus, the region $\bigcup_{(j,l,k)} \text{supp}\, \hat{\psi}_{j,l,k}$ does not cover all of the $\mathbb{R}^2$ and the system $\{\psi_{j,l,k}\}$ does not contain any low-pass element.

To complete the definition of the curvelet system we define the generating low-pass function $\psi_{-1,0,0}$ in the Fourier domain by

$$\hat{\psi}_{-1,0,0}(r, \omega) = W_0(r), \qquad W_0^2(r) := 1 - \sum_{j=0}^{\infty} W^2(2^{-j}r)$$

and complete the curvelet system by all of its translates $\{\psi_{-1,0,k}\}_{k \in \mathbb{Z}^2}$.

Remark. In the spatial domain, the essential support of curvelets is an ellipse which is located near $b_k^{j,l}$ and oriented along the orthogonal direction $\theta_{j,l}^{\perp} = \theta_{j,l} + \pi/2$. The directional localization becomes higher when the scale parameter j increases. Thus, curvelets are highly oriented at fine scales, see Fig. 2.

The index set of the completed curvelet system is now given by

$$\mathcal{I} = \left\{(-1, 0, k) \colon k \in \mathbb{Z}^2\right\} \cup \left\{(j, l, k) \colon j \in \mathbb{N}_0, \ k \in \mathbb{Z}^2, \ -2^{\lceil j/2 \rceil + 1} \leqslant l < 2^{\lceil j/2 \rceil + 1}\right\} =: \mathcal{I}_0 \cup \mathcal{I}_1. \qquad (6)$$

Note that each index $(j, l, k) \in \mathcal{I}$ has a 3 parameter structure, where j denotes the *scale*-parameter, $k = (k_1, k_2)$ is the *location* parameter and l is the *orientation* parameter. The system $\{\psi_{j,l,k}\}_{(j,l,k) \in \mathcal{I}}$ is now complete in the sense that it constitutes a tight frame for $L^2(\mathbb{R}^2)$ [9]. For each $f \in L^2(\mathbb{R}^2)$ there is a curvelet representation

$$f = \sum_{(j,l,k) \in \mathcal{I}} \langle \psi_{j,l,k}, f \rangle \psi_{j,l,k} \tag{7}$$

and a Parseval relation

$$\|f\|^2_{L^2(\mathbb{R}^2)} = \sum_{(j,l,k) \in \mathcal{I}} \left| \langle \psi_{j,l,k}, f \rangle \right|^2.$$

We conclude this section by noting that curvelets can be understood as further development of wavelets. It is well known form this theory that wavelet-like dictionaries, and hence curvelets, provide sparse representations of a large class of functions [11]. This property qualifies curvelets for the use within the framework of the sparse regularization. Even more important is the fact that curvelets offer an optimally sparse representation of functions that are C^2 except for discontinuities along C^2 curves [7]. This means, that curvelets encode edges in a very efficient way and, in this sense, sparse representations of functions with respect to the curvelet dictionary are edge-preserving.

2.2. Curvelet sparse regularization (CSR)

The idea of curvelet sparse regularization is to determine a solution f of the problem (2) which is sparse or compressible with respect to the curvelet frame. To this end, we have to formulate the reconstruction problem in the curvelet domain, i.e., we are interested in the recovery of the curvelet coefficients $c_{j,l,k} = \langle \psi_{j,l,k}, f \rangle$ of f rather than the function f itself. In what follows we assume that the unknown object f can be represented by a finite linear combination of curvelets, i.e., $f = \sum_{n=1}^{N} \langle \psi_n, f \rangle \psi_n$ with $n = n(j, l, k)$ and define the analysis operator T and the synthesis operator T^* as

$$T f = \left\{ \langle \psi_n, f \rangle \right\}_{n=1}^{N}, \qquad T^* c = \sum_{n=1}^{N} c_n \psi_n.$$

Then, the reconstruction problem (2) can expressed in terms of curvelet coefficients via

$$y^\delta = Kc + \eta, \quad K := \mathcal{R}_\Phi T^*. \tag{8}$$

Sparse regularization of curvelet coefficients yields a solution of (8) which is given as a minimizer of the ℓ^1-penalized Tikhonov type functional, i.e.,

$$\hat{c} = \arg \min_{c \in \mathbb{R}^N} \left\{ \frac{1}{2} \|Kc - y^\delta\|^2_{L^2(S^1 \times \mathbb{R})} + \|c\|_{1,w} \right\}, \tag{9}$$

where $\|c\|_{1,w} = \sum_k w_k |c_k|$ denotes the weighted 1-norm with a weight sequence w satisfying $w_k \geqslant w_0 > 0$. A reconstruction for the original problem (2) is then given by applying the synthesis operator to the regularized curvelet coefficients $\hat{c}$, i.e.,

$$\hat{f} = \sum_{n=1}^{N} \hat{c}_n \psi_n. \tag{10}$$

We note that sparse regularization is indeed a regularization method [13], [35, Section 3.3]. Therefore, the computation of a reconstruction by (9) and (10) is stable and favors sparse (or compressible) solutions [20, Section 2]. We will refer to this method by the term *curvelet sparse regularization* or CSR, respectively. In the previous subsection, we have discussed that sparse representation of functions with respect to the curvelet dictionary are edge-preserving. Consequently, curvelet sparse regularization gives rise to an edge-preserving reconstruction method.

In the following proposition we give a general characterization of minimizers of the ℓ^1-penalized Tikhonov functional. Though the proof can be found in [20], we will recall it here for the sake of completeness. To this end, we define the so-called *soft-thresholding operator* $\mathcal{S}_w : \mathbb{R}^N \to \mathbb{R}^N$ by

$$\left(\mathcal{S}_w(x) \right)_k = S_{w_k}(x_k) := \max\{0, |x_k| - w_k\} \operatorname{sgn}(x_k). \tag{11}$$

Proposition 2.1. *The set of minimizers of the ℓ^1-penalized Tikhonov functional*

$$\Psi(c) = \frac{1}{2} \|Kc - y^\delta\|^2_{L^2(S^1 \times \mathbb{R})} + \|c\|_{1,w},$$

is non-empty. Furthermore, each minimizer $\hat{c}$ of Ψ is characterized by

$$\hat{c} = \mathcal{S}_{\gamma w}\big(\hat{c} - \gamma K^*\big(K\hat{c} - y^\delta\big)\big) \tag{12}$$

for any $\gamma > 0$.

Proof. We follow the proof of [20]. Since Ψ is convex and coercive it follows that there is a minimizer $\hat{c}$ of Ψ. We denote by $\partial f(c)$ the subdifferential of a convex function f at x. Each minimizer $\hat{c}$ is characterized by the requirement (cf. [34])

$$0 \in \partial\Psi(\hat{c}) = K^*(K\hat{c} - g) + \partial\|c\|_{1,w}$$

which is equivalent to

$$-K^*(K\hat{c} - g) \in \partial\|\hat{c}\|_{1,w}.$$

Multiplying by $\gamma > 0$ and adding $\hat{c}$ to both sides yields

$$\hat{c} - \gamma K^*(K\hat{c} - g) \in \hat{c} + \gamma\partial\|\hat{c}\|_{1,w} = \big(\mathrm{id} + \gamma\partial\|\cdot\|_{1,w}\big)\hat{c}.$$

Following the arguments in [20] we get that $(\mathrm{id} + \gamma\partial\|\cdot\|_{1,w})^{-1}$ exists and is single valued. Hence, the above inclusion is characterized by the equation

$$\hat{c} = \big(\mathrm{id} + \gamma\partial\|\cdot\|_{1,w}\big)^{-1}\big(\hat{c} - \gamma K^*(K\hat{c} - g)\big).$$

A simple calculation shows that $(\mathrm{id} + \gamma\partial\|\cdot\|_{1,w})^{-1} = \mathcal{S}_{\gamma w}$. $\square$

3. Relation to biorthogonal curvelet decomposition (BCD) for the Radon transform

In this section we will show that if the data is available at a full angular range, i.e., if we are dealing with the ordinary Radon transform $\mathcal{R}$ rather than the limited angle Radon transform $\mathcal{R}_\Phi$, an explicit formula for the minimizer of the ℓ^1-penalized Tikhonov functional (9) can be derived using the biorthogonal curvelet decomposition (BCD) for the Radon transform [6]. This formula is closely related to the BCD based reconstruction which was also proposed in [6]. Afterwards, we will discuss that the curvelet sparse regularization can be understood as a natural generalization of the BCD reconstruction.

3.1. Full angular range

We now briefly recall the definition of the BCD for the Radon transform. For details we refer to [6]. If not otherwise stated, we let $n = n(j, l, k) \in \mathcal{I}$ and denote the curvelet frame by $\{\psi_n\}$. In order to derive the BCD for the Radon transform a pair of frames $\{U_n\}$ and $\{V_n\}$ is constructed for $\mathrm{ran}(\mathcal{R}) \subset L^2(\mathbb{R} \times S^1)$ such that

$$\mathcal{R}\psi_n = 2^{-j}V_n, \qquad \mathcal{R}^*U_n = 2^{-j}\psi_n,$$

and a quasi-biorthogonal relation $\langle V_n, U_{n'}\rangle_{L^2(\mathbb{R}\times S^1)} = 2^{j-j'}\langle\psi_n, \psi_{n'}\rangle$ holds for all $n, n' \in \mathcal{I}$. In particular, there is a L^2-norm equivalence property

$$\sum_{n\in\mathcal{I}}\big|\langle g, U_n\rangle_{L^2(\mathbb{R}\times S^1)}\big|^2 \asymp \|g\|^2_{L^2(\mathbb{R}\times S^1)},$$

for all $g \in \mathrm{ran}(\mathcal{R})$. Similar relations hold for $\{V_n\}$. Using these notations, the BCD of the Radon transform is given by the following reproducing formula

$$f = \sum_{n\in\mathcal{I}} 2^j\langle\mathcal{R}f, U_n\rangle_{L^2(\mathbb{R}\times S^1)}\psi_n, \tag{13}$$

where $f \in L^2(\mathbb{R}^2)$ is assumed to be a finite sum of curvelets $\{\psi_n\}$ [6]. Note that the curvelet coefficients of f are computed from the Radon transform data $\mathcal{R}f$.

We now use the above the frames (U_n), (V_n) and its relations with $\mathcal{R}$ and $\mathcal{R}^*$, respectively, to compute the minimizer of the ℓ^1-penalized Tikhonov functional. We assume that $y^\delta \in \mathrm{ran}(\mathcal{R})$ and denote $\langle\cdot,\cdot\rangle = \langle\cdot,\cdot\rangle_{L^2(\mathbb{R}^2)}$ and $[\cdot,\cdot] = \langle\cdot,\cdot\rangle_{L^2(\mathbb{R}\times S^1)}$. With $y^\delta_n = [y^\delta, U_n]$, we get

$$\big\|\mathcal{R}f - y^\delta\big\|^2_{L^2(\mathbb{R}\times S^1)} \asymp \sum_{n\in\mathcal{I}}\big|[\mathcal{R}f - y^\delta, U_n]\big|^2$$
$$= \sum_{n\in\mathcal{I}}\big|[\mathcal{R}f, U_n] - [y^\delta, U_n]\big|^2$$

$$= \sum_{n \in \mathcal{I}} \left| \langle f, \mathcal{R}^* U_n \rangle - [y^\delta, U_n] \right|^2$$

$$= \sum_{n \in \mathcal{I}} \left| \langle f, 2^{-j} \psi_n \rangle - [y^\delta, U_n] \right|^2$$

$$= \sum_{n \in \mathcal{I}} \left| 2^{-j} c_n - y_n^\delta \right|^2.$$

Using the definition of $\| \cdot \|_{1,w}$ we see that

$$\left\| Kc - y^\delta \right\|_{L^2(\mathbb{R} \times S^1)}^2 + \|c\|_{1,w} \asymp \sum_{n \in \mathcal{I}} \left(\left| y_n^\delta - 2^{-j} c_n \right|^2 + \alpha w_n |c_n| \right),$$

where $K = \mathcal{R}T^*$ and T^* denotes again the synthesis operator with respect to the curvelet frame. We now assume that there is an $\alpha > 0$ such that

$$\hat{c} = \underset{c \in \mathbb{R}^N}{\arg\min} \left\| Kc - y^\delta \right\|_{L^2(\mathbb{R} \times S^1)}^2 + \|c\|_{1,w}$$

$$= \underset{c \in \mathbb{R}^N}{\arg\min} \sum_{n \in \mathcal{I}} \left(\left| 2^{-j} c_n - y_n^\delta \right|^2 + \alpha w_n |c_n| \right). \tag{14}$$

Then, the functional (14) can be minimized by minimizing each term separately. Note that each term in (14) is of the form $|ax - b|^2 + c|x|$ and that its minimum is given by $S_{c/(2a^2)}(b/a)$, where $S_{c/(2a^2)}$ is the soft-thresholding function (cf. (11)) with respect to the threshold $c/(2a^2)$. Therefore, the regularized curvelet coefficients $\hat{c}$ are given by

$$\hat{c}_n = S_{2^{2j-1} \alpha w_n} \left(2^j y_n^\delta \right).$$

So we have proven the following theorem.

Theorem 3.1. *Assume that there is an $\alpha > 0$ satisfying* (14). *Then, the solution of the full angular problem $y^\delta = \mathcal{R}f + \eta$ via curvelet sparse regularization is given by the* (*closed*) *formula*

$$\hat{f} = \sum_{n \in \mathcal{I}} S_{2^{2j-1} \alpha w_n} \left(2^j y_n^\delta \right) \psi_n. \tag{15}$$

The relation between the reproducing formula (13) and (15) is now obvious. If the thresholding parameters $w_n = w_n(\delta)$ in (15) are chosen such that $w_n(\delta) \to 0$ as $\delta \to 0$, i.e., they vanish if there is no noise present in the data, then (15) reduces to (13). On the other hand, if the data is corrupted by noise, then, the curvelet sparse regularized solution is simply a thresholded version of the BCD reproducing formula. The stabilizing character of the curvelet sparse regularization is reflected by the inherent thresholding of the curvelet coefficients (see also (12)).

In [6], a very similar reconstruction rule was derived. Starting form the BCD reproducing formula (13), the authors proposed to use soft-thresholding of curvelet coefficients with a scale dependent threshold τ_j, i.e.,

$$\hat{f} = \sum_{n \in \mathcal{I}} S_{\tau_j} \left(2^j \langle y^\delta, U_n \rangle_{L^2(\mathbb{R} \times S^1)} \right) \psi_n. \tag{16}$$

We see that this formula coincides with (15) for a suitably chosen thresholding sequence $\tau = (\tau_j)$.

Remark. Note the ill-posed nature of the reproducing formula (13). This is evident because the coefficients $2^j \langle \mathcal{R}f, U_\mu \rangle_{L^2(\mathbb{R} \times S^1)}$ corresponding to fine scales (large j) are amplified by the factor 2^j. Since noise is a fine scale phenomenon, there will be very large reconstruction errors when the data is corrupted by noise.

3.2. Limited angular range and limitations of the BCD

We have seen that there is an explicit expression (15) for the CSR reconstruction in the case of full angular tomography. We also noted that the thresholded BCD reconstruction (16) leads (under certain conditions) to the same reconstruction. To extend this observation to the limited angle tomography a biorthogonal curvelet decomposition for the limited angle Radon transform would be needed. To our knowledge there is no such BCD available for the limited angle Radon transform. Consequently, in the case of limited angle tomography, the CSR reconstruction (10) cannot be expressed explicitly as it was done for the full angular range in (15) and the BCD reconstruction of Candès and Donoho [6] cannot be applied in this situation.

In contrast to the BCD method, a reconstruction of the limited angle problem can be computed by our method (CSR). Hence, curvelet sparse regularization can be understood as the natural generalization of the thresholded BCD reconstruction.

Curvelet sparse regularization offers even more flexibility compared to the BCD method. For example, the implementation of the thresholded BCD method is difficult for acquisition geometries which are different from the parallel geometry. This is because the BCD method requires discretization of the functions U_n which live in the Radon domain. The implementation of the curvelet sparse regularization approach, however, does not rely on a specific acquisition geometry. One needs only to implement the system matrix. Moreover, the generalization to higher dimensions is also easier accessible via curvelet sparse regularization approach.

4. Characterization of the limited angle Radon transform

In Section 2 we presented curvelet sparse regularization as our method of choice for the limited angle tomography because it is stable, edge-preserving and flexible. In Proposition 2.1 we noted the existence of a solution and showed that each minimizer of the ℓ^1-penalized Tikhonov functional (9) is given as a fixed point of some operator (cf. (12)). This characterization is generic in the sense that it does not take into account the special structure of the underlying problem, which is in our case the limited angle tomography.

The goal of this section is to give a characterization of the minimizer (9) which is adapted to the setting of limited angle geometry. In the following we will show that, depending on the available angular range, a big portion of the curvelet coefficients of the CSR reconstruction are zero.

We state our main results first and postpone the proofs to the end of this section.

Theorem 4.1. *Let* $0 < \Phi < \pi/2$. *We define the polar wedge* W_Φ *by*

$$W_\Phi = \left\{ \xi \in \mathbb{R}^2 \colon \xi = r(\cos\omega, \sin\omega), \ r \in \mathbb{R}, \ |\omega| \leqslant \Phi \right\}. \tag{17}$$

Moreover, we define the invisible subset of the curvelet index set by

$$\mathcal{I}_\Phi^{\text{invisible}} = \left\{ (j, l, k) \in \mathcal{I} \colon \operatorname{supp} \hat{\psi}_{j,l,k} \cap W_\Phi = \emptyset \right\}, \tag{18}$$

where $\psi_{j,l,k}$ *denotes a curvelet and* $\mathcal{I}$ *is the curvelet index set (cf. Section 2.1). Then,*

$$\mathcal{R}_\Phi \psi_{j,l,k} \equiv 0 \quad \text{for all } (j, l, k) \in \mathcal{I}_\Phi^{\text{invisible}}. \tag{19}$$

The above theorem characterizes a subspace of the kernel of the limited angle Radon transform $\mathcal{R}_\Phi$ in terms of curvelets. As a consequence of Theorem 4.1, we can derive a characterization of CSR reconstructions at a limited angular range.

Theorem 4.2. *Let* $0 < \Phi < \pi/2$, $y^\delta \in \operatorname{ran}(\mathcal{R}_\Phi)$ *and let* $\mathcal{I}_\Phi^{\text{invisible}}$ *be defined by* (18). *Then, each minimizer*

$$\hat{c} \in \operatorname*{arg\,min}_{c \in \mathbb{R}^N} \left\{ \frac{1}{2} \left\| Kc - y^\delta \right\|_{L^2(S^1 \times \mathbb{R})}^2 + \|c\|_{1,w} \right\}$$

satisfies

$$\hat{c}_{j,l,k} = 0 \quad \text{for all } (j, k, l) \in \mathcal{I}_\Phi^{\text{invisible}}.$$

We start to develop the proof of Theorem 4.1 first. To this end, we need some auxiliary results. Though the content of the following lemma is classical, we will give a proof for the sake of completeness.

Lemma 4.3. *Let* $b > 0$ *and let* δ^b *be the function*

$$\delta^b(x) = \frac{1}{2\pi} \int_{-b}^{b} e^{\pm ix\cdot\xi} \, d\xi. \tag{20}$$

Then, $\delta^b \to \delta$ *pointwise in* $\mathcal{S}'(\mathbb{R})$ *as* $b \to \infty$, *i.e., for* $\varphi \in \mathcal{S}(\mathbb{R})$ *we have*

$$\lim_{b \to \infty} \int_{\mathbb{R}} \delta^b(x)\varphi(x) \, dx = \varphi(0). \tag{21}$$

Proof. First note that $\delta^b \in L^2(\mathbb{R})$ and $\hat{\delta}^b(\xi) = \frac{1}{\sqrt{2\pi}} \chi_{[-b,b]}$ in $L^2(\mathbb{R})$. Then, by Plancharel's formula we have

$$\lim_{b \to \infty} \int_{\mathbb{R}} \delta^b(x) f(x) \, dx = \lim_{b \to \infty} \int_{\mathbb{R}} \hat{\delta}^b(\xi) \hat{f}(\xi) \, d\xi = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{f}(\xi) \, d\xi = f(0). \qquad \square$$

The key observation for the proof of Theorem 4.1 is contained in the following lemma.

Lemma 4.4. *For $f \in \mathcal{S}(\mathbb{R}^2)$ and $\xi = (\cos \eta, \sin \eta)$, $\xi^\perp = (-\sin \eta, \cos \eta)$, we have*

$$\int_{\mathbb{R}} f(t \cdot \xi) \, \mathrm{d}t = \int_{\mathbb{R}} \hat{f}(t \cdot \xi^\perp) \, \mathrm{d}t. \tag{22}$$

That is, integration in the spatial domain along a line through the origin corresponds to the integration along a perpendicular line through the origin in the frequency domain.

Proof. The identity (22) follows easily from Lemma 4.4. More specifically, to see (22) in the case $\xi = (1, 0)$ we deduce the assertion as follows:

$$\int_{\mathbb{R}} f(t \cdot \xi) \, \mathrm{d}t = \int_{\mathbb{R}} f(x_1, 0) \, \mathrm{d}x_1 = \frac{1}{2\pi} \int_{\mathbb{R}} \int_{\mathbb{R}^2} \hat{f}(\xi) e^{ix_1 \xi_1} \, \mathrm{d}\xi \, \mathrm{d}x_1$$

$$= \frac{1}{2\pi} \lim_{b \to \infty} \int_{-b}^{b} \int_{\mathbb{R}^2} \hat{f}(\xi) e^{ix_1 \xi_1} \, \mathrm{d}\xi \, \mathrm{d}x_1$$

$$= \frac{1}{2\pi} \lim_{b \to \infty} \int_{\mathbb{R}^2} \hat{f}(\xi) \left(\int_{-b}^{b} e^{ix_1 \xi_1} \, \mathrm{d}x_1 \right) \mathrm{d}\xi$$

$$= \lim_{b \to \infty} \int_{\mathbb{R}} \int_{\mathbb{R}} \hat{f}(\xi_1, \xi_2) \delta^b(\xi_1) \, \mathrm{d}\xi_1 \, \mathrm{d}\xi_2$$

$$= \int_{\mathbb{R}} \left(\lim_{b \to \infty} \int_{\mathbb{R}} \hat{f}(\xi_1, \xi_2) \delta^b(\xi_1) \, \mathrm{d}\xi_1 \right) \mathrm{d}\xi_2$$

$$= \int_{\mathbb{R}} \hat{f}(0, \xi_2) \, \mathrm{d}\xi_2 = \int_{\mathbb{R}} \hat{f}(t \cdot \xi^\perp) \, \mathrm{d}t.$$

Note that the change of the integration order and the interchange of the limit process and integration are allowed due to Fubini's theorem and the dominated convergence theorem, respectively. The general case follows from an orthogonal change of variables. $\square$

We are now able to derive a formula for the Radon transform of curvelets.

Theorem 4.5. *Let $\psi_{j,l,k}$ be a curvelet (cf. (5)) and denote $\theta(\omega) = (\cos \omega, \sin \omega)^\mathsf{T}$. Then,*

$$\mathcal{R}\psi_{j,l,k}(\theta(\omega), s) = 2^{j/4} V\left(\frac{2^{\lceil j/2 \rceil + 1}}{\pi} (\omega + \theta_{j,l}) \right) \sqrt{2\pi} \, \widehat{W}\left(2^j \langle b_k^{j,l}, \theta(\omega + \theta_{j,l}) \rangle - 2^j s \right), \tag{23}$$

where $b_k^{j,l}$ and $\theta_{j,l}$ are defined in Section 2.1.

Proof. Let $\theta := \theta(\omega) = (\cos \omega, \sin \omega)^\mathsf{T}$, $\omega \in [-\Phi, \Phi]$, and $\tau_p f := f(\cdot + p)$, $p \in \mathbb{R}^2$. First note that each curvelet $\psi_{j,l,k}$ is a Schwartz function since its Fourier transform is C^∞ and compactly supported (per definition). Hence, we may apply Lemma 4.4:

$$\mathcal{R}_\Phi \psi_{j,l,k}(\theta, s) = \int_{\mathbb{R}} \psi_{j,l,k}(s\theta + t\theta^\perp) \, \mathrm{d}t = \int_{\mathbb{R}} (\tau_{s\theta} \psi_{j,l,k})(t\theta^\perp) \, \mathrm{d}t = \int_{\mathbb{R}} e^{i\langle s\theta, t\theta \rangle} \hat{\psi}_{j,l,k}(t\theta) \, \mathrm{d}t. \tag{24}$$

At scale 2^{-j}, each curvelet $\psi_{j,l,k}$ is defined via translation and rotation of a generating curvelet $\psi_{j,0,0}$, cf. (5). Using the relation of the Fourier transform and rotation together with the relation $\widehat{\tau_p f}(\xi) = e^{i\langle p, \xi \rangle} \hat{f}(\xi)$ we see that

$$\hat{\psi}_{j,l,k}(\xi) = e^{-i\langle b_k^{j,l}, R_{\theta_{j,l}} \xi \rangle} \hat{\psi}_{j,0,0}(R_{\theta_{j,l}} \xi). \tag{25}$$
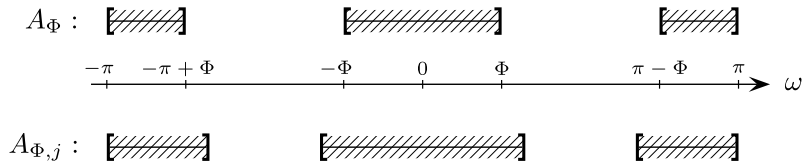
By plugging (25) into (24) we deduce

Fig. 3. The symmetric (visible) angular range of the limited angle Radon transform, A_Φ, and its scale-dependent version, $A_{\Phi,j}$.

$$
\mathcal{R}_\Phi \psi_{j,l,k}(\theta, s) = \int_{\mathbb{R}} e^{ist} e^{-i\langle b_k^{j,l}, R_{\theta_{j,l}}(t\theta)\rangle} \hat{\psi}_{j,0,0}\big(R_{\theta_{j,l}}(t\theta)\big)\, \mathrm{d}t
$$

$$
= \int_{\mathbb{R}} e^{ist} e^{-it\langle b_k^{j,l}, \theta(\omega+\theta_{j,l})\rangle} \hat{\psi}_{j,0,0}(t, \omega+\theta_{j,l})\, \mathrm{d}t
$$

$$
= 2^{-3j/4} V\left(\frac{2^{\lceil j/2\rceil+1}}{\pi}(\omega+\theta_{j,l})\right) \int_{\mathbb{R}} e^{ist} e^{-it\langle b_k^{j,l}, \theta(\omega+\theta_{j,l})\rangle} W\big(2^{-j}t\big)\, \mathrm{d}t
$$

$$
= 2^{j/4} V\left(\frac{2^{\lceil j/2\rceil+1}}{\pi}(\omega+\theta_{j,l})\right) \int_{\mathbb{R}} e^{-ir[2^j\langle b_k^{j,l}, \theta(\omega+\theta_{j,l})\rangle - s]} W(r)\, \mathrm{d}r
$$

$$
= 2^{j/4} V\left(\frac{2^{\lceil j/2\rceil+1}}{\pi}(\omega+\theta_{j,l})\right) \sqrt{2\pi}\, \widehat{W}\big(2^j\langle b_k^{j,l}, \theta(\omega+\theta_{j,l})\rangle - 2^j s\big). \qquad \square
$$

The proof of Theorem 4.1 is now a simple consequence of Theorem 4.5.

Proof of Theorem 4.1. Let $j \in \mathbb{N}$. To abbreviate the notation we let $a_j = \pi^{-1} 2^{\lceil j/2\rceil+1}$ and denote by

$$
A_\Phi := [-\pi, -\pi+\Phi] \cup [-\Phi, \Phi] \cup [\pi-\Phi, \pi]
$$

the symmetric (visible) angular range of the limited angle Radon transform $\mathcal{R}_\Phi$ (cf. Fig. 3).

According to Theorem 4.5, we have to determine all $\theta_{j,l}$ such that $V(a_j(\cdot + \theta_{j,l}))|_{A_\Phi} \equiv 0$. Since supp $V \subset (-1, 1)$ and $\theta_{j,l} \in [-\pi, \pi]$, we have

$$
\theta_{j,l} \notin \big(-a_j^{-1} - \omega, a_j^{-1} - \omega\big) \quad \Rightarrow \quad V\big(a_j(\omega+\theta_{j,l})\big) = 0
$$

for all $\omega \in A_\Phi$. Therefore, by defining the scale-dependent version of A_Φ as (cf. Fig. 3)

$$
A_{\Phi,j} := \big[-\pi, -\pi+\big(\Phi+a_j^{-1}\big)\big] \cup \big[-\big(\Phi+a_j^{-1}\big), \Phi+a_j^{-1}\big] \cup \big[\pi-\big(\Phi+a_j^{-1}\big), \pi\big],
$$

we see that $V(a_j(\cdot + \theta_{j,l}))|_{A_\Phi} \equiv 0$ holds whenever $\theta_{j,l} \notin A_{\Phi,j}$. The assertion follows by defining the invisible index set of curvelet coefficients as

$$
\mathcal{I}_\Phi^{\text{invisible}} = \big\{(j,l,k) \in \mathcal{I}: \theta_{j,l} \notin A_{\Phi,j}\big\}. \qquad \square
$$

We summarize the steps that were needed to prove Theorem 4.1. This procedure is also illustrated in Fig. 4. To evaluate $\mathcal{R}_\Phi \psi_{j,l,k}(\xi, s)$ for a curvelet $\psi_{j,l,k}$, the integration was shifted to the Fourier domain according to Lemma 4.4. This related the value $\mathcal{R}_\Phi \psi_{j,l,k}(\xi, s)$ to the integration of $\hat{\psi}_{j,l,k}$ along the line $L_\xi = \{t\xi^\perp: t \in \mathbb{R}\}$. Because of the limited angular range, the union of all such lines,

$$
W_\Phi = \bigcup_{\eta \in [-\Phi, \Phi]} L_{\xi(\eta)}, \tag{26}
$$

covers not all of the $\mathbb{R}^2$. To prove the assertion, we computed all those curvelet indices (j, l, k) such that

$$
\text{supp } \hat{\psi}_{j,l,k} \cap W_\Phi = \emptyset.
$$

Now, we turn to the proof of Theorem 4.2. This will be a simple consequence of Theorem 4.1 and the following lemma.

Lemma 4.6. *Let $f : \ell^2(\mathcal{I}) \to [-\infty, \infty]$ be defined by $f(x) = \sum_{n \in \mathcal{I}} \varphi(x_n)$, where $\varphi : \mathbb{R} \to \mathbb{R}$ is a convex function such that f is proper. Then, it holds that*

$$
y \in \partial f(x) \quad \Leftrightarrow \quad y_n \in \partial \varphi(x_n) \quad \text{for all } n \in \mathcal{I}. \tag{27}
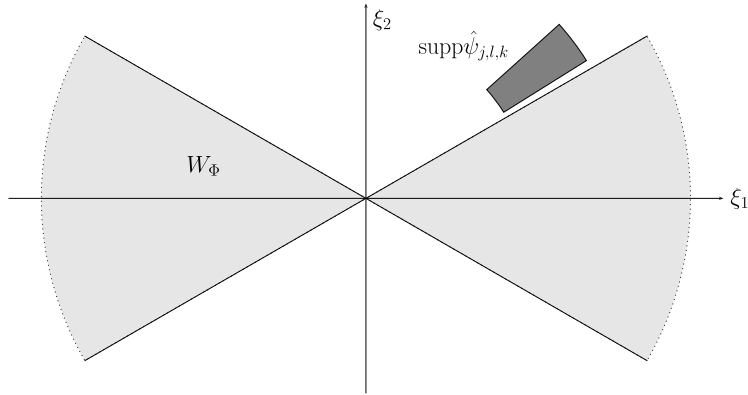$$

Fig. 4. Theorem 4.1 states that a curvelet $\psi_{j,l,k}$ lies in the kernel of the limited angle Radon transform whenever the support of $\psi_{j,l,k}$ lies outside the "visible wedge" $W_\Phi = \{r(\cos\eta, \sin\eta): r \in \mathbb{R}, \ \eta \in [-\Phi, \Phi]\}$. That is, $\mathcal{R}_\Phi \psi_{j,l,k} \equiv 0$ whenever $\operatorname{supp}\hat\psi_{j,l,k} \cap W_\Phi = \emptyset$.

Proof. First note that, since f is proper, for $x \in \ell^2(\mathcal{I})$ we have $\partial f(x) = \emptyset$ if $f(x) = \infty$. In what follows we therefore assume without loss of generality that $f(x) < \infty$.

Suppose $y_n \in \partial\varphi(x_n)$ for all $n \in \mathcal{I}$. Then, by definition of the subgradient we have

$$\forall z_n \in \mathbb{R}: \ \varphi(z_n) \geqslant \varphi(x_n) + y_n(z_n - x_n).$$

Summing over n implies

$$\forall z \in \ell^2(\mathcal{I}): \ \sum_{n \in \mathcal{I}} \varphi(z_n) \geqslant \sum_{n \in \mathcal{I}} \big(\varphi(x_n) + y_n(z_n - x_n)\big) = \sum_{n \in \mathcal{I}} \varphi(x_n) + \langle y, z - x \rangle,$$

which is by definition of f equivalent to $y \in \partial f(x)$. This proves the implication "$\Leftarrow$" of the statement.

On the other hand, if $y \in \partial f(x)$, then $f(z) \geqslant f(x) + \langle y, z - x \rangle$ for all $z \in \ell^2(\mathcal{I})$. In particular, this holds for all $z = x + h e_n$ with $h \in \mathbb{R}$ and $n \in \mathcal{I}$, where $e_n = (\delta_{i,n})_{i \in \mathcal{I}}$ and $\delta_{i,n}$ denotes the Kronecker delta. Therefore we have

$$\forall h \in \mathbb{R} \, \forall n \in \mathcal{I}: \ f(x + h e_n) \geqslant f(x) + \langle y, h e_n \rangle \quad \Leftrightarrow \quad \sum_{i \in \mathcal{I}} \varphi(x_i + \delta_{i,n} h) - \sum_{i \in \mathcal{I}} \varphi(x_i) \geqslant y_n h$$

$$\Leftrightarrow \quad \varphi(x_n + h) - \varphi(x_n) \geqslant y_n h$$

$$\Leftrightarrow \quad y_n \in \partial\varphi(x_n). \quad \square$$

Proof of Theorem 4.2. As in the proof of Proposition 2.1 we see that $\hat c$ fulfills the following relation

$$-K^*\big(K\hat c - y^\delta\big) \in \partial \|\hat c\|_{1,w}. \tag{28}$$

Since $y^\delta \in \operatorname{ran}(\mathcal{R})_\Phi$ we have that $y^\delta = Kc^\delta = \mathcal{R}_\Phi T^* c^\delta$ for some curvelet coefficient vector $c^\delta \in \mathbb{R}^N$. Thus,

$$\hat x := -K^*\big(K\hat c - y^\delta\big) = -K^* K\big(\hat c - c^\delta\big) = -K^*\left(\sum_{j,l,k} \big(\hat c - c^\delta\big)_{j,l,k} \mathcal{R}_\Phi \psi_{j,l,k}\right).$$

By Theorem 4.1, it holds that

$$\hat x_{j,l,k} = 0 \quad \text{for all } (j, k, l) \in \mathcal{I}_\Phi^{\text{invisible}}. \tag{29}$$

Now let $(j, l, k) \in \mathcal{I}_\Phi^{\text{invisible}}$. In view of (28) and (29), Lemma 4.6 implies that $0 \in \partial(w_{j,l,k}|\hat c_{j,l,k}|)$. However, this means that $\hat c_{j,l,k}$ minimizes the function $w_{j,l,k} |\cdot|$ and since $w_{j,l,k} > 0$ we get that $\hat c_{j,l,k} = 0$. $\quad \square$

5. Adapted curvelet sparse regularization

In this section we are going to apply the results from Section 4 to a finite dimensional reconstruction problem. We will show that, in this setting, a significant dimensionality reduction can be achieved. Based on this approach, we will formulate the adapted curvelet sparse regularization (A-CSR).

5.1. Discrete reconstruction problem

We consider the discrete version of the reconstruction problem. To this end, we model f as a finite linear combination of curvelets, i.e.,

$$f = \sum_{n=1}^{N} c_n \psi_n, \tag{30}$$

where $n = n(j, l, k)$ is an enumeration of the curvelet index set $\mathcal{I}$ (cf. Section 2.1) and $N = |\mathcal{I}|$. In practice, the function f is represented by a matrix and the size of this matrix determines the curvelet index set $\mathcal{I}$ (cf. (6)). However, we are not going to address this problem in what follows. In order to perform curvelet decomposition of an image, we will make use of the Matlab library CurveLab version 2.1.2 [4]. In the following, we assume to be given a finite number of measurements $y_m = \mathcal{R}_\Phi f(\theta_m, s_m)$, $1 \leqslant m \leqslant M \in \mathbb{N}$. Using (30), each measurement y_m can be expressed as

$$y_m = \mathcal{R}_\Phi f(\theta_m, s_m) = \sum_{n=1}^{N} c_n \mathcal{R}_\Phi \psi_n(\theta_m, s_m). \tag{31}$$

Now, let us define the so-called *system matrix* K by

$$K_{m,n} = \mathcal{R}_\Phi \psi_n(\theta_m, s_m), \quad 1 \leqslant m \leqslant M, \ n \in \mathcal{I}. \tag{32}$$

Then, the discrete version of the limited angle problem (8) reads

$$y = Kc + \eta. \tag{33}$$

Note that the matrix K is the discrete version of the continuous operator $K = \mathcal{R}_\Phi T^*$. We want to point out, that the reconstruction problem (33) is formulated in terms of all curvelet coefficients $c = Tf$. In order to solve (33), we need to compute $c_{n(j,l,k)}$ for all possible curvelet indices $(j, l, k) \in \mathcal{I}$. Therefore, the dimension of the reconstruction problem is given by $N = |\mathcal{I}|$. In particular, the dimension does not depend on the available angular range in this naive approach. In what follows, we will use the method of curvelet sparse regularization to solve this problem.

5.2. Dimensionality reduction and adapted curvelet sparse regularization (A-CSR)

We first, note that all characterizations that were formulated in Section 4 depend only on the angular range parameter Φ. In turn, this parameter is completely determined by the acquisition geometry. Hence, it is known prior to the reconstruction and can be extracted from the given data by

$$\Phi = \min\{\varphi \colon \exists \varphi_0 \in [-\pi, \pi] \colon \forall 1 \leqslant m \leqslant M \colon \theta_m \in [\varphi_0 - \varphi, \varphi_0 + \varphi]\}, \tag{34}$$

where $\theta_1, \ldots, \theta_M$ denote the projection angles. Having determined Φ, we can use Theorem 4.2 to identify those curvelets that lie in the null space of the limited angle Radon transform $\mathcal{R}_\Phi$. Their index set can be precomputed according to (18) or, equivalently, by

$$\mathcal{I}_\Phi^{\text{invisible}} = \{(j, l, k) \in \mathcal{I} \colon (\cos \theta_{j,l}, \sin \theta_{j,l})^T \notin W_{\Phi,j}\}, \tag{35}$$

where $W_{\Phi,j} = \{\xi \in \mathbb{R}^2 \colon \xi = r(\cos \omega, \sin \omega)^T, \ r \in \mathbb{R}, \ |\omega| < \Phi + \pi 2^{-\lceil j/2 \rceil - 1}\}$ is a polar wedge at scale 2^{-j} and $\theta_{j,l}$ is the orientation of the curvelet $\psi_{j,l,k}$. In what follows, the curvelets $\psi_{n(j,l,k)}$ as well as the curvelet coefficients $c_{n(j,l,k)}$ with indices $(j, l, k) \in \mathcal{I}_\Phi^{\text{invisible}}$ will be called *invisible*[1] from the given angular range. Accordingly, we define *the index set of visible curvelet coefficients* by

$$\mathcal{I}_\Phi^{\text{visible}} = \mathcal{I} \setminus \mathcal{I}_\Phi^{\text{invisible}}. \tag{36}$$

In view of Theorem 4.1, it holds that $K_{m,n} = \mathcal{R}_\Phi \psi_n(\theta_m, s_m) \equiv 0$ for $n \in \mathcal{I}_\Phi^{\text{invisible}}$ and for all $1 \leqslant m \leqslant M$, i.e., those columns of the system matrix K in (32), that correspond to the invisible index set are identified to be zero. Therefore, we may define a new (reduced) system matrix K_Φ with respect to the visible index set by

$$(K_\Phi)_{m,n} = \mathcal{R}_\Phi \psi_n(\theta_m, s_m), \quad 1 \leqslant m \leqslant M, \ n \in \mathcal{I}_\Phi^{\text{visible}}. \tag{37}$$

Such a *reduced system matrix* K_Φ has the size $M \times |\mathcal{I}_\Phi^{\text{visible}}|$. Since $|\mathcal{I}_\Phi^{\text{visible}}| = |\mathcal{I}| - |\mathcal{I}_\Phi^{\text{invisible}}|$, the number of columns is reduced by the quantity $|\mathcal{I}_\Phi^{\text{invisible}}|$. Using the reduced system matrix we formulate the *adapted (or reduced) limited angle problem* as

$$y^\delta = K_\Phi c + \eta. \tag{38}$$

---

[1] We adapted the term *invisible* from [30].

The dimension of the adapted problem is given by $N_\Phi = |\mathcal{I}| - |\mathcal{I}_\Phi^{\text{invisible}}|$. Obviously, this depends on the angular range parameter Φ. From the definition of $\mathcal{I}_\Phi^{\text{visible}}$ it is clear that, as the angular range becomes larger, the number of visible curvelets increases. Hence, the dimension of the adapted problem N_Φ increases as the angular range increases and vice versa. We formulate the *adapted curvelet sparse regularization* (A-CSR) as

$$\hat{c}_\Phi = \underset{c \in \mathbb{R}^{N_\Phi}}{\arg\min} \left\{ \frac{1}{2} \| K_\Phi c - y^\delta \|_2^2 + \| c \|_{1,w} \right\}. \tag{A-CSR}$$

Apparently, the computational amount decreases by using the A-CSR framework instead of the CSR. However, according to Theorem 4.2, the reconstruction quality is preserved. In Section 8, we will present some practical experiments confirming these issues.

Remark. Note that the characterization of Theorem 4.2 may also be applied to the closed form formula (15) by replacing the index of summation $\mathcal{I}$ by $\mathcal{I}_\Phi^{\text{visible}}$. This yields a closed form solution for (A-CSR).

6. Discussion

This section is devoted to the discussion of the results which were presented in the previous section as well as their implications for the practical application of the curvelet sparse regularization.

General angular ranges. So far, we have worked with a symmetric angular range $[-\Phi, \Phi]$ with $0 < \Phi < \pi/2$ which was centered at $\Phi_0 = 0$. The results of Section 4, however, can be easily adapted to a more general situation, where the available angular range $[\Phi_0 - \Phi, \Phi_0 + \Phi]$ is centered around an angle $\Phi_0 \in [-\pi, \pi]$. To this end, let T_{Φ_0} be the translation operator defined by $T_{\Phi_0} \mathcal{R}_\Phi f(\theta, s) = \mathcal{R}_\Phi f(\theta + \Phi_0, s)$. Then, the limited angle Radon transform with respect to a general angular range $[\Phi_0 - \Phi, \Phi_0 + \Phi] \times \mathbb{R}$ is given by $T_{\Phi_0} \mathcal{R}_\Phi$. Theorems 4.1 and 4.2 can be now applied to $T_{\Phi_0} \mathcal{R}_\Phi$, yielding a general index set of invisible curvelet coefficients

$$\mathcal{I}_\Phi^{\text{invisible}} = \left\{ (j, l, k) \in \mathcal{I} : \text{supp}\, \hat{\psi}_{j,l,k} \cap R_{\Phi_0} W_\Phi = \emptyset \right\},$$

where $R_{\Phi_0} W_\Phi = \{ R_{\Phi_0} \xi : \xi \in W_\Phi \}$ is a rotated version of W_Φ.

Computation of the system matrix. In Theorem 4.5 we have derived an expression for the Radon transform of a curvelet $\psi_{j,l,k}$. This expression can be used to compute the entries of the system matrix Φ analytically, if both, the angular window V and the Fourier transform of the radial window $\widehat{W}$ are known analytically. This is useful for practical application since, in this case, the system matrix can be precomputed and needs not to be set up in every iteration of the minimization of the ℓ^1-penalized Tikhonov functional. This may yield an additional speedup of the algorithm.

Additional stabilization of the limited angle problem. Adapting the problem to the limited angular range has an additionally stabilizing effect. This comes from the fact that the reconstruction problem (38) is formulated with respect to visible curvelet coefficients only. In this way, a big portion of the null space of the system matrix (limited angle Radon transform) is excluded from the formulation of the limited angle problem. Therefore, the condition number of the reduced system matrix improves and induces an additional stability.

We want to point out that the formulation of the adapted problem (38) does only depend on null space analysis of the limited angle Radon transform in terms of curvelets (cf. Theorem 4.1). Thus, the adapted limited angle problem (38) is not related to any reconstruction algorithm. Therefore, the additional stabilization will be present if any other method would be used for solving (38). The adapted formulation of the reconstruction problem (38) can be therefore interpreted as preconditioning procedure.

Related work. In [18], the adapted curvelet sparse regularization was introduced using microlocal analysis. There, a qualitative characterization of visible curvelets was derived from the characterization of visible singularities of E.T. Quinto [29] and "the resolution of wavefront set property" of the continuous curvelet transform [8]. These results were stated there without proofs.

We also would like to mention that our approach can be formulated with respect to other "curvelet-like" systems. A very important example in this context is the shearlet system [15]. Shearlets are also highly directional and compactly supported in the Fourier domain and they provide an optimally sparse representation of cartoon-like objects. They also posses the "the resolution of wavefront set" property [24] and there is also a biorthogonal shearlet decomposition of the Radon transform [12]. Therefore, the results of this work hold also for the shearlet system. The proofs can be adopted almost literally in this situation.

7. Implementation of the minimization algorithm

For the minimization of the ℓ^1-penalized Tikhonov functional (9) we implemented a variant of the well-known *iterative soft-thresholding* algorithm [13,2]. This algorithm is given as a fixed point iteration of Eq. (12), namely

$$c^{n+1} = \mathcal{S}_{\tau^n} \big(c^n - s_n K^* (K c^n - y^\delta) \big), \tag{39}$$

where $K = \mathcal{R}_\Phi T^*$, $\mathcal{R}_\Phi$ denotes the limited angle Radon transform and T^* is the synthesis operator for the curvelet frame. We will use this algorithm to compute CSR reconstructions as well as A-CSR reconstructions. Hence, K may denote the full system matrix K (cf. (32)) or the reduced system matrix K_Φ (cf. (37)), respectively. For the computation of the reduced system matrix, we first compute the index set of visible curvelet indices $\mathcal{I}_\Phi^{\text{visible}}$ according to (35), (36). Further we have to note that, instead of computing the system matrix K and storing it in the memory, we implemented the matrix-vector multiplication $c \mapsto Kc = \mathcal{R}_\Phi T^* c$ and its adapted version using the Matlab function `radon` and the CurveLab version 2.1.2 [4].

The iteration (39) consists of a gradient descent step followed by a soft-thresholding with respect to a thresholding sequence $\tau = (\tau_{(j,l,k)})_{(j,l,k) \in \mathcal{I}}$ that may depend on all curvelet parameters (j, l, k). The step length s_n should satisfy

$$0 < \underline{s} \leqslant s_n \leqslant \bar{s} < 2/\|K\|^2, \tag{40}$$

where $\underline{s}, \bar{s} > 0$, cf. [2]. In all of our experiments we will use a constant steplength $s_n = s$, with $s > 0$ satisfying (40).

According to (12), the thresholding sequence τ has to be chosen as $\tau = s \odot w$, where $\odot$ denotes pointwise multiplication of the step length s and the ℓ^1-norm weight sequence w (cf. (9)). The thresholding sequence is a free parameter and has to be selected appropriately because it greatly influences the reconstruction quality. In general, there is no rule how to choose such a thresholding sequence τ. In what follows, we will use two different choice rules for τ. Firstly, we will set τ to be a constant sequence, i.e.,

$$\tau_{(j,l,k)} = \alpha \tag{41}$$

with $\alpha > 0$. In this case the value of α will be set manually for each reconstruction and is based on the visual inspection of the reconstruction quality. Secondly, the thresholding sequence $\tau = (\tau_{(j,l,k)})$ will be chosen adaptively and scale-dependent via

$$\tau_{(j,l,k)} = 2^{j-J} \sigma \sqrt{2 \log_e N_{j,l}}, \tag{42}$$

where σ is the standard deviation of the noise η, $N_{j,l}$ denotes the number of curvelet coefficients at scale 2^{-j} and at orientation $\theta_{j,l}$ and $J \in \mathbb{N}$ is the largest available scale parameter for the image size of interest.[2] This thresholding rule was adapted from [6, Section 6]. To mimic practical situation, we will assume the data to be corrupted by an additive white Gaussian noise, i.e., $y^\delta = y + \eta$ and $\eta \in \mathcal{N}(0, \sigma^2)$, with an unknown standard deviation σ. In order to automatize the reconstruction procedure, we will estimate σ as follows (cf. [27, p. 565])

$$\sigma \approx 1.4826 \cdot \text{MAD}(c_J^n). \tag{43}$$

Above, $\text{MAD}(c_J^n)$ denotes the median of the absolute values of curvelet coefficients c_J^n corresponding to the finest scale 2^{-J} at iteration n. We would like to note that the second choice rule yields a reconstruction procedure that is free of any parameter.

To complete the description of our algorithm we have to specify the choice of an initial guess c^0 for the iteration (39). Since the ℓ^1-penalized Tikhonov functional is convex, the iteration (39) always converges to a global minimizer. Therefore, the choice of an initial guess does not matter in theory. However, in practice it influences the speed of convergence. In all of our experiments we will make the following choices:

$$c^0 \in \{\mathbf{0}, c_{\text{FBP}}\}. \tag{44}$$

Above, $\mathbf{0}$ denotes the zero vector and c_{FBP} denotes the vector of curvelet coefficients of a filtered backprojection reconstruction. To compute c_{FBP}, first, a FBP reconstruction f_{FBP} was computed using the Matlab function `iradon` and, afterwards, the curvelet coefficients c_{FBP} were obtained by computing the curvelet transform of f_{FBP} using CurveLab. A summarized description of our reconstruction procedure is given in Algorithm 1.

Eventually, we would like to note, that all computations in this article that involve the curvelet transform were made using the CurveLab version 2.1.2 [4]. Readers who are interested in the implementation of the fast discrete curvelet transform are referred to [3].

8. Numerical experiments

This section is devoted to the illustration of our results that we presented in previous sections. In the first part of our experiments we will illustrate the visibility of curvelets under the limited angle Radon transform and show how this leads to a dimensionality reduction in the limited angle reconstruction problem. In particular, these experiments are meant to illustrate Theorems 4.1 and 4.2. The second part of our experiments is devoted to reconstructions via CSR, A-CSR and filtered backprojection (FBP). A comparison of these reconstructions will be presented in terms of execution times and reconstruction quality.

---

[2] For an image of size $N \times N$ the maximum scale parameter J was set to be $J = \lceil \log_2 N \rceil - 2$. This is according to the implementation of CurveLab version 2.1.2 [4].

Algorithm 1 Reconstruction algorithm

Require: Noisy data y^δ; Angular range $[-\Phi, \Phi]$; Image size $N \times N$;

1: $\mathcal{I}_\Phi \leftarrow$ visible index set at the angular range $[-\Phi, \Phi]$ according to (35), (36);

2: $K \in \{K, K_\Phi\}$ according to (32), (37);

3: $s \leftarrow$ steplength, such that (40) is satisfied;

4: $c^0 \leftarrow$ initial guess for the iteration according to (44);

5: $\tau_{(j,l,k)} \leftarrow$ choose thresholding sequence according to (41), (42) + (43);

6: $n_{\max} \leftarrow$ maximum number of iterations;

7: $n \leftarrow 0$;

8: **while** $(n \leqslant n_{\max})$ **do**

9:    $c^n \leftarrow \mathcal{S}_{\tau^n}(c^n - sK^*(Kc^n - y^\delta))$ with $\mathcal{S}_\tau$ defined in (11);

10:    $n \leftarrow n + 1$;

11: **end while**

12: $f_{\text{rec}} \leftarrow$ inverse curvelet transform of $c^{n_{\max}}$;
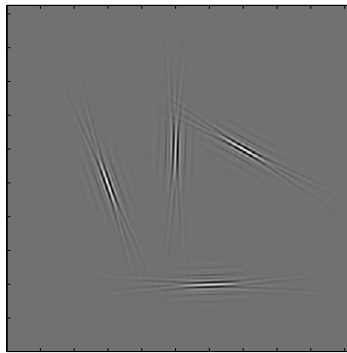
Fig. 5. A Matlab generated 512×512 image of the function given in (45).

8.1. Visibility of curvelets and dimensionality reduction

The aim of our first experiment, is to illustrate the visibility of curvelets under the action of the limited angle Radon transform $\mathcal{R}_\Phi$ for different values of Φ (cf. Theorem 4.1). To this end, we consider the function

$$f = \psi_1 + \psi_2 + \psi_3 + \psi_4, \tag{45}$$

which is given as a linear combination of curvelets ψ_i ($i \in \{1, 2, 3, 4\}$) at a fixed scale 2^{-4} and orientations $\theta_1 = 0°$, $\theta_2 = 20°$, $\theta_3 = 60°$ and $\theta_4 = 90°$. An image (512×512) of this function is shown in Fig. 5. Using the Matlab functions radon and iradon of the Image Processing Toolbox we computed the limited angle Radon transform of f and its inverse at an angular range $[-\Phi, \Phi]$ for $\Phi \in \{35°, 80°\}$. Within the considered angular range the Radon transform was sampled equidistantly with a sampling distance $\Delta\theta = 1°$. The total number of measurements M_Φ was given by $M_{35°} = 51\,759$ and $M_{80°} = 117\,369$.

The results of this experiment are shown in Fig. 6. The first column shows the limited angle Radon transforms $\mathcal{R}_\Phi f$ of f for different values of Φ, whereas the second column shows the inverse Radon transforms $\mathcal{R}_\Phi^{-1} \mathcal{R}_\Phi f$ from the corresponding limited angle data. In the first row we see that only those curvelets are visible in the Radon transform domain which correspond to curvelet orientations $\theta_1 = 0°$ and $\theta_2 = 20°$. The reconstruction from this data shows only two curvelets, namely those with orientations $\theta_1 = 0°$ and $\theta_2 = 20°$. We therefore conclude that

$$\mathcal{R}_{35°}^{-1} \mathcal{R}_{35°} f = \psi_1 + \psi_2.$$

In the second row, we can observe a similar effect. In the Radon transform domain, we can see that another curvelet ψ_3 (corresponding to $\theta_3 = 60°$) becomes visible. This curvelet also appears in the reconstruction and hence

$$\mathcal{R}_{80°}^{-1} \mathcal{R}_{80°} f = \psi_1 + \psi_2 + \psi_3.$$

To explain this effect we compute the set of visible curvelet coefficients according to (35), (36). As a result, we see that

$$\mathcal{I}_{35°}^{\text{visible}} = \{1, 2\} \quad \text{and} \quad \mathcal{I}_{80°}^{\text{visible}} = \{1, 2, 3\}.$$

This is of course in accordance with Theorem 4.1 and we see that the result of Theorem 4.1 immediately translates into practice. As a rule of thumb, we conclude that curvelets having orientations within the angular range $[-\Phi, \Phi]$ are visible to
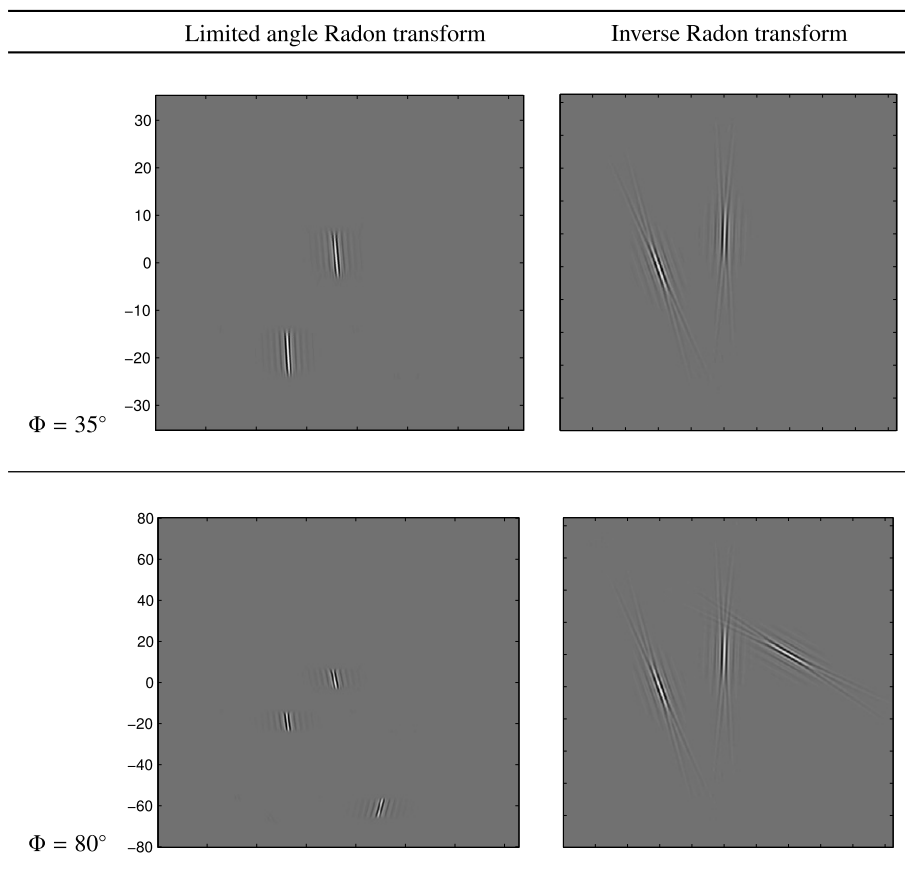
Fig. 6. Limited angle Radon transform $\mathcal{R}_\Phi f$ and the corresponding reconstructions $\mathcal{R}_\Phi^{-1}\mathcal{R}_\Phi f$ of the function given in (45) (cf. Fig. 5) at an angular range $[-\Phi, \Phi]$, $\Phi \in \{35°, 80°\}$, and an angular sampling distance $\Delta\Phi = 1°$. This figure illustrates Theorem 4.1. We can observe that at the angular range $[-35°, 35°]$ (first row) the only visible curvelets are those with orientations $\theta_1 = 0°$ and $\theta_2 = 20°$. However, by enlarging the angular range to $[-80°, 80°]$ (second row) another curvelet with orientation $\theta_3 = 60°$ becomes visible.

the limited angle Radon transform. However, curvelets which correspond to missing directions are not visible to the limited angle Radon transform.

Apparently, the above observation can be generalized as follows: Let $f = \sum_{n \in \mathcal{I}} c_n \psi_n$ be a curvelet expansion of a function f. Then, for a given angular range $[-\Phi, \Phi]$, we can separate the visible and invisible parts of this function by

$$f = \sum_{n \in \mathcal{I}_\Phi^{\text{visible}}} c_n \psi_n + \sum_{n \in \mathcal{I}_\Phi^{\text{invisible}}} c_n \psi_n, = f_{\text{visible}} + f_{\text{invisible}}.$$

Note that this separation depends only on the parameter Φ. According to Theorem 4.1 we have $\mathcal{R}_\Phi^{-1}\mathcal{R}_\Phi f = f_{\text{visible}}$. The reconstruction of f from $\mathcal{R}_\Phi f$ amounts to determining only the visible curvelet coefficients at a given angular range. In accordance to the discussion in Section 5, the reduced dimension of the limited angle problem is given by number of visible curvelets $|\mathcal{I}_\Phi^{\text{visible}}|$. In other words, the dimension of the problem in the curvelet domain is reduced by the number of invisible curvelets $|\mathcal{I}_\Phi^{\text{invisible}}|$. In order to investigate the dimensionality reduction at a particular angular range, we computed $d(\Phi) := |\mathcal{I}_\Phi^{\text{invisible}}|$ for various values of Φ.[3] The plot of $d(\Phi)$ is shown in Fig. 7 for an image of size 256×256. We can observe that the graph of $d(\Phi)$ exhibits the expected behavior: The dimension of the non-adapted problem in the curvelet domain is constant for all angular ranges. This is of course due to the fact that the non-adapted problem is formulated with respect to all curvelet coefficients. Hence, it is independent of Φ. However, the dimension of the adapted problem shows a strong dependence on the available angular range. We can observe a significant dimensionality reduction for any angular range parameter satisfying $\Phi \leqslant 153°$. Moreover, the reduced dimension $d(\Phi)$ seems to exhibit a piecewise constant behavior. The values of $d(\Phi)$ increase stepwise linearly as the angular range increases. The reason for this stepwise structure

---

[3] Note that the dimension $|\mathcal{I}|$ of the reconstruction problem depends on the size of the considered image f.
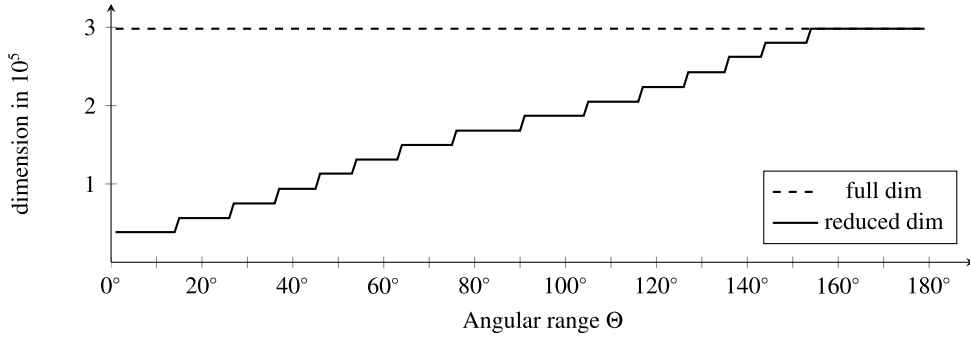
Fig. 7. Dimension of the full problem (33), - - -, and of the adapted problem (38), —, for an image of size 256×256. The plot shows the dependence of the full and reduced dimension on the available angular range $[0, \Theta]$. Since the full problem is formulated in terms of all curvelet coefficients, its dimension is constant for all angular ranges. The adapted problem, however, is formulated only in terms of visible curvelet coefficients. Hence, the reduced dimension depends strongly on the available angular range.

(a) Shepp-Logan head phantom   **(b)** Brainstem   **(c)** Radial pattern
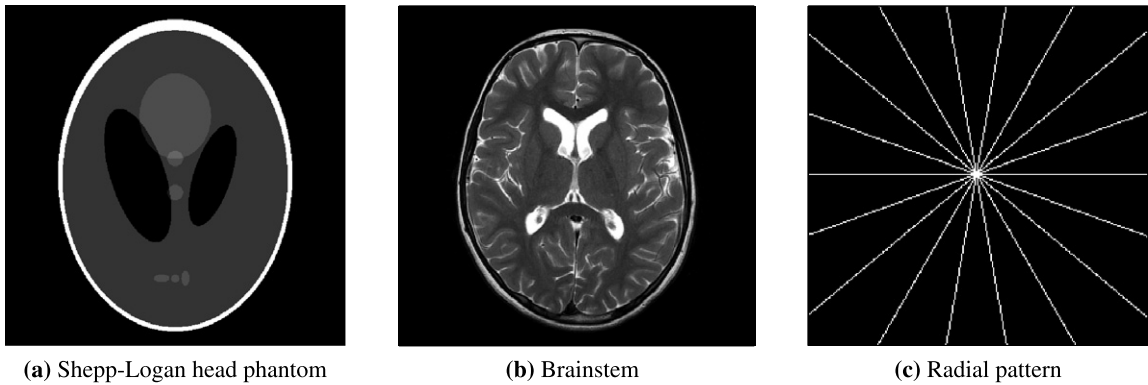
Fig. 8. Original images. See [31] for (b).

lies in the fact that curvelets remain visible as long as $\operatorname{supp} \hat{\psi}_{j,l,k} \cap W_\Phi \neq \emptyset$, see also Fig. 4. The length of one such step therefore corresponds to the length of the support of the angular window V of curvelets at the finest scale 2^{-J}, i.e., to $|\operatorname{supp} V(\frac{2^{\lceil J/2 \rceil + 1}}{\pi} \cdot)|$.

### 8.2. CSR vs. A-CSR: Execution times and reconstruction quality

In the following experiments we are going to investigate execution times and the reconstruction quality of CSR and A-CSR reconstructions at a limited angular range. We will also compare the reconstruction quality of CSR and A-CSR reconstruction to those reconstructions obtained via filtered backprojection.

8.2.1. Experimental setup

In order to minimize the influence of specific image features, we have used three different types of images for our experiments. These images are shown in Fig. 8. We note that all of the following computations were done in Matlab. The generation of the limited angle Radon transform data $y = \mathcal{R}_\Phi f$ was done by the `radon` function of the Matlab Image Processing Toolbox. To mimic practical conditions, the generated data was corrupted by an additive white Gaussian noise, which was generated by the Matlab function `randn`, i.e., $y^\delta = \mathcal{R}_\Phi f + \eta$ with $\eta \in \mathcal{N}(0, \sigma^2)$ and $\sigma = 0.02 \cdot (\max y - \min y)$ (noise level = 2%). The total number of measurements $y^\delta \in \mathbb{R}^M$ was given by $M = L_\Phi \cdot \lceil \sqrt{2} \cdot (N + 2) \rceil$, where L_Φ is the number of angles and $N \times N$ denotes the size of the image f. In all of our experiments, the angular sampling distance was set to $\Delta \Theta = 1°$ and, hence, $L_\Phi = 2\Phi + 1$ if the angular range is given by $[-\Phi, \Phi]$ and $L_\Phi = \Phi + 1$ if the angular range is $[0, \Phi]$. The corresponding noisy limited angle data sets y^δ are shown in Fig. 10. Using this noisy data we computed the CSR, A-CSR and filtered backprojection (FBP) reconstructions at various angular ranges. For the computation of CSR and A-CSR reconstructions, we implemented Algorithm 1 in Matlab. FBP reconstructions were computed using the Matlab function `iradon` of the Image Processing Toolbox.

8.2.2. Execution times

We start by examining the execution times of the CSR reconstructions as well as the A-CSR reconstructions. To this end, several CSR and A-CSR reconstructions of the Shepp–Logan head phantom of size 256×256 were computed at different
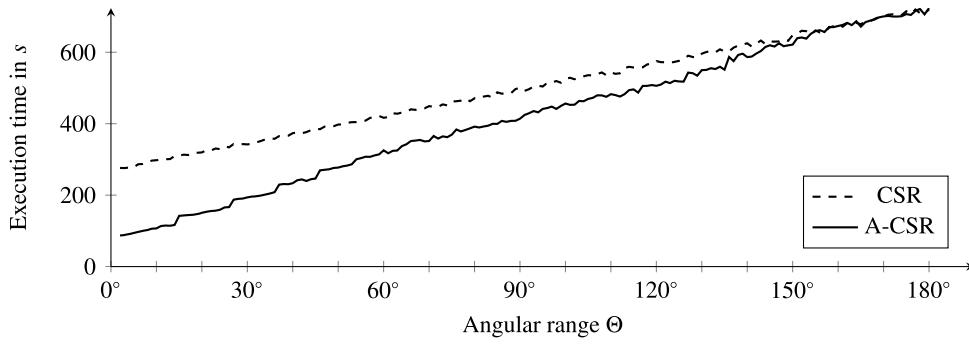
(a) 100 Iterations
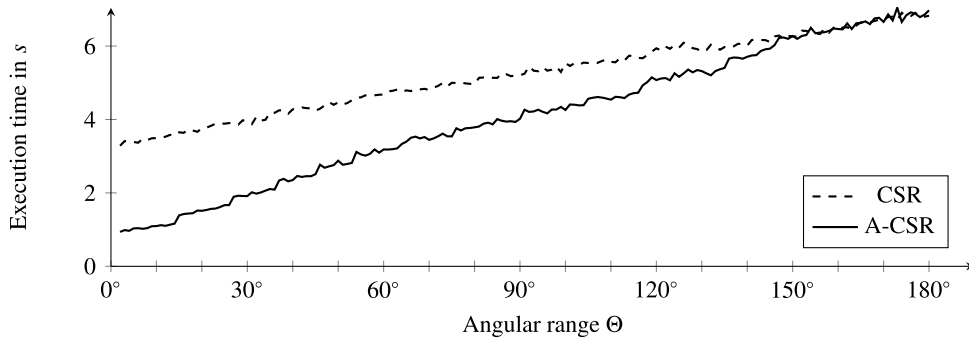
(b) 2 Iterations

Fig. 9. Reconstruction of the Shepp–Logan head phantom of size 256×256 (Fig. 8(a)) at different angular ranges $[0, \Theta]$, $\Theta \in \{1°, \ldots, 180°\}$. The plots show the execution times for CSR and A-CSR reconstruction using 100 iterations (a) and 2 iterations (b) of Algorithm 1. In both cases a significant speedup of the reconstruction procedure A-CSR can be observed.
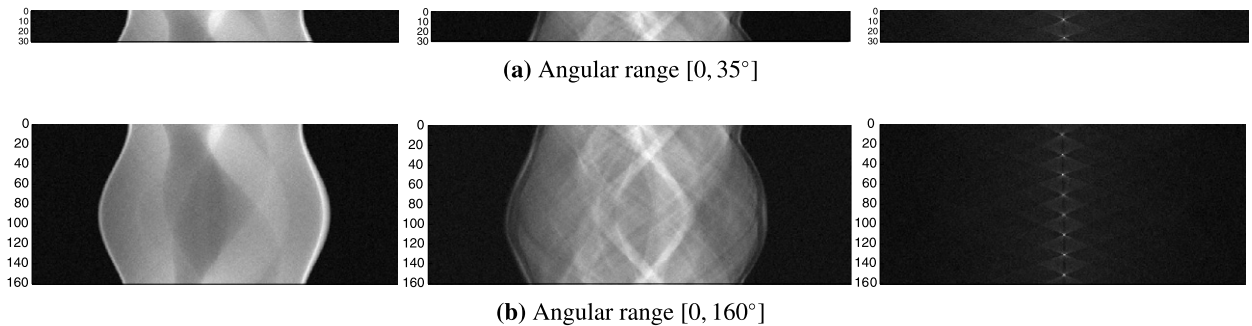

(a) Angular range $[0, 35°]$

(b) Angular range $[0, 160°]$

Fig. 10. Noisy Radon transforms at a noise level of 2%: Shepp–Logan head phantom (left), brainstem (middle) and radial pattern (right), cf. Fig. 8.

angular ranges $[0, \Theta]$. This test was done for two different numbers of iterations of Algorithm 1. In both cases, the angular range parameter Θ was chosen to vary between 1° and 180° (with an angular increment $\Delta\Theta = 1°$). The results of these tests are plotted in Fig. 9. In this figure, the dotted lines (- - -) indicate the execution times of the CSR reconstructions, whereas the solid lines (—) show the execution times of the adapted approach A-CSR. The plots in Fig. 9(a) and (b) show that the dependence of the execution times on the available angular range exhibits a linear behavior. In both cases, the adapted procedure shows a significant speedup, especially, for small angular range parameters $\Theta \leqslant 120°$. This behavior is of course a consequence of the dimensionality reduction and its dependence on the angular range, cf. Fig. 7.

8.2.3. Reconstruction quality

In order to evaluate the reconstruction quality of our method, we performed reconstructions at two different limited angular ranges, namely $[0°, 35°]$ and $[0°, 160°]$, using the methods CSR, A-CSR and FBP. For each computation of CSR and A-CSR reconstruction, we have used an individual set of parameters in Algorithm 1. These parameter sets differ in the number of iterations $n_{\max}$, the initial guess c^0 and the thresholding sequence τ (cf. Section 7). FBP reconstructions were

Table 1

The values MSE(f_{CSR}, $f_{A\text{-}CSR}$) (cf. (46)) corresponding to reconstructions in Figs. 11, 12 and 13. Small values of MSE(f_{CSR}, $f_{A\text{-}CSR}$) indicate similarity of reconstructions f_{CSR} and $f_{A\text{-}CSR}$.

	Shepp–Logan head phantom	Brainstem	Radial pattern
$\Theta = 35°$	$8.29 \cdot 10^{-5}$	$1.14 \cdot 10^{-4}$	$2.45 \cdot 10^{-4}$
	0.0025	0.0024	0.019
$\Theta = 160°$	$1.47 \cdot 10^{-6}$	$2.84 \cdot 10^{-7}$	$1.02 \cdot 10^{-5}$
	$2.05 \cdot 10^{-6}$	$2.49 \cdot 10^{-6}$	$3.95 \cdot 10^{-5}$

computed using default parameters in Matlab's `radon` function. The resulting reconstructions are show in Figs. 11, 12 and 13, where the exact values of the reconstruction parameters are listed in the caption. Each of these figures shows a matrix of images, where the upper half of the matrix contains images that were reconstructed at the angular range $[0°, 35°]$, whereas the lower half shows reconstructions at the angular range $[0°, 160°]$. The original images can be viewed in Fig. 8. The goal of this experiment is twofold: On the one hand, we want to investigate the similarity of CSR and A-CSR reconstructions and, on the other hand, to compare the reconstruction quality of CSR/A-CSR reconstructions to those obtained via FBP.

8.2.3.1. Similarity of CSR and A-CSR reconstructions   First, by visually inspecting the CSR and A-CSR reconstructions in Figs. 11–13 that are located within the same row (these images were generated using the same parameter set), we noticed no difference in reconstruction quality. Thus, we may conclude that curvelet sparse regularization and its adapted version produce reconstructions of the same visual quality when the same parameters are used in Algorithm 1. To make this observation independent of visual perception, we computed the discrepancy between the CSR reconstruction f_{CSR} and the A-CSR reconstruction $f_{A\text{-}CSR}$ with respect to the mean squared error (MSE),

$$\text{MSE}(f_{CSR}, f_{A\text{-}CSR}) = \frac{1}{N^2} \sum_{n,m=1}^{N} \big| f_{CSR}[n,m] - f_{A-CSR}[n,m] \big|^2, \tag{46}$$

where $N \in \mathbb{N}$ corresponds to the size $N \times N$ of the reconstructed images. This quantity gives the average difference between gray values of the CSR and A-CSR reconstructions. The MSE values that correspond to reconstructions in Figs. 11–13 are given in Table 1. Here, we can observe that for all reconstructions, the MSE values vary in the range from 10^{-3} to 10^{-7}. Since the gray values of the original images (cf. Fig. 8) lie between 0 and 1, these MSE values indicate that the average difference between the gray values of the CSR and A-CSR reconstructions is rather small. Thus, this quantitative comparison confirms that the CSR and A-CSR reconstructions are of a similar quality.

However, since MSE values are not equal to zero, there is a difference between CSR and A-CSR reconstructions, depending on the angular range and the particular image. In Table 1, the MSE values at a large angular range are smaller than those at a small angular range. This phenomenon may be explained as follows: The reconstructed sequence of curvelet coefficient c^{CSR} contains visible and invisible components. However, the invisible part of this vector might be not zero after a finite number of iterations and, hence, create a contribution to the MSE values. Since there are more invisible curvelet coefficients at a small angular range, this contribution is larger in the case of a small angular range than in the case of a large angular range. Another source for differences in CSR and A-CSR reconstructions is the estimation of the standard deviation via (43). This is due to the fact that the median of absolute values of all curvelet coefficients might be different from the median of absolute values of the visible curvelet coefficients. The difference in this estimation procedure is again bigger in the case of a small angular range than in the case of a large angular range. Of course, the latter explanation applies only if the generic thresholding rule (42) is used in Algorithm 1. Hence, in this case the MSE values are usually larger. This explains why the MSE values in the second line of Table 1 are significantly larger than those in the first line.

8.2.3.2. Reconstruction quality   Eventually, we compare the reconstruction quality of CSR and A-CSR reconstructions to the quality of filtered backprojection reconstruction. To this end, we will not distinguish between CSR and A-CSR reconstructions (because of their similarity, as noted above). As above, we start by visually inspecting the reconstructions in Figs. 11, 12 and 13. Here, we can observe that the FBP reconstructions contain more noise than reconstructions obtained via curvelet sparse regularization. The visual quality of FBP reconstructions seems to be poor and inferior to the CSR reconstructions. This is particularly evident for reconstructions that were obtained at an angular range $[0°, 160°]$. On the other hand, the visual impression of the CSR and A-CSR reconstructions is better than that of the FBP reconstructions. In particular, the CSR/A-CSR reconstructions are significantly less noisy than the FBP reconstructions. Moreover, we can observe that all details are well preserved and that the edges are clearly visible. This conclusion can be drawn for each parameter set that was used for CSR and A-CSR reconstruction. To verify our visual impression, we have computed the peak signal-to-noise-ratio (PSNR),

$$\text{PSNR}(f_{\text{rec}}) = 10 \cdot \log\left(\frac{1}{\text{MSE}(f, f_{\text{rec}})} \right),$$
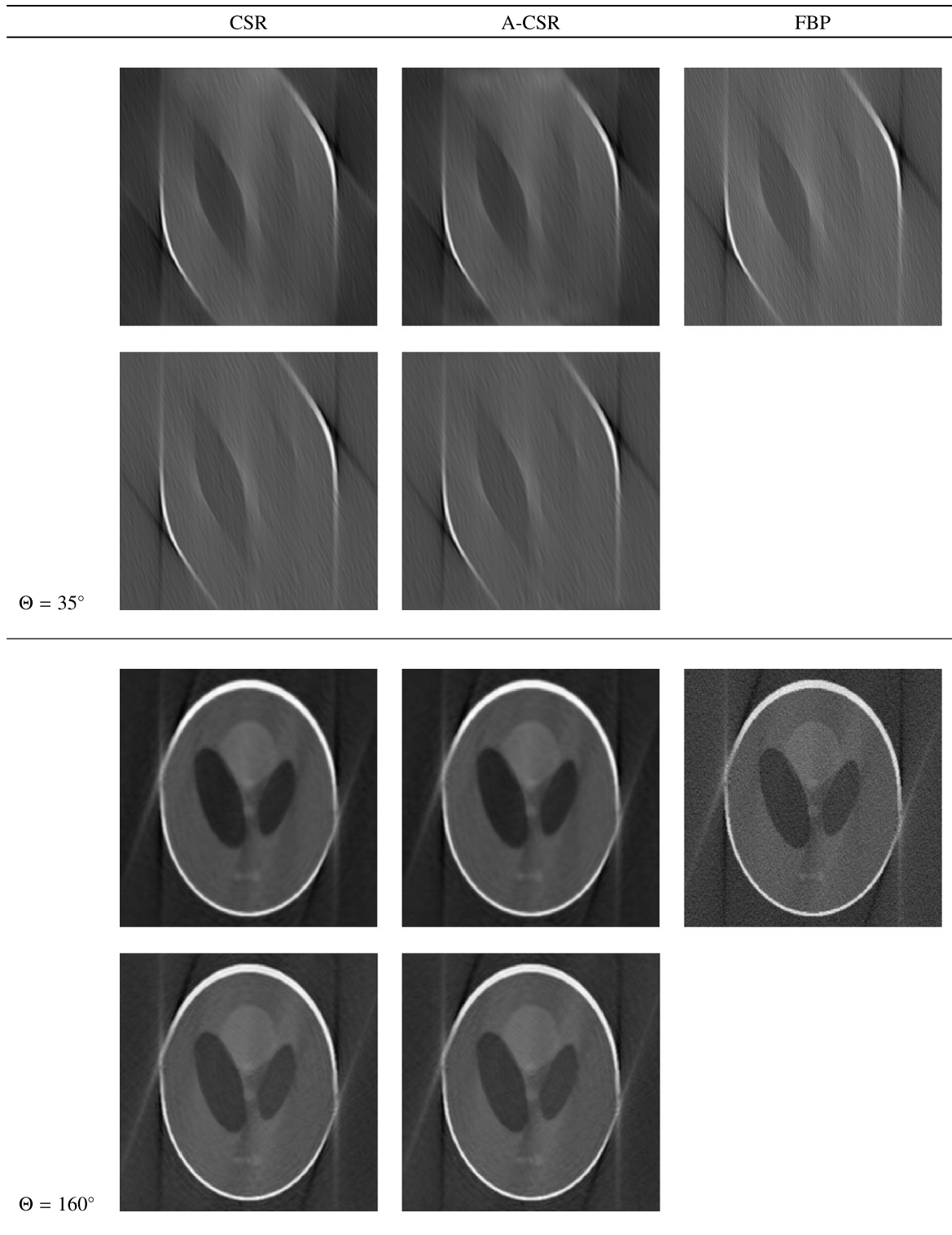
Fig. 11. Reconstructions of the Shepp–Logan head phantom of size 256×256 (Fig. 8(a)) at an angular range $[0, \Theta]$ and noise level 2% by using CSR (left column), A-CSR (middle column) and FBP (right column). The upper half shows reconstructions corresponding to the angular range parameter $\Theta = 35°$, whereas the lower half those reconstructions corresponding to $\Theta = 160°$. CSR as well as A-CSR reconstructions were performed using Algorithm 1 with the following parameters: Row 1: $n_{max} = 300$, $c^0 \equiv 0$, constant threshold $\tau = 1 \cdot 10^{-4}$; Row 2: $n_{max} = 5$, $c^0 \equiv c_{FBP}$, adaptive thresholding rule (42); Row 3: $n_{max} = 300$, $c^0 \equiv 0$, constant threshold $\tau = 1 \cdot 10^{-4}$; Row 4: $n_{max} = 2$, $c^0 \equiv c_{FBP}$, adaptive thresholding rule (42).
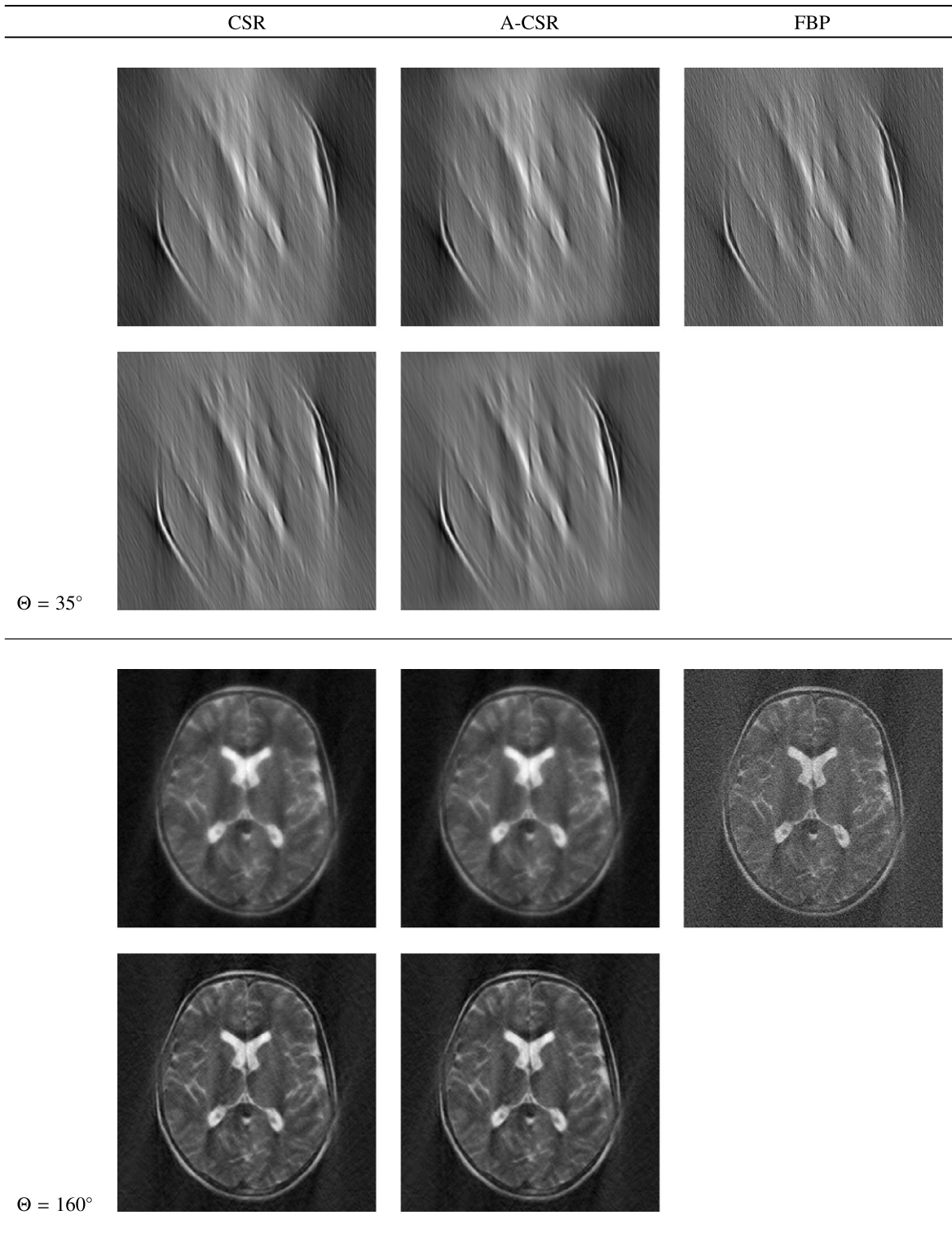
CSR	A-CSR	FBP

$\Theta = 35°$

$\Theta = 160°$

Fig. 12. Reconstructions of the brainstem glioma of size 300×300 (Fig. 8(b)) at an angular range $[0, \Theta]$ and noise level 2% using CSR (left column), A-CSR (middle column) and FBP (right column). The upper half shows reconstructions corresponding to the angular range parameter $\Theta = 35°$, whereas the lower half those reconstructions corresponding to $\Theta = 160°$. CSR as well as A-CSR reconstructions were performed using Algorithm 1 with the following parameters: Row 1: $n_{max} = 300$, $c^0 \equiv 0$, constant threshold $\tau = 1 \cdot 10^{-5}$; Row 2: $n_{max} = 5$, $c^0 \equiv c_{FBP}$, adaptive thresholding rule (42); Row 3: 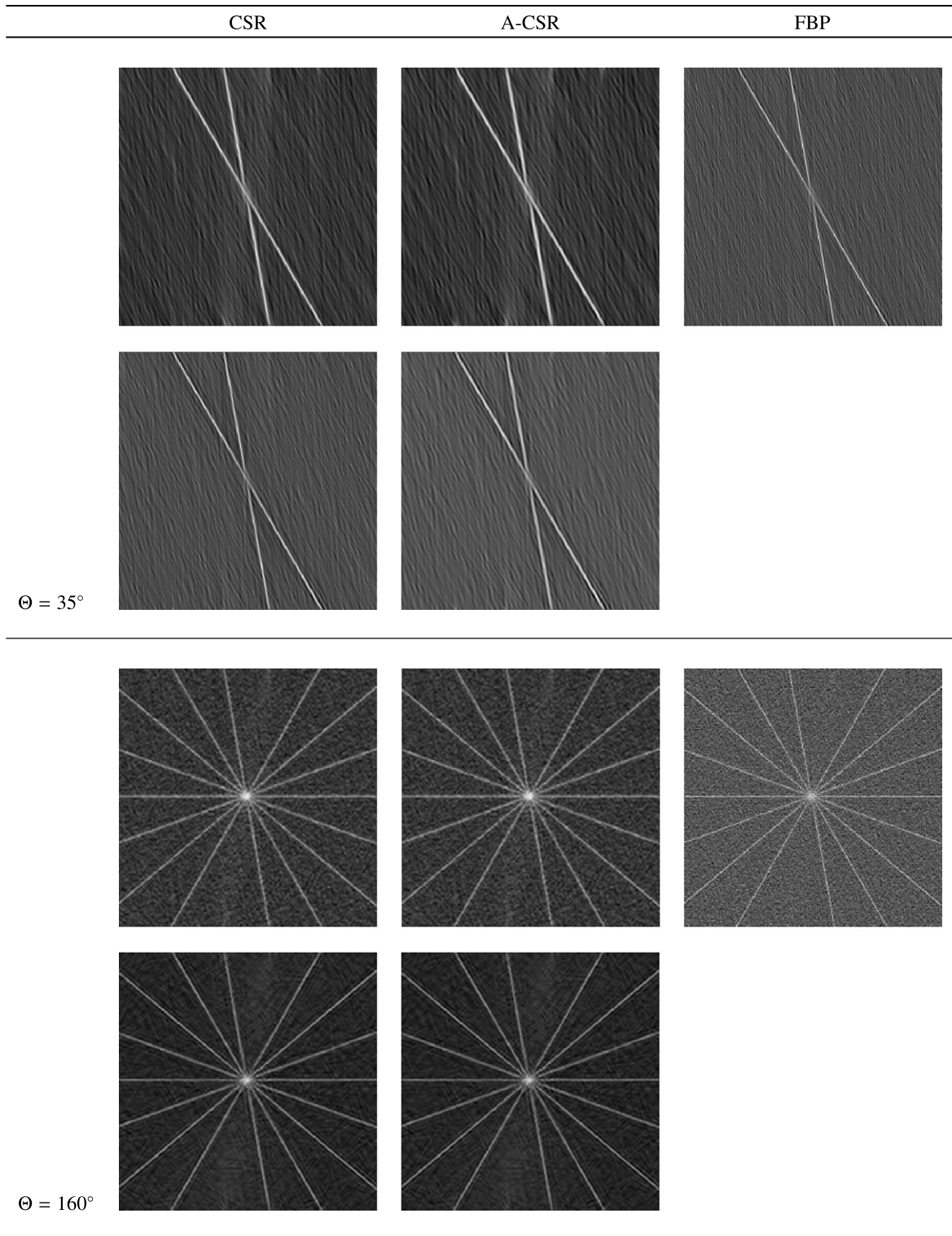$n_{max} = 300$, $c^0 \equiv 0$, constant threshold $\tau = 1 \cdot 10^{-5}$; Row 4: $n_{max} = 2$, $c^0 \equiv c_{FBP}$, adaptive thresholding rule (42).

Fig. 13. Reconstructions of the radial pattern of size 256×256 (Fig. 8(c)) at an angular range $[0, \Theta]$ and noise level 2% using CSR (left column), A-CSR (middle column) and FBP (right column). The upper half shows reconstructions corresponding to the angular range parameter $\Theta = 35°$, whereas the lower half those reconstructions corresponding to $\Theta = 160°$. CSR as well as A-CSR reconstructions were performed using Algorithm 1 with the following parameters: Row 1: $n_{\max} = 40$, $c^0 \equiv 0$, constant threshold $\tau = 35 \cdot 10^{-4}$; Row 2: $n_{\max} = 5$, $c^0 \equiv c_{\text{FBP}}$, adaptive thresholding rule (42); Row 3: $n_{\max} = 170$, $c^0 \equiv 0$, constant threshold $\tau = 35 \cdot 10^{-4}$; Row 4: $n_{\max} = 2$, $c^0 \equiv c_{\text{FBP}}$, adaptive thresholding rule (42).

Table 2
PSNR values of Shepp–Logan head phantom reconstructions corresponding to Fig. 11.

	CSR	A-CSR	FBP
$\Theta = 35°$	15.5723	15.5877	7.5038
	9.6126	9.7110	7.5038
$\Theta = 160°$	21.4573	21.4627	19.6822
	21.1402	21.1456	19.6822

Table 3
PSNR values of Brainstem reconstructions corresponding to Fig. 12.

	CSR	A-CSR	FBP
$\Theta = 35°$	16.0853	16.1068	7.1665
	10.1399	10.4682	7.1665
$\Theta = 160°$	21.0306	21.0319	17.5463
	22.3827	22.3905	17.5463

Table 4
PSNR values of radial pattern reconstructions corresponding to Fig. 13.

	CSR	A-CSR	FBP
$\Theta = 35°$	14.9041	14.8429	6.4456
	10.2064	11.0207	6.4456
$\Theta = 160°$	16.9247	16.9635	12.7875
	17.0377	17.0832	12.7875

where f denotes the original image (cf. Fig. 8) and $f_{rec} \in \{f_{FBP}, f_{CSR}, f_{A\text{-}CSR}\}$ denotes a reconstructed image. The PSNR is a quality measure that is expressed on a logarithmic scale and where large values indicate a good reconstruction quality. The PSNR values corresponding to reconstructions in Figs. 11–13 are listed in Tables 2–4. For every test image and every angular range, we can observe that the PSNR values of curvelet sparse regularizations (CSR as well as A-CSR) are considerably larger than those of the FBP reconstructions. This is in accordance with our visual impression. Therefore, the image quality of CSR and A-CSR reconstructions is considerably better than the quality of FBP reconstructions. This observation applies to every parameter set that we have used in the computation of CSR and A-CSR reconstructions.

However, some difference can be noted when comparing CSR/A-CSR reconstructions that were made using different parameter sets. For our experiments, we have used two sets of reconstruction parameters: *Parameter set 1* consists of a large number of iterations and a constant thresholding sequence (41) together with a relatively poor initial guess $c^0 = 0$. Reconstructions that were computed with such a parameter set are located in rows 1 and 3 of Figs. 11, 12 and 13. Corresponding PSNR values are given in rows 1 and 3 of Tables 2–4. On the other hand, the *parameter set 2* contains only a small number of iterations with the adaptive thresholding sequence (42) and a quite good initial guess $c^0 = c_{FBP}$. These reconstructions are located in rows 2 and 4 of Figs. 11, 12 and 13 and, correspondingly, their PSNR values are listed in rows 2 and 4 of Tables 2–4. Here, we observe that, in contrast to *parameter set 2*, the *parameter set 1* leads to a better quality of CSR/A-CSR reconstructions at a small angular range $[0°, 35°]$. However, at a large angular range $[0°, 160°]$, the opposite is true, cf. Tables 2–4. Nevertheless, both parameter sets produced reconstructions of good quality. The advantage of *parameter set 2* lies in the fact that the thresholding sequence is set automatically (cf. (42)) and has not to be determined in a series of experiments. Moreover, this choice results in a faster computation which is due to a better initial estimate $c^0 = c_{FBP}$.

8.3. Concluding remarks

Numerical experiments in this section demonstrate that the theoretical results of this work directly translate into practice. In Section 8.1 we have illustrated the visibility of curvelets which was theoretically proven in Theorem 4.1 and the dimensionality reduction in the curvelet domain. In Section 8.2, we investigated the performance of the adapted approach (A-CSR). These experiments showed that A-CSR produces reconstructions that have the same quality as CSR reconstructions. This is in accordance with Theorem 4.2 and Section 5. Moreover, we found that the proposed adapted curvelet sparse regularization is stable and edge-preserving. The reconstruction quality is controlled by the reconstruction parameters, such as number of iterations, initial guess and the thresholding sequence. However, the thresholding sequence has the most crucial impact on the reconstruction quality. The choice of this sequence should be made carefully and individually for each reconstruction problem. In practice, this is often done by trial and error. The above experiments show that the adapted choice of the thresholding sequence via (42) provide good quality reconstructions. In addition, the resulting algorithm is free of any parameter.

Another issue that we want to address here is the implementation of the Reconstruction algorithm. We would like to point out that the minimization algorithm (given by (39)) is very simple and our implementation of this procedure is very rudimental. Hence, there is much room for improvements or optimizations. For example, the execution times that are presented in Fig. 9 may be improved by a more elaborate implementation of Algorithm 1 or by using a more sophisticated minimization algorithm.

9. Summary and concluding remarks

In this work we have introduced curvelet sparse regularization as a stable and edge-preserving reconstruction method for the limited angle tomography. In numerical experiments we have shown that the CSR method produces reconstructions that are superior to those obtained via filtered backprojection. In particular, the CSR reconstructions were found to contain much less noise than FBP reconstructions while preserving all edges.

The main part of this work was devoted to the characterization of curvelet sparse regularizations in limited angle tomography. In Section 4, we have given a characterization of limited angle CSR reconstructions in terms of visible and invisible curvelet coefficients. Based on this characterization, an adapted CSR method was formulated. The adaptivity of this approach results from the fact that, depending on the available angular range, the curvelet dictionary can be partitioned into a sub-dictionary of visible curvelets and a sub-dictionary of invisible curvelets. So, by formulating the reconstruction problem only with respect to the visible curvelet sub-dictionary, the problem becomes adapted to the limited angle geometry. Moreover, this entails a significant dimensionality reduction of the original reconstruction problem. This dimensionality reduction can be easily implemented in practice. We implemented this approach in Matlab and investigated its performance in numerical experiments. As a result, we found that the achieved dimensionality reduction is considerable, especially, when the available angular range is small. Consequently, a significant speedup of the reconstruction algorithms was observed. The reconstruction quality of the adapted approach, however, was found to be equal to that of the non-adapted method.

Furthermore, we would like to note that the results of this work can be generalized to the three-dimensional setting. The ideas of this work carry over to this situation, even though, the analysis is more technical in this case.

We conclude this article by emphasizing the role of curvelets in the limited angle tomography and summarize the reasons why they were used in this work: On the one hand, curvelets provide a sparse representation of functions with an optimal encoding of edges. These properties qualify curvelets for the use in sparse regularization and give rise to an edge-preserving reconstruction. On the other hand, curvelets are highly directional. Therefore, they enable a separation of visible and invisible structures of a function which is imaged at a limited angular range. Because of this directionality, curvelets allow us to adapt the problem to the limited angle setting.

References

[1] K. Bredies, K. Kunisch, T. Pock, Total generalized variation, SIAM J. Imaging Sci. 3 (3) (2010) 492–526.
[2] K. Bredies, D.A. Lorenz, Linear convergence of iterative soft-thresholding, J. Fourier Anal. Appl. 14 (5–6) (2008) 813–837.
[3] E. Candès, L. Demanet, D. Donoho, L. Ying, Fast discrete curvelet transforms, Multiscale Model. Simul. 5 (3) (2006) 861–899, http://link.aip.org/link/?MMS/5/861/1.
[4] E. Candès, L. Demanet, D.L. Donoho, L. Ying, CurveLab-2.1.2, http://www.curvelet.org/, 2008.
[5] E.J. Candes, D.L. Donoho, Curvelets and reconstruction of images from noisy radon data, vol. 4119, SPIE, 2000 pp. 108–117, http://link.aip.org/link/?PSI/4119/108/1.
[6] E.J. Candès, D.L. Donoho, Recovering edges in ill-posed inverse problems: optimality of curvelet frames, Ann. Statist. 30 (3) (2002) 784–842, dedicated to the memory of Lucien Le Cam.
[7] E.J. Candès, D.L. Donoho, New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities, Comm. Pure Appl. Math. 57 (2) (2004) 219–266.
[8] E.J. Candès, D.L. Donoho, Continuous curvelet transform. I. Resolution of the wavefront set, Appl. Comput. Harmon. Anal. 19 (2005) 162–197.
[9] E.J. Candès, D.L. Donoho, Continuous curvelet transform. II. Discretization and frames, Appl. Comput. Harmon. Anal. 19 (2) (2005) 198–222.
[10] V. Caselles, A. Chambolle, M. Novaga, The discontinuity set of solutions of the TV denoising problem and some extensions, Multiscale Model. Simul. 6 (3) (2007) 879–894.
[11] C.K. Chui, An Introduction to Wavelets, Wavelet Analysis and Its Applications, vol. 1, Academic Press Inc., Boston, MA, 1992.
[12] F. Colonna, G. Easley, K. Guo, D. Labate, Radon transform inversion using the shearlet representation, Appl. Comput. Harmon. Anal. 29 (2) (2010) 232–250.
[13] I. Daubechies, et al., An iterative thresholding algorithm for linear inverse problems with a sparsity constraint, Comm. Pure Appl. Math. 57 (11) (2004) 1413–1457.
[14] M.E. Davison, The ill-conditioned nature of the limited angle tomography problem, SIAM J. Appl. Math. 43 (2) (1983) 428–448.
[15] G. Easley, D. Labate, W.-Q. Lim, Sparse directional image representations using the discrete shearlet transform, Appl. Comput. Harmon. Anal. 25 (1) (2008) 25–46.

[16] H.W. Engl, M. Hanke, A. Neubauer, Regularization of Inverse Problems, Mathematics and Its Applications, vol. 375, Kluwer Academic Publishers Group, Dordrecht, 1996.
[17] J.M. Fadili, G. Peyré, Total variation projection with first order schemes, IEEE Trans. Image Process. 20 (3) (2011) 657–669.
[18] J. Frikel, A new framework for sparse regularization in limited angle X-ray tomography, in: IEEE International Symposium on Biomedical Imaging: From Nano to Macro, April 2010, pp. 824–827.
[19] J. Frikel, Short communication: dimensionality reduction of curvelet sparse regularizations in limited angle tomography, Proc. Appl. Math. Mechan. (PAMM) 11 (1) (2011) 847–848, http://dx.doi.org/10.1002/pamm.201110412.
[20] R. Griesse, D.A. Lorenz, A semismooth Newton method for Tikhonov functionals with sparsity constraints, Inverse Problems 24 (3) (2008) 035007 (19 pp.) http://stacks.iop.org/0266-5611/24/035007.
[21] J.H. Jørgensen, T.L. Jensen, P.C. Hansen, S.H. Jensen, E.Y. Sidky, X. Pan, Accelerated gradient methods for total-variation-based CT image reconstruction, arXiv:1105.4002v1 [math.NA], 2011.
[22] G.T. Herman, R. Davidi, Image reconstruction from a small number of projections, Inverse Problems 24 (4) (2008) 045011, http://stacks.iop.org/0266-5611/24/i=4/a=045011.
[23] V. Kolehmainen, et al., Statistical inversion for medical X-ray tomography with view radiographs: II. Application to dental radiology, Phys. Med. Biol. 48 (2003) 1465–1490.
[24] G. Kutyniok, D. Labate, Resolution of the wavefront set using continuous shearlets, Trans. Amer. Math. Soc. 361 (5) (2009) 2719–2754.
[25] D.A. Lorenz, D. Trede, Optimal convergence rates for Tikhonov regularization in Besov scales, Inverse Problems 24 (5) (2008) 055010, http://stacks.iop.org/0266-5611/24/i=5/a=055010.
[26] J. Ma, G. Plonka, The Curvelet Transform, IEEE Signal Process. Mag. 27 (2) (2010) 118–133.
[27] S. Mallat, A Wavelet Tour of Signal Processing, The Sparse Way, 3rd edition, Elsevier/Academic Press, Amsterdam, 2009, with contributions from Gabriel Peyré.
[28] F. Natterer, The Mathematics of Computerized Tomography, B.G. Teubner, Stuttgart, 1986.
[29] E.T. Quinto, Singularities of the X-ray transform and limited data tomography in $\mathbb{R}^2$ and $\mathbb{R}^3$, SIAM J. Math. Anal. 24 (5) (1993) 1215–1225.
[30] E.T. Quinto, An introduction to X-ray tomography and Radon transforms, in: The Radon Transform, Inverse Problems, and Tomography, in: Proc. Sympos. Appl. Math., vol. 63, Amer. Math. Soc., Providence, RI, 2006, pp. 1–23.
[31] Radiopedia.org, http://radiopaedia.org/cases/brainstem-glioma, 2010.
[32] M. Rantala, et al., Wavelet-based reconstruction for limited angle X-ray tomography, IEEE Trans. Med. Imag. 25 (2) (February 2006) 210–217.
[33] W. Ring, Structural properties of solutions to total variation regularization problems, Math. Model. Numer. Anal. 34 (4) (2000) 799–810.
[34] R.T. Rockafellar, Convex Analysis, Princeton Math. Ser., vol. 28, Princeton University Press, Princeton, NJ, 1970.
[35] O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, F. Lenzen, Variational Methods in Imaging, Appl. Math. Sci., vol. 167, Springer, New York, 2009.
[36] E.M. Stein, G. Weiss, Introduction to Fourier Analysis on Euclidean Spaces, Princeton Math. Ser., vol. 32, Princeton University Press, Princeton, NJ, 1971.