A comparison of regularization models for few-view CT image reconstruction

Elena Loli Piccolomini[1] (ORCID)

Abstract

In this paper I analyse some regularization models for the reconstruction of X-rays Computed Tomography images from few-view projections. It is well known that the widely used low-cost Filtered Back Projection method is not suitable in case of low-dose data, since it produces images with noise and artifacts. Iterative reconstruction methods based on the model discretization are preferred in this case. However, since the problem has infinite possible solutions and is ill-posed, regularization is necessary to obtain a good solution. Different iterative regularization methods have been proposed in literature, but an organized comparison among them is not available. We compare some regularization approaches in the case of few-view tomography by means of simulated projections from both a phantom and a real image.

Keywords Computed tomography · Regularization · Iterative methods

1 Introduction

X-rays Computed Tomography (CT) [1] is an imaging technique used in different application areas, such as medicine, engineering and arts, to reveal information about the inner structure of an object or a human organ. In particular, motivated by an increasing focus on the potentially harmful effects of X-rays ionizing radiation, a recent trend in medical CT research is to develop safer protocols to reduce the radiation dose. This allows to apply CT techniques to a wider class of medical examinations, including vascular, dental, orthopedic, musculoskeletal, chest and mammographic imaging. Safer protocols are of interest not only for medicine but also for materials science and cultural heritage, to prevent damage to the subject under study, due to excessive radiations.

✉ Elena Loli Piccolomini
  elena.loli@unibo.it

1 Department of Computer Science and Engineering, University of Bologna, Mura Anteo Zamboni, 7, 40126 Bologna, Italy

Springer

One technique to reduce the radiation exposure is to lower the number of X-rays projections. In this case, the protocols are named as *sparse tomography* (or sparse-view, few-view tomography). Figure 1 shows a graphical draft of the acquisition and reconstruction process. In the first column is it sketched the geometry of the acquisition process, where an X-rays source moves along a circular trajectory; in the second column the set of acquired data (also called *sinogram*) is represented; in the third column the reconstructions obtained with the Filtered Backprojection (FBP) method [2] are shown. In the first row the classical full dose CT case is represented, where the source spans the whole circular trajectory; in the second row a *sparse-view full-angle tomography* is considered where a reduced number of views is taken in the whole circular orbit. A different sparse-view geometry using few projections is called *limited-angle tomography* (see the third row of Fig. 1). Here, a further reduction of X-rays scans is made by limiting the source trajectory to a C-shaped path, i.e. by restricting the 360-degrees angular scanning interval to a range smaller than 180 degrees. In some tomographic applications, the human anatomy does not allow a complete circular motion to the X-rays source, thus the use of a reduced range is mandatory and the resulting technique is called *tomosynthesis*. An example is breast imaging, where the breast is in a stationary position between the detector surface and the compression plate. The resulting scans are fast but the sinogram is incomplete and the reconstruction process is quite difficult.

As clearly visibile from Fig. 1 the traditional analytic FBP algorithm produces images with striking artifacts. Hence, regularized iterative methods, which minimize a suitable function modelling the noise, are commonly used in these cases [3] in place of FBP. However, since the reconstruction problem is ill-posed [2] it is necessary to introduce regularization.

Aim of this work is to compare different regularization models for CT image reconstruction. In literature, several works using one of these models and showing its efficiency have been presented, but there isn't any fair comparison between them.

The paper is organized as follows. In Sect. 2 some regularization models for image reconstruction are shortly described; in Sect. 3 numerical experiments using different regularization approaches for 2D image reconstruction from few views are presented and discussed; finally in Sect. 4 we draw some conclusions.

2 Iterative regularized approach

The model-based iterative methods (for a possible classification see [4] and the references therein) solve the linear system obtained from the discretization of both the object and the acquisition device (also called detector). In the following, we present for simplicity the analysis for the 2D case. The extension to 3D is straightforward.

Suppose the discrete object is of size $n_x \times n_y$, the detector of size n_p and the projections are collected from n_θ angles. Hence, the linear system
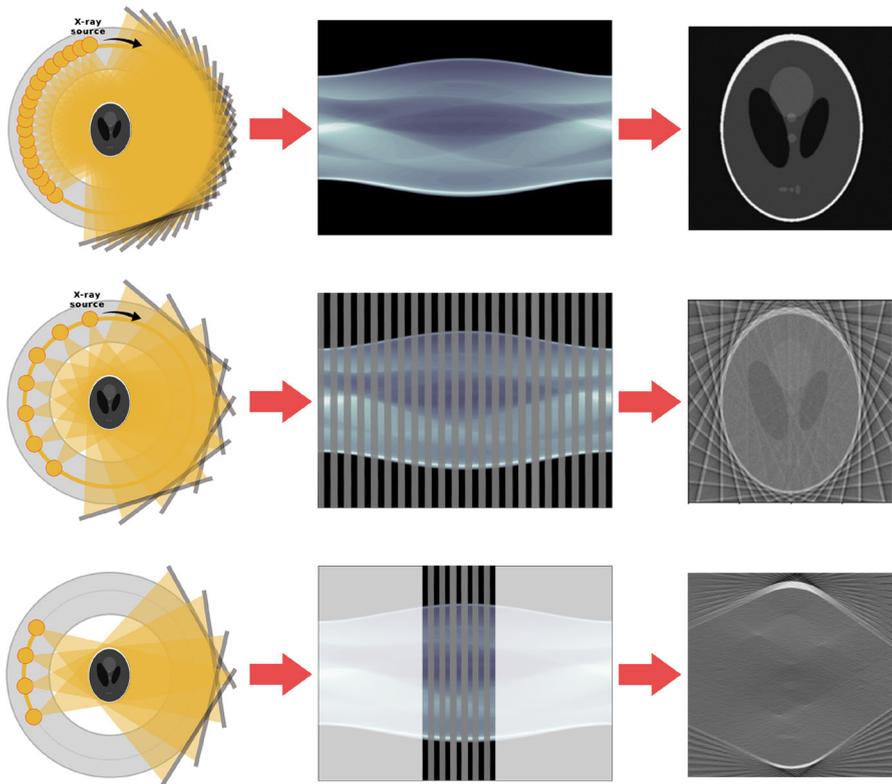
$$Ax = y \tag{1}$$

Fig. 1 Sketches of the tomographic image reconstruction workflow, for full-view, sparse-view full-angle and limited-angle protocols (from top to bottom, respectively). From the different geometries on the left, the acquired projections and the reconstructed image of the Shepp-Logan phantom. The missing portions of sinogram in the sparse-view and limited-angle protocols are depicted in light gray

represents the discretization of the Radon transform [5]. It relates the object $x \in \mathcal{R}^{n_x \times n_y}$ with the sinogram $y \in \mathcal{R}^{n_p \times n_\theta}$, collecting all the projections, through the matrix A, called *projection matrix*. In the case of few-views, $n_p \times n_\theta < n_x \times n_y$, the linear system (1) admits infinite possible solutions.

To overcome this problem, the linear system is usually replaced by a least-squares problem:

$$min_x \|Ax - y\|_2^2. \tag{2}$$

Moreover, since Eq. (1) is the discretization of a mildly ill-posed problem (i.e. the integral Radon equation), then the solution of (2) is sensitive to noise present on the data and regularization is needed [2].

2.1 Regularization by iteration

The simplest form of regularization is *regularization by iteration* [6]. The normal equations of (2) can be solved by a fast iterative method, such as the Conjugate Gradient Least-Squares (CGLS) algorithm, stopped after few iterations, far before convergence [7]. Several criteria have been proposed to suitably stop the iterations before noise enters the solution [8, 9].

2.2 Regularization by function

Another possible, and mostly used, strategy is to use regularization functions which impose prior information on the solution. The so called *model based iterative methods* compute the CT image by solving a constrained or unconstrained minimization problem involving a data-fit function $F(x)$ and a regularization function $R(x)$. Common choices for the a data-fit function $F(x)$ are the least-squares, which models gaussian white noise, and the Kullback-Leibler divergence, modelling Poisson noise. In this work, we always consider the least-squares data-fit:

$$F(x) = \frac{1}{2}\|Ax - b\|_2^2. \tag{3}$$

Concerning the regularization function $R(x)$, there are more possible choices.

The unconstrained formulation of the minimization problem has the form:

$$min_x \|Ax - y\|_2^2 + \lambda^2 R(x), \tag{4}$$

where λ is the regularization parameter balancing the data-fit and regularization terms. A non negative constraint on the solution can also be added as:

$$min_{x\geq 0}\|Ax - y\|_2^2 + \lambda^2 R(x). \tag{5}$$

The most classical regularization approach is the so called Tikhonov regularization, where

$$R(x) = \|Dx\|_2^2 \tag{6}$$

and D is usually chosen as one of the follwing:

1. the identity operator I;
2. the discrete gradient operator G, usually computed with finite differences.

Tikhonov minimization problem (4) can be re-written as:

$$min_x \|\tilde{A}x - \tilde{y}\|_2^2 \tag{7}$$

where:

$$\tilde{A} = \begin{pmatrix} A \\ \lambda D \end{pmatrix} \quad \tilde{b} = \begin{pmatrix} b \\ 0 \end{pmatrix}$$

The problem (7) can be solved by applying the CGLS to the normal equations.

Widely used alternatives to Tikhonov function are sparse promoting regularization functions, such as the l1 norm. In this case,

$$R(x) = \|x\|_1$$

and the differentiable minimization problem (5) is solved here by FISTA method [10].

Another sparse promoting regularizer is the Total Variation (TV), defined as:

$$TV(x) = \sum_{j=1}^{N} \|\nabla x_j\|_2.$$

where $\nabla(\cdot)$ is the discrete gradient operator. The resulting minimization problem (5) is solved again by FISTA. For its edge preserving properties and its effectiveness in removing noise, TV is the most used regularization function in medical CT imaging [11, 12].

A smoothed version of TV is often considered, by introducing a small positive parameter β which makes the TV function differentiable in zero:

$$TV_\beta(x) = \sum_{j=1}^{N} \sqrt{\|\nabla x_j\|_2^2 + \beta^2}$$

In the numerical experiments we have set $\beta = 10^{-3}$. When considering TV_β the Scaled Gradient Projection (SGP) method is used to solve the differentiable problem (5) (see [12] for more details on SGP method on CT image reconstruction).

3 Numerical experiments

Some numerical tests have been executed on both a phantom from the *Tomophantom* package [13] and a real image from the Mayo Clinic data set [14], shown in Fig. 2. For the implementation of the forward and backward projection operators and of some of the model-based iterative reconstruction methods we have used the Python version of the Core Image Library (CIL) [15] (www.ccpi.ac.uk/cil).

We compare the traditional analytic FBP method with the following regularized iterative approaches:

- Regularization by iteration by means of CGLS method stopped after few (8) iterations (hence used as regularization by iteration) and *near convergence* (at 100 iterations). In the following we will name them as CGLS(8) and CGLS(100), respectively;
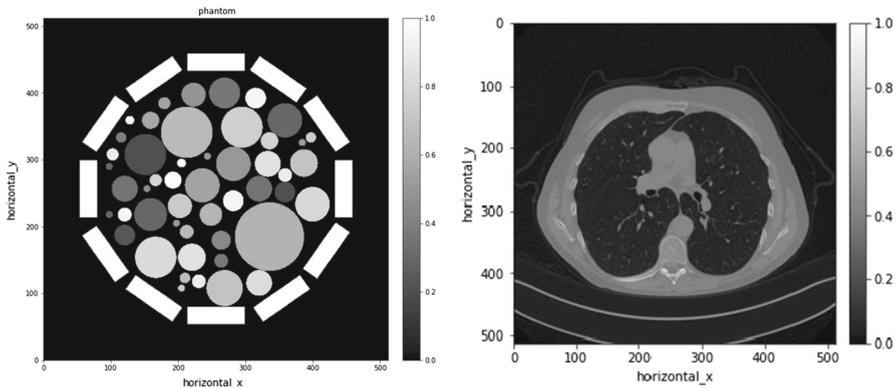
Fig. 2 Ground truth images

- Tikhonov regularization with $D = I$ and $D = G$. We name them as TIK(I) and TIK(G), respectively;
- L1 regularization;
- Total Variation regularization in its smoothed and non-smoothed versions. We name them as TV(s) and TV, respectively.

We evaluate the reconstructions obtained from the different methods by computing the Mean Squared Error (MSE), Mean Absolute Error (MAE) and Peak Signal to Noise Ratio (PSNR) metrics defined as:

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (x - \tilde{x})^2,$$

$$MAE = \frac{1}{N} \sum_{i=1}^{N} |x - \tilde{x}|,$$

$$PSNR = 10 log_{10} \frac{max|x|}{MSE},$$

where $N = n_x \times n_y$ is the number of pixels of the reconstructed image $\tilde{x}$ and x is the ground truth image.

By using the forward projector function available in CIL, we simulate the projections considering 60 angles equally spaced in [0, 180] degrees. In all the experiments reported in this section, the regularization parameter λ is heuristically chosen by trial and error to get the lowest possible MSE value.

3.1 Tests on the phantom image

The first test problem consists in reconstructing the phantom image from 60 noise-free projections. In Fig. 3 we see some reconstructions. The FBP image (top left) is clearly corrupted by streaking artifacts, due to the sparse view and lack of data, which are reduced, but yet visible, in the CGLS(100) (top center). Since we are considering noise-
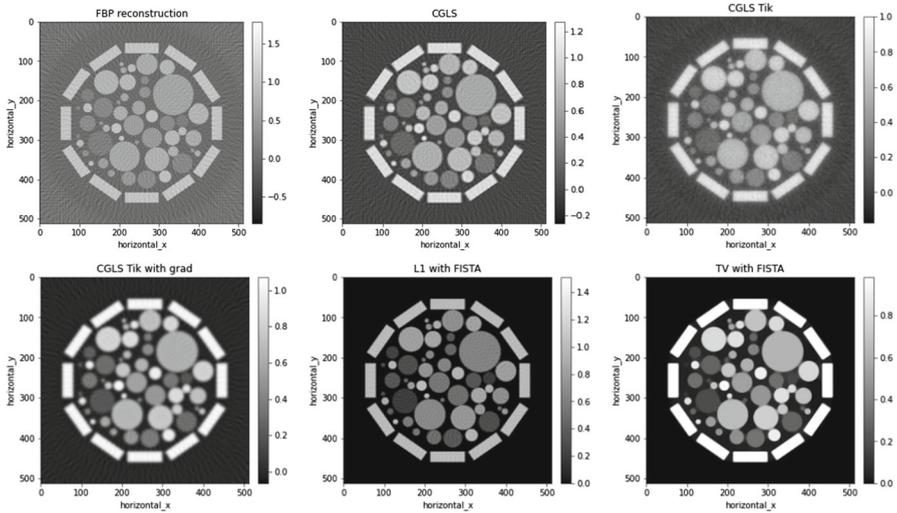
Fig. 3 Reconstructions from noise-free projections. In the top row (from left to right): FBP, CGLS(100),TIK(I). In the bottom row (from left to right): TIK(G), L1, TV

free projections, the CGLS do not show semi-convergence and the images produced by CGLS(100) and CGLS(8) are very similar. In the Tik(I) image (top right) the artifacts are still present whereas the Tik(G) (bottom left) reconstruction is less noisy but slightly out of focus. The L1 regularization (bottom center) has a good contrast but some noise is visible inside the circles. Finally, we appreciate the good accuracy and high quality of the reconstructions obtained with TV regularization (bottom right).

In Fig. 4 we report some plots of a single row of the image. In all the plots, the blue line represents the ground truth. In the top left figure, the ground truth signal is compared with the FBP one, which is highly oscillating as expected. In the top right one, we can see the results obtained with CGLS(8) (orange line), CGLS(100) (green line) and Tik(G) (red line). Tik(G) appears smoother than CGLS. The bottom right graphic confirms that TV regularization is outstanding in recovering the true signal and removing noise.

In the second test problem white gaussian noise with variance 0.01 is added to the previously tested phantom sinogram. Figure 5 contains some reconstructions. When data are affected by noise we can appreciate the difference between CGLS(8) (top left) and CGLS(100) (top center) reconstructions. The noise on the projections produces semi-convergence and, if CGLS is not properly stopped, it also affects the solution, as visibile in the CGLS(100) image. The CGLS(8) image is less noisy, but slightly blurred. Tik(G) (top right) reconstruction is quite blurred and artifacts are visible in the flat regions. We can infer that the 2-norm regularization is not effective in well recovering the object contours from noisy data. In the L1 image (bottom left) the circles have sharp edges but the noise has not been completely removed. The TV method (bottom central) provides a reconstruction with high contrast and no noise but the image looks *cartoon*: the objects are very flat and homogenoeus. This is a well
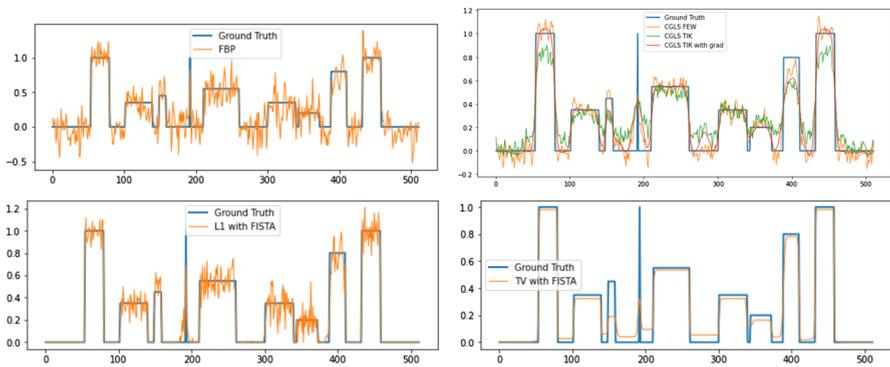
Fig. 4 Plot profile of a row in the phantom test problem (noise-free projections). The blue line is from the true image in all the plots. Top left: FBP reconstruction (orange line); top right: CGLS(8) (orange line), TIK(I) (green line), TIK(G) (red line); bottom left: L1 (orange line); bottom right: TV (orange line)
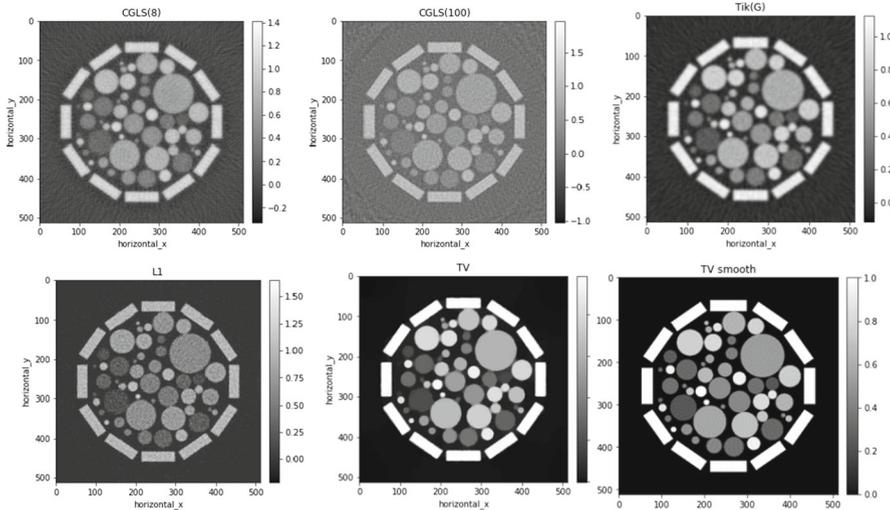

Fig. 5 Reconstructions from noisy projections. In the top row (from left to right): CGLS(100), CGLS(8), TIK(G). In the bottom row (from left to right): L1, TV, TV(s)

known unwanted effect of TV function. Finally, the TV(s) image visually appears as the most accurate reconstruction, similar to the TV one, but less blocky.

Table 1 summarizes the results of these two experiments, reporting the values of the MSE, MAE and PSNR metrics for all the reconstructions obtained on the phantom test image. The first three columns are from noise-free data and the last three columns from noisy sinograms. They are in agreement with the visual results reported in Figs. 3 and 4 and confirm that the TV-based iterative reconstructions (in particular the TV(s) model) outperform the others for all the considered metrics.

Table 1 MSE, MAE and PSNR values for the reconstructions with no noisy projections (first three columns) and noisy projections (last three columns)

Method	MSE	MAE	PSNR	MSE(n)	MAE(n)	PSNR(n)
FBP	2.5 e-02	1.2 e-01	64.00	7.3 e-02	1.2 e-01	59.49
CGLS (100it)	7.4 e-02	5.7 e-01	69.41	1.9 e-02	1.1 e-01	65.31
Tik (I)	7.5e-03	5.6 e-02	69.35	2.0e-02	9.9 e-02	65.06
Tik(G)	1.2e-02	5.4 e-02	67.31	1.2e-02	5.6 e-02	67.26
L1	5.3e-03	3.2e-02	70.89	1.1 e-02	5.3e-02	67.60
TV	4.3e-03	2.7e-02	71.77	4.8e-03	3.1 e-02	71.21
TV(s)	**2.0e-04**	**3.8e-02**	**85.11**	**4.4e-04**	**7.3e-02**	**81.60**

In bold, the best values obtained

3.2 Tests on a chest CT image

The last test problem is created from the real CT image of a chest shown in Fig. 3 by simulating 60 projections in [0, 180] degrees and by adding white noise with variance 0.005. As clearly visible from Fig. 2, the CT image of the chest is quite different from the previously considered phantom. It presents many small details over a dark background and smoothed regions along the borders of the lungs.

Figure 6 shows some reconstructions obtained with the considered methods. In the first row the FBP image (on the left) is clearly damaged by noise; the L1 image (on the right) appears well contrasted with neat object contours. Few noise as speckle is still visible, but it does not alter the objects shape. In the second row, the TV image (on the left) appears smooth and the details inside the lungs are noticeable; however the dark regions look spotty, as effect of the TV function. Also in this test problem, the TV(s) image (bottom right) seems the best reconstruction.

The plots in Fig. 7 confirm the considerations deduced from the images. TV regularization (bottom) suppresses the noise but cannot reproduce the peaks, whereas the L1 reconstruction (top right) is a bit more noisy but it follows the fast changes of the blue line (representing in all the plots the ground truth). The other reconstructions (CGLS and Tikhonov) (top left) are not very accurate.

Finally, Table 2 reports the MSE, MAE and PSNR obtained in this experiment. The best values (highlighted in bold) are again obtained with the TV(s) regularization.

4 Conclusions

In this paper I have investigated the effects of regularization in the reconstruction of X-rays CT images from few views. Due to the lack of data, the linear system deriving from the model discretization has infinite possible solutions. Regularization is necessary not only to suppress noise (as in all the ill-posed inverse problems), but also to choose a suitable solution among the many possible ones.

The most common regularization approaches have been analysed: regularization by iteration and regularization by function, with Tikhonov regularization, Total Variation
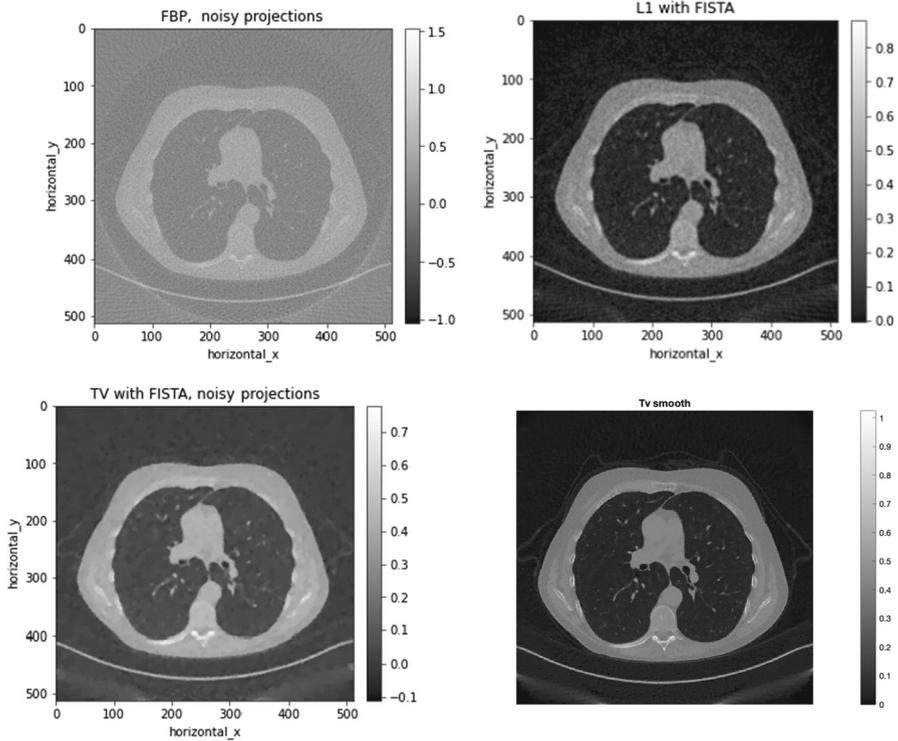
Fig. 6 Reconstructions of the chest image. Top left: FPB; top right: L1; bottom left: TV; bottom right: TV(s)
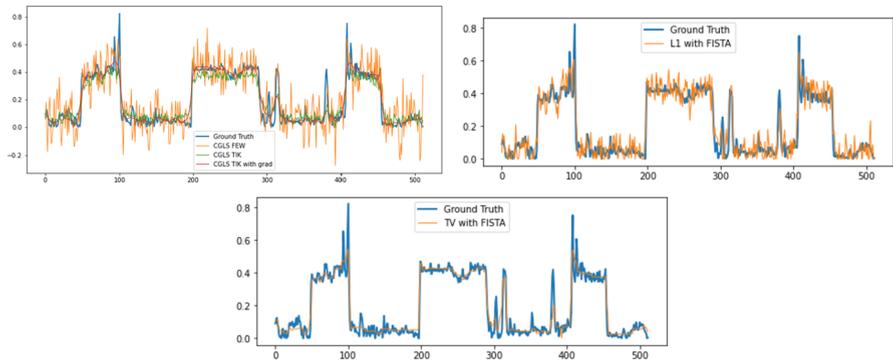

Fig. 7 Plot profile of a row in the chest test problem. The blue line represents the ground truth in all the plots. Top left: CGLS(8) (orange line), TIK(I) (green line), TIK(G) (red line); Top right: L1 (orange line); bottom : TV (orange line)

Table 2 MSE, MAE and PSNR values for the reconstructions of the chest image

Method	MSE(n)	MAE(n)	PSNR(n)
FBP	5.7e-02	1.9e-01	60.5
CGLS(8)	3.1e-03	3.8e-02	73.1
TIK(G)	2.e-03	3.1e-02	73.7
L1	3.6e-03	4.e-02	72.5
TV	1.8e-03	2.5e-02	75.5
TV(s)	**1.0e-03**	**1.8e-02**	**77.9**

and L1 norm. The results obtained on 2D test problems with simulated projections from a sparse geometry show that the images reconstructed by the various regularization approaches are quite different, hence a correct choice of regularization is of primary importance to obtain the image with desired features. For example, TV regularization produces highly contrasted images where the noise is almost completely absent, but the objects are flat and *cartoon*. On the contrary, the images reconstructed with L1 regularization are slightly noisy, but well contrasted and more realistic. TV smoothed regularizer is a good compromise between TV and L1: it produces images with very few noise residual and good contrasted objects.

A very important consideration regards the computational time. The regularization per iteration (obtained by stopping the iterative method after very few iterations) are quite blurred but they are computed in much lower time with respect to the other methods. This is essential in real applications where the time available for the CT exam is limited.

To complete the analysis, in the future I intend to consider 3D experiments and possibly real data. Other regularization techniques will also be tested, such as Generalized Total Variation or Nonlocal Mean Filter, to cite only some. Finally, the role of the regularization parameter will be investigated.

Declarations

Conflict of interest The author declare that they have no conflict of interest.

References

1. Buzug, T.M.: Computed tomography: from photon statistics to modern cone-beam CT. Soc. Nucl. Med. (2009)
2. Bertero, M., Boccacci, P., De Mol, C.: Introduction to inverse problems in imaging. CRC Press, Boca Raton (2021)
3. Beister, M., Kolditz, D., Kalender, W.A.: Iterative reconstruction methods in X-ray CT. Phys. Med. **28**, 94–108 (2012)
4. Graff, C., Sidky, E.: Compressive sensing in medical imaging. Appl. Opt. **54**(8), 23–44 (2015)
5. Radon, J.: On the determination of functions from their integral values along certain manifolds. IEEE Trans. Med. Imaging **5**(4), 170–176 (1986)
6. Hansen, P.C.: Discrete inverse problems: insight and algorithms, (2010)
7. Piccolomini, E.L., Zama, F.: The conjugate gradient regularization method in computed tomography problems. Appl. Math. Comput. **102**(1), 87–99 (1999)
8. Reichel, L., Rodriguez, G.: Old and new parameter choice rules for discrete ill-posed problems. Numer. Algorithms **63**(1), 65–87 (2013)
9. Landi, G., Piccolomini, E.L., Tomba, I.: A stopping criterion for iterative regularization methods. Appl. Numer. Math. **106**, 53–68 (2016)
10. Beck, A., Teboulle, M.: A fast iterative shrinkage-thresholding algorithm for linear inverse problems. SIAM J. Imaging Sci. **2**(1), 183–202 (2009)
11. Piccolomini, E.L., Morotti, E.: A fast total variation-based iterative algorithm for digital breast tomosynthesis image reconstruction. J. Algorithms Comput. Technol. **10**(4), 277–289 (2016)
12. Piccolomini, E.L., Morotti, E.: A model-based optimization framework for iterative digital breast tomosynthesis image reconstruction. J. Imaging **7**(2), 36 (2021)
13. Kazantsev, D., Pickalov, V., Nagella, S., Pasca, E., Withers, P.J.: Tomophantom, a software package to generate 2d–4d analytical phantoms for ct image reconstruction algorithm benchmarks. SoftwareX **7**, 150–155 (2018)
14. McCollough, C.: Tu-fg-207a-04: overview of the low dose ct grand challenge. Med. Phys. **43**(6Part35), 3759–3760 (2016)
15. Jørgensen, J.S., Ametova, E., Burca, G., Fardell, G., Papoutsellis, E., Pasca, E., Thielemans, K., Turner, M., Warr, R., Lionheart, W.R.: Core imaging library-part i: a versatile python framework for tomographic imaging. Philos. Trans. R. Soc. A **379**(2204), 20200192 (2021)