Evaluation of Vectorial Total Generalized Variation Reconstruction of Undersampled CT Acquisitions

J. Wållberg
jwallberg@student.ethz.ch

Abstract

This paper evaluates a CT reconstruction method using a vectorial Total Generalized Variation (TGV) regularization to impose group sparsity on image gradients with the goal of reconstructing natural-looking images from noisy data. The method is compared to Total Variation (TV) and Tikhonov regularized methods as well as standard back projection. It was found that TGV produced the best images in terms of nRMSE and PSNR, in all cases. Furthermore, it was shown that TGV can get rid of the staircase artifacts commonly associated with TV-based reconstruction.

I. Introduction

Reconstructing an image from a CT sinogram is an ill-posed problem due to the discreteness of acquired data. In a perfect setting, x-rays from all possible directions around the object would have been acquired and the image could have been perfectly reconstructed using filtered-backprojection (FBP) [1]. In a real-world setting, however, we can only acquire a finite number of rays leading to the ill-posedness of the reconstruction problem. Furthermore, to decrease acquisition time it can be favorable to employ a sparse acquisition scheme leaving out even further rays. In other cases, only access to a limited angle of the patient is available e.g. in a C-arm scanner.

Even though FBP could theoretically reconstruct an imaged object perfectly, the result of such an algorithm applied to real data will always have various noise and artifacts due to imperfections in the acquisition procedure e.g. electronic noise or movement of the patient during acquisition. Hence, many attempts have been made to look at reconstruction as an iterative process enabling the introduction of prior information about what a good image looks like. This prior can be added to an optimization process as a regularization/penalty term.

In this paper, the performance of vectorial Total Generalized Variation (TGV) regularization is investigated as a way to suppress noise by employing group sparsity on image gradients, hopefully, to end up with more natural-looking images. The results are compared to FBP, Tikhonov regularization, and the simpler method of Total Variation (TV) regularization, which is known for its staircase artifacts.

II. Description of Methods

A. Total Generalized Variation

The aim of vectorial total generalized variation reconstruction is to minimize the following expression

$$\min_{\mathbf{x},\mathbf{V}} \frac{1}{2\lambda}\|\mathbf{M}\mathbf{A}\mathbf{x}-\mathbf{s}\|_2^2+\alpha_0\left\|\left[\nabla_1\mathbf{v}_1,\nabla_2\mathbf{v}_2,\frac{1}{2}(\nabla_1\mathbf{v}_2+\nabla_2\mathbf{v}_1)\right]\right\|_{2,1}+ \alpha_1\|[\nabla_1\mathbf{x},\nabla_2\mathbf{x}]-\mathbf{V}\|_{2,1}, \tag{1}$$

where $\mathbf{s}\in\mathbb{R}^m$ is the observed data, $\mathbf{A}\in\mathbb{R}^{m\times n_1 n_2}$ is the discrete Radon transform, and $\mathbf{M}\in\mathbb{R}^{m\times m}$ an undersampling matrix. By concatenating the different gradients into one vector

and using the $\ell_{2,1}$-norm, group sparsity is imposed on them. Meaning that gradients are penalized together, no matter what gradient direction is considered. The hope with this approach is to get rid of the staircase artifacts associated with the standard TV approach, which independently penalizes each gradient direction.

The optimization in (1) can be equivalently formulated as the following primal problem

$$\begin{aligned} \min_{\mathbf{x},\mathbf{V},\mathbf{Z},\mathbf{Y}} \quad & \frac{1}{2}\|\mathbf{M}\mathbf{A}\mathbf{x}-\mathbf{s}\|_2^2 + \lambda(\alpha_0\|\mathbf{Y}\|_{2,1} + \alpha_1\|\mathbf{Z}\|_{2,1}) \\ \text{s.t.} \quad & \mathbf{Z} = [\nabla_1\mathbf{x}, \nabla_2\mathbf{x}] - \mathbf{V} \\ & \mathbf{Y} = \left[\nabla_1\mathbf{v}_1, \nabla_2\mathbf{v}_2, \frac{1}{2}(\nabla_1\mathbf{v}_2 + \nabla_2\mathbf{v}_1)\right] \end{aligned} \tag{2}$$

enabling the usage of the alternating direction method of multipliers (ADMM) algorithm to numerically solve it [2].

B. ADMM

To the objective function in (2) the ℓ_2-norm of the constraint residuals are added so as to give the augmented Lagrangian

$$\begin{aligned} \mathcal{L}(\mathbf{x},\mathbf{V},\mathbf{Z},\mathbf{Y},\mathbf{U}_Z,\mathbf{U}_Y) = & \frac{1}{2}\|\mathbf{M}\mathbf{A}\mathbf{x}-\mathbf{s}\|_2^2 + \lambda(\alpha_0\|\mathbf{Y}\|_{2,1} + \alpha_1\|\mathbf{Z}\|_{2,1}) \\ & + \rho_Z\|\mathbf{Z} - [\nabla_1\mathbf{x}, \nabla_2\mathbf{x}] + \mathbf{V} + \mathbf{U}_Z\|_2^2 \\ & + \rho_Y\left\|\left[\nabla_1\mathbf{v}_1, \nabla_2\mathbf{v}_2, \frac{1}{2}(\nabla_1\mathbf{v}_2 + \nabla_2\mathbf{v}_1)\right] - \mathbf{Y} + \mathbf{U}_Y\right\|_2^2 \end{aligned}$$

where $\mathbf{U}_Z$ and $\mathbf{U}_Y$ are the scaled dual variables $\mathbf{W}_Z/\rho_Z$ and $\mathbf{W}_Y/\rho_Y$. And $\mathbf{W}_Z$, $\mathbf{W}_Y$ are the dual variables for the $\mathbf{Z}$ and $\mathbf{Y}$ constraints respectively. This formulation yields the corresponding ADMM update steps as

$$\begin{aligned} \mathbf{x}^{k+1} &\leftarrow \arg\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{M}\mathbf{A}\mathbf{x}-\mathbf{s}\|_2^2 + \rho_Z\left\|\mathbf{Z}^k - [\nabla_1\mathbf{x}, \nabla_2\mathbf{x}] + \mathbf{V}^k + \mathbf{U}_Z^k\right\|_F^2 \\ \mathbf{V}^{k+1} &\leftarrow \arg\min_{\mathbf{V}} \rho_Z\left\|\mathbf{Z}^k - \left[\nabla_1\mathbf{x}^{k+1}, \nabla_2\mathbf{x}^{k+1}\right] + \mathbf{V} + \mathbf{U}_Z^k\right\|_F^2 + \rho_Y\left\|\left[\nabla_1\mathbf{v}_1, \nabla_2\mathbf{v}_2, \frac{1}{2}(\nabla_1\mathbf{v}_2 + \nabla_2\mathbf{v}_1)\right] - \mathbf{Y}^k + \mathbf{U}_Y^k\right\|_F^2 \\ \mathbf{Z}^{k+1} &\leftarrow \arg\min_{\mathbf{Z}} \lambda\alpha_1\|\mathbf{Z}\|_{2,1} + \rho_Z\left\|\mathbf{Z} - \left[\nabla_1\mathbf{x}^{k+1}, \nabla_2\mathbf{x}^{k+1}\right] + \mathbf{V}^{k+1} + \mathbf{U}_Z^k\right\|_F^2 \\ \mathbf{Y}^{k+1} &\leftarrow \arg\min_{\mathbf{Y}} \lambda\alpha_0\|\mathbf{Y}\|_{2,1} + \rho_Y\left\|\left[\nabla_1\mathbf{v}_1^{k+1}, \nabla_2\mathbf{v}_2^{k+1}, \frac{1}{2}(\nabla_1\mathbf{v}_2^{k+1} + \nabla_2\mathbf{v}_1^{k+1})\right] - \mathbf{Y} + \mathbf{U}_Y^k\right\|_F^2 \\ \mathbf{U}_Z^{k+1} &\leftarrow \mathbf{U}_Z^k + \mathbf{Z}^{k+1} - \left[\nabla_1\mathbf{x}^{k+1}, \nabla_2\mathbf{x}^{k+1}\right] + \mathbf{V}^{k+1} \\ \mathbf{U}_Y^{k+1} &\leftarrow \mathbf{U}_Y^k + \left[\nabla_1\mathbf{v}_1^{k+1}, \nabla_2\mathbf{v}_2^{k+1}, \frac{1}{2}(\nabla_1\mathbf{v}_2^{k+1} + \nabla_2\mathbf{v}_1^{k+1})\right] - \mathbf{Y}^{k+1}. \end{aligned} \tag{3}$$

The $\mathbf{x}$-step is a convex optimization problem that was solved by the conjugate gradients method with a maximum iteration limit of 100, and tolerance of 1e-8. Likewise, $\mathbf{V}$ is updated by conjugate gradients using the same stopping criteria as above. The $\mathbf{Z}$- and $\mathbf{Y}$-updates are row-wise separable and performed using the ℓ_2 proximal mapping

$$\text{prox}_{2\lambda\alpha_i/\rho_i,\ell_2}(\mathbf{a}) = \left(1 - \frac{2\lambda\alpha_i}{\rho_i\|\mathbf{a}\|_2}\right)_+ \mathbf{a}. \tag{4}$$

The gradient operators were given by finite difference approximations implemented using `numpy.roll` gaining better speed compared to using sparse matrices [3]. The ADMM iterations were stopped if the difference in the cost function (1) for consecutive iterations were lower than 1e-2 with a maximum limit of 100 iterations.

C. Comparison of methods

The TGV reconstruction method was compared to three other methods. The first one is the regular filtered backprojection algorithm, which simply highpass-filters the sinogram data before backprojecting it.

The second method was Generalized Tikhonov regularized reconstruction which is given as the solution to the following optimization problem

$$\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{MAx} - \mathbf{s}\|_2^2 + \lambda\|\nabla\mathbf{x}\|_2^2. \tag{5}$$

This problem can be solved analytically and is efficiently implemented using the conjugate gradients method.

The third method is TV regularization. For this method, the first derivative penalty is given as an ℓ_1 penalty to enforce a Laplacian prior on the image gradient distribution. It is the solution to the following problem

$$\min_{\mathbf{x}} \frac{1}{2}\|\mathbf{MAx} - \mathbf{s}\|_2^2 + \lambda\|\nabla\mathbf{x}\|_1. \tag{6}$$

As in the case of TGV, this is solved using ADMM.

D. Parameter selection

The TGV-ADMM algorithm comes with five parameters in total. Effectively they consist of four free parameters since the loading costs α_0 and α_1 are balanced via their quotient α_0/α_1. Hence, the following parameters $\lambda, \alpha_0/\alpha_1$, ρ_Z, ρ_Y are to be chosen. We let $\rho_Z = \rho_Y$, and $2\alpha_1 = \alpha_0$ in accordance with previous TGV reconstruction applications [4],[5]. The ρ's are boosted by 5% per iteration for faster convergence, as suggested for the Method of Multipliers in [6]. The penalty parameter λ is determined using a greedy search, optimizing for nRMSE (see section III-C). For comparison reasons, we let ρ in the TV-ADMM scheme equal ρ_Z and ρ_Y motivating the equality $\rho_Z = \rho_Y = \rho$. Moreover, the same penalty parameter λ is used for TGV as well as for TV and Tikhonov regularization.

III. Experimental setting

A. Data simulation

The dataset used is knee MRI acquisitions consisting of 1000 images of size 211×211 pixels. All reconstructions were computed on a Mac M1 chip with 16 GB RAM.

To simulate CT acquisitions, the discrete Radon transform, matrix $\mathbf{A}$, is applied to an image to which noise $\sim N(0, \xi\sigma_{gt}\mathbf{I})$ is then added. Here, σ_{gt} is the ground truth sinogram standard deviation, and $\xi \in [0, 0.1]$ is a parameter determining the amount of noise. After applying noise, a sparse view undersampling pattern $\mathbf{M}$ is multiplied with the noisy sinogram, yielding the resulting simulated undersampled sinogram as

$$\mathbf{s} = \mathbf{M}(\mathbf{Ax} + \text{noise}) \tag{7}$$

where $\mathbf{x}$ is the ground truth MRI image. The undersampling matrix $\mathbf{M}$ depends on the acceleration rate R.

B. Choosing penalty parameter

To compare the performance of the different methods for different noise levels ξ, a search over relevant penalty parameters λ using TGV reconstruction with nRMSE as a measure of reconstruction error is conducted. This is important for achieving better performance of the penalized models since higher noise generally requires more regularization. The considered noise levels are $\xi = [0 : 0.1 : 0.01]$ and the search for the optimal parameter starts at 0 and increases until the error reaches a minimum.

TABLE I
RECONSTRUCTION METRICS CORRESPONDING TO THE RECONSTRUCTIONS PRESENTED IN FIGURE 5.

Method	nRMSE (%)	PSNR (dB)	SSIM
FBP	66.6	21.7	0.391
Tikhonov	32.7	30.2	0.728
TV	30.1	30.6	0.717
TGV	25.3	32.4	0.776

C. Measuring reconstruction quality

Three different measures are used to evaluate the quality of reconstruction. These are the normalized Root Mean Squared Error (nRMSE) given by

$$\text{nRMSE} = \text{MSE}/\text{std}$$

where MSE is the Mean Squared Error and std is the pixel intensity standard deviation of the ground truth. The second measure is the Peak-Signal-to-Noise ratio, given by

$$\text{PSNR} = 20 \cdot \log_{10} \frac{\text{MAX}_{\text{gt}}}{\text{RMSE}}$$

where RMSE is the (not normalized) Root Mean Squared Error and MAX_{gt} is the maximum possible intensity value of the ground truth. Since the data is given by MRI images, the pixel intensities have no direct interpretation. Hence the maximal value of the ground truth is used as a proxy for MAX_{gt}. The third and last measure is the Structural Similarity score (SSIM) defined by

$$\text{SSIM} = \frac{(2\mu_x\mu_{gt} + c_1)(2\sigma_{x,gt} + c_2)}{(\mu_x^2 + \mu_{gt}^2 + c_1)(\sigma_x^2 + \sigma_{gt}^2 + c_2)}$$

where μ is the mean pixel intensity, and σ^2 the corresponding pixel intensity variances/covariances. The constants c_1 and c_2 are given by $c_1 = (0.01L)^2$ and $c_2 = (0.03L)^2$ where L is the dynamic range of the pixel intensities.

All scores are calculated over the square patch inscribed in a circle whose diameter equals the image side length. The reason for this is that FBP images reconstructed using Scipy's built-in function lack information outside of this circle.

IV. Results

Figure 1 shows the convergence of the ADMM iterations for a general TGV reconstruction. It can be seen that the primal residuals first decay to zero and after that, the full cost function converges.

In Figure 2, the average reconstruction time using ADMM is shown. The result from the grid search for optimal λ is shown in Figure 3 for three different image sizes. Figure 4 shows the ρ versus reconstruction time trade-off plot.

Typical reconstruction results are shown in Figure 5, where a small patch is extracted for analysis of finer details. The corresponding reconstruction metrics are listed in Table I.

The average PSNR and nRMSE for 500 reconstructions are showcased in Figure 6 as a function of the sampling rate $1/R$. It can be seen that TGV reconstruction is superior to the other methods in all cases. Reconstructions from three, four, five, and six times undersampled data are shown in Figure 7.

V. Discussion

A. Convergence of ADMM

The convergence properties of the ADMM iterations were found to be heavily dependent on the value of the parameter ρ. For larger values, less time was needed to reach the convergence criterion whereas the optimal cost function value was increased, as indicated in Figure 4. The

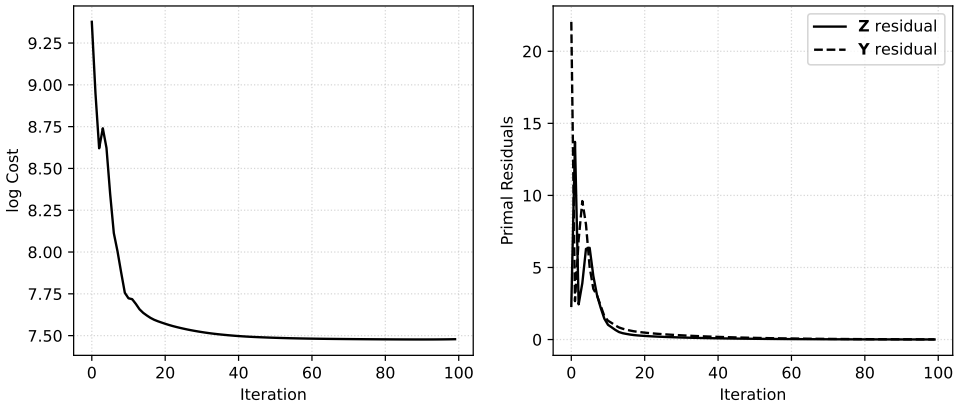


Fig. 1. Convergence of the TGV ADMM iterations. The left plot depicts the current value of the (log) cost function in (1) as a function of iteration number. The right plot shows the residuals of the constraints given in the primal problem in (2), also as a function of the current iteration.

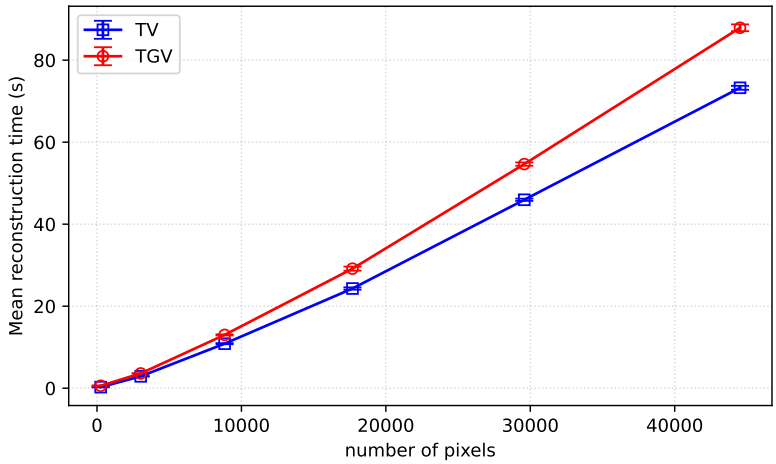


Fig. 2. Average reconstruction time for 25 different acquisitions of the ADMM iterations for TV and TGV regularization as a function of the number of pixels. The error bars indicate the sample standard deviation.

aforementioned figure shows that reconstruction time and cost exhibit a somewhat inverse relationship suggesting that ρ is a parameter that balances accuracy and the runtime of the algorithm. Using this plot, the value of ρ used in all other simulations was set to 50 for optimal balance between reconstruction time and cost.

Larger values of ρ force the primal residuals to be smaller as seen in the $\mathbf{V}$-, $\mathbf{Z}$- and $\mathbf{Y}$-updates in (3), therefore the cost function quickly become independent of these during the iterations and only varies with the data-fidelity term. This behavior can be seen in Figure 1 where the residuals quickly decay and thereafter the full cost function converges. Hence, higher values of ρ augment the suppression of the residuals, and faster convergence of the cost function is achieved. However, as seen in Figure 4, this will generally result in a less optimal minimum.

The time complexity of the ADMM solver can be seen to be linear in Figure 2, which is in accordance with the theoretical result [7]. The TGV-ADMM algorithm is, as expected, slower than the TV-ADMM algorithm since more calculations are needed per ADMM iteration. The graph also displays typical reconstruction times for the two methods. For the full size images, the TGV reconstruction takes an average of 90 seconds whereas roughly 70 seconds is needed for the TV reconstruction.

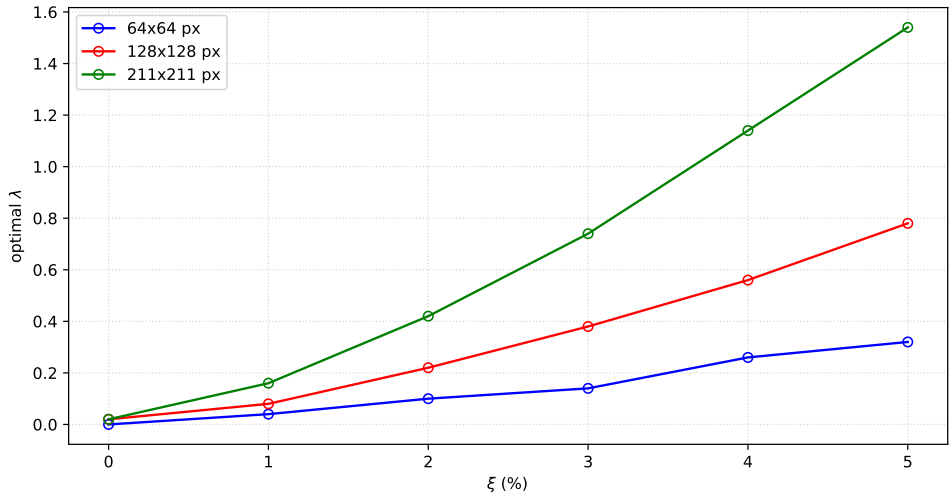


Fig. 3. Optimal $\boldsymbol{\lambda}$ as a function of noise level $\boldsymbol{\xi}$ for three different image side lengths: 64, 128, and 211 pixels respectively.

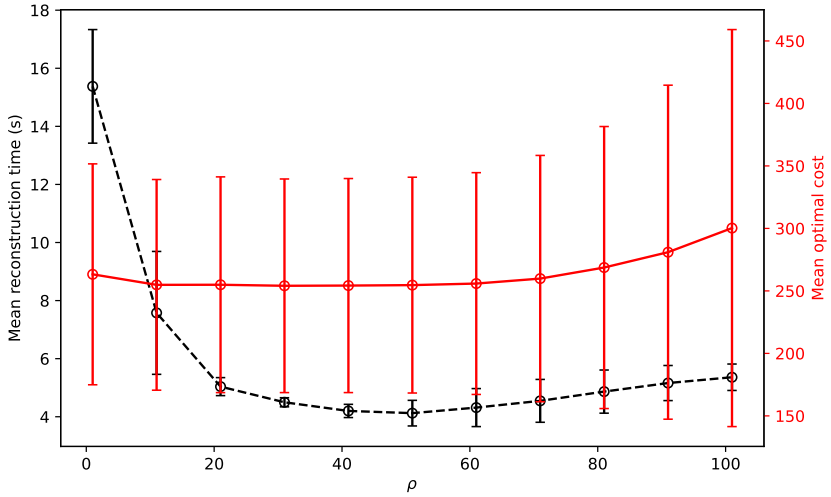


Fig. 4. Mean TGV-ADMM reconstruction time (left y-axis) and optimal cost function value (right y-axis) as a function of $\boldsymbol{\rho}$. The bars indicate the standard deviation of the respective metric for 200 reconstructions of size $\mathbf{64 \times 64}$.

B. Comparison of methods

Comparing the reconstruction metrics in Table I we surprisingly see that the Tikhonov image got a higher SSIM score than the TV reconstruction. However, looking at the actual reconstructions in Figure 5 we see that sharp edges in the Tikhonov reconstruction are not preserved, as they are in the TV (and TGV) case. To catch this difference in performance one might need to turn to another metric. Another reason for this result may be that the used λ was optimized for TGV performance and hence it may not be the optimal value for the TV method for the given noise level.

The TGV method gives rise to a more natural-looking image compared to the TV reconstruction that suffers from staircase artifacts. But on the other hand, sharp edges look sharper in the TV reconstruction due to its favoring of piecewise constant values compared to the TGV case,

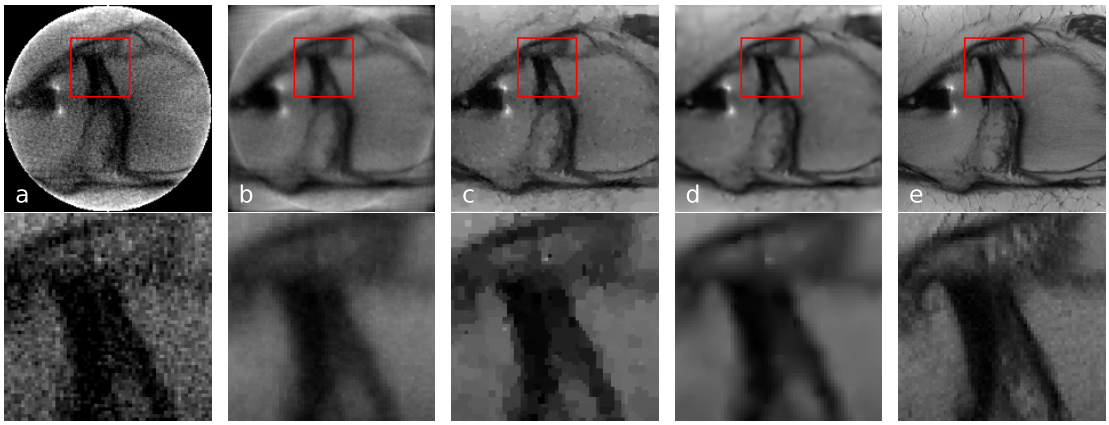
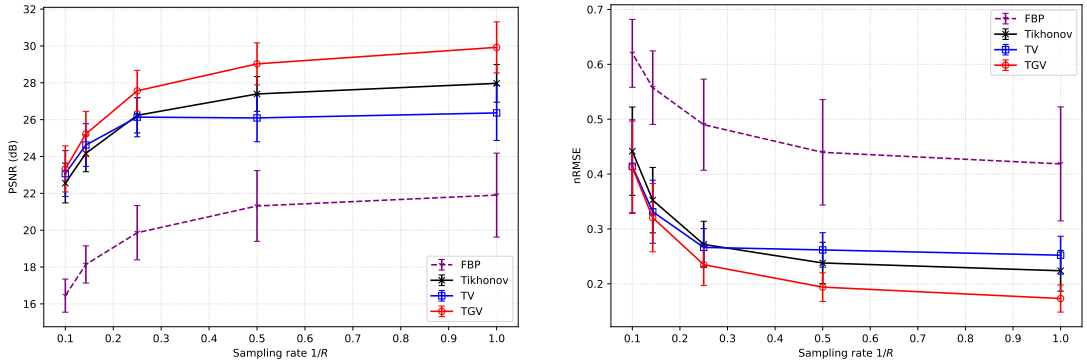


Fig. 5. Reconstruction results using an acceleration rate of $\boldsymbol{R = 2}$ and noise level $\boldsymbol{\xi = 0.05}$ of (a) filtered backprojection (FBP), (b) Tikhonov regularization, (c) TV regularization, and (d) TGV regularization together with (e) the ground truth image.



(a) PSNR (b) nRMSE

Fig. 6. Mean PSNR (a) and nRMSE (b) using $\boldsymbol{\xi = 0.05}$ for FBP, Tikhonov regularization, TV regularization, and TGV regularization as a function of sampling rate $\boldsymbol{1/R}$. Each dot represents the average score calculated for 500 acquisitions of size $\boldsymbol{64 \times 64}$ per noise level and method. The bars correspond to the sample standard deviations.

where they look slightly smeared out.

Even though the TGV regularized image looks the best from a full field of view, it can be seen that some noise is augmented (white dot in the upper middle part of the TGV patch in fig. 5) that is not present in the ground truth. This is of course not desirable since it can fool radiologists that it is a lesion. This dot is also clearly visible in the TV image as well, whereas, in the Tikhonov image, it is more smoothed out.

In Figure 6, for low sampling rates ($1/R < 0.15$), the three different regularized methods do not seem to vary much in terms of reconstruction PSNR/nRMSE. Moreover, it is clear that FBP is not competitive with the other methods at all (not surprisingly since it does not remove any noise). For higher sampling rates, the TGV regularized method outperforms the others. Interestingly, the PSNR/nRMSE values for $1/R = 0.5$ is not very different from that achieved using full sampling, for the regularized methods. This suggests that it may not be expensive to drop from a dense view to a double-undersampled acquisition in terms of these metrics, using such regularization. Using a fully sampled TV or Tikhonov reconstruction is similar to a four times undersampled TGV reconstruction on average, as indicated by Figure 6.

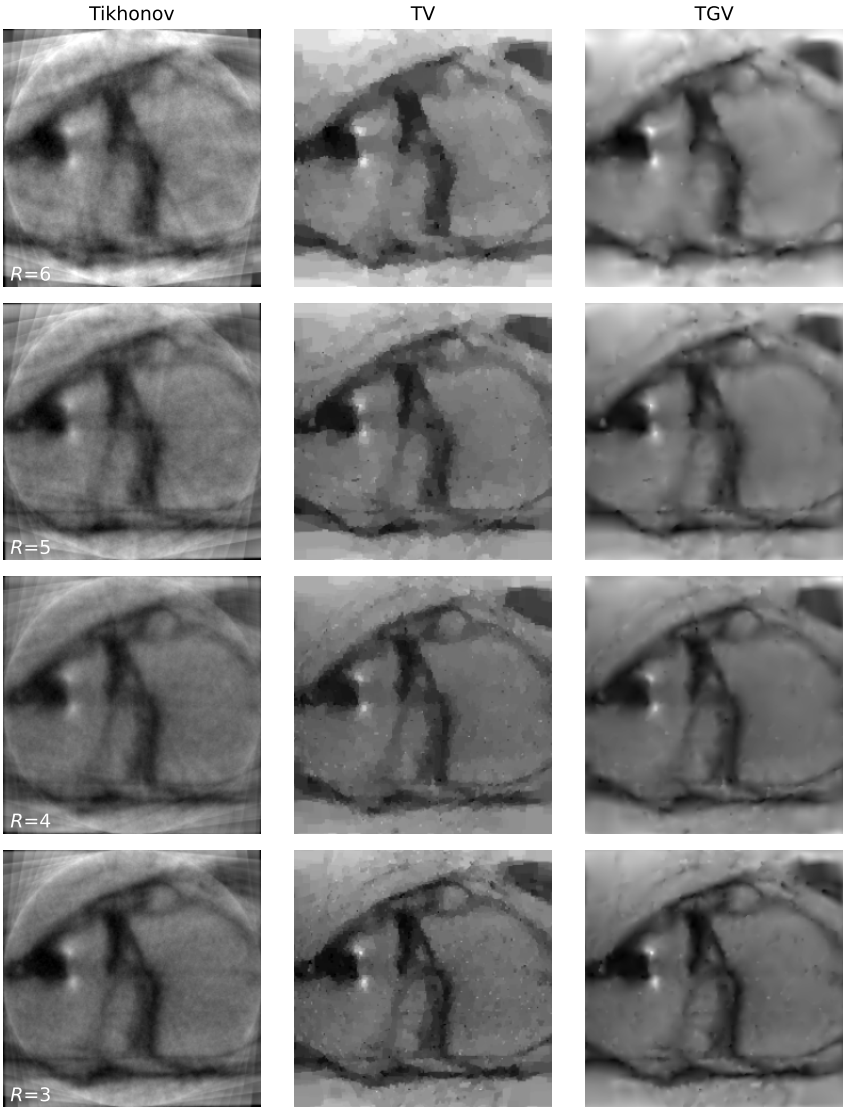


Fig. 7. Reconstructions using $\boldsymbol{\xi} = \mathbf{0.05}$ for Tikhonov regularization, TV regularization, and TGV regularization for different undersampling rates $\boldsymbol{R}$.

In Figure 7 it can be seen that for high acceleration rates, all reconstructions start degrading. The Tikhonov images show more streaking artifacts, the TV images show fewer isles of constant value and the TGV images become more smeared out.

C. Improvements & difficulties

In this study, only the cost function as a whole was used as a stopping criterion for the ADMM iterations. As described in [2], one could instead bound primal and dual residuals to ensure convergence since these in turn bound the value of the cost function. How this difference affects the results of this study is yet to be investigated.

Also mentioned in [2] is the use of dynamic updates of ρ during the ADMM iterations to achieve faster convergence. In this study, the values of ρ were statically boosted with 5% per iteration, but more advanced dynamic updates are to be investigated in further studies.

When searching for optimal λ's for different noise levels ξ, only the nRMSE score was used as an optimization objective. Upon visual inspection, this would tend to over-regularize the reconstructions producing very smeared-out images. Instead one could have used a combination

of different reconstruction scores and/or investigated how a different stopping criterion (as mentioned above) would have affected this outcome. Another major fault in this part of the study was that only one image was used to search for the optimal values in Figure 3. Since the dataset had quite some image variability depending on the slice, one could argue that it would have made more sense to search for an optimal λ over the whole dataset for each noise level. But due to the long reconstruction time, this was not feasible.

The long reconstruction time was the major difficulty of this study. Especially when performing larger scale simulations over parts of the dataset for computation of various metrics as shown in Figures 2 and 6.

A take-home message from this experiment is that it is hard to come up with appropriate values for the regularization parameter depending on the noise level, as well as a good value of ρ that works well for a variety of images. This is because one has to monitor both the reconstruction error as well as the time-consuming optimization all at once. Another thing found was that the value of ρ affected the convergence time of the ADMM algorithm. Lastly, it was interesting to find that three regularized methods performed similarly for low sampling rates, whereas the TGV method outperformed the others for more densely sampled data.

References

[1] *X-ray Computed Tomography: Advanced Lab Course No. 79*. "[Online; accessed 31-May-2023]". URL: https://www.ph.tum.de/academics/org/labs/fopra/docs/userguide-79.en.pdf.

[2] Stephen Boyd et al. "Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers". In: *Foundations and Trends® in Machine Learning* 3.1 (2011), pp. 13–24. DOI: http://dx.doi.org/10.1561/2200000016. URL: https://stanford.edu/~boyd/papers/pdf/admm_distr_stats.pdf.

[3] *numpy.roll*. "[Online; accessed 25-May-2023]". URL: https://numpy.org/doc/stable/reference/generated/numpy.roll.html.

[4] Florian Knoll et al. "Second order total generalized variation (TGV) for MRI". In: *Magnetic Resonance in Medicine* 65.2 (2011), pp. 480–491. DOI: https://doi.org/10.1002/mrm.22595. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/mrm.22595. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.22595.

[5] Florian Knoll et al. "Parallel imaging with nonlinear reconstruction using variational penalties". In: *Magnetic Resonance in Medicine* 67.1 (2012), pp. 34–41. DOI: https://doi.org/10.1002/mrm.22964. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/mrm.22964. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/mrm.22964.

[6] R. Tyrrell Rockefellar. "Monotone operators and the proximal point algorithm". In: *SIAM Journal on Control and Optimization* 14 (1976), p. 880.

[7] Ryan Tibshirani. *Lecture 21: Convex Optimization*. "[Online; accessed 27-May-2023]". 2016. URL: https://www.stat.cmu.edu/~ryantibs/convexopt-F16/scribes/admm-scribed.pdf.