



Nonlinear material decomposition using a regularized iterative scheme based on the Bregman distance

Juan F P J Abascal, Nicolas Ducros, Françoise Peyrin

► To cite this version:

Juan F P J Abascal, Nicolas Ducros, Françoise Peyrin. Nonlinear material decomposition using a regularized iterative scheme based on the Bregman distance. 2017. hal-01621265v1

HAL Id: hal-01621265

<https://hal.science/hal-01621265v1>

Preprint submitted on 23 Oct 2017 (v1), last revised 14 Nov 2018 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Nonlinear material decomposition using a regularized iterative scheme based on the Bregman distance

JFPJ Abascal, N Ducros, F Peyrin

Univ Lyon, INSA-Lyon, Université Claude Bernard Lyon 1, UJM-Saint Etienne,
CNRS, Inserm, CREATIS UMR 5220, U1206, Lyon, France

E-mail: juan.abascal@creatis.insa-lyon.fr

October 2017

Abstract. In this paper, we address the resolution of material decomposition, which is a nonlinear inverse problem encountered in spectral computerized tomography (CT). The problem is usually solved in a variational framework but, due to the nonlinearity of the forward operator, the objective function may be nonconvex and standard approaches may fail. Regularized iterative schemes based on the Bregman distance have been suggested for improving global convergence properties. In this work, we analyse the convexity of the material decomposition problem and propose a regularized iterative scheme based on the Bregman distance to solve it. We evaluate our Bregman iterative algorithm and compare it to a regularized Gauss-Newton (GN) method using data simulated in a realistic thorax phantom.

First, we prove the existence of a convex set where the usual data fidelity term is convex. Interestingly, this set includes zero, making it a good initial guess for iterative minimisation schemes. Using numerical simulations, we show that the data fidelity term can be nonconvex for large values of the decomposed materials. Second, the proposed Bregman iterative scheme is evaluated in different situations. It is observed to be robust to the selection of the initial guess, leading to the global minimum in all tested examples while the GN method fails to converge when the initial guess is not well chosen. Moreover, it is found to avoid the selection of the regularization parameter for little extra computation.

In conclusion, we have provided a suitable initialization strategy to solve the nonlinear material decomposition problem using convex optimization methods and evaluated a Bregman iterative scheme for this problem. The improvement in global convergence of Bregman iterative scheme combined with other interesting properties of the Bregman distance appears as a compelling strategy for nonlinear inverse problems.

Keywords: Bregman distance, convexity, spectral computerized tomography

1. Introduction

We consider regularized methods for solving nonlinear inverse problems. Successful methods for linear inverse problems have been generalized for the case of nonlinear inverse problems. However, in general, objective functions involving nonlinear operators are nonconvex, so classical optimization procedures may fail. Regularized iterative procedures have been proposed to ensure that a global minimizer exists. The idea behind these methods is to replace the original regularized problem with a sequence of problems, each of which satisfies stronger convergence properties [1, 2].

A popular and efficient iterative regularization procedure for the ℓ_1 - and total variation (TV)-problems is based on the use of the Bregman distance and has been widely studied for linear measurements [3, 4, 5, 6, 7, 8]. Replacing a convex regularization functional with the Bregman distance associated to this functional leads to an iterative scheme that converges to the solution of the constrained optimization problem. The use of the Bregman distance provides additional appealing properties. For instance, it has been shown that minimizing the Bregman distance associated to TV can recover the contrast lost that may appear when solving the TV problem [9]. The split Bregman method proposed for image processing and compressed sensing combines the Bregman distance with a splitting technique to efficiently deal with non differentiable convex functional, such as total variation [7, 10, 11]. It has been successfully applied to solve the total variation problem in image denoising, image inpainting, and compressed sensing [7, 12, 13, 14, 15, 16, 17]. The use of Bregman distance for the TV minimization subject to nonlinear measurements enjoys further benefits but it has not been widely explored [14, 18]. In [19], Bregman iterative schemes based on classical reconstruction methods were proposed to improve global convergence for nonlinear inverse problems.

In this work, we are concerned with the nonlinear material decomposition problem that appears in spectral computerized tomography (CT). Spectral CT is gaining increasing attention due to the recent development of energy discriminating photon-counting detectors that simultaneously count photons and resolve their energy [20, 21]. This technology addresses the limitations of conventional CT and provides material decomposition capabilities, which allows, for instance, characterizing lesions for breast imaging [22] and atherosclerotic plaque [23, 24]. The spectral CT material reconstruction problem can be solved by dividing it in to two separated steps: a nonlinear material decomposition step, which takes places on the Radon domain, and a linear tomographic reconstruction step. We focus on the material decomposition problem and adopt a model-based approach which models source spectrum and energy responses of the detector [25]. For alternative approaches to material reconstruction see the following works [26, 27, 28, 29, 30, 31]. We remark that spectral CT material reconstruction differs from recent works that proposed regularization approaches for the linear tomographic problem and that regularized on the image domain [14, 32, 33, 34].

Most previous studies that addressed spectral CT have ignored nonconvexity or adopted some assumptions to make the problem convex [35]. For dual-energy CT and

for the negative loglikelihood functional, a hybrid first and second order variational approach was proposed to provide a more robust algorithm [36]. Recently, a regularized Gauss-Newton (GN) method with specific regularization applied to each material was applied to solve the material decomposition problem [37]. This method provided significantly superior results than the non regularized maximum-likelihood method that is commonly used in practice [25].

The goal of this work is to investigate convexity for the material decomposition inverse problem and to propose a regularized Bregman iterative scheme to solve this problem. To this goal we build from the previously proposed regularized GN method [37] and modify it in order to minimise the Bregman distance associated to the same regularization functional. We put particular emphasis to study convexity and to provide a suitable initialization strategy for convex optimization algorithms. In addition, we assess the benefits of a Bregman iterative algorithm to provide a global minimum by studying its dependence on the initial guess.

The paper is organized as follows. In section 2 we present the formulation of the forward and inverse problem. In section 3 we show that the data fidelity term is convex in a defined set and that the zero solution belongs to this region. In section 4 we describe the numerical approaches adopted to solve this problem, the regularized Bregman iterative scheme and a regularized Gauss-Newton method. In section 5, we show the numerical results. First, we analyse convexity using numerical examples. Then, we assess the Bregman iterative scheme in comparison to the Gauss-Newton approach using a numerical thorax phantom. Section 6 contains the discussion, further work and conclusion.

2. Material decomposition in spectral CT

2.1. Forward problem

Material decomposition requires the inversion of the nonlinear system [37]

$$s = \mathcal{F}(a) \tag{1}$$

where s represents the measured photon counts, a represents the unknown projected mass densities, and \mathcal{F} models both the light-matter interaction within the object and the measurement process. In spectral imaging, the measured photon counts can be written as $s = (s_1, \dots, s_I)$ where s_i , $1 \leq i \leq I$, represents the photon counts measured in the i -th energy bin of the detector. Moreover, we assume that the unknown projected mass densities can be written as $a = (a_1, \dots, a_M)$ where a_m , $1 \leq m \leq M$, is the projected mass density image of the m -th material in the object.

For a detector having P pixels, we have $s_i = (s_i^1, \dots, s_i^P)$ and $a_m = (a_m^1, \dots, a_m^P)$, where s_i^p (resp. a_m^p) represents the photon counts (resp. projected mass density) in the p -th pixel of the detector. Assuming that no inter-pixel spectral mixing occurs, the

system (1) is equivalent to

$$s_i^p = \mathcal{F}_i^p(a_1^p, \dots, a_M^p), \quad 1 \leq i \leq I \text{ and } 1 \leq p \leq P \quad (2)$$

The nonlinear mapping $\mathcal{F}_i^p : \mathbb{R}^M \rightarrow \mathbb{R}$ can be modelled by [25]

$$\mathcal{F}_i^p(a_1^p, \dots, a_M^p) := \int_{\mathcal{E}} n_0(E) d_i(E) \exp \left[- \sum_{m=1}^M a_m^p \tau_m(E) \right] dE \quad (3)$$

where $\mathcal{E} = [0, E_{\max}]$ is the energy range delivered by the X-ray tube, $n_0(E)$ is the X-ray source energy spectrum, $d_i(E)$ is the energy-dependent response function of the i -th bin of the detector, and $\tau_m(E)$ is the energy-dependent mass attenuation of the m -th constituent of the object.

2.2. Inverse problem

Within a variational framework, material decomposition is formulated as the following optimisation problem

$$\min_a C(a) = \min_a L(s^\delta, a) + \tilde{\alpha} R(a) \quad (4)$$

where s^δ represents the noisy measurements, L is a data fidelity term, R is a regularization term encoding *a priori* knowledge about the solution, and $\tilde{\alpha}$ is the regularisation parameter. Assuming that the measurements are Poisson distributed, the following approximation of the discretised Kullback-Leibler divergence is considered

$$L(s^\delta, a) = \frac{1}{2} \|\mathcal{F}(a) - s^\delta\|_W^2, \quad (5)$$

where $\|x\|_W$ is the weighted norm defined by $\|x\|_W^2 = x^\top W^\top W x$. Here, the weight matrix W is chosen as

$$W = \text{diag} \left(\frac{1}{\max(\sqrt{s^\delta}, \epsilon)} \right), \quad (6)$$

where ϵ is a small parameter that is added to the denominator to avoid division by zeros.

It is important to note that although the regularization functional R is convex the cost function C in (4) can be nonconvex whenever the data fidelity term L is non convex, which usually occurs with nonlinear measurements.

2.3. Gauss Newton method

A wide variety of approaches can be used to solve the nonlinear problem (4) [38, 39]. Among them the Gauss Newton method has fast (quadratic) convergence and has been recently shown to be very effective for material decomposition [36, 37]. The Gauss Newton update is computed as

$$a_{n+1} = a_n + \lambda_n H^{-1}(a_n) C'(a_n) \quad (7)$$

where λ_n is the step length, C' is the gradient of C , and H is an approximation for the Hessian of C , which is given by

$$H(a_n) = \mathcal{F}'(a_n)^\top W^\top W \mathcal{F}'(a_n) + \tilde{\alpha} R''(a_n) \quad (8)$$

where \mathcal{F}' denotes the Jacobian of \mathcal{F} and R'' the Hessian of R . The step length is usually found by a line search strategy, i.e., minimizing the cost function along the descent direction provided by $H^{-1}(a_n)C'(a_n)$.

3. Analysis of the convexity of the problem

The convexity of the material decomposition problem is analysed deriving the following properties for the Hessian of the data fidelity functional. For simplicity, the analysis is conducted considering one pixel only (i.e. for $P = 1$) and the index p has been dropped. Since the data fidelity term L is separable with respect to pixels, the results bellow can be generalised to any pixel number (i.e., $P > 1$).

Definition 1. Let $g_i(a, E)$ be the number of incident photons seen at energy E in the i -th energy bin of the detector. It is given by

$$g_i(a, E) = n_0(E)d_i(E) \exp \left[- \sum_{m=1}^M a_m \tau_m(E) \right] \quad (9)$$

Proposition 1. The Hessian of L is given by

$$H(a) = - \sum_i w_i^2 (s_i - \mathcal{F}_i(a)) T^\top G_i(a) T + \mathcal{F}'^\top W^\top W \mathcal{F}' \quad (10)$$

where $T = (\tau_{j,m}) \in \mathbb{R}^{J \times M}$ is the mass attenuation matrix defined by $\tau_{j,m} = \tau_m(j\Delta E)$, $1 \leq j \leq J$ and $G_i(a) = \text{diag}(g_i(a, E_1), \dots, g_i(a, E_J)) \Delta E$.

Proof. The functions $\mathcal{F}_i: \mathbb{R}^M \rightarrow \mathbb{R}$, $1 \leq i \leq I$ are twice continuously differentiable and the Hessian of L takes the following matrix form [40]

$$H(a) = - \sum_i w_i^2 (s_i - \mathcal{F}_i(a)) H_i(a) + \mathcal{F}'^\top W^\top W \mathcal{F}' \quad (11)$$

where H_i denotes the Hessian of \mathcal{F}_i . It can be shown that $H_i(a) = T^\top G_i(a) T$ (see section Appendix A), which completes the proof. \square

Proposition 2. The data fidelity functional L is convex on the (possibly nonconvex) set \mathcal{A} defined by

$$\mathcal{A} = \{a \in \mathbb{R}^M \mid \mathcal{F}_i(a) \geq s_i, \forall i\} \quad (12)$$

Proof. While $\mathcal{F}'^\top W^\top W \mathcal{F}'$ is clearly positive semi-definite, some or all the terms $w_i^2 (s_i - \mathcal{F}_i(a)) T^\top G_i(a) T$ can contribute against convexity. Since $G_i(a)$ is a diagonal matrix with positive elements, $T^\top G_i(a) T$ is positive definite. If $s_i - \mathcal{F}_i(a)$ is negative for all i , then $H(a)$ is the sum of positive definite matrices, hence it is positive definite. \square

Proposition 3. *Let a_m^{true} be the solution. The data fidelity functional L is convex on the convex set \mathcal{B} defined by*

$$\mathcal{B} = \{a \in \mathbb{R}^M \mid a_m \leq a_m^{\text{true}}, \forall m\} \quad (13)$$

Proof. Noting that all $\mathcal{F}_i : a_m \mapsto \mathcal{F}_i(a_m)$ are monotonically decreasing, choosing $a_m \leq a_m^{\text{true}}$ for all m guaranties that $\mathcal{F}_i(a) \geq \mathcal{F}_i(a^{\text{true}})$. \square

Remark 4. *The particular point $a = 0$ lies within the convex region, i.e., $0 \in \mathcal{B}$.*

Proof. Physically, solutions are positive. \square

4. Resolution of the problem

4.1. Iterative regularization scheme based on the Bregman distance

The Bregman distance generalizes the concept of distance for convex regularization functionals R [41]. The Bregman distance associated with the functional R for two points a and \tilde{a} is defined as

$$D_\xi^R(a, \tilde{a}) = R(a) - R(\tilde{a}) - \langle \xi, a - \tilde{a} \rangle, \quad (14)$$

where ξ is a subgradient of R at \tilde{a} .

The use of Bregman distance became popular after its application to construct iterative regularization schemes for total variation regularization [3]. Since then, it has been widely applied to image restoration, deconvolution and compressed sensing problems with linear measurement operators and with convex regularization functional [4, 5, 7, 9, 13, 42]. A generalization of the Bregman distance for nonconvex regularization functionals has been also suggested [43]. However, its application to solve problems with nonlinear measurements is less known. In [19] authors proposed to replace the problem (4) by a sequence of iterative problems

$$a_{k+1} = \arg \min_a \frac{1}{2} \|\mathcal{F}(a) - s^\delta\|_W^2 + \alpha D_{\xi_k}^R(a, a_k), \quad (15)$$

$$\xi_{k+1} = \xi_k - \alpha^{-1} (W\mathcal{F}'(a_{k+1}))^\top (W\mathcal{F}(a_{k+1}) - Ws^\delta), \quad (16)$$

where $\alpha > 0$ is an hyperparameter, which is *not* a regularization parameter and that can be chosen as a large constant. Each of the subproblems (15) is an optimization problem that has the same form as (4) but where the regularization functional R is replaced by the Bregman distance between the unknown, a , and its previous estimate, a_k . Equation (16) provides an update of the subgradient derived from the first optimality condition of (15), which enforces the data fidelity term iteratively [7].

The Bregman iterative scheme (15-16) constitutes then a sequence of subproblems where each of them has the same complexity as the original problem (4), but there is a substantial gain, as each subproblem is locally convex for sufficiently large values of α [19]. This provides a scheme to find a global minimiser for nonconvex functional associated to nonlinear operator equations.

An additional modification to the iterative scheme (15-16) was proposed in [19] to ensure local convexity for total variation regularization. A zero-order Tikhonov term $\frac{\kappa}{2}\|a\|_2^2$ is added to the original total variation regularization functional R to ensure local convexity around the current estimate a_k . The value of κ must be sufficiently small in order not to modify the properties of the regularization functional R . Substituting the Bregman distance (14) into (15), the iterative scheme is rewritten as

$$\begin{aligned} a_{k+1} &= \min_a \Theta_k(a) \\ &= \min_a \frac{1}{2}\|\mathcal{F}(a) - s^\delta\|_W^2 + \alpha R(a) - \alpha \langle \xi_k, a \rangle + \alpha \frac{\kappa}{2}\|a\|_2^2, \end{aligned} \quad (17)$$

$$\xi_{k+1} = \xi_k - \alpha^{-1}(W\mathcal{F}'(a_{k+1}))^\top (W\mathcal{F}(a_{k+1}) - Ws^\delta). \quad (18)$$

The minimization problem (17) can be efficiently solved using a Gauss Newton algorithm and (18) is the update of the Bregman iteration [3, 7]. Therefore, the iterative scheme (17-18) is referred to as GN-Bregman (GNB).

4.2. Algorithm stopping criterion

Note that GNB algorithm numerically involves two iterative procedures: the (outer) Bregman iteration indexed by k and the (inner) Gauss Newton iteration that we denote with the index n . We define N_k the total number of GN iterations required to minimise Θ_k and $N_{\text{tot}} = \sum_k N_k$ the total number of GN iterations of the GNB algorithm. The stopping criterion for the minimisation of Θ_k (17) is based on the cost function decrease, i.e., the algorithm is stopped when $(\Theta_k(a^{n-1}) - \Theta_k(a^n))/\Theta_k(a^{n-1}) < \theta$, and the stopping criterion for the Bregman iterative scheme is the discrepancy principle, i.e., the algorithm is stopped when $\frac{1}{2}\|\mathcal{F}(a^k) - s^\delta\|_W^2 < \text{tol}$ [7, 19].

4.3. Regularization functional

In this work, we consider the separable regulariser proposed in [37] and defined by

$$R(a) = \sum_m R_m(a_m), \quad (19)$$

where R_m is the regularization functional for the m -th material. This choice allows to incorporate *a-priori* information for each material independently. Here, we consider Tikhonov regularization to promote smooth materials, i.e.,

$$R_m(a_m) = \|D^\ell a_m\|_2^2 \quad (20)$$

where D^ℓ represents the ℓ -th spatial derivative operator, and a smooth approximation of total variation to promote piecewise constant materials, i.e.,

$$R_m(a_m) = \sqrt{(Da_m)^2 + \epsilon^2} - \epsilon. \quad (21)$$

where ϵ is a small hyperparameter. Although only differentiable regularisers are used, Bregman iterative algorithms are also valid for the space of functions of bounded variations and other convex sparse regularization functional that uses ℓ_1 -norm to impose sparsity in a given orthonormal basis, such as wavelet transform or framelets.

5. Numerical simulations and results

5.1. Numerical phantom data

We used numerical simulations to illustrate the (non) convexity of the problem and evaluate the proposed algorithms. Different datasets were simulated using realistic X-ray tube and photon-counting detector with four energy bins (see figure 1) and considering numerical phantoms made of soft tissue, bone and gadolinium (Gd). This is a relevant example in K-edge imaging for characterizing atherosclerotic plaque [23, 24]. The linear mass attenuation coefficients for the three materials are provided in figure 1c.

Two numerical phantoms were created: a one-pixel phantom (only used to study convexity) and a realistic thorax phantom built from a segmented thoracic-abdominal CT scan [44, 45]. For the realistic phantom, a 611×167 -pixel detector is simulated, leading to $P = 102,037$. Two projection images for two different projection angles were evaluated. For each view, the dataset comprised $I = 4$ projection images corresponding to the different energy bins and $M = 3$ unknowns projected mass density images were recovered. Increasing levels of Poisson noise were considered by lowering the number of incident photons from 10^8 to 10^5 . The datasets are publicly available from [37].

We have used specific regularization for each material as described in section 4.3. We used first- and second-order Tikhonov regularization for bone and soft tissue, respectively, and smooth total variation regularization for the gadolinium image.

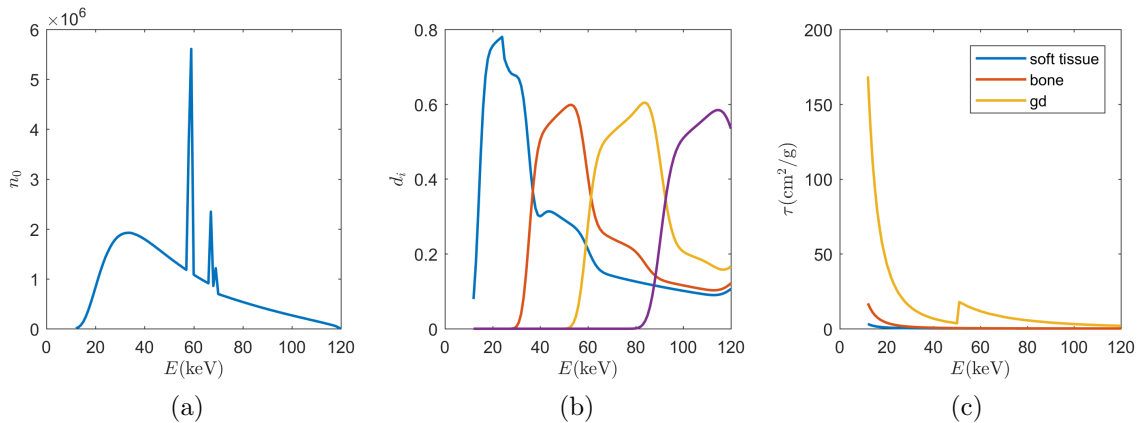


Figure 1. Forward model parameters used in the numerical simulations. (a) Source spectrum $n_0(E)$. (b) Detector response function $d_i(E)$ for each of the $I = 4$ energy bins. (c) Linear mass attenuation coefficients $\tau_m(E)$ for each of the $M = 3$ materials.

5.2. Numerical study of the data fidelity convexity

We have shown in section 3 that the problem is convex in a certain region of the solution space (proposition 3) and that the zero mass density image belongs to this region (remark 4). This is a relevant result as it can serve as a guide to initialize convex optimization algorithms. Nevertheless, there may be regions of the solution space where

the data fidelity term is nonconvex. In this section, we examine convexity numerically by computing the data fidelity term (5) for two typical numerical phantoms. Figure 2 shows the data fidelity term for a one-pixel mass density image a while figure 3 displays the data fidelity term for the multiple-pixel thorax phantom. In the latter simulation, we consider only constant images for the three materials, i.e., the same value is assigned to the entire image for each material. This has been done for simplification but it is actually relevant as constant images are commonly used as initial guesses.

These examples show that the zero mass density image belongs to the convex region and that the problem is nonconvex for large values of the three materials. It can be noticed that the data fidelity term becomes asymptotically flat for increasing values of the three materials, i.e., $\lim_{a \rightarrow \infty} L(a) = \frac{1}{2} \|s^\delta\|_W^2$.

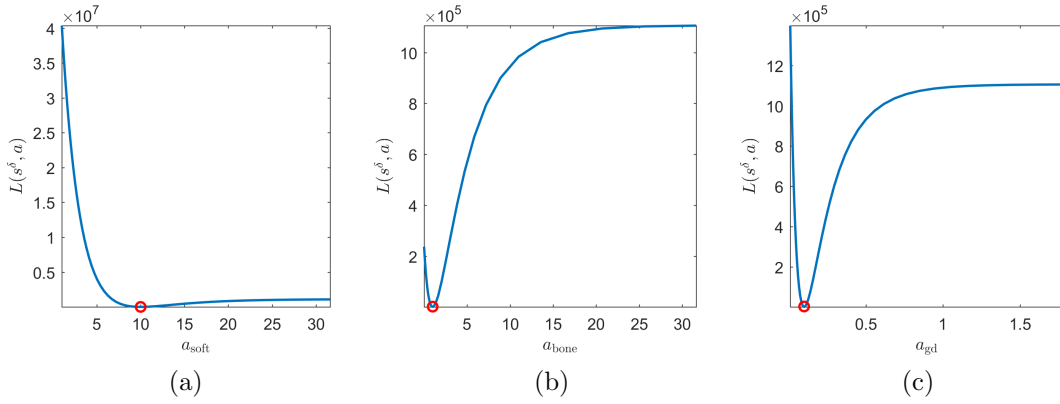


Figure 2. One dimensional plot of the data fidelity term for a one-pixel image $a \in \mathbb{R}^M$. The ground truth solution is set to $a_{\text{soft}}^{\text{true}} = 10 \text{ g/cm}^3$, $a_{\text{bone}}^{\text{true}} = 1 \text{ g/cm}^3$, and $a_{\text{gd}}^{\text{true}} = 0.1 \text{ g/cm}^3$ (see red circle). Data fidelity is displayed versus the value of one material while the other two materials remain fixed and equal to the ground truth solution.

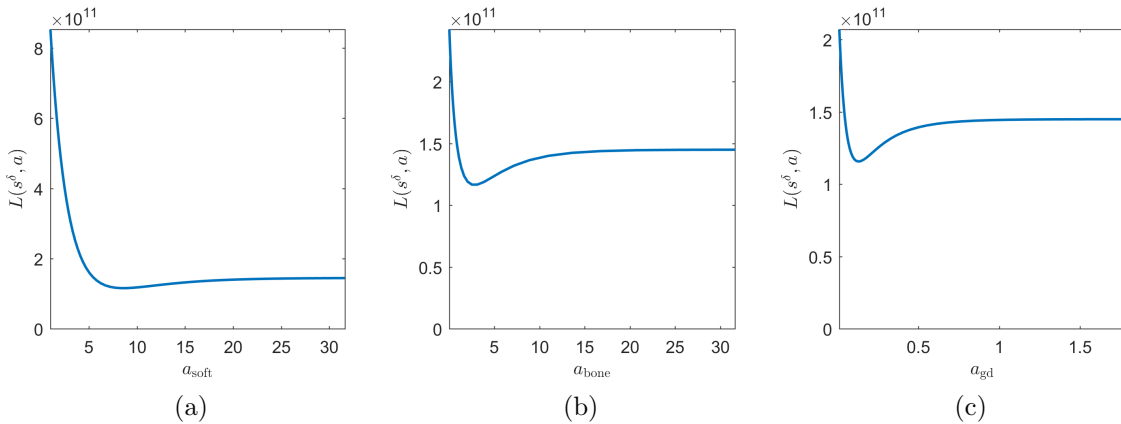


Figure 3. Data fidelity term for the realistic thorax phantom where the three materials are set to a constant image $a_{\text{soft}}^{\text{true}} = 10 \text{ g/cm}^3$, $a_{\text{bone}}^{\text{true}} = 1 \text{ g/cm}^3$, and $a_{\text{gd}}^{\text{true}} = 0.1 \text{ g/cm}^3$. Data fidelity is displayed versus the value of one material while the other two materials remain fixed and equal to (a) $a_{\text{bone}} = 0.15 \text{ g/cm}^3$ and $a_{\text{gd}} = 0.007 \text{ g/cm}^3$, (b) $a_{\text{soft}} = 4 \text{ g/cm}^3$ and $a_{\text{gd}} = 0.007 \text{ g/cm}^3$, and (c) $a_{\text{soft}} = 4 \text{ g/cm}^3$ and $a_{\text{bone}} = 0.15 \text{ g/cm}^3$.

Figure 4 displays the two characteristic sets \mathcal{A} and \mathcal{B} that are introduced in proposition 2 and 3, respectively, for the one-pixel mass density phantom (same as figure 2). The colormap represents the number of energy bins $1 \leq i \leq I$ satisfying $\mathcal{F}_i(a) \geq s_i$. The solution is set to $a_{\text{soft}}^{\text{true}} = 10 \text{ g/cm}^3$, $a_{\text{bone}}^{\text{true}} = 1 \text{ g/cm}^3$, and $a_{\text{gd}}^{\text{true}} = 0.1 \text{ g/cm}^3$ (see red circle). The data is computed using the parameters plotted in figure 1. The set \mathcal{A} is depicted in yellow (all bins satisfy $\mathcal{F}_i(a) \geq s_i$) while the set \mathcal{B} is indicated by the grey dashed line.

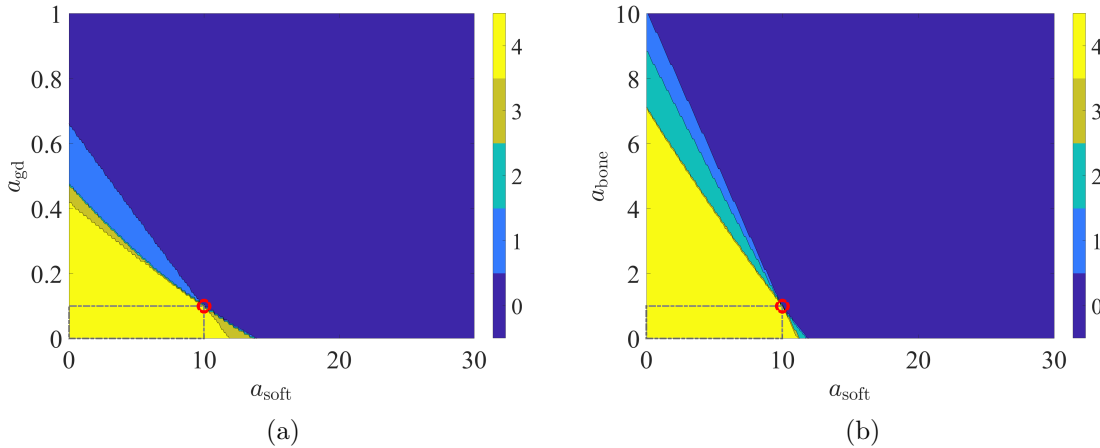


Figure 4. Characteristic sets included in the solution space. The colormap represents the number of energy bins $1 \leq i \leq I$ satisfying $\mathcal{F}_i(a) \geq s_i$. The solution is set to $a_{\text{soft}}^{\text{true}} = 10 \text{ g/cm}^3$, $a_{\text{bone}}^{\text{true}} = 1 \text{ g/cm}^3$, and $a_{\text{gd}}^{\text{true}} = 0.1 \text{ g/cm}^3$ (see red circle). The data is computed using the parameters plotted in figure 1. The set \mathcal{A} is depicted in yellow (all bins satisfy $\mathcal{F}_i(a) \geq s_i$) while the set \mathcal{B} is indicated by the grey dashed line. Sets are displayed versus the value of two materials while the other material remain fixed and equal to the ground truth solution.

5.3. Convergence of GNB

Solution error and data fidelity versus the number of iterations are shown in figure 5 for GNB for different values of α . GNB leads to an optimal solution as long as α is sufficiently large, where the larger α the more iterations it requires [7, 13]. For these results, we used $\kappa = 10^{-6}$, initial guess $a^0 = 0$ and tolerance $tol = 1.12 \cdot 10^5$. As a comparison, we also provide the solution error for GN for the value of the regularization parameter ($\tilde{\alpha} = 0.23$) that led to the lowest solution error. For GN we set initial guess to $a^0 = 0$ and tolerance to $\theta = 10^{-4}$, which corresponded to a value of the data fidelity term of $1.12 \cdot 10^5$. GN converged in $N_{\text{tot}} = 16$ iterations for $\tilde{\alpha} = 0.23$ (see figure 5(b)). GNB solves a sequence of subproblems, so it needs a larger number of iterations. It needed $N_{\text{tot}} = 28$ and $N_{\text{tot}} = 40$ for $\alpha = 2$ and $\alpha = 10$, respectively.

Figure 6(a) shows the solution error versus the regularization parameter $\tilde{\alpha}$ for GN and versus the parameter α for GNB. While GNB provides the same solution for all α 's larger than 0.5, GN depends dramatically on the choice of α . Although GN converged in only $N_{\text{tot}} = 16$ iterations (for $\tilde{\alpha} = 0.23$, see figure 5(b)), extra decompositions for

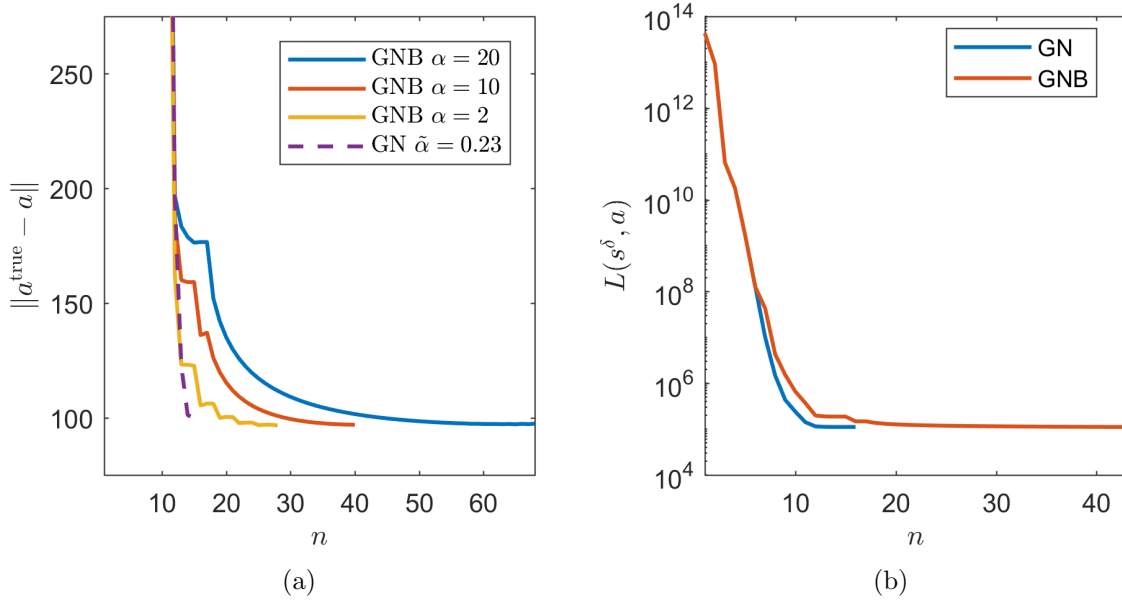


Figure 5. Convergence of the algorithms. (a) Solution error versus the number of Gauss Newton iterations (n) for GNB (full line) and for GN (dotted line). (b) Data fidelity decrease for GNB ($\alpha = 10$) and for GN ($\tilde{\alpha} = 0.23$). We used $a^0 = 0$ as initial guess.

different regularization parameters are required. On the contrary, GNB can be used with a large $\tilde{\alpha}$ chosen *a priori*.

Figure 6(b) shows that the number of Bregman iterations is almost constant with α and equal to $k = 15$ in the cases investigated here. Figure 6(c) shows that the total number of Gauss Newton iterations N_{tot} is linearly related to the parameter α .

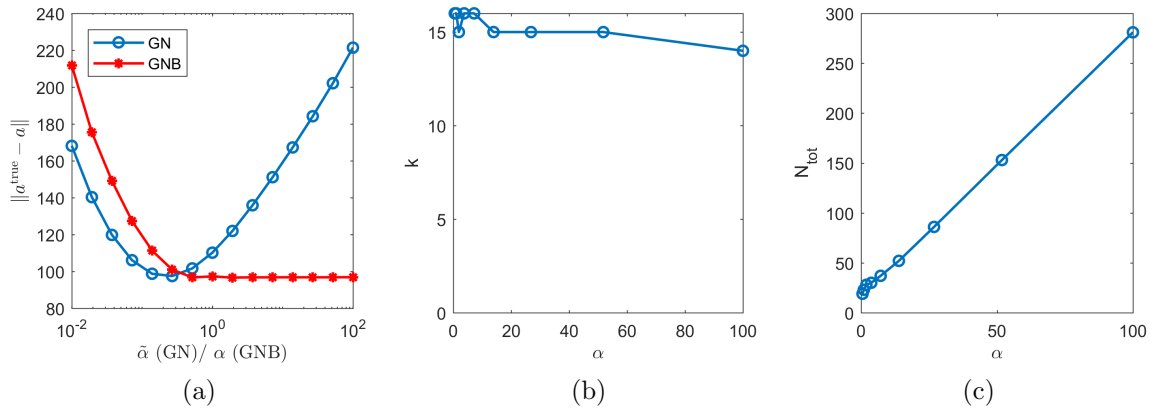


Figure 6. Influence of the hyperparameter α . (a) Solution error for GN and GNB. (b) Total number of Bregman iterations for GNB. (c) Total number of Gauss Newton iterations for GNB.

Figure 7 shows the number of Gauss Newton iterations N_k needed to minimize Θ_k , for different values of α . An important remark is that GNB required nearly only one iteration $N_k = 1$ to solve each subproblem for $k > 1$, so it adds just few iterations with

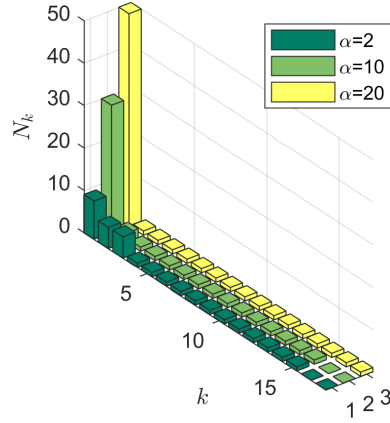


Figure 7. Number of Gauss Newton iterations, N_k , needed at each Bregman iteration, k , for different values of α .

respect to GN.

Figure 8 shows decomposed images by GN and GNB. Second row in figure 8 gives decomposed images by GN for an optimal regularization parameter $\tilde{\alpha} = 0.23$. Third to fifth rows are the decomposed images by GNB for different number of iterations, setting $\alpha = 10$, $\kappa = 1 \cdot 10^{-6}$ and $a^0 = 0$. Third and fourth rows correspond to an intermediate and final solution of the first subproblem ($k = 1$). The solution of the first subproblem provides a good approximation of the decomposed materials but images are smooth given the large value of α . Further Bregman iterations ($k = 25$) are needed to find the optimal solution (bottom row in figure 8).

We evaluated two different strategies for the initialization of each subproblem (17) for GNB (figure 9). Using the solution found in the previous subproblem as initial guess in the next subproblem, i.e., setting $a^{k,n=0} = a^{k-1,N_{k-1}}$, led to an optimal solution in $N_{\text{tot}} = 120$ iterations for $\alpha = 100$. Initializing each Bregman iteration to a zero image, i.e., setting $a^{k,n=0} = 0$, led to larger solution error and diverged quickly after few updates of the Bregman distance.

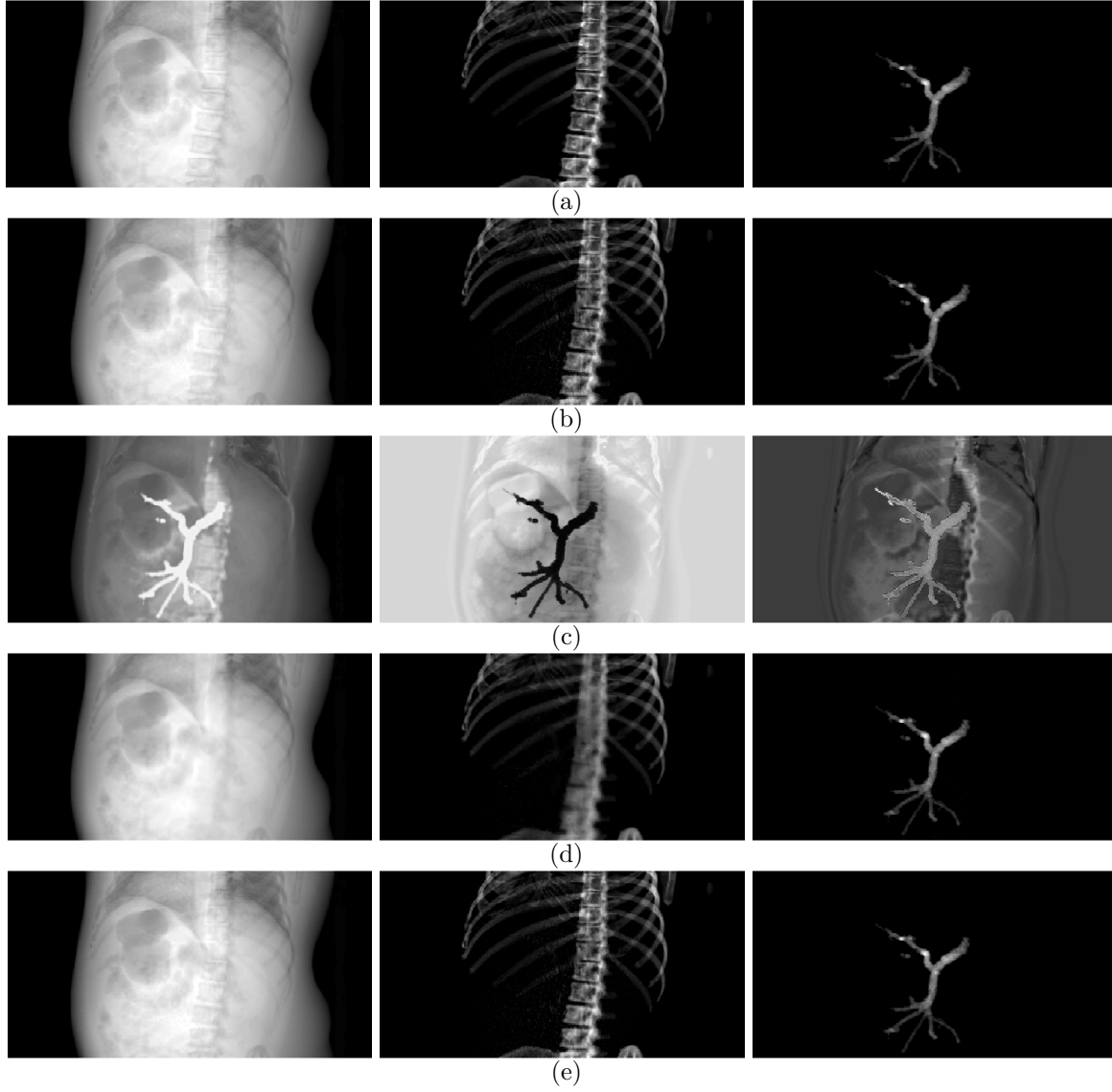


Figure 8. (a) Reference image for the thorax phantom comprising three materials: soft tissue, bone and portal vein marked with gadolinium. A particular projection view at angle of 60° is considered. (b) Material images decomposed by GN for $\tilde{\alpha} = 0.23$ and $a^0 = 0$. (c)-(e) Material images decomposed by GNB at different iterations: (c) $k = 1$, $n = 5$, $N_{\text{tot}} = 5$, (d) $k = 1$, $n = 15$, $N_{\text{tot}} = 15$ and (e) $k = 25$, $n = 1$, $N_{\text{tot}} = 44$, which is the optimal decomposition with GNB. k indicates the (outer) Bregman iteration, n the (inner) Gauss Newton iterations, and N_{tot} is the total number of Gauss-Newton iterations. We set $\alpha = 10$, $\kappa = 1 \cdot 10^{-6}$ and $a^0 = 0$.

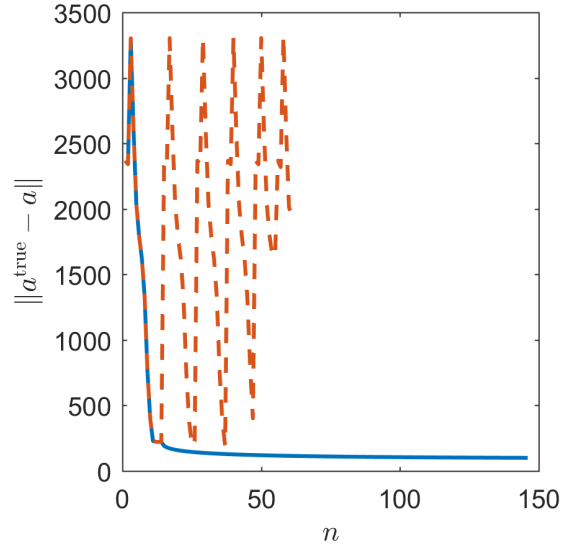


Figure 9. Influence of the initialization of the Bregman iteration. The solution error is plotted for two types of initialization for the minimization of Θ_k : using the solution found at the previous Bregman iteration (full line) and using a zero image for all Bregman iteration (dotted line). We used $\alpha = 100$, $\kappa = 1 \cdot 10^{-6}$ and $a^{k=0, n=0} = 0$.

5.4. Influence of the initial guess

In previous sections the initial guess was set to a zero constant image for all materials, which belongs to the convex part of the data fidelity term. To assess now the effect of the Bregman distance on global convergence we set the initial guess to large values in order to make sure this belong to the nonconvex region ($a_0 = (10, 10, 10)$). GNB converged to the optimal solution independently of the initial guess. On the contrary, GN did not converge, stopping at the first iteration as a descent direction could not be found (figure 10).

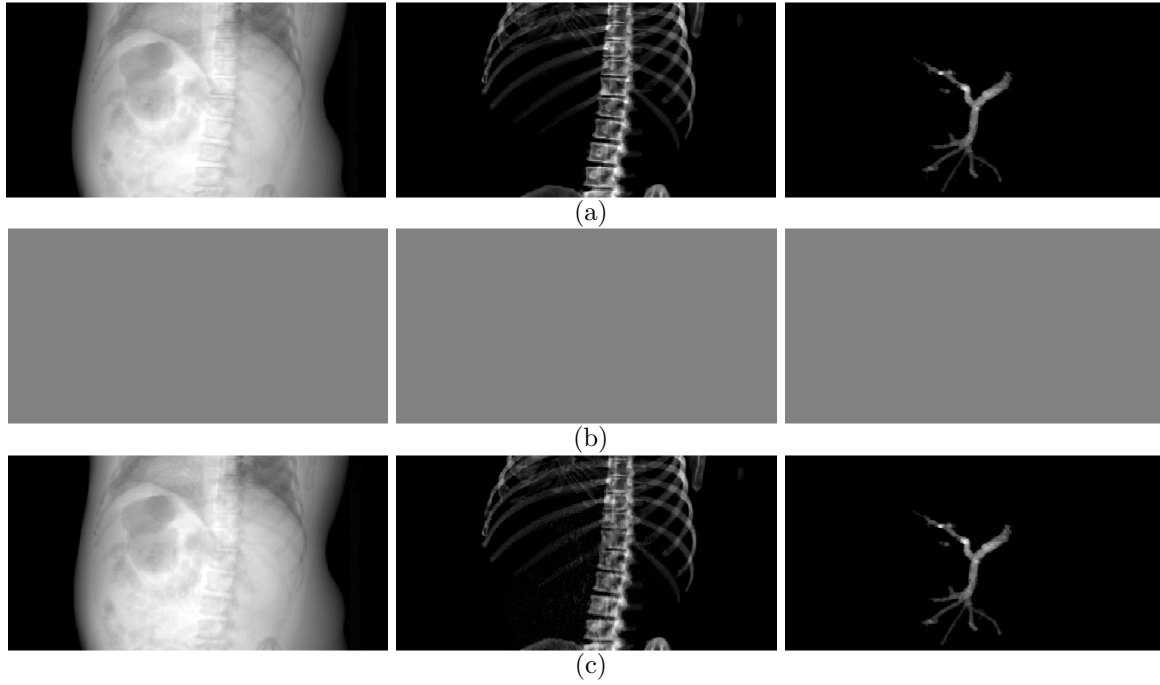


Figure 10. Influence of the initialization of the Gauss-Newton iteration. (a) Reference image and decomposed images by using (b) GN with $\tilde{\alpha} = 0.23$ and (c) GNB with $\alpha = 10$ and $\kappa = 10^{-6}$. The initial guess was $a^{k=0, n=0} = 10$ for both methods.

GNB presented a dependence on κ when the initial guess belonged to the nonconvex region. In this case, κ was critical as the Hessian was ill-conditioned, which occurred for a large value of the initial guess. We used $\kappa = 10^{-6}$. When the initial guess belonged to the convex region, GNB led to same results independently of the value of κ as long as it was small enough $\kappa \leq 10^{-4}$.

5.5. Influence of noise and phantom

In spectral CT, an important problem is the low photon count in each energy bin corresponding to noisy data. To illustrate the impact of noise on GNB, figure 11 shows decomposed material images by GNB for different number of incident photons from $n_0 = 10^8$ to $n_0 = 10^5$. The lower the number of incident photons the more regularized the solution. The tolerance θ was set to $1.3 \cdot 10^5$, $1.5 \cdot 10^5$ and $1.8 \cdot 10^5$ for $n_0 = 10^7$, $n_0 = 10^6$ and $n_0 = 10^5$, respectively.

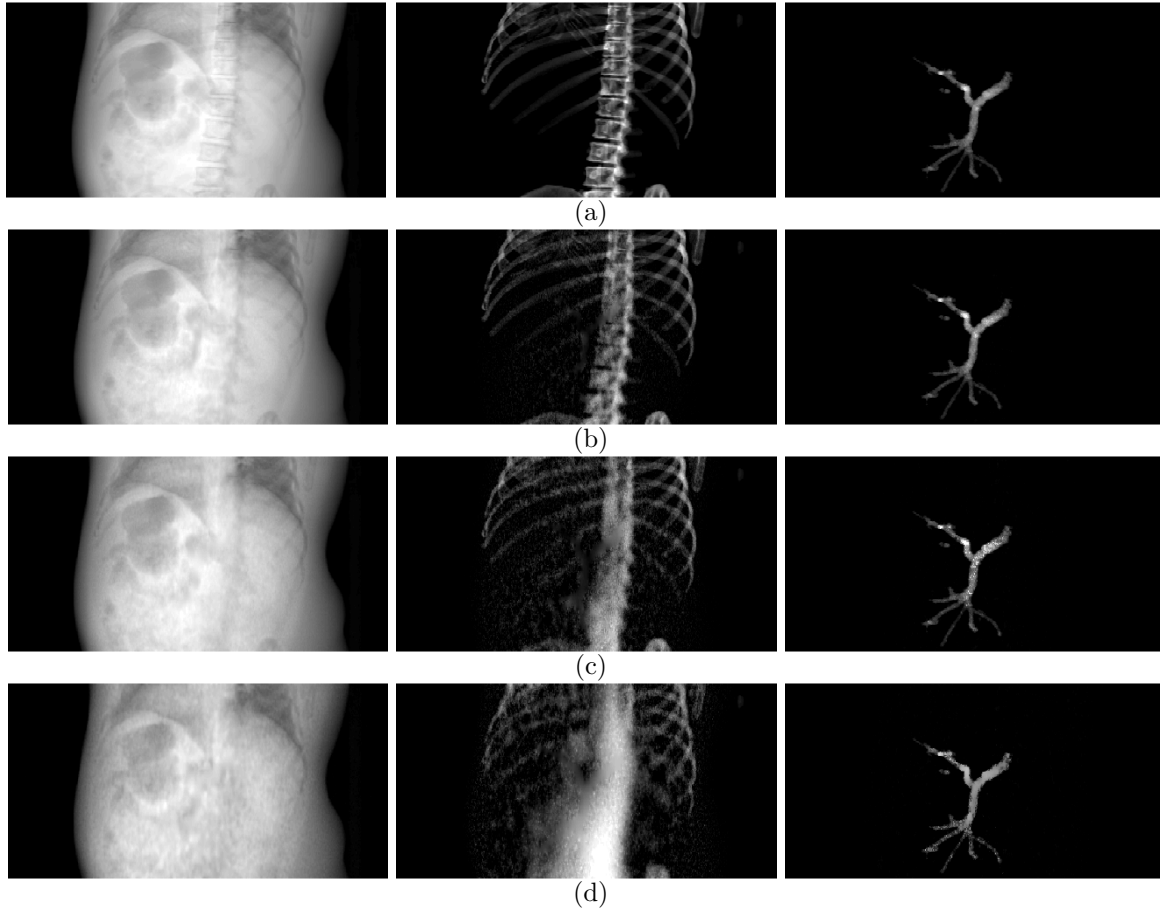


Figure 11. Material decomposed images by GNB for different number of photons n_0 . (a) Reference images and decomposed images for (b) $n_0 = 10^7$, (c) $n_0 = 10^6$ and (d) $n_0 = 10^5$. We used the following parameters: $\alpha = 10$, $\kappa = 10^{-6}$, and initial guess $a^0 = 0$.

We also show the decompositions obtained for a different projection angle (here 180°) in figure 12.

6. Discussion

We proposed and assessed a regularized iterative scheme based on the Bregman distance (GNB) for the solution of the material decomposition problem in spectral CT, which is a nonlinear inverse ill-posed problem and for which we consider the minimization of a

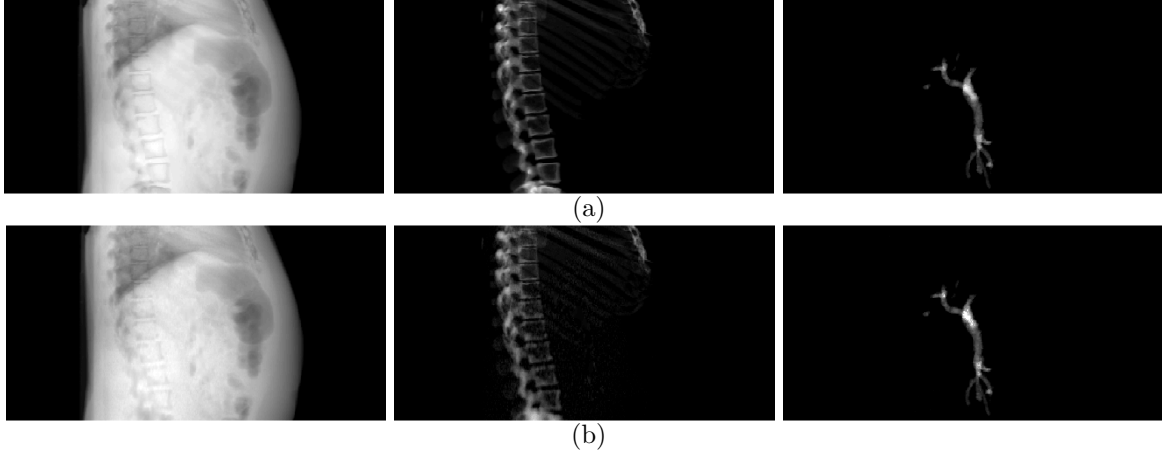


Figure 12. Material decomposed images provided for a different projection view (for an angle of 180°). (a) The ground truth. (b) Decomposed data provided by GNB in the bottom row. The measurement vector is simulated for $n_0 = 10^8$ photons and the decomposition is performed using the following parameters: $\alpha = 10$, $\kappa = 10^{-6}$, and initial guess $a=0$.

regularized functional. First, we proved that the problem is convex in a region of the solution space and that the zero solution belongs to this region. Then, using numerical examples we showed that the problem is nonconvex for large values of the decomposed materials.

Using a numerical thorax phantom, we showed that zero initialization is a suitable initialization strategy for convex optimization methods such as a regularized Gauss-Newton algorithm method [37]. GN led to an optimal solution as long as the initial guess belonged to the convex region of the data fidelity term. On the contrary, GN led to a wrong solution when the initial guess is chosen far from the convex region.

We then assessed a regularized iterative scheme based on the Bregman distance. The proposed method has some advantages with respect to more standard approaches. It is robust to the choice of the initial guess, as it led to global minimum even when the initial guess was in the nonconvex region. On the contrary, Gauss-Newton method failed in this case. This is a relevant feature of the Bregman distance, which makes each subproblem locally convex around the current estimate. This makes Bregman iterative algorithms more robust but at the price of requiring more iterations than an equivalent Gauss-Newton method. Nevertheless, we found that the number of iterations required was not very large: around ten Gauss-Newton iterations to solve the first subproblem ($k = 1$) but generally only one iteration for $k \geq 2$. This may indicate that subproblems are approximately linear when close to the solution. Reconstructed images show that the first subproblem leads to a good estimate of the decomposed materials and then further Bregman iterations are used to improve the solution. This could be then further exploited to reduce computationally complexity, for instance, by using a linearization of the data fidelity term for $k > 1$ as suggested in [19]. In addition, standard approaches such as Gauss-Newton method requires optimal selection of the

regularization parameter using the L-curve or similar method [46, 47]. In this work, we selected the solution provided by GN method that led to the lowest solution error by assessing a large number of regularization parameters. As Bregman iterative methods do not require optimal selection of the regularization parameter, the extra iterations needed by these methods may not translate in much larger computational demand than standard regularization methods. In addition, we found a linear relation between the number of total iterations and the parameter α , which has been previously shown for linear measurement operators [7, 13]. Hence, α can be tuned to decrease the number of required iterations. The smaller α the lower the number of iterations but a very low value may be less robust.

Initialization of each Bergman iteration for GNB was critical. Using the solution found at the previous iteration led to robust convergence. On the other hand, a zero initialization strategy for each subproblem required a large number of iterations to solve each subproblem and became unstable as the number of Bregman iterations increased. This can be explained by the fact that the problem is nonlinear and that as the number of Bregman iterations increases, the solution is further away from the zero initial guess. Nevertheless, for linear problems, zero initialization was a valid strategy for Bregman iterative methods in a wide variety of applications [13, 42].

In [19] authors included an extra ℓ_2 -term (weighted by κ in (17)) to make the functional associated to total variation convex on each subproblem. Here we found a dependence on κ . When the initial guess was in the nonconvex region, this term was essential to reach to a global minimum. An optimal solution was found by using a sufficiently small value of κ ($\kappa \leq 10^{-4}$). The effect of κ can be explained by the fact that when the initial guess is in the nonconvex region the Hessian is ill-conditioned, where the nonconvex region in this problem corresponds to large values of the materials. For the case where the initial guess was in the convex region, this term did not play a role and any small value of κ (including $\kappa = 0$) led to the same solution.

In conclusion, we have studied convexity for the material decomposition problem and showed that an iterative scheme based on the Bregman distance had a positive effect to find a global minimum. In addition to this, the Bregman distance presents other benefits. For instance, it can efficiently minimize nonsmooth functional such as total variation, which has been widely exploited for linear measurements [7]. Another benefit of using the Bregman distance is to correct for the lost of contrast that affects total variation [9]. This is crucial for quantitative imaging techniques such as spectral CT. In future work we will investigate these other benefits of the Bregman distance for solving the material decomposition problem in spectral CT.

Acknowledgments

Authors thank Bernhard Brendel and Ewald Roessl for fruitful discussions and suggestions on the material decomposition problem.

This project has received funding from the European Union's Horizon 2020 research

and innovation programme under the Marie Skłodowska-Curie grant agreement N° 701915. We also acknowledge the support of the ANR project SALTO (ANR-17-CE19-0011-01). It was also performed within the framework of the LabEx PRIMES (ANR-11-LABX-0063) of University de Lyon. This project has received funding from the EU's H2020 research and innovation program under the grant agreement N° 633937.

Appendix A. Hessian of the forward mapping

Let $H_i(a)$ be the Hessian matrix of the operator \mathcal{F}_i . It is defined as

$$(H_i)_{m,m'} = \frac{\partial^2 \mathcal{F}_i}{\partial a_m \partial a_{m'}}. \quad (\text{A.1})$$

The operator \mathcal{F}_i is continuously twice differentiable and its second derivative is given by

$$\frac{\partial^2 \mathcal{F}_i}{\partial a_m \partial a_{m'}}(a) = \int_{\mathcal{E}} \tau_m(E) \tau_{m'}(E) g_i(a, E) \, dE \quad (\text{A.2})$$

where $g_i(a, E)$ is given by (9). Discretising the energy interval $\mathcal{E} = [0, E_{\max}]$ with an energy step ΔE , the integral (A.2) can be approximated by

$$\frac{\partial^2 \mathcal{F}_i}{\partial a_m \partial a_{m'}}(a) = \sum_j \tau_m(E_j) \tau_{m'}(E_j) g_i(a, E_j) \Delta E. \quad (\text{A.3})$$

Hence, it can be noted that the Hessian takes the following matrix form

$$H_i(a) = T^\top G_i(a) T, \quad (\text{A.4})$$

where $G_i(a) = \text{diag}(g_i(a, E_1), \dots, g_i(a, E_J)) \Delta E$ is the vector of photons incident on the detector, and $T = (\tau_{j,m}) \in \mathbb{R}^{J \times M}$ is the mass attenuation matrix defined by $\tau_{j,m} = \tau_m(E_j)$.

References

- [1] Ronny Ramlau and Gerd Teschke. Tikhonov replacement functionals for iteratively solving nonlinear operator equations. *Inverse Probl.*, 21(5):1571–1592, 2005.
- [2] Ronny Ramlau and Gerd Teschke. A Tikhonov-based projection iteration for nonlinear ill-posed problems with sparsity constraints. *Numer. Math.*, 104(2):177–203, 2006.
- [3] Stanley Osher, Martin Burger, Donald Goldfarb, Jinjun Xu, and Wotao Yin. An Iterative Regularization Method for Total Variation-Based Image Restoration. *Multiscale Modeling & Simulation*, 4(2):460–489, 2005.
- [4] Wotao Yin, Stanley Osher, Donald Goldfarb, and Jerome Darbon. Bregman Iterative Algorithms for l1-Minimization with Applications to Compressed Sensing. *SIAM J. Imaging Sci.*, 1(1):143–168, 2008.
- [5] Jian-Feng Cai, Stanley Osher, and Zuowei Shen. Linearized Bregman iterations for compressed sensing. *Math. Comput.*, 78(267):1515–1536, 2009.

- [6] Jian-Feng Cai, Stanley Osher, and Zuowei Shen. Split Bregman Methods and Frame Based Image Restoration. *Multiscale Modeling & Simulation*, 8(2):337–369, 2010.
- [7] Tom Goldstein and Stanley Osher. The Split Bregman Method for L1-Regularized Problems. *SIAM J. Imaging. Sci.*, 2(2):323–343, 2009.
- [8] Stanley Osher, Yu Mao, Bin Dong, and Wotao Yin. Fast linearized Bregman iteration for compressive sensing and sparse denoising. *Commun Math Sci.*, 8(1):93–111, 2010.
- [9] Martin Burger and Stanley Osher. A Guide to the TV Zoo. *Lecture Notes in Mathematics*, 2090:1–70, 2013.
- [10] Erny Esser. Applications of Lagrangian-based alternating direction methods and connections to split Bregman. Technical report, 2009.
- [11] Chunlin Wu and Xue-Cheng Tai. Augmented Lagrangian Method, Dual Methods, and Split Bregman Iteration for ROF, Vectorial TV, and High Order Models. *SIAM J. Imaging. Sci.*, 3(3):300–339, 2010.
- [12] S Setzer, G Steidl, and T Teuber. Deblurring Poissonian images by split Bregman techniques. *J. Vis. Commun. Image Represent.*, 21(3):193–199, 2010.
- [13] J F P J Abascal, J Chamorro-Servent, J Aguirre, S Arridge, T Correia, J Ripoll, J J Vaquero, and M Desco. Fluorescence diffuse optical tomography using the split Bregman method. *Med. Phys.*, 38(11):6275–6284, 2011.
- [14] Hao Gao, Hengyong Yu, Stanley Osher, and Ge Wang. Multi-energy CT based on a prior rank, intensity and sparsity model (PRISM). *Inverse Probl.*, 27(11):115012, 2011.
- [15] Gerlind Plonka and Jianwei Ma. Curvelet-Wavelet Regularized Split Bregman Iteration for Compressed Sensing. *Int. J. Wavelets Multiresolution*, 09(01):79–110, 2011.
- [16] J F P J Abascal, P Montesinos, E Marinetto, J Pascau, and M Desco. Comparison of total variation with a motion estimation based compressed sensing approach for self-gated cardiac cine MRI in small animal studies. *PLoS One*, 9(10), 2014.
- [17] Juan F P J Abascal, Manuel Desco, and Juan Parra-Robles. Incorporation of prior knowledge of the signal behavior into the reconstruction to accelerate the acquisition of MR diffusion data. *IEEE Trans. Med. Imaging*, (in press), 2017.
- [18] A Gholami, H R Siahkoochi, W Navidi, L Tenorio, and W Yin. Regularization of linear and non-linear geophysical ill-posed problems with joint sparsity constraints. *Geophys. J. Int.*, 180(2):871–882, 2010.
- [19] Markus Bachmayr and Martin Burger. Iterative total variation schemes for nonlinear inverse problems. *Inverse Probl.*, 25(10):105004, 2009.
- [20] Katsuyuki Taguchi and Jan S Iwanczyk. Vision 20/20: Single photon counting x-ray detectors in medical imaging. *Med. Phys.*, 40(10):100901, 2013.
- [21] Bjorn J Heismann, Bernhard T Schmidt, and Thomas Flohr. *Spectral computed tomography*. SPIE, 2012.
- [22] Huanjun Ding, Bo Zhao, Pavlo Baturin, Farnaz Behroozi, and Sabee Molloy. Breast tissue decomposition with spectral distortion correction: A postmortem study. *Med. Phys.*, 41(10):101901, 2014.
- [23] David P Cormode, Ewald Roessl, Axel Thran, Torjus Skajaa, Ronald E Gordon, Jens-Peter Schlomka, Valentin Fuster, Edward A Fisher, Willem J M Mulder, Roland Proksa, and Zahi A Fayad. Atherosclerotic Plaque Composition: Analysis with Multicolor CT and Targeted Gold Nanoparticles1. *Radiology*, 256(3):774–782, 2010.
- [24] L Boussel, P Coulon, A Thran, E Roessl, G Martens, M Sigovan, and P Douek. Photon counting spectral CT component analysis of coronary artery atherosclerotic plaque samples. *Br. J. Radiol.*, 87(1040):20130798, 2014.
- [25] J P Schlomka, E Roessl, R Dorscheid, S Dill, G Martens, T Istel, C BÄumer, C Herrmann, R Steadman, G Zeitler, A Livne, and R Proksa. Experimental feasibility of multi-energy photon-counting K-edge imaging in pre-clinical computed tomography. *Phys. Med. Biol.*, 53(15):4031, 2008.

- [26] Bo Zhao, Hao Gao, Huanjun Ding, and Sabee Molloy. Tight-frame based iterative image reconstruction for spectral breast CT. *Med. Phys.*, 40(3):031905, 2013.
- [27] Y Long and J A Fessler. Multi-Material Decomposition Using Statistical Image Reconstruction for Spectral CT. *IEEE Trans. Med. Imaging*, 33(8):1614–1626, 2014.
- [28] Darin P Clark and Cristian T Badea. Spectral diffusion: an algorithm for robust material decomposition of spectral CT data. *Phys. Med. Biol.*, 59(21):6445–6466, 2014.
- [29] Mats Persson, Ben Huber, Staffan Karlsson, Xuejin Liu, Han Chen, Cheng Xu, Moa Yveborg, Hans Bornefalk, and Mats Danielsson. Energy-resolved CT imaging with a photon-counting silicon-strip detector. *Phys. Med. Biol.*, 59(22):6709, 2014.
- [30] Kyungsang Kim, Jong Chul Ye, William Worstell, Jinsong Ouyang, Yothin Rakvongthai, Georges El Fakhri, and Quanzheng Li. Sparse-view spectral {CT} reconstruction using spectral patch-based low-rank penalty. *IEEE Trans. Med. Imaging*, 34(3):748–760, 2015.
- [31] Mengheng Touch, Darin P Clark, William Barber, and Cristian T Badea. A neural network-based method for spectral distortion correction in photon counting x-ray CT. *Phys. Med. Biol.*, 61(16):6132, 2016.
- [32] O Semerci, Ning Hao, M E Kilmer, and E L Miller. Tensor-Based Formulation and Nuclear Norm Regularization for Multienergy Computed Tomography. *IEEE Trans. Image Process.*, 23(4):1678–1693, 2014.
- [33] Yan Xi, Yi Chen, Rongbiao Tang, Jianqi Sun, and Jun Zhao. United Iterative Reconstruction for Spectral Computed Tomography. *IEEE Trans. Med. Imaging*, 34(3):769–778, 2015.
- [34] David S Rigie and Patrick J La Rivière. Joint reconstruction of multi-channel, spectral CT data via constrained total nuclear variation minimization. *Phys. Med. Biol.*, 60(5):1741–1762, 2015.
- [35] Jiulong Liu and Hao Gao. Material reconstruction for spectral computed tomography with detector response function. *Inverse Probl.*, 32(11):114001, 2016.
- [36] Bernhard Brendel, Frank Bergner, Kevin Brown, and Thomas Koehle. Penalized likelihood decomposition for dual layer spectral CT. In *Proc. 4th Intl. Mtg. on image formation in X-ray CT*, 2016.
- [37] Nicolas Ducros, Juan Felipe Perez-Juste Abascal, Bruno Sixou, Simon Rit, and Françoise Peyrin. Regularization of nonlinear decomposition of spectral x-ray projection images. *Med. Phys.*, 44(9):e174–e187, 2017.
- [38] Jorge Nocedal and Stephen J Wright. *Numerical optimization*. Springer, 2006.
- [39] Simon Arridge and John Schotland. Optical tomography: forward and inverse problems. *Inverse Probl.*, 25(12):123010, 2009.
- [40] S A Vavasis. *Nonlinear Optimization*. Princeton University Press, 1991.
- [41] L M Bregman. The relaxation method of finding the common point of convex sets and its application to the solution of problems in convex programming. *USSR Comput. Math.*, 7(3):200–217, 1967.
- [42] J F P J Abascal, M Abella, E Marinetto, J Pascau, and M Desco. A novel prior- and motion-based compressed sensing method for small-animal respiratory gated CT. *PLoS One*, 11(3), 2016.
- [43] Markus Grasmair. Generalized Bregman distances and convergence rates for non-convex regularization methods. *Inverse Probl.*, 26(11):115014, 2010.
- [44] 3D Image Reconstruction for Comparison of Algorithm Database.
- [45] R Kéchichian, S Valette, M Desvignes, and R Prost. Shortest-Path Constraints for 3D Multiobject Semiautomatic Segmentation Via Clustering and Graph Cut. *IEEE Trans. Image Process.*, 22(11):4224–4236, 2013.
- [46] Per Christian Hansen and Dianne Prost O’Leary. The Use of the L-Curve in the Regularization of Discrete Ill-Posed Problems. *SIAM J. Sci. Comput.*, 14(6):1487–1503, 1993.
- [47] J F P J Abascal, S R Arridge, R H Bayford, and D S Holder. Comparison of methods for optimal choice of the regularization parameter for linear electrical impedance tomography of brain function. *Physiol. Meas.*, 29(11), 2008.