# Robust artefact reduction in tomography using Student's t data fitting

Folkert Bleichrodt*, Tristan van Leeuwen†, K. Joost Batenburg*‡§
*Centrum Wiskunde & Informatica
Science Park 123, 1098 XG Amsterdam, The Netherlands
Email: {F.Bleichrodt,K.J.Batenburg}@cwi.nl
†Utrecht University
Budapestlaan 6, 3584 CD Utrecht, The Netherlands
Email: T.vanLeeuwen@uu.nl
‡University of Antwerp
iMinds-Vision Lab, Universiteitsplein 1, B-2610 Antwerp, Belgium
§Leiden University, Mathematical Institute
Niels Bohrweg 1, 2333 CA Leiden, The Netherlands

***Abstract*—Algebraic methods are popular for tomographic image reconstruction from limited data. These methods typically minimize the Euclidean norm of the residual of the corresponding linear equation system. The underlying assumption of this approach is that the noise has a Gaussian distribution. However, in cases where large outliers are present in the projection data, *e.g.*, due to defective camera pixels, photon starvation from metal implants etc., the equation system is not consistent and the reconstruction will be fitted to these outliers, resulting in artefacts in the reconstruction.**

**In this paper we use a penalty function for the residual that is based on the maximum likelihood estimate from the Student's t distribution, which assigns a smaller penalty to outliers. No preprocessing is required to locate the outliers. We demonstrate the effectiveness of this approach on a 3D cone-beam simulated dataset for a series of perturbations in the projection data. Our results suggest that artefacts due to metal objects, defective camera pixels, or corrupted (randomized) projection images can be suppressed by using algebraic reconstruction methods in combination with the Student's t penalty function.**

## I. Introduction

Tomography is a technique for reconstructing a 3D volume from 2D projection images, such as X-rays obtained in CT-scanners. A 3D reconstruction can be obtained from the projection images by solving an inverse problem. In algebraic reconstruction methods a linear system of equations is solved that represents a discretization of the Radon transform [1, 2]:

$$\boldsymbol{W}\boldsymbol{x} = \boldsymbol{p}. \tag{1}$$

The *projection matrix* $\boldsymbol{W} \in \mathbb{R}^{M\times N}$ relates pixel values in the tomographic reconstruction $\boldsymbol{x} \in \mathbb{R}^N$ (*gray values*) to discrete detector measurements $\boldsymbol{p} \in \mathbb{R}^M$. In experiments the projections are perturbed by an unknown noise vector $\boldsymbol{\epsilon}$,

$$\tilde{\boldsymbol{p}} = \boldsymbol{p} + \boldsymbol{\epsilon}.$$

Most algebraic methods such as SIRT, CGLS or LSQR [3, 4, 5] optimize the consistency of the reconstruction in the Euclidean norm, known as least squares solution:

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}} \frac{1}{2}\|\boldsymbol{W}\boldsymbol{x} - \tilde{\boldsymbol{p}}\|_2^2. \tag{2}$$

It is well known that this approach is equivalent to finding the maximum likelihood estimate (MLE) of $\boldsymbol{x}$ under the assumption that the error term or noise $\boldsymbol{\epsilon}$ is Gaussian distributed [6]. However, the $\ell_2$-norm assigns a heavy penalty to outliers in the projection data. Outliers may arise due to acquisition problems ranging from hardware problems to physical effects such as scattering or photon starvation due to high density particles [7]. Because these errors are heavily penalized by the $\ell_2$-norm, the solution of eq. (2) will be fitted to these outliers, producing artefacts in the reconstruction.

In this paper we propose the use of algebraic methods combined with the Student's t penalty function to solve the reconstruction problem in eq. (1). The Student's t distribution has heavy-tails meaning that outliers in the noise are penalized less compared to the $\ell_2$-norm. Therefore the Student's t MLE of the reconstruction should be influenced less by such outliers.

Many methods for artefact reduction are aimed to remove or suppress outliers from the projection data [8, 9, 10, 11], which rely heavily on the accuracy of segmentation techniques to locate outliers. By minimizing the Student's t penalty of the data-fit there is no need for segmentation and therefore the method is not biased by the result of a segmentation step.

We explain the method for finding the Student's t MLE of the reconstruction in section II. Subsequently, results are presented for a series of 3D cone-beam simulation experiments for reduction of several kinds of artefacts in section III. Finally, we discuss the results and conclude the paper in section IV.

## II. Methods

In general, maximum likelihood estimation of $\boldsymbol{x}$ in eq. (1) gives rise to a maximization problem

$$\max_{\boldsymbol{x}} \rho(\boldsymbol{W}\boldsymbol{x} - \tilde{\boldsymbol{p}}),$$

where $\rho(\cdot)$ is the probability density function (PDF) of the probability distribution of the noise $\boldsymbol{\epsilon}$. In practice, the problem

is posed as a minimization problem by taking the $-\log$:

$$\min_{\boldsymbol{x}} -\log \rho(\boldsymbol{W}\boldsymbol{x} - \tilde{\boldsymbol{p}}).$$

The resulting estimate $\widehat{\boldsymbol{x}}$ can be interpreted as the most likely solution of eq. (1) under the assumption that the noise is indeed distributed according to $\rho$. When $\rho$ represents the Gaussian PDF, this leads to the conventional least squares formulation, eq. (2). When the data contain large outliers, the Gaussian assumption is violated and a different PDF has to be employed. A possible choice is the multivariate Student's t distribution

$$\rho(\boldsymbol{r}) \propto \prod_i \left(1 + r_i^2/\nu\right)^{-(\nu+1)/2},$$

where $\nu$ is the variance. Such an assumption on the noise allows for large outliers to be present in the residual, whereas under a Gaussian assumption large outliers are extremely unlikely and thus the reconstruction will aim to fit them.

The *penalty* derived from the Student's t distribution is

$$p(\boldsymbol{r}) = \sum_i \log(1 + r_i^2/\nu), \tag{3}$$

and a graph is shown in fig. 1a. The maximum likelihood estimate is now obtained by solving

$$\min_{\boldsymbol{x}} p(\boldsymbol{W}\boldsymbol{x} - \tilde{\boldsymbol{p}})$$

using Newton's method [12]. This leads to an iterative method of the form

$$\boldsymbol{x}^{(k+1)} = \boldsymbol{x}^{(k)} + \alpha_k \boldsymbol{s}^{(k)},$$

where $\alpha_k$ is the stepsize, determined by a backtracking line-search and $\boldsymbol{s}^{(k)}$ is obtained by solving

$$\boldsymbol{W}^T \boldsymbol{H}^{(k)} \boldsymbol{W} \boldsymbol{s}^{(k)} = -\boldsymbol{W}^T \boldsymbol{g}^{(k)}. \tag{4}$$

Here, the gradient $\boldsymbol{g}^{(k)}$ and diagonal matrix $\boldsymbol{H}^k$ are given in terms of the residual $\boldsymbol{r}^{(k)} = \boldsymbol{W}\boldsymbol{x}^{(k)} - \tilde{\boldsymbol{p}}$ as

$$g_i^{(k)} = \frac{2r_i}{\nu + r_i^2},$$

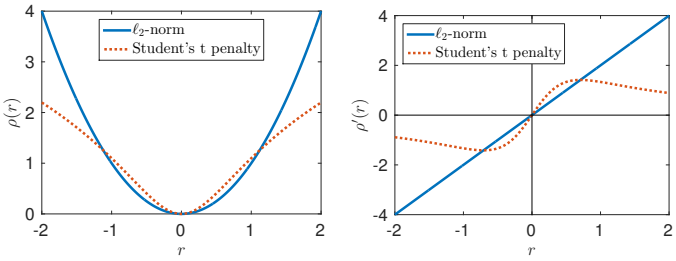
and

$$h_{ii}^{(k)} = \frac{2}{\nu + r_i^2}.$$

Note that we can use any algebraic method to solve eq. (4), but in our case we chose the CG method. In effect, the algorithm repeatedly performs a reconstruction with a weighted residual, where the weight $\left(\nu + r_i^2\right)^{-1}$ down-weights large residuals.

If we look at the so-called *influence-function* [13] of eq. (3) in fig. 1b which is defined by the gradient, it is clear that the influence of large residuals $r^2 \gg \nu$ is small. However, for $r^2 < \nu$ the influence behaves similar to a least squares penalty. The role of $\nu$ can be seen as tuning parameter to indicate the magnitude of outliers. This parameter can be adjusted automatically [14], however, in our experiments we estimate the parameter empirically.

From this point forward we will refer to the methods for MLE estimation using least squares and Student's t penalties as LSQR-MLE and ST-MLE respectively, where we use the method LSQR for minimizing the $\ell_2$-norm.




(a) Penalty functions (b) Influence-functions

Fig. 1: Least squares and Student's t penalty functions and corresponding influence-functions with $\nu = 0.5$.

(a) 3D rendering (b) slice

Fig. 2: (a) 3D rendering of the Shepp–Logan head phantom with a wedge cut out of the sample; (b) central slice of size $256 \times 256$.

## III. Experiments and Results

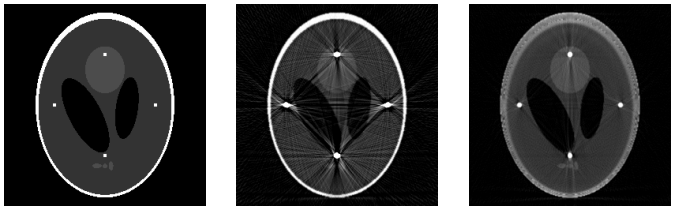
In these simulation experiments we consider a 3D Shepp–Logan head phantom of size $256 \times 256 \times 256$ of which a central slice is shown in fig. 2b. We used the ASTRA tomography toolbox [15] to generate 180 projection images with $1^\circ$ angular separation using the cone-beam geometry. The detector has a size of $284 \times 284$ pixels and was positioned in the origin. The projection matrix is generated on-the-fly by the GPU back end of the toolbox using a slice interpolation kernel [16].

In the following sections we will discuss several distortions or perturbations in the projection images that cause severe artefacts in the reconstruction and we compare a least squares approach to data fitting using the Student's t penalty function.

### A. Metal artefact reduction

In this experiment we consider the 3D Shepp–Logan head phantom with six small dense particles that represent metal implants (density is 10 times that of the outer "skull" region). A single slice is shown in fig. 3a, the six particles form the vertices of an octahedron.

In the area of the detector where the metal implants are projected the data becomes corrupted due to beam hardening, scatter and photon starvation. For this experiment we focus on the effects of photon starvation. In the projection data we simulated a saturation due to photon starvation by setting the region corresponding to the metal objects to a constant, large value. The effect of this missing or corrupted data if we apply



(a) slice of phantom (b) LSQR-MLE (c) ST-MLE

Fig. 3: Metal particles Shepp–Logan head phantom and corresponding least squares fit and Student's t fit.

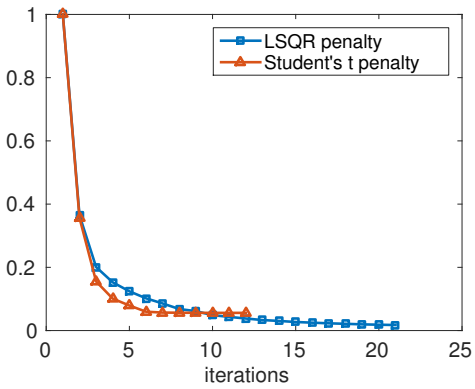



Fig. 4: Convergence of the $\ell_2$-norm compared to Student's t penalty. These are relative residuals.

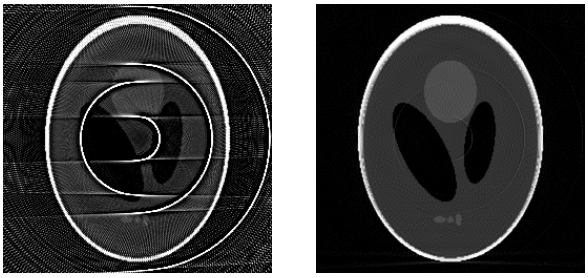
LSQR-MLE is shown in fig. 3b. Usually, these regions in the projection data are either ignored or filled in by interpolation or inpainting techniques [8, 10, 11]. These methods rely on sophisticated segmentation techniques in order to locate the metal implants.

We show a convergence plot in fig. 4 of both penalty functions. This figure shows that the ST-MLE method converges rapidly compared to LSQR-MLE. Note, however, that the ST-MLE method requires solving of eq. (4) in each iteration and is therefore significantly more costly. In all of the following experiments, the ST-MLE method converges in approximately 10 iterations.

Our proposed method ST-MLE is able to suppress most of the artefacts, as shown in fig. 3c, while still reconstructing the metal implants without needing to locate the outliers in the projection images. There is an underestimation of the gray value of the skull area, however, visually the reconstruction is very useful for detecting also smaller details, such as the three ellipses below the bottom metal particle. Moreover, the ST-MLE solution can be used initially to obtain a better segmentation of the metal particles.

### B. Defective camera pixels

In the second experiment we simulate the effect of defective camera pixels. We assume that the detector has several "dead" detector pixels which measure no photons at all. This produces bright pixels in the projection images that are constant between



(a) LSQR-MLE (b) ST-MLE

Fig. 5: Defective camera pixels lead to semicircular reconstruction artefacts. The Student's t solution is much less affected by these artefacts.

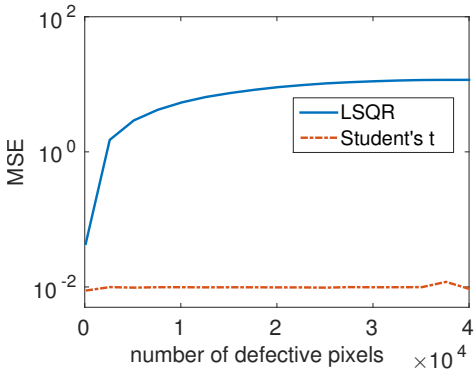



Fig. 6: Mean squared error of the reconstruction compared to the ground truth for an increasing number of dead detector pixels.

projections. The uncorrected projection data will produce ring artefacts which are typically removed by inpainting of dead pixels [9].

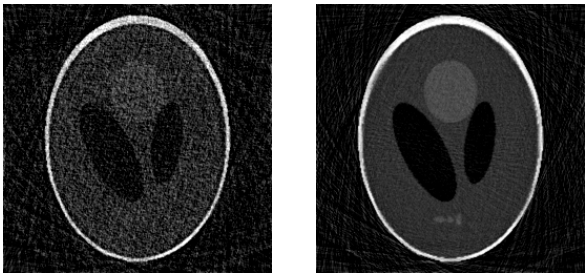
We simulated a dataset with 100 randomly selected dead pixels which we set to a constant value of two times the maximum value of the projection data. The least squares solution is shown in fig. 5a. The artefacts are severe, but the Student's t approach in fig. 5b is able to remove the artefacts almost completely.

In fig. 6 we show the effect of increasingly many dead pixels on the mean squared error of the reconstruction compared to the ground truth. Surprisingly, even if the number of dead pixels is close to $50\%$ of the total number of detector pixels the ST-MLE solution does not seem to be influenced by this missing data.

### C. Randomized projection images

In the final experiment we created a dataset of which we replaced 50 from the 180 projections by completely random images (white noise) with average intensity similar to the other projection images. Although this is not a very realistic dataset, we want to see how far we can stress our ST-MLE method and see if it can ignore such inconsistent data.

The LSQR-MLE solution is shown in fig. 7a, which is very noisy due to the randomized projections. The ST-MLE solution



(a) LSQR-MLE (b) ST-MLE

Fig. 7: LSQR and Student's t fit for dataset with 50 randomized projection images out of 180 total projection images.

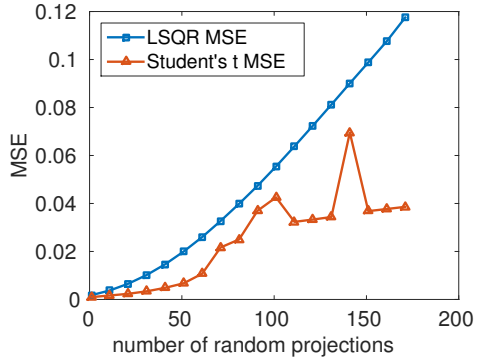



Fig. 8: Mean squared error of the reconstructions for an increasing number of random projections replacing the original projections.

(fig. 7b) suffers far less from the random projections and only shows mild noise. There are some streak artefacts because the projection images in these directions are missing, but this is expected.

We also compared LSQR-MLE and ST-MLE on datasets with an increasing number of random projections. Of course we cannot expect that the ST-MLE solution will be unaffected by this as was the case in the previous experiment, because we are essentially removing projections. However, the result shown in fig. 8 indicates that the ST-MLE method is beneficial for each of these dataset and is a large improvement over the least squares solution.

## IV. Discussion and conclusions

In this paper we have discussed the Student's t penalty function that can be used in combination with Newton's optimization approach to produce the maximum likelihood estimate of the tomographic reconstruction problem eq. (1) corresponding to the Student's t distribution. In our experiments we have seen that perturbations introduced in the projection data due to hardware problems or photon starvation from metal implements is significantly reduced using our proposed method ST-MLE when compared to algebraic reconstruction methods that minimize the Euclidean norm of the residual (LSQR-MLE). In contrast to other methods for artefact reduction, there is no need to locate outliers in the projection data by segmentation methods. Therefore, the ST-MLE method can be applied effectively without any pre-processing steps. Moreover, the Student's t penalty can be used in combination with other reconstruction algorithms and image priors and has other potential use cases such as artefact reduction from diffraction effects.

## Acknowledgment

This research was supported by the Netherlands Organisation for Scientific Research (NWO), programme 639.072.005. Networking support was provided by the EXTREMA COST Action MP1207.

## References

[1] F. Natterer and F. Wübbeling, *Mathematical methods in image reconstruction*. SIAM, 2001.

[2] A. C. Kak and M. Slaney, *Principles of computerized tomographic imaging*. Society for Industrial and Applied Mathematics, 2001.

[3] A. Björck, *Numerical methods for least squares problems*. SIAM, 1996.

[4] J. Gregor and T. Benson, "Computational analysis and improvement of SIRT," *IEEE Transactions on Medical Imaging*, vol. 27, no. 7, pp. 918–924, 2008.

[5] C. C. Paige and M. A. Saunders, "Lsqr: an algorithm for sparse linear equations and sparse least squares," *ACM Transactions on Mathematical Software (TOMS)*, vol. 8, no. 1, pp. 43–71, 1982.

[6] W. H. Press, *Numerical recipes 3rd edition: The art of scientific computing*. Cambridge university press, 2007.

[7] F. E. Boas and D. Fleischmann, "CT artifacts: Causes and reduction techniques," *Imaging in Medicine*, vol. 4, no. 2, pp. 229–240, 2012.

[8] J. Gu, L. Zhang, G. Yu, Y. Xing, and Z. Chen, "X-ray CT metal artifacts reduction through curvature based sinogram inpainting," *Journal of X-ray Science and Technology*, vol. 14, no. 2, pp. 73–82, 2006.

[9] Y. Pan, F. De Carlo, and X. Xiao, "Ring artifact removal for micro-tomography in synchrotron radiation," in *SPIE Medical Imaging*, International Society for Optics and Photonics, 2012, pp. 831 329–831 329.

[10] W. J. Veldkamp, R. M. Joemai, A. J. van der Molen, and J. Geleijns, "Development and validation of segmentation and interpolation techniques in sinograms for metal artifact suppression in CT," *Medical physics*, vol. 37, no. 2, pp. 620–628, 2010.

[11] G. Wang, D. L. Snyder, J. O'Sullivan, and M. Vannier, "Iterative deblurring for CT metal artifact reduction," *Medical Imaging, IEEE Transactions on*, vol. 15, no. 5, pp. 657–664, 1996.

[12] J. Nocedal and S. Wright, *Numerical optimization*, 2nd. Springer Verlag, 2006, pp. 245–262.

[13] F. R. Hampel, E. M. Ronchetti, P. J. Rousseeuw, and W. A. Stahel, *Robust statistics: The approach based on influence functions*. John Wiley & Sons, 2011, vol. 114.

[14] A. Y. Aravkin and T. Van Leeuwen, "Estimating nuisance parameters in inverse problems," *Inverse Problems*, vol. 28, no. 11, p. 115 016, 2012.

[15] W. J. Palenstijn, K. J. Batenburg, and J. Sijbers, "The ASTRA tomography toolbox," in *13th International Conference on Computational and Mathematical Methods in Science and Engineering, CMMSE*, 2013.

[16] P. Joseph, "An improved algorithm for reprojecting rays through pixel images," *Medical Imaging, IEEE Transactions on*, vol. 1, no. 3, pp. 192–196, 1982.