COMPRESSIVE COMPUTED TOMOGRAPHY IMAGE RECONSTRUCTION WITH DENOISING MESSAGE PASSING ALGORITHMS

Alessandro Perelli, Mike E. Davies

Institute for Digital Communications (IDCOM)
University of Edinburgh, EH9 3JL, UK

ABSTRACT

In this paper we address the compressive reconstruction of images from a limited number of projections in order to reduce the X-ray radiation dose in Computed Tomography (CT) while achieving high diagnostic performances. Our objective is to study the feasibility of applying message passing Compressive Sensing (CS) imaging algorithms to CT image reconstruction extending the algorithm from its theoretical domain of i.i.d. random matrices. Exploiting the intuition described in [1] of employing a generic denoiser in a CS reconstruction algorithm, we propose a denoising-based Turbo CS algorithm (D-Turbo) and we extend the application of the denoising approximate message passing (D-AMP) algorithm to partial Radon Projection data with a Gaussian approximation of the Poisson noise model. The proposed CS message passing approaches have been tested on simulated CT data using the BM3D denoiser [2] yielding an improvement in the reconstruction quality compared to existing direct and iterative methods. The promising results show the effectiveness of the idea to employ a generic denoiser Turbo CS or message passing algorithm for reduced number of views CT reconstruction.

Index Terms— Computed Tomography, Radon Transform, Approximate Message Passing, Turbo Compressed Sensing

1. INTRODUCTION

Nowadays, X-ray computed tomography (CT) is widely used as a 3-D imaging technique in materials science or for medical applications such as the reconstruction of inner organs of patients from an extensive collection of X-ray projection images. In medical imaging one of the main challenges is to reduce the X-ray dose absorbed by the patient together with the scan time while obtaining high quality images at a minimum radiation dose level. The reduction of projection data leads to benefits such as reducing patient dose, reducing scan time, and improving time-resolution in CT. Classical reconstruction algorithms such as filtered back-projection (FBP) require the linear system to be sufficiently over-determined, i.e. the number of angles needed is more than the number of pixels along a line. It is well-known that traditional FBP-based reconstruction methods produce significant artifacts when applied to a limited number of angles [3], [4] and algorithms incorporating prior information on the sample are needed to improve the quality of the reconstruction. Various approaches have been proposed, from Algebraic Reconstruction Technique (ART) [5] to convex relaxation [6] and fast iterative FBP. Compressed sensing (CS) has sparked a tremendous amount of research activity, because it performs image processing using fewer samples. Many compressive imaging algorithms have been proposed in the literature, however existing compressive imaging algorithms may either not achieve good reconstruction quality or not be fast enough. Therefore, in this paper, we focus on a variation of fast algorithms based on graphical models belonging to the family of message passing to improve over the prior art. AMP is an iterative signal reconstruction algorithm that performs scalar denoising within each iteration, and proper selection of the denoising function used within AMP is needed to obtain better reconstruction quality [7]. The state evolution of AMP with Gaussian sensing matrices is shown to be consistent with that derived using the replica method. However, the situation is different in the CT context since the distortion is not Gaussian and the entries of the partial Radon transform are not random.

1.1. Main Contribution

The objective of this paper is to develop message passing algorithms for image reconstruction in the context of sparse projection measurements CT. In particular, we exploit both the intuition in [1] to use a generic denoiser in the AMP framework (D-AMP) with the Radon operator. To tackle the problem that the Radon in a non tight operator, a preconditioning step within the D-AMP is performed [8]; as a specific denoiser the BM3D algorithm has been used. Furthermore, we incorporate in the preconditioner a weighted matrix to represent the popular Gaussian approximation of the Poisson noise model in CT [9]. Finally an alternative message passing algorithm D-Turbo CS is proposed which is based on the idea of using a specific denoiser in the Turbo CS framework [10].

The paper is organised as follows. In the next section a review of the Radon projection forward model for paral-

978-0-9928626-3-3/15/$31.00 ©2015 IEEE

lel beam X-ray CT is described. The extension of denoising approximate message passing applied with the Radon operator in described in Section 4. In Section 5 the proposed denoising-based Turbo CS (D-Turbo) framework with partial Radon Matrix measurement is described. Finally, Section 6 presents a simulated evaluation and comparison of the denoising message passing algorithms on a CT scan.

2. FORWARD RADON MODEL

The goal of CT is to determine the attenuation coefficients $\mu(\overrightarrow{s}) \in L_2(\mathbb{R}^2)$, where $\overrightarrow{s}$ represents the spatial location. The mono-energetic X-ray CT projection process is modelled by the Beer's law:

$$I_i = I_0 e^{\left(-\int_{L_i} \mu(\overrightarrow{s})dl\right)} + \epsilon_i \ , \ i = 1, \ldots, N_d \qquad (1)$$

where I_i is the mean detected intensity, I_0 is the initial intensity of a X-ray beam, ϵ_i is the room background and mean scatter radiation which will not taken into account in the following formulation, L_i is a line path of the ray through the object and N_d is the total number of rays.
One X-ray CT projection provides N_d line integral values of $\mu(\overrightarrow{s})$ for each ray, which are equivalent to the following forward model with post-log data y_i of the mean measurement I_i:

$$y_i = \log\left(\frac{I_0}{I_i}\right) = \int_{L_i} \mu(\overrightarrow{s})dl \ , \ i = 1, \ldots, N_d \qquad (2)$$

Under the assumption of parallel-beam the operator coincides with the Radon projection $\mathcal{P} : L_2(\mathbb{R}^2) \rightarrow L_2([0, \pi] \times \mathbb{R})$:

$$\mathcal{P}\left\{x(\overrightarrow{s})\right\}(\theta, \xi) = \int_{\mathbb{R}^2} x(s)\delta(\xi - s_1 \cos\theta - s_2 \sin\theta)ds \quad (3)$$

with $\overrightarrow{s} = (s_1, s_2) \in \mathbb{R}^2$. Considering the Fourier slice theorem using the Radon Operator is equivalent to

$$\mathcal{P}\left\{x(\overrightarrow{s})\right\}(\theta, \xi) = \int_{\mathbb{R}^2} X(\xi\cos\theta, \xi\sin\theta)e^{2\pi i\xi\theta}d\theta \qquad (4)$$

where $X(\xi\cos\theta, \xi\sin\theta)$ is the 2D Fourier transform of $x(s)$ in polar coordinates. The Radon operator is a non tight frame so according to the Filtered Back Projection theorem [11] we need to precondition the operator $\mathcal{P}^*\mathcal{V}\mathcal{P} = \mathcal{I}$ where $\mathcal{V}$ is the operator that applies the ramp filter to each projection

$$\mathcal{V} = \mathcal{I} \otimes (\mathcal{F}_1^{-1}\mathcal{D}(|\xi|)\mathcal{F}_1) \qquad (5)$$

where $\mathcal{I} \otimes \mathcal{F}_1$ takes the 1D DFT of each projection view of the sinogram and $\mathcal{D}$ defines the diagonal polar Fourier-space operator. The Non Uniform Fourier Transform (NUFFT) [12] has been exploited to evaluate a discretised version of (3).
To develop a Fourier-based projector for fan-beam geometries, we can use the well-known relation between parallel-beam and fan-beam coordinates $r = R\sin\beta, \ \theta = \alpha + \beta$ where R is the source to rotation center distance, α the angle of the source relative to one axis, and β is the angle of the ray relative to the source.
The statistical X-ray CT noise model consists in modelling the pre-log observation data by a Poisson model of a discretised version of (1):

$$Y_i \sim \text{Poisson}\{\hat{y}_i\} = \text{Poisson}\left\{I_0 e^{\left(-\int_{L_i}\mu(\overrightarrow{s})dl\right)}\right\}$$

To process the reconstruction, it is necessary to discretise a continuous object $\mu(\overrightarrow{s})$ by a basis function expansion. The image domain discretization leads to the following forward projection for (2):

$$y_i = \log\left(\frac{I_0}{I_i}\right) \approx \sum_{j=1}^{N_d} x_j \int_{L_i} \gamma_j(\overrightarrow{s})dl = \sum_{j=1}^{N_d} p_{ij}x_j = [\mathbf{P}x]_i$$

where $\gamma_j(\overrightarrow{s})$ is the basis function. In this work we are interested in the CT imaging using a limited number of projections therefore, given the input $N \times N$ pixel image x, the problem is equivalent to the undetermined linear system where we define the sensing matrix as the partial Radon Transform matrix, where the full matrix has N columns equal to the number of detectors, $\mathbf{\Phi} = \mathbf{SP} \in \mathbb{R}^{M\times N}$ and M rows ($M < N$), with $\mathbf{S}$ the subselection matrix $y = \mathbf{\Phi}x + w$ where w is the additive noise with Poisson distribution.
The post-log observation data y_i has been modelled by a weighted linear additive Gaussian noise model [13], which comes from quadratically approximating the negative log likelihood

$$y_i = [\mathbf{SP}x]_i + w_i \ , \ w_i \sim \mathcal{N}(0, \sigma_i^2)$$

where the noise variance is $\sigma_i^2 = y_i$.

3. DENOISING APPROXIMATE MESSAGE PASSING

D-AMP is an iterative signal reconstruction algorithm based on a graphical model approximation; consider the model of y where the signal distribution follows $x_i \sim f_X$ and the noise is i.i.d. Gaussian. The entries of the measurement matrix $\mathbf{\Phi}$ are i.i.d. $\mathcal{N}(0, \frac{1}{M})$ distributed, and thus the columns of the matrix have unit l_2-norm, on average. D-AMP proceeds iteratively according to

$$\begin{aligned} x^{t+1} &= D_{\hat{\sigma}^t}(\mathbf{\Phi}^T r^t + x^t) \\ r^t &= y - \mathbf{\Phi}x^t + \frac{1}{R} r^{t-1} D'_{\hat{\sigma}^t}(\mathbf{\Phi}^T r^{t-1} + x^{t-1}) \\ (\sigma^t)^2 &= \frac{\|r^t\|_2^2}{R} \end{aligned} \qquad (6)$$

where $R = M/N$ represents the measurement rate, $D_{\hat{\sigma}^t}(\cdot)$ is a denoising function at the t-th iteration and $D'_{\hat{\sigma}^t}$ denotes the divergence of the denoiser with respect to the estimated

noise level σ^t which is well approximated using Monte Carlo technique [1]. At each iteration D-AMP generates an estimate for x that can be represented as the true x corrupted by Gaussian noise, i.e. $x^{t+1} \approx x + v^t$ where $v_i^t \sim \mathcal{N}(0, \sigma_t^2)$. The so called Onsager term in r^t yields the D-AMP to behave like a self-tuned algorithm. The power of D-AMP is that it can exploit a huge number of image denoisers to enhance the reconstruction performance.

4. RADON TRANSFORM IN D-AMP

We now show the image denoising can be performed within D-AMP with the Radon operator. The intuition comes from using D-AMP incorporating the filtered back projection within each iteration as in other iterative FBP algorithms.
The convergence of AMP algorithm depends on the singular values of the matrix $\mathbf{\Phi}$ [8]. Thus, to speed up the convergence of the algorithm, one way is to precondition the system. Using the forward model (1) it has been shown that the singular value decomposition of the Radon matrix follows the ramp function in the frequency domain. This property has been exploited to incorporate the preconditioner in the D-AMP iterative algorithm. Furthermore to account for the noise model we also incorporate the weight matrix, $\mathbf{W} = \mathrm{diag}\{w_i\}$, into the preconditioner.
Let $\vartheta_x = (\mathbf{WV})^{\frac{1}{2}} x$ where $\mathbf{V}$ is the discrete form of Eq. (5). Therefore the compressed measurement becomes $y = \mathbf{SP}(\mathbf{VW})^{-\frac{1}{2}}\vartheta_x + z$, where $\mathbf{SP}(\mathbf{VW})^{-\frac{1}{2}}$ is our new matrix in (6) and ϑ_x is the corresponding input signal. Let us express the D-AMP iterations (6) for settings where the matrix is $\mathbf{SP}(\mathbf{VW})^{-\frac{1}{2}}$,

$$\vartheta_x^{t+1} = D_{\hat{\sigma}^t}((\mathbf{SP}(\mathbf{VW})^{-\frac{1}{2}})^T r^t + \vartheta_x^t) \quad (7)$$

$$\begin{aligned} r^t &= y - (\mathbf{SP}(\mathbf{VW})^{-\frac{1}{2}})\vartheta_x^t + \quad (8)\\ &\quad \frac{1}{R} r^{t-1} D'_{\hat{\sigma}^{t-1}}((\mathbf{SP}(\mathbf{VW})^{-\frac{1}{2}})^T r^{t-1} + \vartheta_x^{t-1}) \\ &= y - \mathbf{SP}x^t + \frac{1}{R} r^{t-1}\vartheta_x^t \\ &\quad + \frac{1}{R} r^{t-1} D'_{\hat{\sigma}^{t-1}}((\mathbf{SP}(\mathbf{VW})^{-\frac{1}{2}})^T r^{t-1} + \vartheta_x^{t-1}) \quad (9) \end{aligned}$$

the input of the denoiser $D_{\hat{\sigma}^t}$ becomes

$$\begin{aligned} (\mathbf{SP}(\mathbf{VW})^{-\frac{1}{2}})^T r^t + \vartheta_x^t &= (\mathbf{SP})^T(\mathbf{VW})^{\frac{1}{2}} r^t + \vartheta_x^t = \\ &= (\mathbf{SP})^T(\mathbf{VW})^{\frac{1}{2}} r^t + (\mathbf{WV})^{\frac{1}{2}} x^t \\ &= (\mathbf{VW})^{\frac{1}{2}} q^t \end{aligned}$$

where q^t is the noisy image at the iteration t.
While D-AMP is only guaranteed to converge with i.i.d. sensing Gaussian matrices, one of the main aims of this work is to show how successfully message passing algorithms can work in solving undetermined system away from the ideal case similar to other popular CS recovery algorithms such as L_1 or Iterative Hard Thresholding (IHT).

5. DENOISING-BASED TURBO CS

The next proposed algorithm is based on the turbo principle in iterative decoding [14]. The message passing scheme of the D-Turbo algorithm is illustrated in Fig. 1 and it is composed of two stages with a loopy message passing.

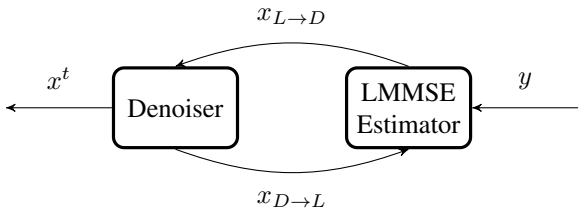

Fig. 1. Denoising-based Turbo CS message passing scheme

The first stage is the Linear Minimum Mean Square Error Estimator (LMMSEE) which produces an extrinsic estimate of x based on the observation y. Exploiting the extinsic formulation of the LMMSE [14], we can obtain at the t-th iteration the following expression for the messages $x_{L\to D}^t$ and $\eta_{L\to D}^t$ which are respectively the mean value and variance of the estimate:

$$\begin{aligned} x_{D\to L}^t &= \eta_{D\to L}^t \left(\frac{x^t}{(\sigma^t)^2} - \frac{x_{L\to D}^{t-1}}{\eta_{L\to D}^{t-1}} \right) \\ x_{L\to D}^t &= \eta_{L\to D}^t \left(\frac{x_{LMMSE}^t}{\eta_{LMMSE}^t} - \frac{x_{D\to L}^{t-1}}{\eta_{D\to L}^{t-1}} \right) \\ &= x^t + \tfrac{N}{M}\mathbf{P}^T\mathbf{VWS}^T\left(y - \mathbf{SP}x^{t-1} + \right. \\ &\quad \left. + \mathbb{E}\left[D_{\eta_{L\to D}^t}(x_{L\to D}^{t-1}) \right] \right) \\ \eta_{D\to L}^t &= \frac{\mathbb{E}\| D_{\eta_{L\to D}^t}(x_{L\to D}^t)\|^2}{1 - \mathbb{E}\| D_{\eta_{L\to D}^t}(x_{L\to D}^t)\|^2 \cdot \eta_{L\to D}^{t-1}} \\ \eta_{L\to D}^t &= \left(\frac{N}{M} - 1 \right) \eta_{D\to L}^{t-1} + (\sigma^t)^2 \end{aligned}$$

where $\mathbb{E}[\cdot]$ is the expectation. The vector $x_{L\to D}^t$ is used as the input for the denoiser $x^t = D_{\eta_{L\to D}^t}(x_{L\to D}^t)$. It is important to point out that the vector $x_{L\to D}^t$ has a similar structure as the equivalent vector q^t in the AMP message passing and also it incorporates the preconditioning of the Radon operator. The main difference between AMP and Turbo CS is related to the derivation of the so called Onsager term.

6. RESULTS

To compare the different CT reconstruction strategies for reduced number of views, the projection dataset from a real phantom in Fig. 2(a) was simulated in Matlab with a fan-beam X-ray CT geometry [15]. The number of views is 90 evenly spanned on an orbit of $180°+$ fan angle. The detector arrays are on an arc concentric to the X-ray source with a

distance of 949 mm and the distance from the rotation center to the X-ray source is 541 mm as depicted in Fig. 2(b). The detector cell spacing is 1.0239 mm and FOV $\sim$ 25 cm. A quarter detector offset is also included to reduce aliasing. The reconstructed image is of 512×512 array size and the Poisson noise with $I_0 = 10^5$ has been included in the simulation.
The existing image denoising algorithm BM3D has been used as the denoiser in both D-AMP and D-Turbo CS since it produces the state-of-art recovery for natural images [2].

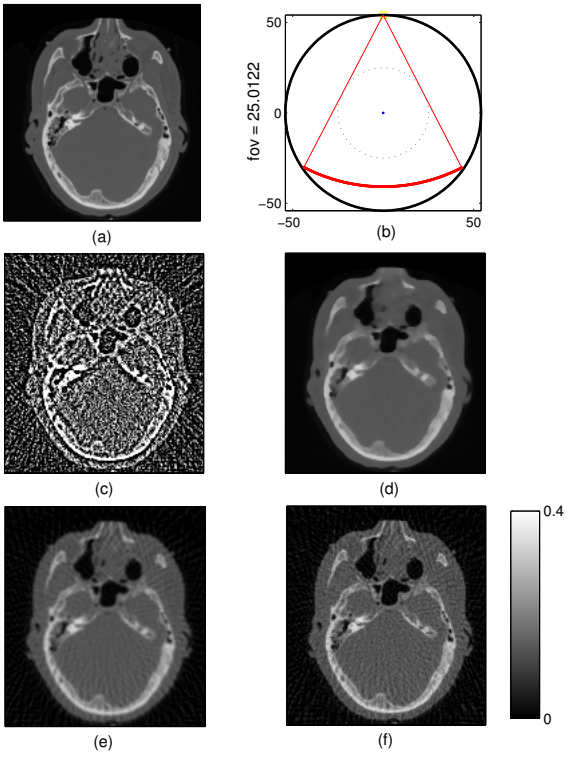

Fig. 2. CT image reconstruction: (a) Real image used for CT scan simulation, (b) Fan-Beam simulated geometry, (c) FBP recovery, (d) PWLS-PCG recovery (e) BM3D-Turbo CS recovery, (f) BM3D-AMP recovery

In Fig. 2 are shown the image recovery results for Filtered Back Projection (FBP), D-AMP and D-Turbo CS. The PWLS, which represents the state of the art for CT reconstruction, refers to the minimization of the following cost function

$$\arg\min_x \|y - SPx\|_W^2 + \beta R(x)$$

through the preconditioning Conjugate Gradient (PCG) algorithm. $R(x)$ is a TV-like regularization $R(x) = \sum_{i=1}^{K_n} \eta_i(x - x_i)$ where $\eta_i(x)$ is an edge-preserving potential function. From Fig. 3(e)-(f) and the zoomed details in Fig. 4 it can be seen that both BM3D-AMP and BM3D-Turbo CS yield a better reconstruction of the structure of the image than PWLS-PCG which thens to over smooth features due to the TV-like regularization even with different values of $\beta > 1$.

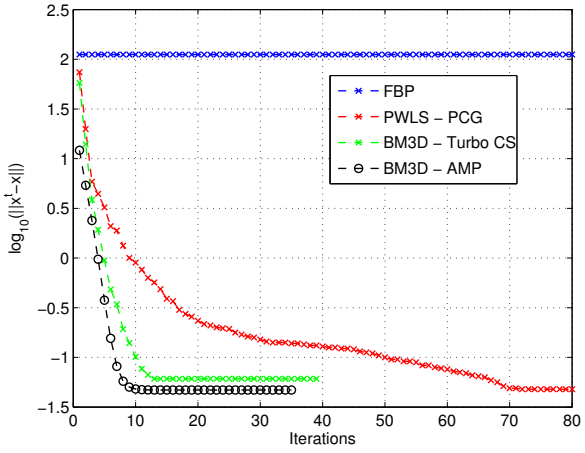

Fig. 3. Error $\log \|x^t - x\|$ at each iteration for the tested recovery algorithms: the stopping criterion has been manually selected when the algorithm reached the plateau.

In Fig. 3 it can be seen that D-Turbo and D-AMP converge in approximately 10 iterations while PWLS-PCG requires 70 iterations. Moreover, the D-Turbo CS and D-AMP outperform the iterative PWLS-PCG in terms of computation time. The computation times, PSNR and complexity analysis are summarised in Table 1. The algorithm complexity of both D-AMP and D-Turbo CS is mainly constituted by the denoiser BM3D applied at each iteration. Finally both D-Turbo CS and D-AMP achieve comparable results in terms of PSNR with low computational time respect to PWLS-PCG. In addition D-AMP leads to an higher flexibility in choosing the appropriate denoiser according to the image and noise model and in balancing the trade-off between complexity of each iteration and recovery resolution.

7. CONCLUSIONS

In this paper two message passing type algorithms for CT reconstruction with fewer angular projection measurements have been proposed and tested. The CS recovery algorithms are based on the idea of embedding generic denoisers into the message passing framework and have been adapted to include suitable preconditioning for the CT imaging task. The results show an improvement in reconstruction of the fine detail in comparison with a state of the art model based iterative reconstruction technique. Further analysis will include a detailed comparison of the Radon-based algorithms with the performance predicted by state evolution [1] and a comparison between the two message passing approaches D-AMP and D-Turbo CS.

Table 1. PSNR and computational time for an 512×512 image reconstruction

Algorithms	**PSNR** [dB]	**Time / Iterations**	**Complexity**
BM3D-Turbo CS	10.8	2.7 min 10 iterations	2D NUFFT $O(2(\alpha N)^2 \log(\alpha N))+$ $O(N_r N_\theta + 3MN)$ + BM3D
BM3D-AMP	11.05	3.1 min 10 iterations	2D NUFFT $O(2(\alpha N)^2 \log(\alpha N))+$ $O(N_r N_\theta + 4MN)$ + BM3D
PWLS-PCG	11.3	8.5 min 70 iterations	2D NUFFT $O(2(\alpha N)^2 \log(\alpha N))+$ $O(N_r N_\theta + 2MN)$
FBP	3.2	0.5 min	2D NUFFT $O(2(\alpha N)^2 \log(\alpha N) + N_r N_\theta)$

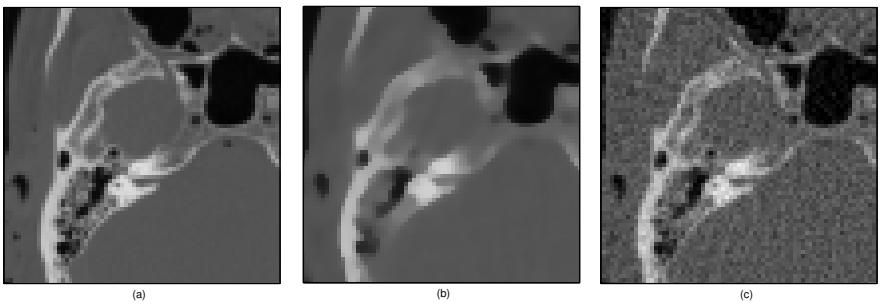

Fig. 4. Image detail: (a) Original image, (b) PWLS-PCG, (c) BM3D - AMP

REFERENCES

[1] C. A. Metzler, A. Maleki, and R. Baraniuk, "From denoising to compressed sensing," *arXiv*, July 2014.

[2] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. on Image Processing*, vol. 16, no. 8, pp. 2080–2095, 2007.

[3] X. Pan, E. Y. Sidky, and M. Vannier, "Why do commercial CT scanners still employ traditional, filtered back-projection for image reconstruction?," *Inverse problems*, vol. 25, no. 12, pp. 123009, 2009.

[4] E. Gouillart, F. Krzakala, M. Mézard, and L. Zdeborov, "Belief-propagation reconstruction for discrete tomography," *Inverse Problems*, vol. 29, no. 3, 2013.

[5] R. Gordon, R. Bender, and Herman G. T., "Algebraic reconstruction techniques (art) for three-dimensional electron microscopy and X-ray photography," *Journal of Theoretical Biology*, 1970.

[6] E. Y. Sidky and X. Pan, "Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization," *Physics in Medicine and Biology*, vol. 53, 2008.

[7] C. Guo and M. E. Davies, "Near optimal compressed sensing without priors: Parametric SURE approximate message passing," *IEEE Trans. on Signal Processing*, vol. 63, no. 8, 2015.

[8] S. Rangan, P. Schniter, and A. Fletcher, "On the convergence of approximate message passing with arbitrary matrices," in *IEEE International Symposium on Information Theory (ISIT)*, 2014, pp. 236–240.

[9] J. Nuyts, B. De Man, J. A. Fessler, W. Zbijewski, and F. J. Beekman, "Modelling the physics in the iterative reconstruction for transmission computed tomography," *Physics in medicine and biology*, vol. 58, no. 12, 2013.

[10] J. Ma, X. Yuan, and L. Ping, "Turbo compressed sensing with partial DFT sensing Matrix," *IEEE Signal Processing Letter*, vol. 22, no. 2, pp. 158–161, 2015.

[11] S. Matej, J.A. Fessler, and I.G. Kazantsev, "Iterative tomographic image reconstruction using Fourier-based forward and back-projectors," *IEEE Trans. on Medical Imaging*, vol. 23, no. 4, pp. 401–412, 2004.

[12] J. A. Fessler and B. P. Sutton, "Nonuniform fast fourier transforms using min-max interpolation," *IEEE Trans. on Signal Processing*, vol. 51, no. 222, 2003.

[13] J. Wang, T. Li, H. Lu, and Z. Liang, "Penalized weighted least-squares approach to sinogram noise reduction and image reconstruction for low-dose x-ray computed tomography," *IEEE Trans. on Medical Imaging*, vol. 25, no. 10, pp. 1272–1283, 2006.

[14] M. Tuchler, R. Koetter, and AC. Singer, "Turbo equalization: principles and new results," *IEEE Trans. on Communications*, vol. 50, no. 5, pp. 754–767, 2002.

[15] J. Fessler, "CT Toolbox," web.eecs.umich.edu/ fessler/.