DTU Library

Joint image reconstruction method with correlative multi-channel prior for x-ray spectral computed tomography
Paper

Kazantsev, Daniil; Jørgensen, Jakob Sauer; Andersen, Martin S; Lionheart, William R B; Lee, Peter D; Withers, Philip J

[Link back to DTU Orbit](#)

PAPER • OPEN ACCESS

Joint image reconstruction method with correlative multi-channel prior for x-ray spectral computed tomography

View the article online for updates and enhancements.

Joint image reconstruction method with correlative multi-channel prior for x-ray spectral computed tomography

Daniil Kazantsev[1,2] , **Jakob S Jørgensen[1,3,5]** ,
Martin S Andersen[4] , **William R B Lionheart[3]** , **Peter D Lee[1,2]**
and Philip J Withers[1,2]

[1] The Manchester X-Ray Imaging Facility, School of Materials, The University of Manchester, Manchester, M13 9PL, United Kingdom
[2] The Manchester X-Ray Imaging Facility, Research Complex at Harwell, Didcot, Oxfordshire, OX11 0FA, United Kingdom
[3] School of Mathematics, The University of Manchester, Manchester, M13 9PL, United Kingdom
[4] Department of Applied Mathematics and Computer Science, Technical University of Denmark, Lyngby, DK-2800, Denmark

E-mail: daniil.kazantsev@manchester.ac.uk

Abstract

Rapid developments in photon-counting and energy-discriminating detectors have the potential to provide an additional spectral dimension to conventional x-ray grayscale imaging. Reconstructed spectroscopic tomographic data can be used to distinguish individual materials by characteristic absorption peaks. The acquired energy-binned data, however, suffer from low signal-to-noise ratio, acquisition artifacts, and frequently angular undersampled conditions. New regularized iterative reconstruction methods have the potential to produce higher quality images and since energy channels are mutually correlated it can be advantageous to exploit this additional knowledge. In this paper, we propose a novel method which jointly reconstructs all energy channels while imposing a strong structural correlation. The core of the proposed algorithm is to employ a variational framework of parallel level sets to encourage joint smoothing directions. In particular, the method selects reference channels

[5] Part of the work was carried out while the author was with the Department of Applied Mathematics and Computer Science, Technical University of Denmark.

from which to propagate structure in an adaptive and stochastic way while preferring channels with a high data signal-to-noise ratio. The method is compared with current state-of-the-art multi-channel reconstruction techniques including channel-wise total variation and correlative total nuclear variation regularization. Realistic simulation experiments demonstrate the performance improvements achievable by using correlative regularization methods.

Keywords: multi-spectral, image reconstruction, structural regularization, inverse problems, total variation, materials science, x-ray imaging

(Some figures may appear in colour only in the online journal)

1. Introduction

1.1. New energy-discriminating detectors enabling spectral CT

Conventional x-ray imaging entails a polychromatic x-ray source (i.e. a beam having a wide spectrum of energies) utilizing detectors that count photons without any energy discrimination. This increases the intensity and photon count, but results in non-linear attenuation leading to 'beam-hardening' artifacts [1–5]. During propagation of a poly-energetic beam in matter, low-energy photons are absorbed more easily than high-energy photons resulting in a shift of the x-ray spectrum toward the higher energies. This affects the assumed linearity of Beer's law and biases the reconstructions with 'beam hardening' artifacts [4]. These artifacts can be avoided or reduced using a monochromator, which filters the beam allowing only a very narrow wavelength band. However, it is only practical in synchrotrons due to the dramatically higher flux, whereas current laboratory sources require a full spectrum (white-beam) to produce sufficient signal-to-noise ratio (SNR) in an acceptable time. At the image processing side one can apply a calibration procedure to raw projection data using polynomial fitting towards the approximate linear absorption model (linearization) [1–3]. With a larger number of materials involved, higher orders polynomials must be estimated leading to an ill-posed multidimensional optimization problem. With recent developments in energy-discriminating detectors, one can now turn the presence of the polychromatic beam into an advantage, increasing the information that can be obtained by white-beam CT.

Energy-discriminating photon counting detectors (PCDs) [6] can enable 3D *chemical imaging* [7, 8], as x-ray attenuation of each material element depends on the energy, i.e. different elements can be clearly distinguished when looking across the full energy spectrum (e.g. distinct K-edges [9]). K-edge imaging deals with the localization of a sudden increase in the attenuation coefficient. The characteristic absorption peak is a unique footprint of a material and it happens when a photon energy just above the binding energy of the K shell electron of the atoms interacting with the photons.

Energy-discriminating detectors allow one to capture information about the chemical composition, vastly improving the ability to distinguish different materials within a sample. The new generation of spectral detectors (also called called hyper-spectral) can achieve energy resolution less than 1 keV resulting in hundreds of energy channels [10]. This is different to medical imaging spectral CT where much coarser spectral resolution is available (normally dual and up to five energy channels [11]). Such fine spectral resolution is required when the K-edges of materials are closely spaced [8]. The significant disadvantage of hyper-spectral detectors is that they normally operate at low count rates. To overcome this, longer exposure

times can be used, however, is not always possible and can result in underexposed (noisy) and angular under-sampled datasets. New effective reconstruction strategies are needed to address this challenge and ensure more accurate spatio-spectral K-edge identification. To this end the present work proposes a new correlative reconstruction method which addresses the low data SNR by encouraging joint structures across the channels.

1.2. Spectral CT reconstruction approaches and the proposed method

Existing reconstruction techniques for spectral tomography can be classified into two categories: *material-decomposed* reconstruction methods [12–15] and *multi-channel* reconstruction methods [16–18]. The former approach estimates material-decomposed sinograms or basis-material maps from raw spectral measurements directly. Some recent material decomposition techniques seek statistical correlations in the data in order to increase the accuracy of the decomposition process [12, 13, 15]. The basis-material sinograms obtained are reconstructed in order to obtain material-specific images. In this paper, we are interested in the latter approach, when data are reconstructed directly without any special rearranging steps.

For the case of spectral image reconstruction, we adopt the framework of *parallel level sets* which has been applied previously to vector-valued images [19] and bimodal imaging [20–23]. In some cases, one modality image (so-called *reference*) is known and used to regularize the reconstruction process of another modality [21, 22]. Alternatively, one can try to solve a joint reconstruction problem where both modalities or channels are reconstructed simultaneously supporting each other during the process [20, 23–25].

In this paper we adopt the latter approach and generalize it to an arbitrary large number of channels. In our case, all channels are reconstructed jointly [23, 24] and the reference channel is selected from the previous iteration of the reconstruction algorithm. An additional contribution to this generalization is given with a novel probabilistic approach to select a suitable reference channel. Priority is given to a random channel with a higher SNR based on the geometric mean of the flux. The reference channel is selected by a random draw from the data-specific probability mass function. This is motivated by the problem of non-uniform distribution of noise variances across the energy domain. The proposed approach is heuristic, however, by providing extensive numerical simulations we conclude that the stochastic selection of the reference channel is crucial to ensure stability of the method.

In order to demonstrate the advantages of the proposed technique we perform realistic numerical experiments with a variation of Rigie's total nuclear variation (TNV) method [18] and also compare with the classical channel-wise total variation (TV) regularization. Our results show that the proposed method is a competitive correlative technique for multi-channel reconstruction.

The paper is organized as follows. In section 2 we discuss the forward reconstruction problem setting and existing reconstruction methods with correlative multi-channel regularization techniques. In section 3 we look into regularization using parallel level sets and propose novel methods. In section 4 we discuss the general reconstruction framework; details of verification given in appendix. In section 5 we provide details on synthetic data generation process and all numerical results are introduced. Discussion and conclusions are given in section 6.

2. Forward problem setting and existing reconstruction methods

2.1. Poly-energetic x-ray CT measurement model

Given a poly-energetic x-ray source, the spectral version of Beer's law can be expressed as [1, 2]

$$\Lambda_i(E) = \sigma_i(E) \exp\left(-\int_{L_i} \mu(\boldsymbol{r}, E)\, \mathrm{d}l\right), \quad i = 1, \ldots, M,$$ (1)

where M is the total number of rays/measurements, E is the energy, $\boldsymbol{r} \in \mathbb{R}^3$ defines the spatial position, μ is the energy-dependent attenuation coefficient, σ_i is the energy-dependent intensity flux of the x-ray source associated with ray i, and $\Lambda_i(E)$ is the spectrum of the x-ray beam incident on the detector. By introducing a parameterization $\mu(\boldsymbol{r}, E) = \sum_{j=1}^{N} \mu_j(E)\chi_j(\boldsymbol{r})$, where N denotes the number of pixels, and defining $A_{ij} = \int_{L_i} \chi_j(\boldsymbol{r})\, \mathrm{d}l$, we arrive at the model

$$\Lambda_i(E) = \sigma_i(E) \exp\left(-\sum_{j=1}^{N} A_{ij}\mu_j(E)\right), \quad i = 1, \ldots, M.$$

For basis functions $\chi_j(\boldsymbol{r})$ we use a pixel basis such that A_{ij} is the length of the intersection of the ith ray and the jth pixel.

The measurement model depends on the detector. For hyper-spectral detectors with high energy resolution, it is common to consider a simple discretization of the energy into K separate channels [7, 8], corresponding to K energies $E_1, \ldots, E_K$, i.e. the x-ray intensity associated with the ith ray and the kth channel is given by

$$\Lambda_i(E_k) = \sigma_i(E_k) \exp\left(-\sum_{j=1}^{N} A_{ij}\mu_j(E_k)\right), \quad i = 1, \ldots, M, \quad k = 1, \ldots, K.$$ (2)

The model (2) is an approximation, and it tends to be more accurate with finer spectral discretization.

For photon counting detectors, it is common to assume that the measurements are Poisson distributed with a parameter $\Lambda(E)$. This assumption leads to the measurement model

$$Y_{ik} \sim \mathrm{Poiss}\{\Lambda_i(E_k)\}, \quad i = 1, \ldots, M, \ k = 1, \ldots, K.$$ (3)

If we define $X_{jk} = \mu_j(E_k)$ and $S_{ik} = \sigma_i(E_k)$, we arrive at the discrete linear model

$$B_{ik} = -\ln\left(\frac{Y_{ik}}{S_{ik}}\right) \approx \sum_{j=1}^{N} A_{ij}X_{jk}.$$ (4)

To simplify notation, we now define a matrix $\boldsymbol{X} \in \mathbb{R}^{N \times K}$ with elements X_{jk}, and $\boldsymbol{x} = \mathbf{vec}(\boldsymbol{X}) \in \mathbb{R}^{NK}$ denotes the vectorized image (i.e. $\boldsymbol{x}$ is obtained by stacking the columns of $\boldsymbol{X}$). Moreover, $\boldsymbol{X}_k$ denotes the kth column of $\boldsymbol{X}$. Similarly, $\boldsymbol{Y} \in \mathbb{R}^{M \times K}$ is a matrix with measurements Y_{ik}, $\boldsymbol{B} \in \mathbb{R}^{M \times K}$ is the matrix with elements B_{ik}, and $\boldsymbol{y} = \mathbf{vec}(\boldsymbol{Y})$ and $\boldsymbol{b} = \mathbf{vec}(\boldsymbol{B})$ denote vectorized versions of $\boldsymbol{Y}$ and $\boldsymbol{B}$, respectively.

Using this notation, (4) can be expressed as a system of linear equations

$$\boldsymbol{b} \approx \bar{\boldsymbol{A}}\boldsymbol{x},$$ (5)

where $\bar{\boldsymbol{A}} = \boldsymbol{I}_{K \times K} \otimes \boldsymbol{A}$ and $\boldsymbol{A} \in \mathbb{R}^{M \times N}$ is a sparse projection matrix, $\otimes$ is the Kronecker product, and $\boldsymbol{I}_{K \times K}$ is the identity matrix of order K.

2.2. Objective function in the multi-channel reconstruction

Conventional direct image reconstruction methods are generally not suitable for spectral data recovery as they rely on an oversimplified data model. In contrast, iterative image

reconstruction (IIR) methods can be used which able to account for various noise models, angular under-sampled conditions, and other data irregularities [26]. Furthermore, the regularity of the solutions sought can be imposed through a penalty or *regularization* term [27]. The regularization can abstractly exploit image sparsity or smoothness, or can be formulated in a way that it fits the expectations directly.

We formulate all IIR methods considered in the present paper in a common framework with the generic objective function

$$x^{\star} = \arg\min_{x} \left\{ \mathcal{F}(x) + \beta\,\mathcal{G}(x) \right\}, \tag{6}$$

where $\mathcal{F} : \mathbb{R}^{NK} \to \mathbb{R}_+$ is the data misfit term which can be based on the assumptions of the acquired data noise characteristics or account for more substantial data inconsistencies [26]. The penalty term $\mathcal{G} : \mathbb{R}^{NK} \to \mathbb{R}_+$ defines a regularization penalty with $\beta > 0$ being the (scalar) regularization parameter balancing the trade-off between the two terms.

We now consider the data misfit term. From a Bayesian perspective, the function $\mathcal{F}$ should be the negative log-likelihood function associated with the measurement model (3) [28–30]. Alternatively, by employing a quadratic approximation to the log-likelihood function [31], we arrive at the weighted least-squares objective function

$$\mathcal{F}(x) = \frac{1}{2}\|b - \bar{A}x\|_W^2 = \frac{1}{2}(b - \bar{A}x)^\top W(b - \bar{A}x) \tag{7}$$

where $W = \mathrm{diag}(y)$ is a diagonal matrix with the measurements on its diagonal (3). The data term in equation (7) is sometimes called the *penalized weighted least squares* (PWLS) model and is frequently used in image reconstruction applications [13, 17, 18, 32].

2.3. Existing multi-channel regularization methods

The total variation (TV) penalty has been successful as a first-choice regularizer due to its edge-preserving qualities [33, 34]. In this work we use the discrete (non-smoothed) isotropic TV defined for a (single-channel) image $v \in \mathbb{R}^N$ as:

$$\mathrm{TV}(v) = \sum_{j=1}^{N} \|D_j v\|_2, \tag{8}$$

where D_j is a $2 \times N$ matrix such that $D_j v$ is a finite-difference approximation of the gradient of v at pixel j. In our implementation we use forward differences with Neumann boundary conditions.

In application to the case of spectral CT, we use the *channel-wise* convex TV penalty defined as the sum of the TV norm of each channel:

$$\mathcal{G}_{\mathrm{TV}}(x) = \sum_{k=1}^{K} \mathrm{TV}(X_k). \tag{9}$$

The disadvantage of the channel-wise TV (9) is its inability to employ high SNR channels to aid with the recovery of the problematic channels (e.g. high energy channels). Therefore, we seek an effective and efficient way to correlate channels and consequently provide an additional support to overcome higher levels of noise and a loss of resolution of the problematic channels.

Recently, the TNV penalty which enforces strong correlation between channels has been considered by Rigie *et al* [18]. It consists of using the nuclear norm of the Jacobian matrix

of a multi-channel image, therefore penalizing the singular values of the Jacobian [35–37]. The SVD decomposition of the Jacobian matrix results in eigenvalues and eigenvectors which are encouraged to be aligned favoring a low-rank Jacobian. The limiting case of the rank-one matrix means that all gradient vectors are linearly dependent and thus aligned in the same orientation. Rigie *et al* [18] found that TNV performed well in medical spectral CT application with five energy channels. It is unclear whether the same holds when the number of spectral channels is much higher.

The Jacobian matrix is obtained by applying the finite difference operator at pixel j to all channels simultaneously, i.e. $\boldsymbol{D}_j\boldsymbol{X}$ for $j = 1, \ldots, N$. The TNV penalty is based on the Schatten matrix norm of order $p = 1$, which corresponds to the ℓ_1 norm of the vector that contains the singular values of the matrix [36]. We write the TNV penalty as:

$$\mathcal{G}_{\text{TNV}}(\boldsymbol{x}) = \sum_{j=1}^{N} \|\boldsymbol{D}_j\boldsymbol{X}\|_*, \tag{10}$$

where $\| \cdot \|_*$ denotes the nuclear norm, i.e. the sum of the singular values.

There are other correlative methods which also require TNV to be applied. In a prior rank, intensity and sparsity model (PRISM) [16], the low-rank matrix (penalized by TNV) corresponds to the stationary background of the spectral dimension and the sparse matrix (penalized by TV or ℓ_1 norm) represents dynamic features, e.g. channels closer to the K-edge. The PRISM model, although theoretically appealing, involves the singular value decomposition (SVD) of a large Casorati matrix, the size of which scales with the number of image pixels/voxels and spectral channels. A similar method has been proposed to identify a low-rank background in the unfoldings of the multidimensional tensor [17]. The significant advantage of the Rigie's method [18] over PRISM-like models, is that the Jacobian matrix is substantially smaller than the Casorati matrix and the general problem can be decoupled pixelwise (note the sum in (10)).

3. Structural multi-channel regularization using parallel level sets

3.1. Directional TV regularization with known reference

Another group of correlative techniques is based on the structure-driven diffuson-based variational methods [20–22, 29, 30]. Here we introduce a multi-channel generalization of the parallel level sets approach [19] which has been successfully applied to bi-modal reconstruction [20–22].

One can consider a two-channel case where $\boldsymbol{v}$ is a channel to be regularized and a second reference channel $\boldsymbol{z}$ is either *known* (already estimated and fixed) or unknown (i.e. to be estimated). In an image reconstruction framework, the former approach is usually convex and easier to solve [21, 22], while the latter approach can be non-convex [20] and more complicated. Let us consider the case when the reference $\boldsymbol{z}$ is known.

Since the parallel level sets framework focuses on structural information, $\boldsymbol{v}$ and $\boldsymbol{z}$ channels may have completely different intensities and contrast. The channels, however, are assumed to share the majority of structural information such as edges due to being acquired for the same object. The so-called directional TV regularizer (dTV) is introduced as

$$\text{dTV}(\boldsymbol{v}, \boldsymbol{z}) = \sum_{j=1}^{N} \|\boldsymbol{P}_z\boldsymbol{D}_j\boldsymbol{v}\|_2, \tag{11}$$

where

$$
\boldsymbol{P}_z = \begin{cases} \boldsymbol{I}_{2\times 2} - \dfrac{\boldsymbol{D}_j zz^\top \boldsymbol{D}_j^\top}{z^\top \boldsymbol{D}_j^\top \boldsymbol{D}_j z} & \boldsymbol{D}_j z \neq 0 \\ \boldsymbol{I}_{2\times 2} & \boldsymbol{D}_j z = 0. \end{cases} \tag{12}
$$

The regularization function (11) is the *directional diffusion* of the channel $\boldsymbol{v}$, given a known reference $\boldsymbol{z}$. The measure of a structural similarity of $\boldsymbol{v}$ to $\boldsymbol{z}$ satisfies the following inequality [19]:

$$
0 \leqslant \|\boldsymbol{P}_z \boldsymbol{D}_j \boldsymbol{v}\|_2^2 = \|\boldsymbol{D}_j \boldsymbol{v}\|_2^2 - \left(\frac{\boldsymbol{v}^\top \boldsymbol{D}_j^\top \boldsymbol{D}_j \boldsymbol{z}}{\|\boldsymbol{D}_j \boldsymbol{z}\|_2} \right)^2 \leqslant \|\boldsymbol{D}_j \boldsymbol{v}\|_2^2. \tag{13}
$$

It follows that $\|\boldsymbol{P}_z \boldsymbol{D}_j \boldsymbol{v}\|_2 = \|\boldsymbol{D}_j \boldsymbol{v}\|_2$ when $\boldsymbol{v}^\top \boldsymbol{D}_j^\top \boldsymbol{D}_j \boldsymbol{z} = 0$ (i.e. when the gradient vectors $\boldsymbol{D}_j \boldsymbol{v}$ and $\boldsymbol{D}_j \boldsymbol{z}$ are orthogonal) and when either $\boldsymbol{D}_j \boldsymbol{v} = \boldsymbol{0}$ or $\boldsymbol{D}_j \boldsymbol{z} = \boldsymbol{0}$. Similarly, we have that $\|\boldsymbol{P}_z \boldsymbol{D}_j \boldsymbol{v}\|_2 = 0$ when $\boldsymbol{D}_j \boldsymbol{v}$ and $\boldsymbol{D}_j \boldsymbol{z}$ are parallel.

The natural convex extension of dTV to multi-channel case with K channels will be to minimize the sum: $\sum_{k=1}^K \mathrm{dTV}(\boldsymbol{X}_k, \boldsymbol{Z}_k)$, assuming that all references $\boldsymbol{Z}_k = z_1, \ldots, z_K$ are known *a priori*. Being a sum of convex dTV terms, this is still a convex functional and the reconstruction problem would remain convex. The problem, however, is that the references $\boldsymbol{Z}_k$ are unknown.

3.2. Adaptation to multiple channels with unknown reference

Joint reconstruction methods, which update simultaneously two or more channels, can be a powerful tool enabling high-quality solutions for all channels. These methods use the unknown reference (normally estimated from the previous iteration) while updating the current channel. Following the recent research on joint reconstruction techniques [23, 24], we can generalize our multi-channel reconstruction problem.

Let us assume that for each channel k we have unknown reference channel $z(k)$, which we can estimate during a joint reconstruction process. In this case, one can write the following expression for the multi-channel dTV penalty:

$$
\mathcal{G}_{\mathrm{dTV}}(\boldsymbol{x}) = \sum_{k=1}^K \mathrm{dTV}(\boldsymbol{X}_k, \boldsymbol{z}(k)). \tag{14}
$$

Here we propose an adaptive approach in which the reference channel is selected as a 'good' reconstructed channel from the previous iteration of the chosen iterative algorithm. Specifically, when updating the kth channel during iteration s, the reference is selected according to:

$$
z^{[s]}(k) = \frac{1}{|R^{[s]}(k)|} \sum_{\ell \in R^{[s]}(k)} \boldsymbol{X}_\ell^{[s-1]}, \tag{15}
$$

where $\boldsymbol{X}_\ell^{[s-1]}$ is the ℓth channel from the previous iteration and $R^{[s]}(k)$ is a subset of channel indices over which to average.

The proposed correlative regularizer is named as dTV-p (where 'p' stands for probabilistic). The method takes $R^{[s]}(k)$ to be a set containing a single random sample channel drawn according to the discrete probability distribution described below. In one s-iteration, reconstruction of k-channels are likely to choose different references. Also, across iterations, a single channel is likely to choose different references. In figure 1, we graphically explain the selection procedure which takes into account energy-variant noise characteristics. The
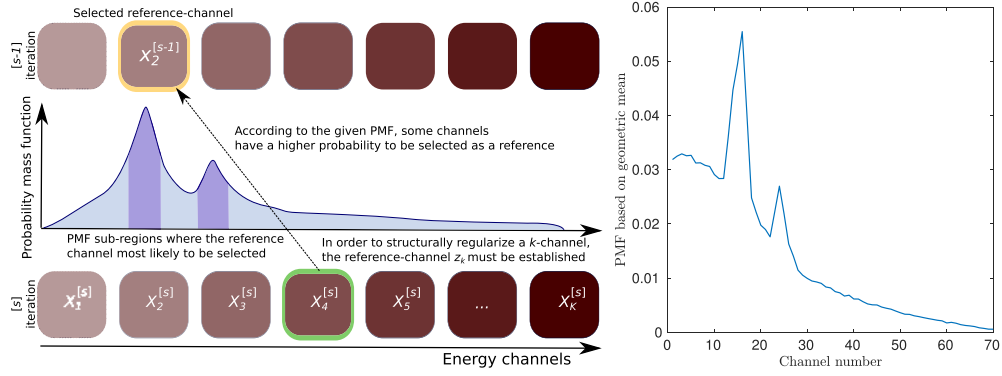
Figure 1. The reference-channel $z_k^{[s]} = X_{k^\star}^{[s-1]}$ is selected by randomly drawing a sample (index $k^\star \in \{1, \ldots, K\}$ of a channel) from the given probability mass function (PMF). The PMF specific to our experiment is depicted in the right figure.

main idea is to update the kth channel by selecting a reference channel based on some global *measure of goodness*. In this paper, we present a measure of goodness which is formulated through a *probability mass function (PMF)*. As a heuristic, we propose to define a PMF based on the channel-wise geometric mean of the estimated SNR. Specifically, it follows from the quadratic model (7) that the SNR associated with ray i and channel k is

$$\frac{[AX_k]_i}{Y_{ik}^{-1/2}} \approx B_{ik} \sqrt{Y_{ik}},$$

hence the channel-wise geometric mean of the approximate SNR is

$$\rho_k = \left(\prod_{i=1}^{M} B_{ik} \sqrt{Y_{ik}} \right)^{1/M}.$$

Based on this, we define a PMF over the K channels such that the probability of selecting channel $k^\star$ is given by

$$\bar{\rho}_{k^\star} = \frac{\rho_{k^\star}}{\sum\limits_{k=1}^{K} \rho_k}. \tag{16}$$

A channel with a high SNR is more likely to be chosen as a reference than a channel with a low SNR.

After the PMF has been obtained for the set of spectral measurements, it remains unchanged throughout the algorithm and is used to select reference images for all channels. Specifically, at s-iteration, new draws of reference channels are initiated from the previous iteration $X^{[s-1]}$. If we denote a reference index by $k^\star$ it means that $k^\star$ depends both on the iteration s and a channel to be reconstructed k, i.e. $k^\star = k^\star(s, k)$.

We write the dTV-p regularizer explicitly as:

$$\mathcal{G}_{\text{dTV-p}}(\boldsymbol{x}^{[s]}) = \sum_{k=1}^{K} \text{dTV}\left(X_k^{[s]}, X_{k^\star(s,k)}^{[s-1]} \right). \tag{17}$$

Note that the regularizer has $\boldsymbol{x}^{[s]}$ argument as it is inherently dependent on the previous iterate.

One can also select a single channel with the highest SNR as a reference for all channels, i.e. the 16th channel in our case (see figure 1 (right)). From our numerical experiments (we omit producing them in the paper, see more discussion in section 6), this deterministic selection does not provide satisfactory results and the reconstruction quality is poor for all channels.

There is a conventional assumption that the adjacent channels may have a similar structure [28] and therefore selecting closer references to k might be a better choice rather than stochastic drawing. In order to test this hypothesis we also introduce the deterministic version of dTV, dTV-d. It consists in taking a mean of several channels (from the previous iteration) in the neighborhood of k:

$$\mathcal{G}_{\text{dTV-d}}(\boldsymbol{x}^{[s]}) = \sum_{k=1}^{K} \text{dTV} \left(\boldsymbol{X}_k^{[s]}, \; \frac{1}{5} \sum_{\ell=k-2}^{k+2} \boldsymbol{X}_\ell^{[s-1]} \right). \tag{18}$$

The number of averaged channels, here chosen to 5, could vary, however, we do not focus here on establishing the optimal number. In fact, we demonstrate later that the dTV-p can provide substantial improvements in terms of reconstruction quality compared to the more obvious deterministic choice of a reference channel used in dTV-d.

We emphasize that while the proposed regularizers are motivated from the general optimization problem (6), they no longer fit strictly within this framework. This is due to the choice of a reference image from the previous algorithm iteration. One possible convex approach which also fits (6), will be to reconstruct projection data mean over all energies as a fixed (known) reference (see more discussion in section 6).

While heuristic in nature, it is a natural and adaptive approach to generalize the dTV framework with a known channel (see section 3.1) to the more challenging case of multiple unknown channels. Our intention is to demonstrate empirically the reconstruction improvements enabled hereby, while leaving further in-depth analysis for future work.

4. Implementation aspects

4.1. Proximal operators framework and the reconstruction algorithm

In order to efficiently solve the general reconstruction problem (6) we will apply the proximal operator methodology to decompose our cost function on simpler problems [38, 39]. The general form we consider allows us to use the FISTA (fast iterative shrinkage-thresholding algorithm) [40] as a general method for solving all problems, only using different proximal operators to implement different regularizations.

For the data fidelity step in FISTA method we use the gradient of the PWLS function in equation (7), which can be written as

$$\nabla \mathcal{F}(\boldsymbol{v}) = \bar{\boldsymbol{A}}^\top \boldsymbol{W} (\bar{\boldsymbol{A}} \boldsymbol{v} - \boldsymbol{b}). \tag{19}$$

With this data misfit, the smallest Lipschitz parameter L ensuring convergence can be found as

$$L = \|\bar{\boldsymbol{A}}^\top \boldsymbol{W} \bar{\boldsymbol{A}}\|_2. \tag{20}$$

We compute this value without forming the matrix product by using the power method.

The proximal operator step in FISTA is given generally as

$$\text{prox}_{\beta/L}[\mathcal{G}](\boldsymbol{v}) = \arg \min_{\boldsymbol{u}} \left\{ \frac{\beta}{L} \mathcal{G}(\boldsymbol{u}) + \frac{1}{2} \|\boldsymbol{u} - \boldsymbol{v}\|_2^2 \right\}. \tag{21}$$

Table 1. Phantom materials, their chemical material proportions and mass densities.

Material No.	Title	Base material	Mass density ($g\,cm^{-3}$)
1 (a)	Quartz	SiO_2	2.65
2 (b)	Pyrite	Fe(46.6%) + S(53.4%)	5.01
3 (c)	Galena	Pb(86.6%) + S(13.4%)	7.6
4 (d)	Gold	Au	19.25

This can be interpreted as a denoising problem with the respective choices of regularization function $\mathcal{G}_{TV}$ of equation (9), $\mathcal{G}_{TNV}$ of equation (10), $\mathcal{G}_{dTV\text{-}p}$ of equation (17) and $\mathcal{G}_{dTV\text{-}d}$ of equation (18), all with regularization parameter β/L.

All inner proximal operator steps are solved using first-order primal-dual algorithms [38, 39]. The TV and dTV-related proximal steps are solved using the fast gradient projection (FGP) method [41], more details can be found in [21, 42]. The TNV proximal step is solved using the primal-dual hybrid gradient (PDHG) method [43], more details on the implementation are given in [37].

The general FISTA algorithm has been implemented in MATLAB with the time-consuming proximal operators to specify TV, TNV and dTV regularizers respectively implemented in OpenMP/CUDA to accelerate computations.

We note that in every evaluation of the dTV proximal operator for channel k in iteration s we select a reference channel as described in the previous section from the reconstructed channels of the *previous* iteration.

While channelwise TV and TNV are convex regularizers for which the described FISTA method converges to the global minimizer, the dTV methods do not necessarily share this behavior due to the adaptive choice of a reference channel.

5. Numerical results

5.1. Synthetic multi-material phantom and data generation process

To test our reconstruction algorithms we design a synthetic phantom consisting of four materials: quartz, pyrite, galena, and gold (see table 1). This particular selection of materials has been inspired by an actual geological study of a rock sample [7].

The concentration of chemical elements in the composite phantom approximates the real sample composition (see figure 2). Specifically, the presence of larger amounts of quartz and pyrite and much smaller amounts of highly attenuating galena and gold (see the mass density values in table 1). The geometrical shapes of materials in the phantom have been chosen to complement properties of the proposed piecewise-constant (TV-based) penalties. To build the phantom in figure 2 we use the `TomoPhantom` software [44], which allows composing various modular 2D–4D phantoms analytically.

Now we explain how multi-spectral tomographic projection data are generated from the phantom in figure 2. To obtain realistic spectral projection data we use the following three software packages: `Spektr` [45], `PhotonAttenuation` [46], and `ASTRA-toolbox` [47]. The `Spektr` package is used to generate an x-ray spectrum $q(E)$ according to the given tube potential ($E = 120$ kVp in our case) with a tungsten anode target [45]. The obtained source spectrum has been normalized $\tilde{q}(E)$ and multiplied by the photon flux chosen as $I_0 = 4 \times 10^4$. See the full 1–120 keV spectrum in figure 3(a). Skipping the lowest and highest energies we
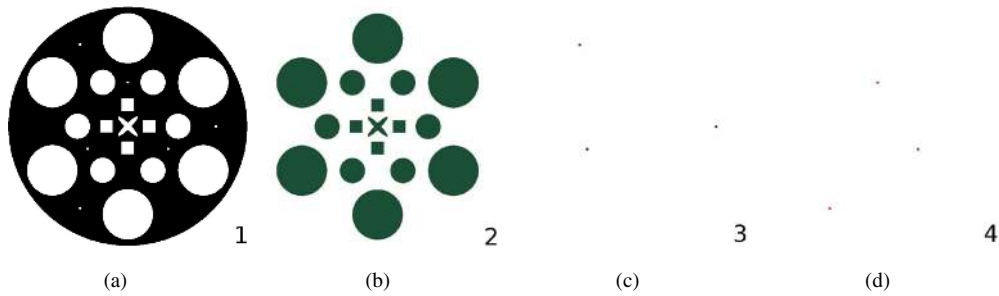
Figure 2. Numerical phantom consisting of 4 materials: Quartz (a), Pyrite (b), Galena (c), and Gold (d). The background is air. The phantom is inspired by a real mineralized ore sample [7]. (a) Quartz (No.1). (b) Pyrite (2). (c) Galena (3). (d) Gold (4).
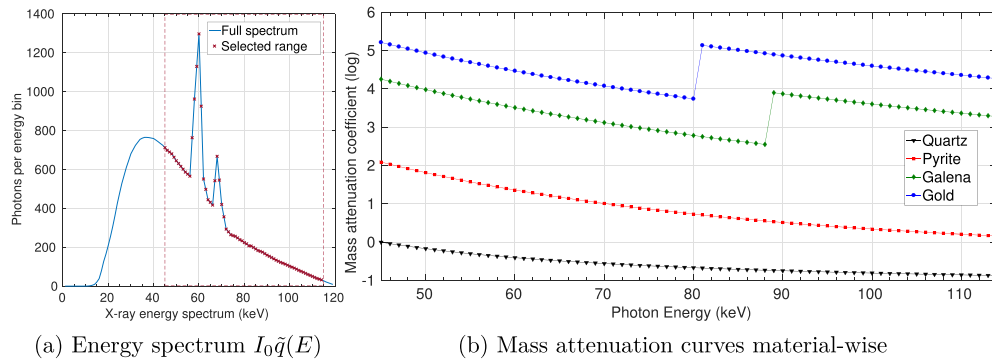

(a) Energy spectrum $I_0\tilde{q}(E)$ (b) Mass attenuation curves material-wise

Figure 3. (a) The full x-ray energy spectrum and the selected range (45–114 keV) of energies used for data modeling. The selected spectrum is discretized into 70 energy bins; (b) A plot of mass attenuation coefficients (MACs) for each material in table 1 for the selected photon energies. The step changes in MACs correspond to the K-edges of gold (80.725 keV) and galena (88.005 keV), respectively. (a) Energy spectrum $I_0\tilde{q}(E)$. (b) Mass attenuation curves material-wise.

select the energy range of 45–114 keV which corresponds to the real multi-spectral scanner settings [7]. We discretize the spectrum $I_0\tilde{q}(E)$ into 70 energy bins (see the selected region in figure 3(a)).

In figure 3(b), we present a plot of mass attenuation coefficients (MACs) for each of the investigated materials. Since we have 70 energy bins in the energy range of 45–114 keV the step equals 1 keV. Two distinct K-edges are visible for gold (80.725 keV) and galena (88.005 keV) materials. Notably, the K-edges are less than 8 keV apart, so cruder energy binning can lead to the loss of spectral resolution and an ambiguity in the identification of unique K-edges. Therefore for some cases, fine energy discretization is crucial to ensure a correct material classification based on the K-edges [7].

The `PhotonAttenuation` software is used to model the photon attenuation process through the designed phantom. With a unique energetic signature for a chosen chemical element, one can obtain the mass attenuation coefficients specific to each considered material (see figure 3(b)).

In order to generate multi-spectral projection data avoiding the 'inverse crime' [48], we use the `ASTRA-toolbox` package [47]. We set the fan-beam scanning geometry, which replicates the cone-beam geometry used in [7] for the central slice of 3D volume. The width of the
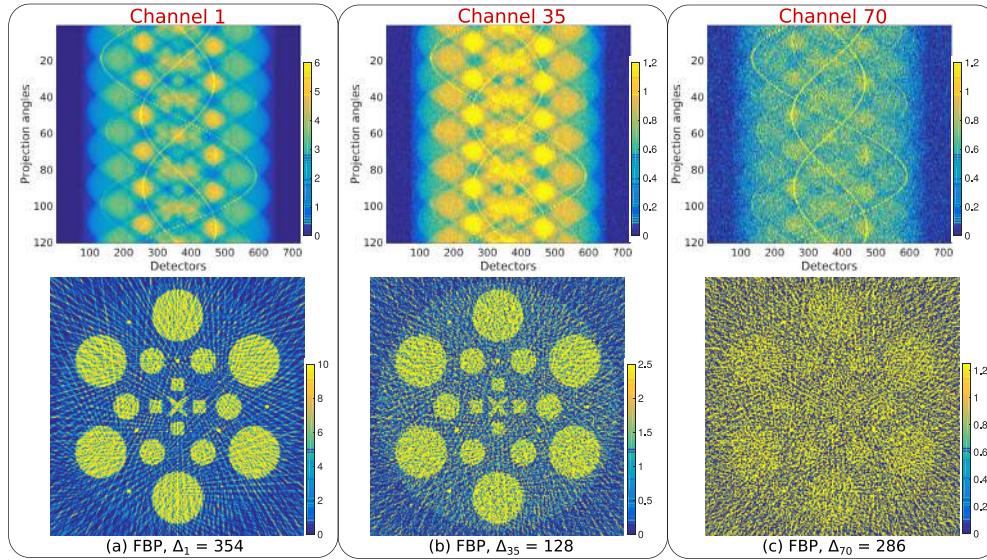
Figure 4. Upper row: three selected sinograms with 120 projection angles and 724 detectors, (a) channel no. 1 (45 keV), (b) no. 35 (79 keV) and (c) no. 70 (114 keV). Note different attenuation characteristics, noise levels, and the grayscale levels. The bottom row shows corresponding FBP reconstructions. (a) FBP, $\Delta_1 = 354$. (b) FBP, $\Delta_{35} = 128$. (c) FBP, $\Delta_{70} = 286$.

reconstruction domain is set to be 1 cm, the distance from the source to the rotation center is 3 cm, the distance from the source to the detector is 5 cm, and the width of the detector array is 2 cm. The pixel size of the reconstructed domain is 512×512 pixels (the data have been simulated on a larger grid of 1024×1024 pixels), the number of projection angles is set to 120. Poisson distributed noise has been added to the projection data according to the energy-variant spectrum $I_0 \tilde{q}(E)$.

In figure 4 (upper row) we demonstrate sinograms for three energy channels which are taken from the discretized spectrum in figure 3(b). Spanning the full spectral range, they reflect differences in the grayscale values, noise levels, and also the pronunciation of features with respect to different energies. In the bottom row we demonstrate the FBP reconstructions of the selected sinograms. Note that for lower energies (e.g. channel no. 1), the beam has been absorbed substantially in quartz. The FBP reconstruction is noisy and shows many streak artifacts due to poor angular sampling. For mid-range energies, the quartz and pyrite materials (see figure 2) are strongly pronounced in the sinogram, while highly attenuated materials are less visible. The FBP reconstruction of the 35th channel in figure 4(b) is substantially noisier (the error is lower because intensity is overall lower for the channel). For higher energies, gold and galena traces in the sinogram are visible in figure 4(c), however quartz and pyrite are obscured in higher levels of noise.

5.2. Quantitative reconstruction quality assessment

To *globally* quantify the quality of obtained reconstructions we use the root mean square error (RMSE) Δ_Σ averaged over all channels given as:

$$\Delta_\Sigma(\boldsymbol{x}, \hat{\boldsymbol{x}}) = \frac{1}{K} \sum_{k=1}^{K} \Delta_k \left(\boldsymbol{X}_k, \hat{\boldsymbol{X}}_k \right), \tag{22}$$

where for (single-channel) images $\boldsymbol{v}, \hat{\boldsymbol{v}} \in \mathbb{R}^N$ we have the channel-wise RMSE:

$$\Delta_k(\boldsymbol{v}, \hat{\boldsymbol{v}}) = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (v_i - \hat{v}_i)^2} \times 100, \tag{23}$$

where $\hat{\boldsymbol{v}}$, $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{X}}_k$ refer to exact images and we compute the error over the entire image. The global measure Δ_Σ conveniently simplifies some optimization tasks and represents the general measure of reconstruction quality over all channels.

The second quality metric is a channel-wise mean structural similarity index (MSSIM) [49] which is defined as:

$$\text{MSSIM}(\boldsymbol{v}, \hat{\boldsymbol{v}}) = \frac{1}{N} \sum_{i=1}^{N} \frac{(2\mu(v_i)\mu(\hat{v}_i) + C_1) + (2\sigma(v_i\hat{v}_i) + C_2)}{(\mu^2(v_i) + \mu^2(\hat{v}_i) + C_1)(\sigma^2(v_i) + \sigma^2(\hat{v}_i) + C_2)}, \tag{24}$$

where μ and σ are the mean intensity and the standard deviation of a 8×8 patch centered at the ith pixel. The variable $\sigma(v\hat{v})$ denotes the cross-correlation and $C_{1,2}$ are small constants to avoid division by zero. The MSSIM metric is a different quality metric from RMSE (23) as it considers image degradation as a visually perceived change in structural information. The MSSIM metric fits well to our problem where problematic energy channels can have significant structural degradation. The MSSIM value equals 1 if images are identical and -1 if they anti-correlated, i.e. $\text{MSSIM}(\boldsymbol{v}, \hat{\boldsymbol{v}}) \in [-1, 1]$.

5.3. Selection of the optimal regularization parameters

The proper choice of the regularization parameter in (6) is very important for the successful reconstruction of all energy channels with the uniform spatial resolution. It is known that noise variance can vary substantially across energy space (see figure 4). To the authors' knowledge, two approaches currently exist to deal with the issue.

The first approach requires data variances σ_k^2 of all energy channels so a 'noise-balancing' procedure can be implemented on sinograms as $\boldsymbol{b_k}/\sigma_k^2$, see [18]. After data have been noise-balanced, one can apply a constant regularization parameter to reconstruct all channels. The disadvantage of this technique is that it requires a change of data space and therefore modification of the original reconstruction problem. Additional difficulty lies in the estimation of variances, which for a larger number of channels can be nontrivial and time-consuming to obtain.

An alternative approach of dealing with spectrally-varying variances and signal amplitudes will be to introduce energy-variant regularization parameters [13]. This approach does not modify the original reconstruction problem but can result in difficulties of estimating *spectral equalization multipliers* (SEM) for a multi-channel dataset. Additional difficulty of using spectrally-variant regularization parameters lies in estimating them for channel-correlated priors such as dTV and TNV. In order to simplify our presentation in this paper we use a scalar regularization parameter (energy-invariant) for all channels.

Here we present the strategy we used to choose optimal regularization parameters for the methods. Since our multi-channel optimization problem is high dimensional for correlative reconstruction methods, we use the global Δ_Σ error (22) to assess how well the whole sequence of channels is reconstructed. In order to establish regularization parameters for all methods we perform the following tasks. We use 30 values of regularization parameters to
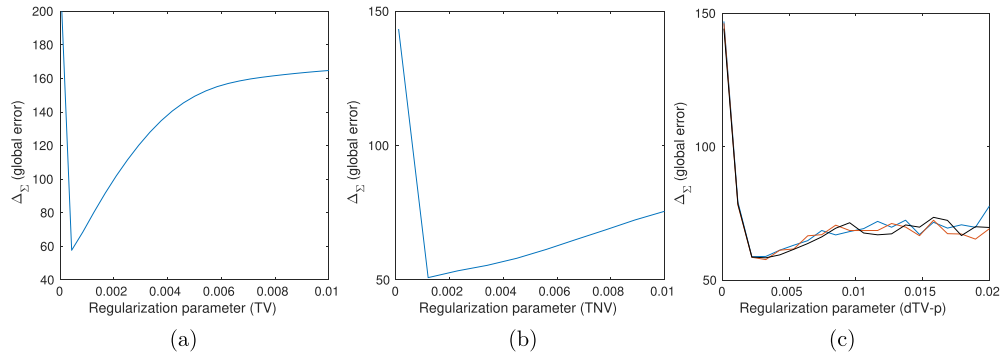
Figure 5. Optimally selected regularization parameters for TV, TNV, and dTV-p (three realizations) penalties. Using the global Δ_Σ error, the optimal regularization parameters can be established. (a) TV, $\beta_{\mathrm{opt}} = 4.2 \times 10^{-4}$. (b) TNV, $\beta_{\mathrm{opt}} = 1.2 \times 10^{-3}$. (c) dTV-p, $\beta_{\mathrm{opt}} = 3.2 \times 10^{-3}$.
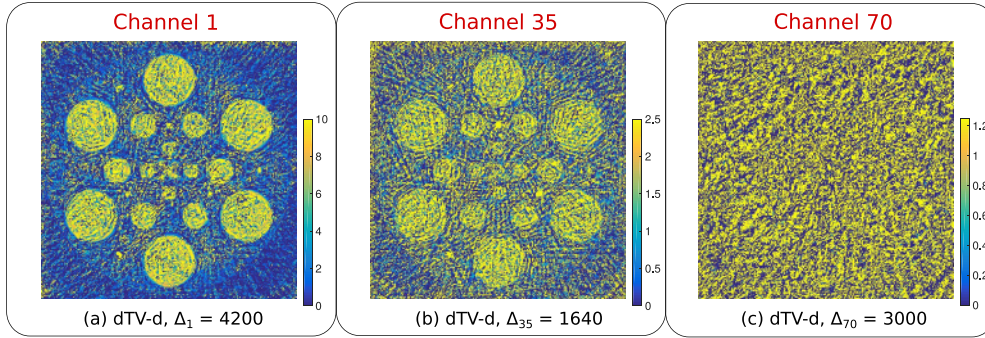

Figure 6. Reconstruction using the dTV-d method for β_{opt}. (a) dTV-d, $\Delta_1 = 4200$. (b) dTV-d, $\Delta_{35} = 1640$. (c) dTV-d, $\Delta_{70} = 3000$.

select the optimal one β_{opt} for which the global error is minimal. Note that the global error is calculated for the final outer iteration of the FISTA algorithm.

In figure 5 we demonstrate how regularization parameters for all methods have been selected. Notably a well-defined global minimum is present for TV and TNV methods. The estimation of the optimal β_{opt} for the dTV-p method is more complicated since it includes a probabilistic component (see section 3). To minimize possible deviations due to the random selection of a reference channel in iterations, we perform three optimization instances for each unique value of β. In figure 5(c) one can identify the general trend of three realizations for the dTV-p method. Based on this the optimal parameter can be established. Now using optimal parameters one can reconstruct the whole sequence of channels and calculate Δ_k errors (23) and (24) with respect to each channel. This provides more detailed information how well each channel is reconstructed.

5.4. Reconstruction with optimally selected regularization parameters

Before we proceed with comparisons with optimally selected regularization parameters we will demonstrate that the deterministic dTV-d method (18) fails to reconstruct variance-unbalanced datasets (see figure 6). Such a poor recovery can be explained by using consistently very
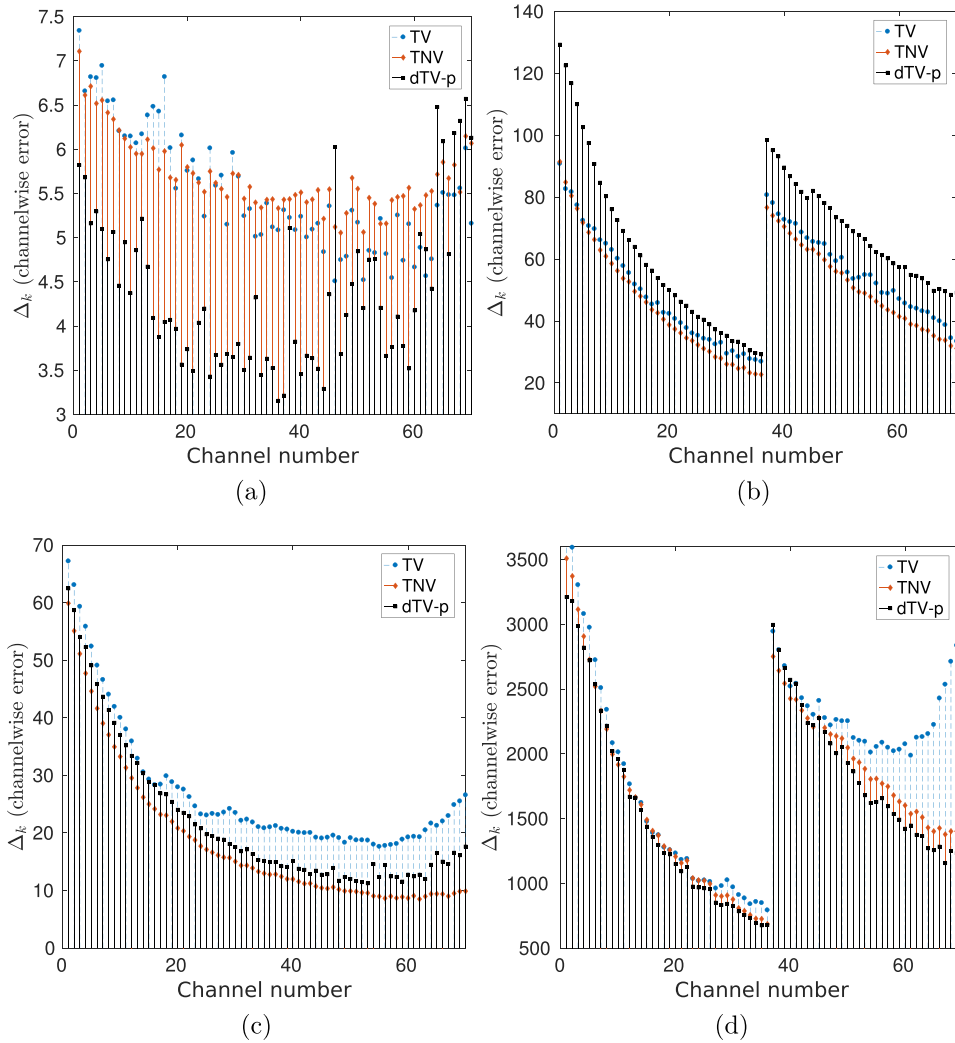
Figure 7. Reconstruction errors for 70 channels for TV, TNV and dTV-p methods with respect to different ROIs. (a) Background ROI. (b) Quartz ROI. (c) Pyrite ROI. (d) Gold and galena ROI.

low SNR references (e.g. from high energy channels) throughout the reconstruction process. Therefore the error from low SNR channels accumulates, amplifies and propagates into high SNR channels as it can be seen in (a) and (b) in figure 6.

In figure 7 we show channel-wise errors for TV, TNV, and dTV-p methods with respect to different ROIs: background, quartz, pyrite, and galena and gold (see figure 2). In figure 8 we also produce channel-wise MSSIM values (24) and in figure 9–11 we show reconstructions, image errors (squared difference between the reconstructed image and the exact phantom) and binary segmentations for the TV, TNV, and dTV-p methods, respectively. In this work we used binary thresholding for which the optimal threshold is manually selected channel-wise. The same thresholds used consistently for each reconstruction method which ensures fair comparison. Also in figure 12 we show zoomed image errors for TNV and dTV-p methods for 70th channel. In the text below we give a general overview of results while refer to different figures.
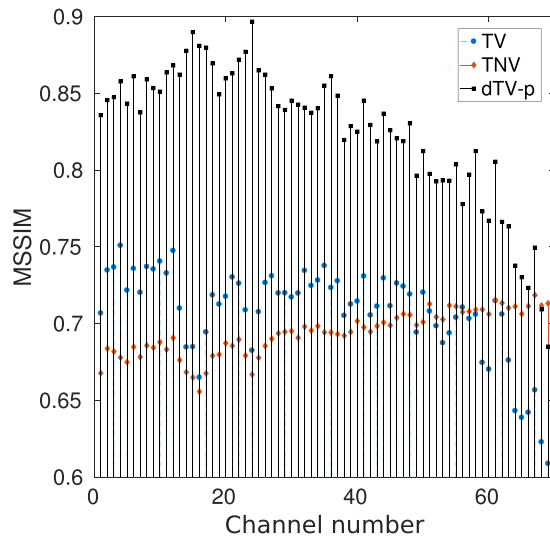
Figure 8. MSSIM values for 70 channels for TV, TNV and dTV-p methods.

In figure 7(a) one can see that for ROI-background the dTV-p method generally outperforms other methods. Indeed, it is noted in the reconstructed images (see figure 9) that dTV-p reconstructs inner areas of objects better than other methods. Figure 7(b), however, shows the highest RMS errors for the dTV-p method and a small advantage of TNV over TV. We believe that high errors in dTV-p recovery are due to outliers of high intensity which exist on the boundaries of the reconstructed objects (see figure 12). These high-intensity localized pixels can contribute significantly to the overall error while visually may not be so obvious (see figure 11). For ROI-pyrite (see figure 7(c)), the dTV-p method performs only slightly worse than TNV. Meanwhile, for galena and gold ROI (see figure 7(d)), the dTV-p method outperforms (especially for higher energy channels) all other methods, including TNV.

Additionally in figures 7(b) and (d) one can notice a jump in quartz and gold and galena channel-wise RMSE at channel 37 (81 keV). We ascribe this to the K-edge of gold at 80.725 keV after which attenuation is strongly increased. This reduces the SNR of the measured data substantially, and consequently all methods observe a drop in the performance exactly at this point. Furthermore, on close inspection a second jump at channel 45 (89 keV) can be seen for some methods. This agrees with the lead (galena) K-edge at 88.005 keV. It also makes sense that it is quartz and gold and galena errors that show this K-edge dependence, since the gold and galena particles are embedded in the quartz phase, whereas the background and pyrite phases are unaffected.

We also believe that in our case the RMSE does not reflect well the visual perception of the reconstructed images. For instance one can focus on pyrite reconstruction in 70th channel using TV (see figure 9) and compare it to dTV-p (see figure 11). Most of central objects are totally missing (smoothed) from the TV-reconstructed image giving the pronounced image error and poor segmentation. The dTV-p image, however, is reconstructed relatively well when all pyrite features are clearly visible. Still, the RMSE values in figure 7(c) infer the opposite.

Therefore after slightly ambiguous results based on channel-wise RMSE, we are interested to obtain more global measure of structural similarity provided by MSSIM (24). In figure 8 we present MSSIM values for three methods in comparison. We see that the dTV-p outperforms other methods substantially (only last two channels are slightly worse than with TNV).
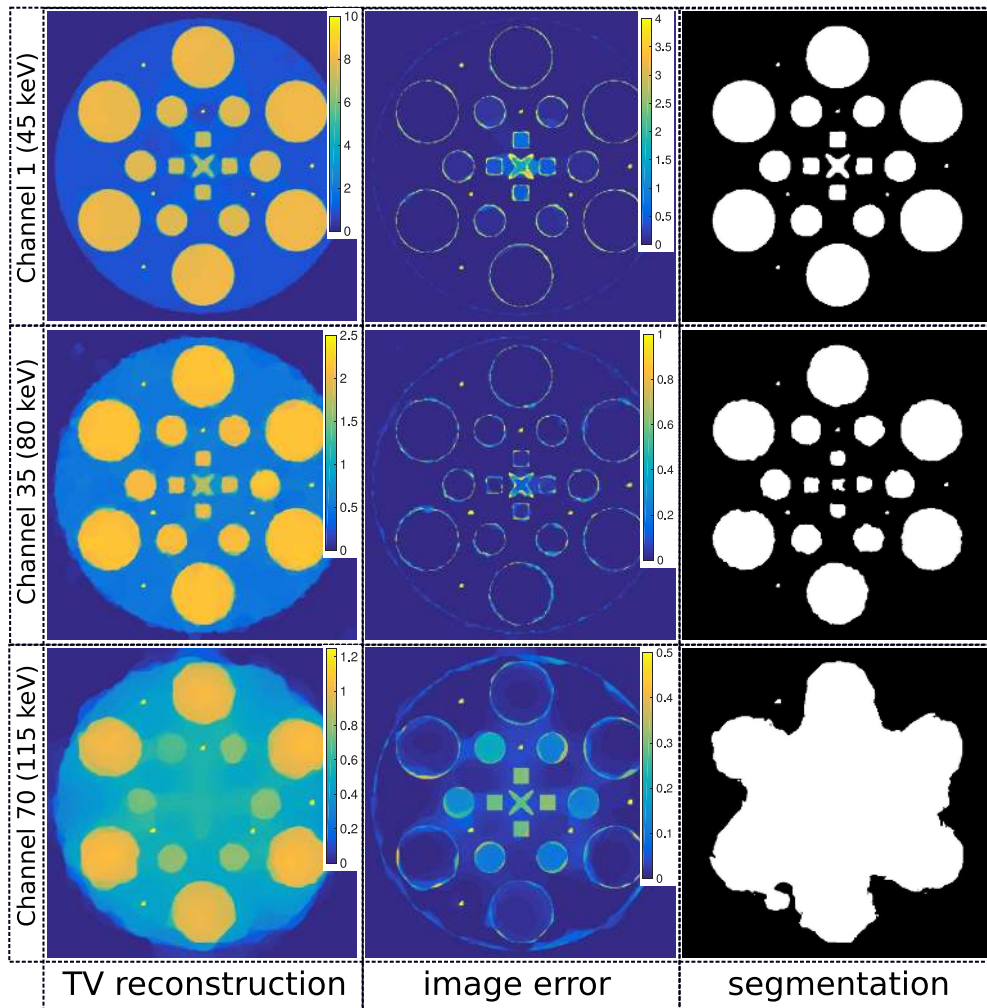
Figure 9. Reconstruction, image error and segmentation for TV method with β_{opt}.

Notably, TV and dTV-p have a similar trend, i.e. when SNR reduces the MSSIM values also diminish. Interestingly, until the 50th channel channel-wise TV slightly outperforms TNV, but after that the method quickly deteriorates.

In figure 10 one can see the result of reconstruction and segmentation for the TNV method. One can notice the significant improvement in terms of better spatial resolution and less noise in the reconstructed images using the correlative method. The image errors appears to be lower and segmentation results are better than with channel-wise TV (see figure 9). Note, however, the quality of segmentation of the 70th channel. It is clearly better than with TV, but misclassified areas are still abundant.

In figure 11 we show results with the proposed dTV-p method. One can see that the quality of reconstructions is high and image errors are visually low. Notably the segmentation of the 70th channel is much better with dTV-p than with TNV. Nevertheless, to understand the reason of higher RMSE values for dTV-p we show image errors with magnification in figure 12. One can see that with dTV-p (c) some erroneous pixels (or small clusters of pixels) on
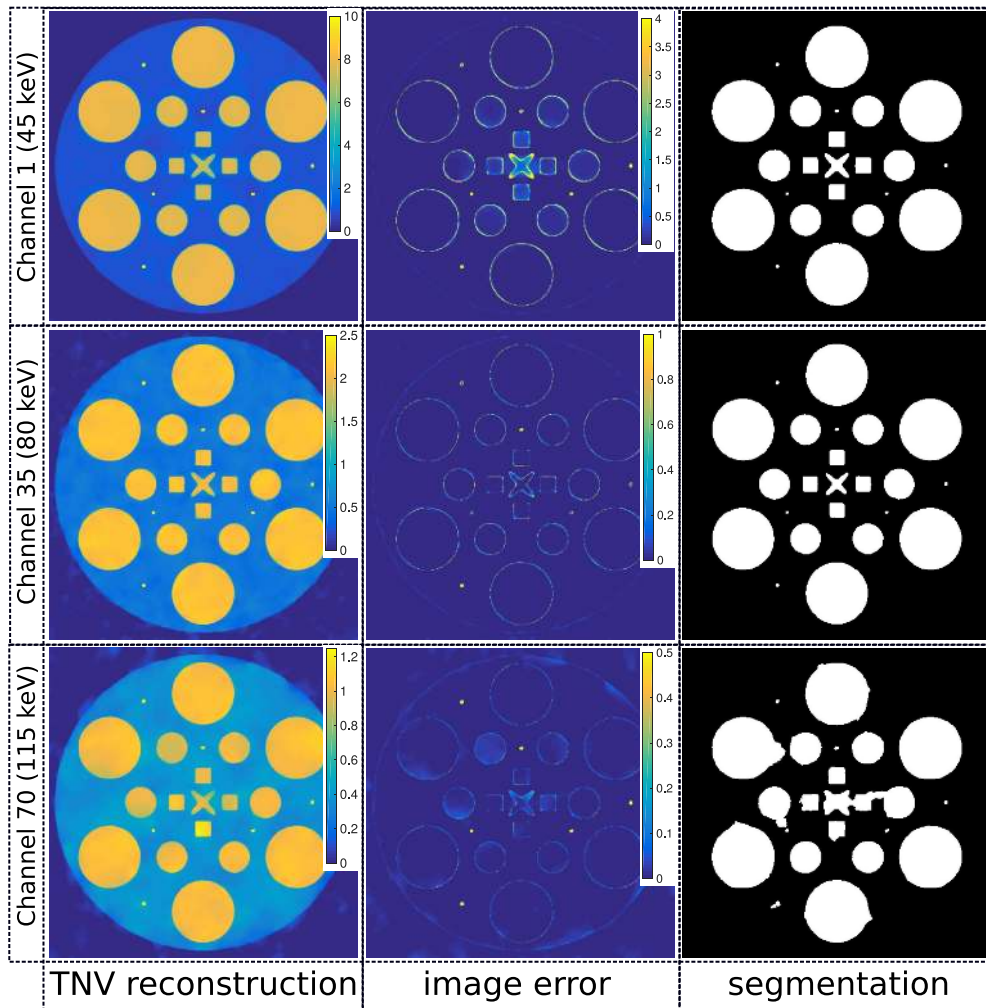
Figure 10. Reconstruction, image error and segmentation for TNV method with β_{opt}.

boundaries of objects are more abundant than with TNV reconstruction and also more intense. The presence of those perturbations is what results in high RMSE values.

The presence of outliers can be due to insufficient resistibility of the dTV-p method to handle errors in the selected reference channel. If the given reference has a low SNR or some artifacts, there is a fair chance that for some pixel elements, the direction of dTV smoothing will be incorrect. As with dTV-d case, this can affect the reconstruction significantly (see figure 6). Substantially more reliable behavior of dTV-p can be explained by favoring the selection of high-SNR references. More detailed insight into the problem is given in section 6.

To further confirm the competitiveness of the proposed method, in figure 13 we present MAC plots for four materials reconstructed with the FBP, TV, TNV, and dTV-p methods. The plots show MAC curves for one randomly selected pixel per material plotted across the whole energy domain. As expected, in the FBP curves one can see major perturbations for quartz and pyrite and somewhat minor for galena and gold. The channel-wise TV method significantly reduces noise perturbations in quartz and pyrite, however strong deviation to lower
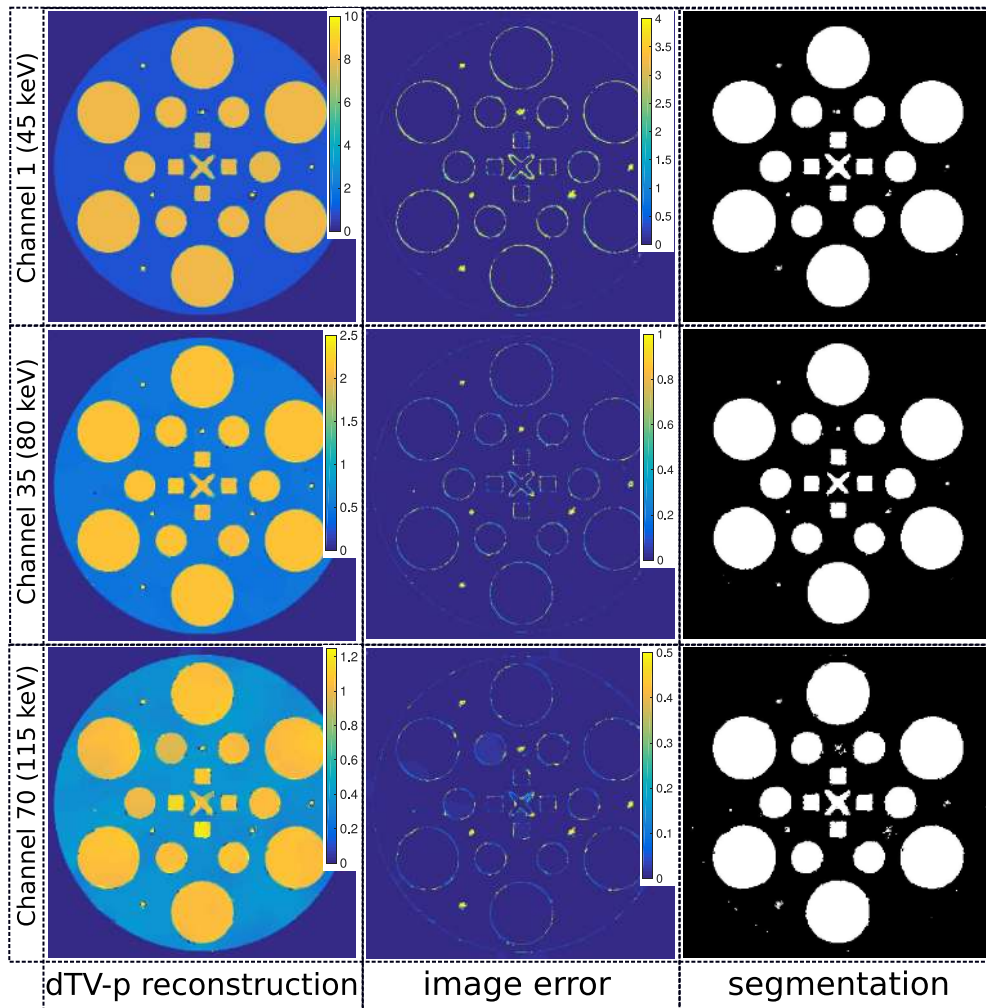
Figure 11. Reconstruction, image error and segmentation for dTV-p method with β_{opt}.

MAC values is visible for gold and galena right after K-edge. The correlative priors provide substantial improvement for all materials. The TNV method, however has more perturbations for quartz than the dTV-p method. Again, this confirms our hypothesis that only due to selective erroneous pixels the resulting error can be higher for the dTV-p method over TNV. The K-edge curves (galena and gold) are well recovered with both correlative priors.

5.5. Assessing the robustness of the dTV-p method to the selection of the reference

Because the proposed dTV-p method is probabilistic, one needs to establish how it responds to different realizations (runs) with a fixed set of parameters. In figure 14 we demonstrate the distribution of errors for 10 realizations (complete reconstructions of all channels). The reconstructions are performed for the same value of the optimal regularization parameter and errors are given for the final iteration.
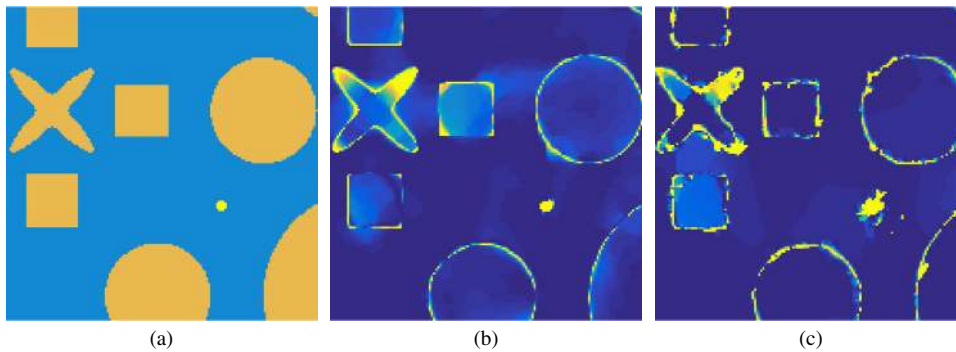
(a) (b) (c)

Figure 12. The magnified error images for TNV (b) and dTV-p (c) methods. The main source of errors for the dTV-p method is the clustered pixels-outliers on the edges of the objects. Note that the inner regions of largest objects are recovered better with dTV-p. (a) Zoomed phantom. (b) TNV image error. (c) dTV-p image error.

From figure 14 one can see that the 'spread' of errors is very small and the number of outliers is quite insignificant. Notably the outliers lie very close to the median values. The average difference between the maximum and the minimum Δ_k error per channel is less than 2. Compare to the general spread of errors, this is a small deviation indeed. This result confirms that the dTV-p method is globally robust to random selection of a reference channel and does not deviate substantially.

6. Discussion

In this paper, we presented a one possible approach to select a PMF, based on which the reference channel is selected for the dTV-p method. As we have demonstrated through numerical experiments, the correct choice of a PMF is crucial to obtain quantifiable results. One can choose a unique PMF based on information within the *k*th channel. In this case, the PMF is not fixed as it is currently implemented, but may be variable to further improve the quality of reconstructions. Nevertheless, this approach can introduce an additional level of uncertainty how the algorithm will converge globally. However, it is exactly the stochastic selection of the reference channel ensures the stable behavior of the method. The deterministic dTV-d method fails to reconstruct because it uses neighboring channels, which for some energies can be severely distorted by noise. Therefore the dTV-d method propagates errors into high SNR channels and amplifies bias in the subsequent iterations.

Related to the choice of suitable PMF there is also the question of how energy variant multipliers or spectral equalization multipliers (SEM) should be selected. In this paper we use a constant SEM, however since the variance and amplitude vary substantially across the energy space one should account for that. In future work, we will design and test a suitable SEM which will help to distribute the regularization strength according to some rule.

In section 3.2, another deterministic method was mentioned. It is based on selecting the highest SNR reference for all channels. Although we do not provide results in the paper we tested this approach and the obtained reconstruction quality was poor. This method has shown a similar performance to the dTV-d method. This further confirms the hypothesis that the stochastic component is highly important to obtain higher quality of the reconstruction.

Furthermore, in section 3.2 we briefly discussed a possibility of generating a reference by reconstructing of projection data mean over the spectral dimension. This approach can
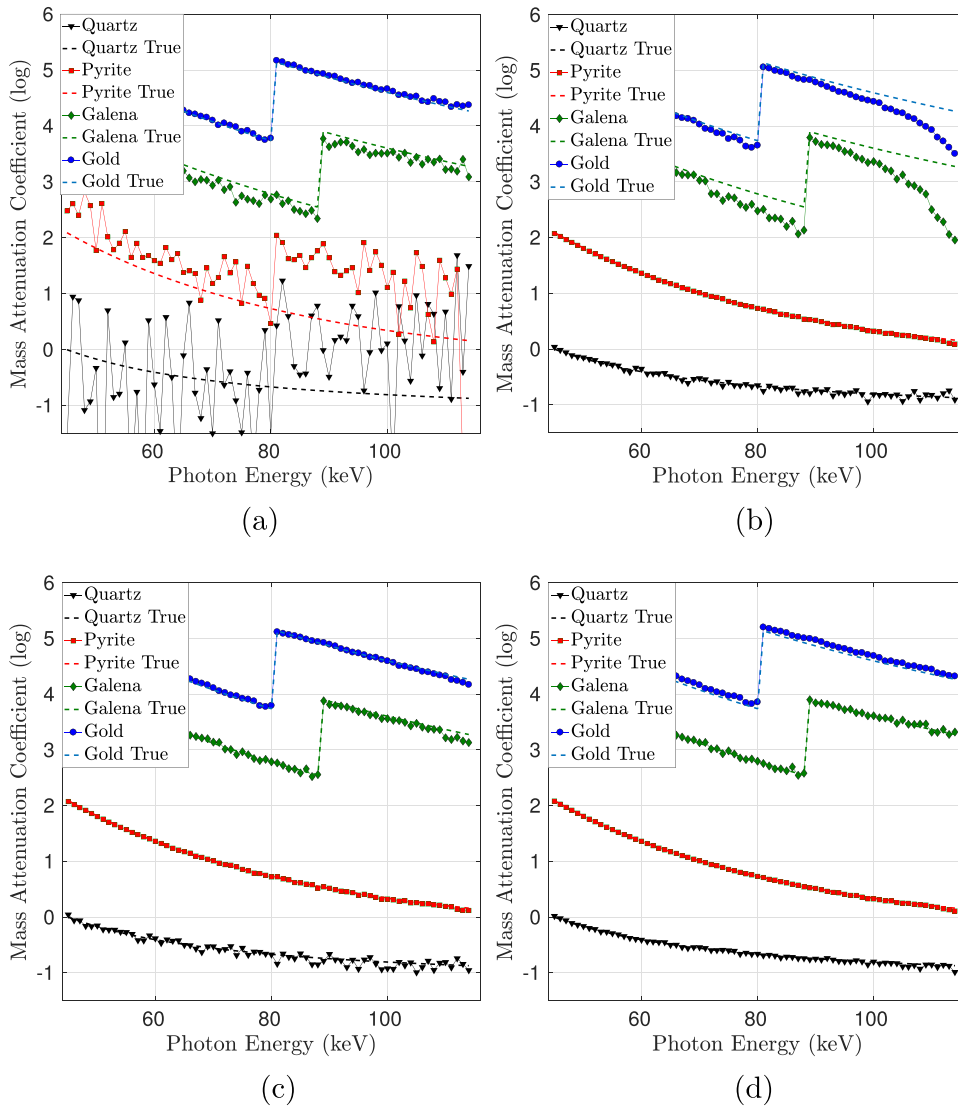
Figure 13. MACs plots for FBP, TV, TNV, and dTV-p reconstruction methods. The dTV-p method produces the best fit for all materials. (a) FBP. (b) TV. (c) TNV. (d) dTV-p.

potentially improve SNR of the reconstructed image by increasing the total number of photon counts. In this case the information about the photon-energy will be lost affecting the linearity of Beer's law. Also when the energy resolution is fine (e.g. hundreds of energy-channels available) and the spectrum range is broad, this can lead to severe beam-hardening artifacts and possibly the disappearance of some subtle structures through linear averaging. We have also preliminarily tested this approach using synthetic data and the results were promising and close to the quality of the proposed dTV-p method. However, before drawing conclusions how robust this approach is, one needs to apply this method to real multi-channel data and experiment with different energy-discretization scales and energy ranges. This will be in the scope of future research.
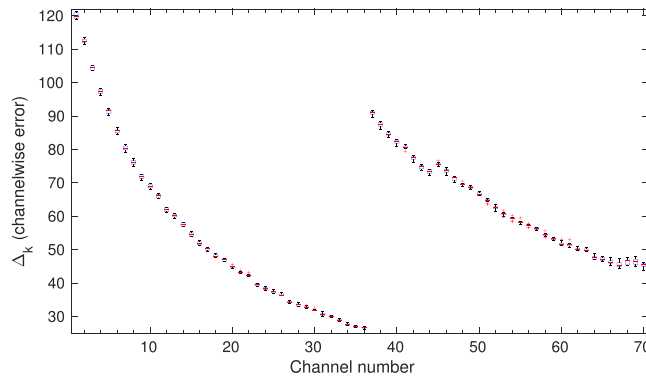
Figure 14. The distribution of errors for the dTV-p method based on 10 realizations (complete reconstructions of all channels) with optimally selected regularization parameter $\beta_{\text{opt}} = 3.2 \times 10^{-3}$. The central mark indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers, and the outliers are plotted individually using the '+' symbol.

In section 5.4 we mentioned that the dTV-p method can be sensitive to noise and locally produce slightly inferior reconstructions to TNV. Unfortunately, it is a common problem of gradient driven local smoothing methods which exploit unregularized gradient field [29, 30, 36]. One needs to ensure more robust to noise calculation of the gradient. This can be done, for example, through direct smoothing of the TV-norm as in [19, 33] or using the non-local gradient methods. Potentially, the choice of the threshold for the TV-norm smoothing can depend on the data variance. We will investigate this problem in future.

Another important issue is the separability (with respect to energies) of the proposed regularizer and a potential to improve global convergence of the algorithm. In this work we use the smallest Lipschitz parameter *L* (20) which ensures global convergence of the algorithm solving the generic optimization problem (6). This partially has been done to overcome the non-separability of the TNV penalty. However, both TV and dTV are separable and therefore the optimal *L* can be calculated channel-wise. This can potentially accelerate the convergence of the algorithm. Furthermore, the dTV-p regularization step can be realized in parallel, which makes it suitable for the implementation on CPU and GPU computing clusters.

Regarding the computation times, the proposed method has almost equivalent to TV computing efficiency and memory footprint. Compare to TNV, the dTV-p method generally requires less than 2–3 times number of inner iterations (see appendix) to reach an accurate solution. The dTV-p method is also faster in the current CPU implementation, therefore for the same amount of iterations it updates all channels approximately two times faster then TNV. However, there is a space to accelerate TNV implementation as well.

7. Conclusions

In this paper we presented a novel iterative reconstruction algorithm for multi-channel x-ray computed tomography. We showed that the structural correlation embedded into the regularization term can result in significant improvements in image quality of the reconstructed images. Furthermore, we make an important observation that the information collected from the adjacent channels is not always reliable for multi-channel problems. We proposed and implemented a probabilistic technique to select a reference channel which is based on the knowledge about the signal-to-noise ratio of the flux. This prior was demonstrated to be a

strong competitor to state-of-the-art correlative regularizers, such as total nuclear variation. The method is computationally efficient, simple to implement, and can be easily parallelized.

Acknowledgments

Appendix

A.1. Algorithm verification and selection of number of algorithm iterations

To ensure correctness of our implementations we verify the computed reconstructions against reference solutions computed using the established and highly reliable CVX modeling system for convex optimization [50, 51]. Our motivation for this verification step is firstly to ensure we assess the methods considered in terms of the specified optimization solution and not an inaccurate algorithm-dependent early approximation which may substantially skew comparative results. Secondly, to aid in determining a sufficient yet practical number of inner and outer iterations to run.

To test single channel TV we define a 64×64 pixel Shepp–Logan test image and generate a parallel-beam test problem with 36 projections over $180°$ using the `paralleltomo` function of the AIR Tools MATLAB package [52, 53]. With the regularization parameter $\beta = 2$ and configured to its best precision setting, CVX produces the highly accurate reconstruction in figure A1, top left, with the resulting optimal objective value of 613.648 756.

We now run our FISTA implementation for the same problem using all combinations of 50, 250, 2000 outer and 10, 50, 250 inner iterations. In the bottom left of figure A1 we plot the RMSE (equation (23)) with respective to the CVX reference solution (and not the original image) for each choice of number of inner iterations as function of number of outer iterations. The bottom right plot shows the relative error of the final objective value with respect to the CVX reference objective value plotted in the same way. We conclude that our FISTA implementation for TV-regularized reconstruction works since both quantities approach zero when the number of inner and outer iterations are increased. The three images to the top right are the FISTA reconstructions for 50/10, 250/50 and 2000/250 outer/inner iterations with the CVX reference solution subtracted and displayed in a very narrow grayscale to highlight differences. Each image title states the final FISTA objective value and RMSE, compare to CVX. Based on the difference images and plot we choose to use 250 outer and 50 inner iterations in the main numerical experiments as a compromise between achieving an accurate solution and reasonable computing time.

Similarly, for TNV we use a small test problem but with 3 channels instead of just one. It was established that our implementation could accurately reproduce the CVX TNV solution, however a larger number of inner iterations is required. For this reason in the reconstruction experiments we use 250 outer and 200 inner iterations for TNV.
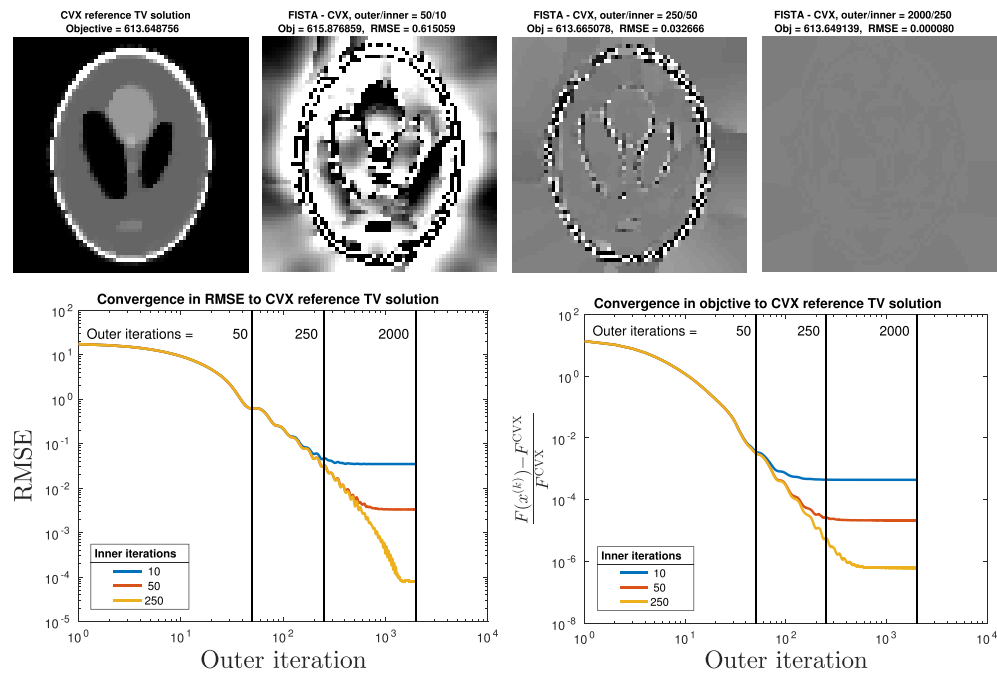
Figure A1. Top left: reference TV reconstruction obtained by CVX software, final objective value given in title. Grayscale range: $[0, 0.5]$. Top middle-right: difference images of FISTA reconstructions and CVX reconstruction. Number of outer and inner FISTA iterations, objective function value and RMSE are given in each title. All difference images shown in same narrow color range $[-5 \times 10^{-4}, 5 \times 10^{-4}]$ to highlight differences. Bottom left: RMSE as function of outer FISTA iterations for three choices of inner iteration numbers, verifying that the FISTA reconstruction converge to the CVX reference solution. Bottom right: same for relative error of FISTA objective value, verifying that the FISTA reconstructions converge in objective.

While the dTV problem is not convex and hence does not admit a CVX implementation, we tested the dTV implementation on a 2-channel case with the reference channel fixed as a blank image. In this case the dTV regularizer reduces to the standard TV and we confirmed numerically that the dTV implementation in this case reproduced the CVX TV solution. Therefore for both TV and dTV methods in this paper we use 250 outer and 50 inner iterations.

ORCID iDs

Daniil Kazantsev ⓘ https://orcid.org/0000-0003-4529-5837
Jakob S Jørgensen ⓘ https://orcid.org/0000-0001-9114-754X
Martin S Andersen ⓘ https://orcid.org/0000-0002-4654-3946

References

[1] Buzug T M 2008 *Computed Tomography: from Photon Statistics to Modern Cone-Beam CT* (Berlin: Springer)

[2] Herman G T 2009 *Fundamentals of Computerized Tomography: Image Reconstruction from Projections* (London: Springer)

[3] Davis G, Jain N and Elliott J 2008 A modelling approach to beam hardening correction *Proc. SPIE* **7078** 70781E

[4] Blumensath T and Boardman R 2015 Non-convexly constrained image reconstruction from nonlinear tomographic x-ray measurements *Phil. Trans. R. Soc.* A **373** 20140393

[5] Van de Casteele E, Van Dyck D, Sijbers J and Raman E 2004 A model-based correction method for beam hardening artefacts in x-ray microtomography *J. X-Ray Sci. Technol.* **12** 43–57

[6] Shikhaliev P M 2008 Energy-resolved computed tomography: first experimental results *Phys. Med. Biol.* **53** 5595–613

[7] Egan C K, Jacques S D M, Wilson M D, Veale M C, Seller P, Beale A M, Pattrick R A D, Withers P J and Cernik R J 2015 3D chemical imaging in the laboratory by hyperspectral x-ray computed tomography *Sci. Rep.* **5** 15979

[8] Egan C K, Jacques S D M, Wilson M D, Veale M C, Seller P, Pattrick R A D, Withers P J and Cernik R J 2017 3D elemental mapping of materials and structures by laboratory scale spectroscopic x-ray tomography *J. Phys.: Conf. Ser.* **849** 012013

[9] Roessl E and Proksa R 2007 K-edge imaging in x-ray computed tomography using multi-bin photon counting detectors *Phys. Med. Biol.* **52** 4679

[10] Wilson M D, Dummott L, Duarte D D, Green F H, Pani S, Schneider A, Scuffham J W, Seller P and Veale M C 2015 A 10 cm × 10 cm CdTe spectroscopic imaging detector based on the HEXITEC ASIC *J. Instrum.* **10** P10011

[11] Heismann B J, Schmidt B T and Flohr T 2012 *Spectral Computed Tomography* (Bellingham, WA: SPIE)

[12] Schirra C O, Roessl E, Koehler T, Brendel B, Thran A, Pan D, Anastasio M A and Proksa R 2013 Statistical reconstruction of material decomposed data in spectral CT *IEEE Trans. Med. Imaging* **32** 1249–57

[13] Sawatzky A, Xu Q, Schirra C O and Anastasio M A 2014 Proximal ADMM for multi-channel image reconstruction in spectral x-ray CT *IEEE Trans. Med. Imaging* **33** 1637–368

[14] Barber R F, Sidky E Y, Schmidt T G and Pan X 2016 An algorithm for constrained one-step inversion of spectral CT data *Phys. Med. Biol.* **61** 3784

[15] Schmidt T, Barber R and Sidky E 2017 A Spectral CT method to directly estimate basis material maps from experimental photon-counting data *IEEE Trans. Med. Imaging* **36** 1808–19

[16] Gao H, Yu H, Osher S and Wang G 2011 Multi-energy CT based on a prior rank, intensity and sparsity model (PRISM) *Inverse Problems* **27** 115012

[17] Semerci O, Hao N, Kilmer M E and Miller E L 2014 Tensor-based formulation and nuclear norm regularization for multienergy computed tomography *IEEE Trans. Image Process.* **23** 1678–93

[18] Rigie D S and La Rivière P J 2015 Joint reconstruction of multi-channel, spectral CT data via constrained total nuclear variation minimization *Phys. Med. Biol.* **60** 1741

[19] Ehrhardt M J and Arridge S R 2014 Vector-valued image processing by parallel level sets *IEEE Trans. Image Process.* **23** 9–18

[20] Ehrhardt M J, Thielemans K, Pizarro L, Atkinson D, Ourselin S, Hutton B F and Arridge S R 2014 Joint reconstruction of PET-MRI by exploiting structural similarity *Inverse Problems* **31** 015001

[21] Ehrhardt M J and Betcke M M 2016 Multicontrast MRI reconstruction with structure-guided total variation *SIAM J. Imaging Sci.* **9** 1084–106

[22] Bathke C, Kluth T, Brandt C and Maaß P 2017 Improved image reconstruction in magnetic particle imaging using structural *a priori* information *Int. J. Magn. Part. Imaging* **3** 1–10

[23] Knoll F, Holler M, Koesters T, Otazo R, Bredies K and Sodickson D K 2017 Joint MR-PET reconstruction using a multi-channel image regularizer *IEEE Trans. Med. Imaging* **36** 1–16

[24] Rasch J, Brinkmann E M and Burger M 2017 Joint reconstruction via coupled Bregman iterations with applications to PET-MR imaging *Inverse Problems* **34** 014001

[25] Moeller M, Brinkmann E M, Burger M and Seybold T 2014 Color Bregman TV *SIAM J. Imaging Sci.* **7** 2771–806

[26] Kazantsev D, Bleichrodt F, van Leeuwen T, Kaestner A, Withers P J, Batenburg K J and Lee P D 2017 A novel tomographic reconstruction method based on the robust Student's t function for suppressing data outliers *IEEE Trans. Comput. Imaging* **3** 682–93

[27] Vogel C R 2002 *Computational Methods for Inverse Problems* vol 23 (Philadelphia, PA: SIAM)

[28] Gürsoy D, Biçer T, Almer J D, Kettimuthu R, Stock S R and De Carlo F 2015 Maximum *a posteriori* estimation of crystallographic phases in x-ray diffraction tomography *Phil. Trans.* A **373** 20140392
[29] Kazantsev D, Lionheart W R, Withers P J and Lee P D 2014 Multimodal image reconstruction using supplementary structural information in total variation regularization *Sens. Imaging* **15** 97
[30] Kazantsev D, Ourselin S, Hutton B F, Dobson K J, Kaestner A P, Lionheart W R, Withers P J, Lee P D and Arridge S R 2014 A novel technique to incorporate structural prior information into multi-modal tomographic reconstruction *Inverse Problems* **30** 065004
[31] Sauer K and Bouman C 1993 A local update strategy for iterative reconstruction from projections *IEEE Trans. Signal Process.* **41** 534–48
[32] Ramani S and Fessler J A 2012 A splitting-based iterative algorithm for accelerated statistical x-ray CT reconstruction *IEEE Trans. Med. Imaging* **31** 677–88
[33] Rudin L I, Osher S and Fatemi E 1992 Nonlinear total variation based noise removal algorithms *Physica* D **60** 259–68
[34] Chambolle A 2004 An algorithm for total variation minimization and applications *J. Math. Imaging Vis.* **20** 89–97
[35] Holt K M 2014 Total nuclear variation and jacobian extensions of total variation for vector fields *IEEE Trans. Image Process.* **23** 3975–89
[36] Lefkimmiatis S, Roussos A, Maragos P and Unser M 2015 Structure tensor total variation *SIAM J. Imaging Sci.* **8** 1090–122
[37] Duran J, Moeller M, Sbert C and Cremers D 2016 Collaborative total variation: a general framework for vectorial TV models *SIAM J. Imaging Sci.* **9** 116–51
[38] Parikh N and Boyd S 2014 Proximal algorithms *Found. Trends Optim.* **1** 127–239
[39] Beck A 2017 *First-Order Methods in Optimization* vol 25 (Philadelphia, PA: SIAM)
[40] Beck A and Teboulle M 2009 A fast iterative shrinkage-thresholding algorithm for linear inverse problems *SIAM J. Imaging Sci.* **2** 183–202
[41] Beck A and Teboulle M 2009 Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems *IEEE Trans. Image Process.* **18** 2419–34
[42] Xu Q, Yang D, Tan J, Sawatzky A and Anastasio M A 2016 Accelerated fast iterative shrinkage thresholding algorithms for sparsity-regularized cone-beam CT image reconstruction *Med. Phys.* **43** 1849–72
[43] Esser E, Zhang X and Chan T F 2010 A general framework for a class of first order primal-dual algorithms for convex optimization in imaging science *SIAM J. Imaging Sci.* **3** 1015–46
[44] Kazantsev D, Pasca E, Nagella S and Pickalov V 2018 TomoPhantom v.1.0. Software to generate 2D-4D analytical phantoms and their Radon transforms for image processing (Version 1.0) *Zenodo* (https://doi.org/10.5281/zenodo.1215759)
[45] Siewerdsen J H, Waese A M, Moseley D J, Richard S and Jaffray D A 2004 Spektr: a computational tool for x-ray spectral analysis and imaging system optimization *Med. Phys.* **31** 3057–67
[46] Tuszynski J 2006–2016 PhotonAttenuation—Software for modeling of photons passing through different materials https://uk.mathworks.com/matlabcentral/fileexchange/12092-photonattenuation
[47] van Aarle W, Palenstijn W J, De Beenhouwer J, Altantzis T, Bals S, Batenburg K J and Sijbers J 2015 The ASTRA toolbox: a platform for advanced algorithm development in electron tomography *Ultramicroscopy* **157** 35–47
[48] Kaipio J and Somersalo E 2007 Statistical inverse problems: discretization, model reduction and inverse crimes *J. Comput. Appl Math.* **198** 493–504
[49] Wang Z, Bovik A C, Sheikh H R and Simoncelli E P 2004 Image quality assessment: from error visibility to structural similarity *IEEE Trans. Image Process.* **13** 600–12
[50] Grant M and Boyd S 2014 CVX: MATLAB software for disciplined convex programming, version 2.1 http://cvxr.com/cvx
[51] Grant M and Boyd S 2008 Graph implementations for nonsmooth convex programs *Recent Advances in Learning and Control* (*Lecture Notes in Control and Information Sciences*) ed V Blondel *et al* (Berlin: Springer) pp 95–110
[52] Hansen P C and Saxild-Hansen M 2012 AIR tools—a MATLAB package of algebraic iterative reconstruction methods *J. Comput. Appl Math.* **236** 2167–78
[53] Hansen P C and Jørgensen J S 2017 AIR tools II: algebraic iterative reconstruction methods, improved implementation *Numer. Algorithms* https://doi.org/10.1007/s11075-017-0430-x