



**HAL**  
open science

# Error-Splitting Forward Model for Iterative Reconstruction in X-ray Computed Tomography and application with Gauss-Markov-Potts prior

Camille Chapdelaine, Ali Mohammad-Djafari, Nicolas Gac, Estelle Parra

► **To cite this version:**

Camille Chapdelaine, Ali Mohammad-Djafari, Nicolas Gac, Estelle Parra. Error-Splitting Forward Model for Iterative Reconstruction in X-ray Computed Tomography and application with Gauss-Markov-Potts prior. IEEE Transactions on Computational Imaging, 2019, 5 (2), pp.317 - 332. 10.1109/TCI.2018.2885432 . hal-01948680

**HAL Id: hal-01948680**

**<https://hal.science/hal-01948680>**

Submitted on 19 Jan 2020

**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.

# Error-Splitting Forward Model for Iterative Reconstruction in X-ray Computed Tomography and application with Gauss-Markov-Potts prior

Camille Chapdelaine, Ali Mohammad-Djafari, Nicolas Gac, and Estelle Parra

**Abstract**—In order to enhance image quality for controlling the interior of a volume in industry, model-based iterative reconstruction (MBIR) methods in 3D X-ray Computed Tomography (CT) have shown their efficiency compared to analytical reconstruction methods. MBIR methods enforce a prior model on the volume to reconstruct and make fusion of the information contained in the prior model and the projection data. The projections have many uncertainties which have very different origins in 3D CT. They are taken into account in MBIR methods but, despite of their different origins, they are in general gathered in only one term in the forward model. In this paper, we propose to derive a new forward model by adding a further error term in the Poisson statistics of photon-count, corresponding to the deviation of the monochromatic model with respect to the actual polychromacy of the rays. A Taylor expansion of the Poisson log-likelihood leads us to a new algebraic forward model accounting for two terms of uncertainties : we call it the error-splitting forward model. Different prior models are assigned to each of distinguished uncertainties : measurement uncertainties are modeled as Gaussian, while linear model uncertainties are modeled as heavy-tailed to bring robustness to the reconstruction process. We give strategies to fix the parameters of the error-splitting forward model. Next, we use it in a full MBIR method with Gauss-Markov-Potts prior model on the volume, in order to reconstruct piecewise-constant objects for non-destructive testing in industry. Compared to the conventional forward model, we show in our experiments that the use of the error-splitting forward model with Gauss-Markov-Potts prior combines better robustness and accuracy.

**Index Terms**—3D Computed Tomography, non-destructive testing, iterative reconstruction, forward model, error-splitting, Bayesian approach, Gauss-Markov-Potts.

## I. INTRODUCTION

NON-destructive testing (NDT) in industry relies on X-ray Computed Tomography (CT) for virtually reconstructing a volume and assessing its healthiness. The reconstruction is performed by sending X-rays through the volume from a source to a detector. According to Beer-Lambert's law [1], the intensity of the rays decreases depending on the materials encountered in the volume. The measurements of this decrease in a logarithmic scale are called the projections. The

projections are modeled as the X-ray transform (XT) of the object. Under some strict conditions, the XT can be inverted to perform the reconstruction [2].

Nevertheless, inverting XT is an ill-conditioned problem [3]. In order to obtain usable reconstructions for NDT, filtered backprojection (FBP) methods, such as FDK [4], need a large amount of complete projections. Since the 1970s, algebraic reconstruction techniques have been developed based on the discretized expression of XT

$$\mathbf{g} = \mathbf{H}\mathbf{f} + \boldsymbol{\zeta} \quad (1)$$

where  $\mathbf{g} \in \mathbb{R}^M$  contains all the projections,  $\mathbf{f} \in \mathbb{R}^N$  contains the voxels' attenuation coefficients of the volume and  $\boldsymbol{\zeta}$  is a term accounting for uncertainties.  $\mathbf{H}$  is the projection operator of which each coefficient  $H_{ij}$  gives the intersecting length between ray  $i$  and voxel  $j$  [5] according to XT forward model. Its adjoint  $\mathbf{H}^T$  is called the backprojection operator. First algebraic reconstruction methods such as ART [6], SIRT [7] or SART [8], try to match the data  $\mathbf{g}$  and the output  $\mathbf{H}\mathbf{f}$  of the forward model. As a result, they are very sensitive to the noise. The actual strength of algebraic reconstruction techniques lies in their ability to enforce a prior model  $\mathcal{M}$  on the volume to reconstruct in order to regularize the reconstruction problem [3] : this leads to model-based iterative reconstruction (MBIR) methods. Compared to FBP methods, MBIR methods have the advantage to significantly enhance the reconstruction quality particularly when the number of projections is low or when their angle distribution is limited [9].

Many prior models have been proposed in order to take advantage from the specificities of the volume to reconstruct. For retrieving a reconstruction divided in compact and smooth regions, the volume can be modeled as having minimum total variation (TV) [10]. In 2D and 3D, joint reconstruction and segmentation using a Potts model [11] have been considered [12]–[14]. Other prior models focus on sparsity in a wavelet domain [15], [16] or in a learnt dictionary representation [17], [18]. Prior model  $\mathcal{M}$  very often has hyperparameters  $\boldsymbol{\theta}$  which can be difficult to tune. Bayesian MBIR methods enable to estimate optimal hyperparameters jointly with the volume

$$p(\mathbf{f}, \boldsymbol{\theta} | \mathcal{M}) = p(\mathbf{f} | \boldsymbol{\theta}; \mathcal{M})p(\boldsymbol{\theta} | \mathcal{M}) \quad (2)$$

as it is done in [14], [16].

Due to the huge dimension of the data and the unknowns, the main computational burden of MBIR methods lies in projection and backprojection operations  $\mathbf{H}$  and  $\mathbf{H}^T$ . This computational cost can be compensated by the use of parallel

C. Chapdelaine is with Laboratoire des Signaux et des Systèmes (L2S) (CNRS–CentraleSupélec–Université Paris-Sud–Université Paris-Saclay) and SAFRAN SA, Safran Tech, Pôle Technologie du Signal et de l'Information (e-mail : camille.chapdelaine@safrangroup.com). A. Mohammad-Djafari and N. Gac are with Laboratoire des Signaux et des Systèmes (L2S) (CNRS–CentraleSupélec–Université Paris-Sud–Université Paris-Saclay). E. Parra is with SAFRAN SA, Safran Tech, Pôle Technologie du Signal et de l'Information. This work is supported by CIFRE grant n° 2016/0188 from French Agence Nationale de la Recherche et de la Technologie (ANRT).

computing on Graphical Processor Units (GPU) [19] and momentum techniques [20], [21]. Efficient models for  $\mathbf{H}$  and  $\mathbf{H}^T$  are a key aspect to ensure speed and accuracy of MBIR methods [22], [23]. The reliability of the forward model (1) also depends on the prior model for uncertainties  $\zeta$ . To derive the prior for uncertainties  $\zeta_i$  for each ray  $i$  modeled as monochromatic, the expression of the Poisson statistics of photon count has been considered as [20] :

$$p(I_i|\mathbf{f}, r_i) = \mathcal{P}\left(I_0 e^{-[\mathbf{H}\mathbf{f}]_i} + r_i\right). \quad (3)$$

In (3),  $I_0$  denotes the mean number of photons sent by the source and  $r_i$  denotes the mean number of background events [20]. These background events can be for instance Compton scatter interactions undergone by X-ray photons inside the detector [24]. The background events mean  $r_i$  is precomputed before the acquisition. It is taken into account in the linearization of the reconstruction problem :

$$g_i = \ln\left(\frac{I_0}{I_i - r_i}\right) \forall i. \quad (4)$$

As a result, uncertainties  $\zeta$  can be shown to be well modeled by a Gaussian distribution

$$p(\zeta_i|v_{\zeta_i}) = \mathcal{N}(\zeta_i|0, v_{\zeta_i}), \forall i. \quad (5)$$

thanks to a second order Taylor series expansion of the Poisson statistics (3) of photons counts [25]–[27]. Inverses of variances  $v_{\zeta}$  weight the data and translate how the actual physical phenomena occurring in each measurement are far from or near the ideal physics described by Beer-Lambert's law. These variances can be pre-computed [27] or optimized during the reconstruction [14], [16].

Although the commonly-used forward model (1) has given good results in 3D X-ray CT when combined with a prior model on the volume, the expression of the mean of photon counts in equation (3) does not take into account all the phenomena involved in 3D X-ray CT. Further phenomena correspond to object-dependent uncertainties, such as scattering inside the volume, but also absorption of low energy photons occurring in presence of some materials such as metal [28]. These further phenomena can be taken into account by considering the actual polychromacy of X-rays. The polychromatic model is obtained by integrating the attenuation term in (3) over all the energy spectrum :

$$p(I_i|\mathbf{f}(E), r_i) = \mathcal{P}\left(\int_E I_0(E) e^{-[\mathbf{H}\mathbf{f}(E)]_i} dE + r_i\right). \quad (6)$$

Based on this model, MBIR methods aiming at compensating beam-hardening address the reconstruction problem by estimating the density of material in each voxel, which is part of attenuation coefficients' expression [29], [30]. This estimation is preferred to the one of attenuation coefficients because density of material only depends on voxel location [29], [30] : hence, the estimation of one reconstruction per energy is avoided. Nevertheless, the integration in polychromatic model (6) hinders the linearization of the reconstruction problem done for the monochromatic model by formula (4). Consequently, the use of polychromatic model (6) adds complexity to MBIR methods which are already computationally-demanding

due to projection and backprojection operations. For this reason, even recently, polychromatic model (6) has only been used for 2D reconstruction problems [29], [30].

In order to tackle this problem for reconstructing 3D images for NDT by MBIR methods, in this paper, we propose to make a tradeoff between the accuracy of polychromatic model (6) and the simplicity of monochromatic model (3). Since monochromatic model (3) is an approximation of polychromatic model (6), we suggest to add an unknown error term  $n_i$  in monochromatic model (3), which corresponds to the approximation error with respect to the polychromatic model :

$$p(I_i|\mathbf{f}, n_i, r_i) = \mathcal{P}\left(I_0 e^{-[\mathbf{H}\mathbf{f}]_i} + n_i + r_i\right), \forall i. \quad (7)$$

Furthermore, unknown error  $n_i$  enables to handle other physical phenomena described in [24], which would be too complex to model in a MBIR method.

Based on expression (7), we propose to derive a new linear forward model accounting for uncertainties, different from (1), in order to use it in a full MBIR method and to compare results with those obtained with the usual forward model (1). In section II, we perform a second order Taylor series expansion from Poisson statistics (7), in order to show that an error-splitting forward model of the form

$$\begin{cases} \mathbf{g} = \mathbf{g}_0 + \boldsymbol{\epsilon} \\ \mathbf{g}_0 = \mathbf{H}\mathbf{f} + \boldsymbol{\xi} \end{cases} \quad (8)$$

can be used for iterative reconstruction in 3D X-ray CT. Next, we assign different priors to uncertainties  $\boldsymbol{\epsilon}$  and  $\boldsymbol{\xi}$ . Uncertainties  $\boldsymbol{\epsilon}$  are called the measurement uncertainties and are modeled as Gaussian, while uncertainties  $\boldsymbol{\xi}$  are called linear model uncertainties and are more difficult to model. In order to introduce robustness with respect to outliers, we assign a heavy-tailed prior to  $\boldsymbol{\xi}$ , which is in this paper a generalized Student- $t$  distribution [31]. Both priors for  $\boldsymbol{\epsilon}$  and  $\boldsymbol{\xi}$  introduce parameters about which we give strategies to fix them. Then, in section III, we propose a general MBIR algorithm for maximizing the joint posterior distribution of the unknowns and compare its cost to Joint Maximization A Posteriori (JMAP) with usual forward model (1). To implement this algorithm, conditional distributions of the unknowns in the error-splitting forward model are derived. The proposed algorithm is valid for any prior model on the volume. In order to reconstruct piecewise-constant volumes for NDT in industry, we propose then to combine the error-splitting forward model and Gauss-Markov-Potts prior model on the volume, as described in section IV. We present results on real data in section V and conclusions in section VI. The results compare the uses of usual forward model (1) and of the proposed forward model, with Gauss-Markov-Potts prior model on the volume. Compared to usual forward model (1), we show that, with Gauss-Markov-Potts prior, the use of the error-splitting forward model leads to more robust and accurate reconstructions.

## II. ERROR-SPLITTING FORWARD MODEL

### A. Derivation of the error-splitting forward model

We consider a cone-beam acquisition process shown in figure 1 : a volume  $\mathbf{f}$  discretized in  $N$  voxels is placed

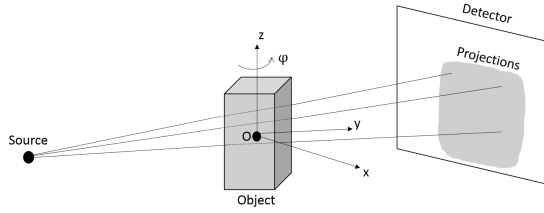


Fig. 1: Cone beam 3D X-ray acquisition of the projections

between a source and a detector so X-rays sent from the source cross the volume. After passing through the volume, the rays reach the cells of the detector which measure their intensities. These measurements form an image with  $N_u \times N_v$  pixels corresponding to the cells of the detector. The object is rotated around  $z$ -axis by a projection angle  $\phi$  in order to acquire  $N_\phi$  projections of the object.

The photon count  $I_i$  for projection  $i$  is modeled as a Poisson process of which the statistics is given by equation (7), where  $r_i$  is the mean of background events, and  $n_i$  is the approximation error induced by the monochromatic model with respect to polychromatic model (6). Contrary to background events mean  $r_i$  which is precomputed, error  $n_i$  is unknown because of its dependance on both the acquisition and controlled volume  $\mathbf{f}$ . Thus, error  $n_i$  cannot be pre-computed and taken into account in the linearization of the reconstruction problem in (4). We have

$$-\ln(p(I_i|\mathbf{f}, n_i, r_i)) = \ln(I_i!) + I_0 e^{-[\mathbf{H}\mathbf{f}]_i} + n_i + r_i - I_i \ln \left( I_0 e^{-[\mathbf{H}\mathbf{f}]_i} + n_i + r_i \right). \quad (9)$$

After reparametrization (4), a quadratic approximation from second order Taylor series expansion of (9) is derived in appendix A. It leads to

$$-\ln(p(g_i|\mathbf{f}, n_i, r_i)) \approx \ln(I_i!) + I_i + n_i - I_i \ln(I_i + n_i) + \frac{a_i}{2} \left( g_i - [\mathbf{H}\mathbf{f}]_i + \frac{b_i}{a_i} \right)^2 - \frac{b_i^2}{2a_i} \quad (10)$$

where

$$\begin{cases} a_i = (I_i - r_i) \left( 1 - \frac{I_i(r_i + n_i)}{(I_i + n_i)^2} \right) \\ b_i = \frac{n_i(I_i - r_i)}{I_i + n_i} \end{cases}. \quad (11)$$

By denoting

$$\xi_i = -\frac{b_i}{a_i} = -\frac{n_i(I_i + n_i)}{(I_i + n_i)^2 - I_i(r_i + n_i)} \quad (12)$$

and

$$v_{\epsilon_i} = \frac{1}{a_i}, \forall i, \quad (13)$$

we have the following expression for the log-likelihood

$$\ln(p(\mathbf{g}|\mathbf{f}, \mathbf{n}, \mathbf{r})) = C - \frac{1}{2} \|\mathbf{g} - \mathbf{H}\mathbf{f} - \boldsymbol{\xi}\|_{\mathbf{V}_\epsilon}^2 \quad (14)$$

where  $\mathbf{V}_\epsilon = \text{diag}[v_\epsilon]$ ,  $\mathbf{v}_\epsilon = (v_{\epsilon_i})_i$  and

$$C = -\sum_{i=1}^M \left[ \ln(I_i!) + I_i + n_i - I_i \ln(I_i + n_i) - \frac{b_i^2}{2a_i} \right] \quad (15)$$

does not include  $\mathbf{f}$ . In the log-likelihood, term  $\boldsymbol{\xi}$  appears in the data-fidelity term as a bias which is unknown due to error  $n_i$ . By introducing intermediate projections

$$g_{0_i} = [\mathbf{H}\mathbf{f}]_i + \xi_i, \forall i, \quad (16)$$

we are able to write a new forward model

$$\begin{cases} g_i = g_{0_i} + \epsilon_i & \text{with } p(\epsilon_i|v_{\epsilon_i}) = \mathcal{N}(\epsilon_i|0, v_{\epsilon_i}), \forall i \\ g_{0_i} = [\mathbf{H}\mathbf{f}]_i + \xi_i \end{cases} \quad (17)$$

which leads to the same log-likelihood than (14), up to an additive constant. The vector form of this forward model is given by equation (8) : we see that it splits the uncertainties into two terms :  $\epsilon$  and  $\boldsymbol{\xi}$ . That is why we call it the *error-splitting forward model*. Since uncertainties  $(\epsilon_i)_i$  are related to measurements  $g_i$ , we call  $\epsilon$  the *measurement uncertainties*. Concerning uncertainties  $(\xi_i)_i$ , since it corresponds in (14) to a bias between actual and theoretical projections  $\mathbf{g}$  and  $\mathbf{H}\mathbf{f}$ , we call  $\boldsymbol{\xi}$  the *linear model uncertainties*.

In sections II-B and II-C, we deal with which priors to assign to uncertainties  $\epsilon$  and  $\boldsymbol{\xi}$ .

### B. Measurement uncertainties

As we have seen in section II-A, the measurement uncertainties  $(\epsilon_i)_i$  in

$$g_i = g_{0_i} + \epsilon_i \quad (18)$$

are zero-mean Gaussian :

$$p(\epsilon_i|v_{\epsilon_i}) = \mathcal{N}(\epsilon_i|0, v_{\epsilon_i}), \forall i. \quad (19)$$

with  $v_{\epsilon_i}$  given by (13). We see that  $v_{\epsilon_i}$  is the inverse of  $a_i$  which depends on  $n_i$  in (11). Since  $n_i$  is unknown, we jointly estimate  $v_{\epsilon_i}$  with the volume. In order to have analytically tractable update formulae,  $v_{\epsilon_i}$  is assigned a conjugate prior which is an Inverse-Gamma distribution

$$\begin{aligned} p(v_{\epsilon_i}|\alpha_{\epsilon_0}, \beta_{\epsilon_0}) &= \mathcal{IG}(v_{\epsilon_i}|\alpha_{\epsilon_0}, \beta_{\epsilon_0}) \\ &= \frac{\beta_{\epsilon_0}^{\alpha_{\epsilon_0}}}{\Gamma(\alpha_{\epsilon_0})} v_{\epsilon_i}^{-\alpha_{\epsilon_0}-1} \exp \left[ -\frac{\beta_{\epsilon_0}}{v_{\epsilon_i}} \right], v_{\epsilon_i} > 0, \forall i, \end{aligned} \quad (20)$$

where parameters  $\alpha_{\epsilon_0}$  and  $\beta_{\epsilon_0}$  are fixed. We will discuss about the choice of these parameters in section II-D.

### C. Linear model uncertainties

As we have seen in section II-A, linear model uncertainties  $\boldsymbol{\xi}$  are a bias which can be separated from measurement uncertainties  $\epsilon$  by introducing intermediate projections

$$\mathbf{g}_0 = \mathbf{H}\mathbf{f} + \boldsymbol{\xi}. \quad (21)$$

From

$$\mathbf{g} = \mathbf{g}_0 + \epsilon, \quad (22)$$

projections  $\mathbf{g}_0$  appear as the projections noisier by the measurement uncertainties. For this reason, we call  $\mathbf{g}_0$  the "true" projections. In order to estimate volume  $\mathbf{f}$  from  $\mathbf{g}_0$  taking  $\boldsymbol{\xi}$  into account, we need to assign a prior to linear model uncertainties  $\boldsymbol{\xi}$ . This is not obvious because, contrary to  $\epsilon$ , this prior cannot be deduced from the second order Taylor series expansion derived in section II-A. Since  $\boldsymbol{\xi}$  is difficult to

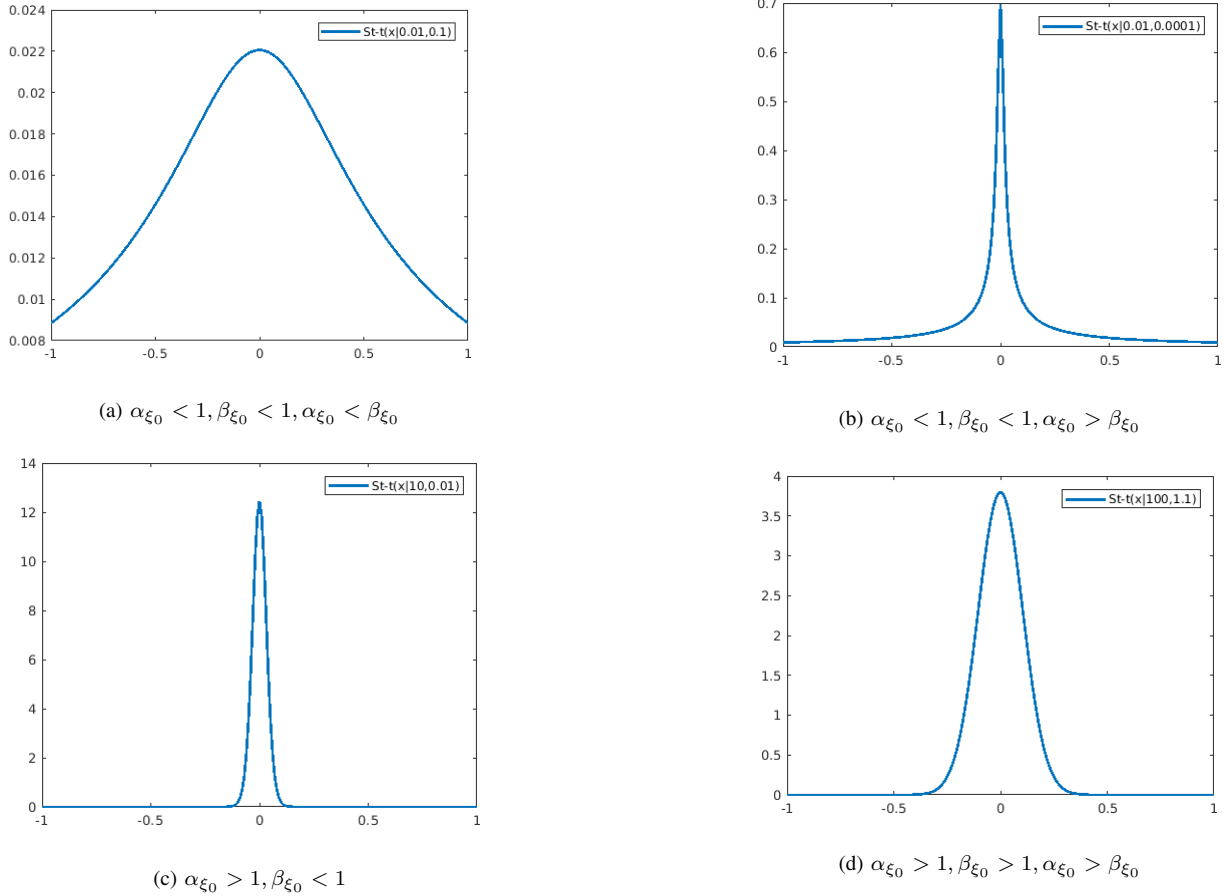


Fig. 2: Probability density function of  $St_g(\cdot|\alpha_{\xi_0}, \beta_{\xi_0})$  for several values of  $(\alpha_{\xi_0}, \beta_{\xi_0})$

physically model due to its high dependency on the controlled volume, we prefer to assign a heavy-tailed prior to it. The use of a heavy-tailed prior enables to model that possibly high  $\xi_i$  can be present in projections  $\mathbf{g}_0$ . Consequently, compared, for instance, to a Gaussian prior, it brings more robustness to the estimation of  $\mathbf{f}$  from  $\mathbf{g}_0$ .

Among heavy-tailed distributions, zero-mean normal variance mixtures have appealing properties for simple computations in linear inverse problems [32]. Among them, generalized Student- $t$  distribution  $St_g$  [31]

$$p(\xi_i|\alpha_{\xi_0}, \beta_{\xi_0}) = St_g(\xi_i|\alpha_{\xi_0}, \beta_{\xi_0}), \forall i, \quad (23)$$

corresponds to a mixture on the variances between a zero-mean Gaussian and an Inverse-Gamma distributions

$$\begin{cases} p(\xi_i|v_{\xi_i}) = \mathcal{N}(\xi_i|0, v_{\xi_i}) \\ p(v_{\xi_i}|\alpha_{\xi_0}, \beta_{\xi_0}) = \mathcal{IG}(v_{\xi_i}|\alpha_{\xi_0}, \beta_{\xi_0}) \end{cases}, \forall i. \quad (24)$$

By introducing hidden variances  $v_{\xi}$  in the reconstruction problem, generalized Student- $t$  distribution can be easily dealt with thanks to simple calculations involving normal distribution for  $\xi$  and conjugate Inverse-Gamma prior for  $v_{\xi}$ . Parameters  $\alpha_{\xi_0}$  and  $\beta_{\xi_0}$  are fixed according to the strategy presented in section II-D.

#### D. Fixing the parameters of the error-splitting forward model

From (19) and (20), since we have assigned a conjugate prior on variances  $v_{\epsilon}$ , the prior for measurement uncertainties is in fact a normal variance mixture

$$\begin{cases} p(\epsilon_i|v_{\epsilon_i}) = \mathcal{N}(\epsilon_i|0, v_{\epsilon_i}) \\ p(v_{\epsilon_i}|\alpha_{\epsilon_0}, \beta_{\epsilon_0}) = \mathcal{IG}(v_{\epsilon_i}|\alpha_{\epsilon_0}, \beta_{\epsilon_0}) \end{cases}, \forall i, \quad (25)$$

which is a generalized Student- $t$  distribution as for the linear model uncertainties. This is a problem because  $\epsilon$  has to be Gaussian according to (17). In order to see how to actually assign a Gaussian prior to measurement uncertainties keeping  $v_{\epsilon}$  unknown, we write the marginal prior of  $\epsilon_i$  with respect to  $v_{\epsilon_i}$

$$p(\epsilon_i|\alpha_{\epsilon_0}, \beta_{\epsilon_0}) = \int_0^{+\infty} \mathcal{N}(\epsilon_i|0, v_{\epsilon_i}) \mathcal{IG}(v_{\epsilon_i}|\alpha_{\epsilon_0}, \beta_{\epsilon_0}) dv_{\epsilon_i} \quad (26)$$

which reads [32]

$$p(\epsilon_i|\alpha_{\epsilon_0}, \beta_{\epsilon_0}) \propto \left(1 + \frac{\epsilon_i^2}{2\beta_{\epsilon_0}}\right)^{-(\alpha_{\epsilon_0} + \frac{1}{2})} \quad (27)$$

where normalizing constants are omitted. From formula (27), we see that, if we fix  $\beta_{\epsilon_0}$  such that

$$\left|\frac{\epsilon_i^2}{\beta_{\epsilon_0}}\right| \ll 1 \quad \forall i, \quad (28)$$

the following approximation holds :

$$p(\epsilon_i | \alpha_{\epsilon_0}, \beta_{\epsilon_0}) \Big|_{\frac{\epsilon_i^2}{\beta_{\epsilon_0}} \ll 1} \sim \exp \left[ - \left( \alpha_{\epsilon_0} + \frac{1}{2} \right) \frac{\epsilon_i^2}{2\beta_{\epsilon_0}} \right] \quad (29)$$

and means that  $\epsilon_i$  can actually be modeled as zero-mean Gaussian

$$p(\epsilon_i | \alpha_{\epsilon_0}, \beta_{\epsilon_0}) \Big|_{\frac{\epsilon_i^2}{\beta_{\epsilon_0}} \ll 1} \sim \mathcal{N} \left( \epsilon_i \mid 0, \frac{\beta_{\epsilon_0}}{\alpha_{\epsilon_0} + \frac{1}{2}} \right), \forall i. \quad (30)$$

Thus, generalized Student- $t$  distribution is sufficiently flexible to assign Gaussian prior for measurement uncertainties  $\epsilon$  and heavy-tailed prior for linear model uncertainties  $\xi$ .

For fixed  $\beta_{\epsilon_0}$  complying with constraint (28), we derive in appendix B a formula to automatically set  $\alpha_{\epsilon_0}$  :

$$\alpha_{\epsilon_0} = M \times \frac{\beta_{\epsilon_0}}{\|\mathbf{g}\|_2^2} \times \left( 1 + 10 \frac{SNR}{10} \right) - \frac{1}{2}, \quad (31)$$

where  $SNR$  corresponds to a prior on the *Signal-to-Noise Ratio* (SNR). In this work,  $SNR$  is a parameter which is fixed manually by the user based on physical intuition.

Concerning parameters  $\alpha_{\xi_0}$  and  $\beta_{\xi_0}$  for linear model uncertainties, their fixation tunes the tail of generalized Student- $t$  distribution in order to introduce robustness as motivated in section II-C. Figure 2 shows the probability density function of  $\mathcal{St}_g(\cdot | \alpha_{\xi_0}, \beta_{\xi_0})$  for different values of  $\alpha_{\xi_0}$  and  $\beta_{\xi_0}$ . We see that generalized Student- $t$  distribution is heavy-tailed for  $\alpha_{\xi_0} < 1$ . On the contrary, the tail is short for  $\alpha_{\xi_0} > 1$ , which is not good to enforce robustness. Hence, we fix  $\alpha_{\xi_0} < 1$ . For  $\alpha_{\xi_0} < 1$ , we see that  $\mathcal{St}_g(\cdot | \alpha_{\xi_0}, \beta_{\xi_0})$  is concentrated around 0 for  $\beta_{\xi_0} < 1$  and  $\alpha_{\xi_0} > \beta_{\xi_0}$  : this makes this setting appropriate when linear model uncertainties are high in a small number of measurements. In the other case, fixing  $\alpha_{\xi_0} < \beta_{\xi_0}$  and  $\beta_{\xi_0} < 1$  makes the tail very heavy : it is more appropriate when linear model uncertainties are high in all the data. Consequently, we fix  $\beta_{\xi_0} < \alpha_{\xi_0} < 1$  when scattering and absorption are low, and  $\alpha_{\xi_0} < \beta_{\xi_0} < 1$  when they are high.

Now we have explained strategies to fix the parameters, we detail in section III a reconstruction algorithm using the error-splitting forward model.

### III. JOINT MAXIMIZATION A POSTERIORI (JMAP) USING THE ERROR-SPLITTING FORWARD MODEL

#### A. Joint Maximization A Posteriori by alternate optimization

The error-splitting forward model is summarized by equations

$$\begin{cases} \mathbf{g} = \mathbf{g}_0 + \epsilon \\ \mathbf{g}_0 = \mathbf{H}\mathbf{f} + \xi \end{cases} \quad (32)$$

The model developed in section II for splitted uncertainties :

$$\begin{cases} p(\epsilon_i | v_{\epsilon_i}) = \mathcal{N}(\epsilon_i | 0, v_{\epsilon_i}) \\ p(v_{\epsilon_i} | \alpha_{\epsilon_0}, \beta_{\epsilon_0}) = \mathcal{IG}(v_{\epsilon_i} | \alpha_{\epsilon_0}, \beta_{\epsilon_0}) \\ p(\xi_i | v_{\xi_i}) = \mathcal{N}(\xi_i | 0, v_{\xi_i}) \\ p(v_{\xi_i} | \alpha_{\xi_0}, \beta_{\xi_0}) = \mathcal{IG}(v_{\xi_i} | \alpha_{\xi_0}, \beta_{\xi_0}) \end{cases} \quad \forall i, \quad (33)$$

is denoted by  $\mathcal{S}$ . We consider a general prior model  $\mathcal{M}$  on the volume with hyperparameters  $\theta$  which are estimated jointly with the other unknowns. Figure 3 summarizes the general

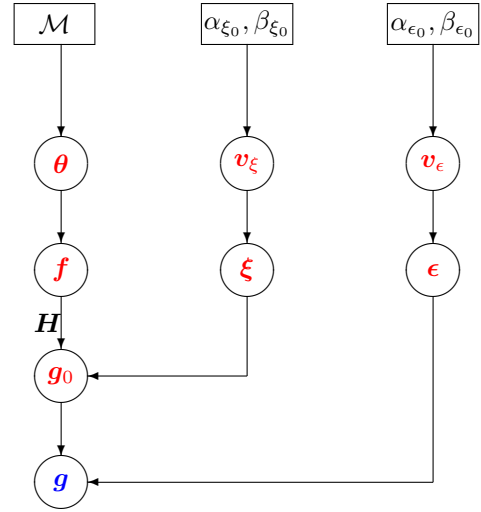


Fig. 3: General hierarchical model including the error-splitting forward model

hierarchical model when the error-splitting forward model is used : data  $\mathbf{g}$  are the combination of "true" projections  $\mathbf{g}_0$  and of measurement uncertainties  $\epsilon$  with variances  $v_{\epsilon}$ . Projections  $\mathbf{g}_0$  result from volume  $\mathbf{f}$  (through projection operation  $\mathbf{H}$ ), and from linear model uncertainties  $\xi$  with variances  $v_{\xi}$ . At last, volume  $\mathbf{f}$  is described by prior model  $\mathcal{M}$  with hyperparameters  $\theta$ . Hyperparameters  $\theta$  can be assigned a prior distribution depending on  $\mathcal{M}$ , while variances  $v_{\xi}$  and  $v_{\epsilon}$  have Inverse-Gamma priors with parameters  $(\alpha_{\xi_0}, \beta_{\xi_0})$  and  $(\alpha_{\epsilon_0}, \beta_{\epsilon_0})$  respectively. The joint posterior distribution of the unknowns is given by Bayes's rule :

$$p(\mathbf{g}_0, v_{\epsilon}, \mathbf{f}, v_{\xi}, \theta | \mathbf{g}; \mathcal{S}, \mathcal{M}) \propto p(\mathbf{g} | \mathbf{g}_0, v_{\epsilon}) p(\mathbf{g}_0 | \mathbf{f}, v_{\xi}) p(v_{\epsilon} | \alpha_{\epsilon_0}, \beta_{\epsilon_0}) p(\mathbf{f} | \theta; \mathcal{M}) p(v_{\xi} | \alpha_{\xi_0}, \beta_{\xi_0}) p(\theta | \mathcal{M}). \quad (34)$$

Due to their computational cost which grows as the dimension of the unknowns increases, MCMC methods are unaffordable for large 3D volumes. For this reason, we choose to perform a joint maximization a posteriori (JMAP)

$$(\hat{\mathbf{g}}_0, \hat{v}_{\epsilon}, \hat{\mathbf{f}}, \hat{v}_{\xi}, \hat{\theta}) = \arg \max_{(\mathbf{g}_0, v_{\epsilon}, \mathbf{f}, v_{\xi}, \theta)} \{ p(\mathbf{g}_0, v_{\epsilon}, \mathbf{f}, v_{\xi}, \theta | \mathbf{g}; \mathcal{S}, \mathcal{M}) \} \quad (35)$$

by an alternate optimization algorithm

$$\mathbf{g}_0 \leftarrow \arg \max_{\mathbf{g}_0} \{ p(\mathbf{g}_0, v_{\epsilon}, \mathbf{f}, v_{\xi}, \theta | \mathbf{g}; \mathcal{S}, \mathcal{M}) \} \quad (36a)$$

$$v_{\epsilon} \leftarrow \arg \max_{v_{\epsilon}} \{ p(\mathbf{g}_0, v_{\epsilon}, \mathbf{f}, v_{\xi}, \theta | \mathbf{g}; \mathcal{S}, \mathcal{M}) \} \quad (36b)$$

$$\mathbf{f} \leftarrow \arg \max_{\mathbf{f}} \{ p(\mathbf{g}_0, v_{\epsilon}, \mathbf{f}, v_{\xi}, \theta | \mathbf{g}; \mathcal{S}, \mathcal{M}) \} \quad (36c)$$

$$v_{\xi} \leftarrow \arg \max_{v_{\xi}} \{ p(\mathbf{g}_0, v_{\epsilon}, \mathbf{f}, v_{\xi}, \theta | \mathbf{g}; \mathcal{S}, \mathcal{M}) \} \quad (36d)$$

$$\theta \leftarrow \arg \max_{\theta} \{ p(\mathbf{g}_0, v_{\epsilon}, \mathbf{f}, v_{\xi}, \theta | \mathbf{g}; \mathcal{S}, \mathcal{M}) \} \quad (36e)$$

Criterion  $\mathcal{L}$  to maximize is derived from joint posterior distribution (34), from which constants are dropped :

$$\begin{aligned} \mathcal{L}(\mathbf{g}_0, \mathbf{v}_\epsilon, \mathbf{f}, \mathbf{v}_\xi, \boldsymbol{\theta}) &= -\frac{1}{2}\|\mathbf{g} - \mathbf{g}_0\|_{\mathbf{V}_\epsilon}^2 - \frac{1}{2}\|\mathbf{g}_0 - \mathbf{H}\mathbf{f}\|_{\mathbf{V}_\xi}^2 \\ &- \sum_{i=1}^M \left[ \left( \alpha_{\epsilon_0} + \frac{3}{2} \right) \ln(v_{\epsilon_i}) + \frac{\beta_{\epsilon_0}}{v_{\epsilon_i}} \right] + \ln(p(\mathbf{f}|\boldsymbol{\theta}; \mathcal{M})) \\ &- \sum_{i=1}^M \left[ \left( \alpha_{\xi_0} + \frac{3}{2} \right) \ln(v_{\xi_i}) + \frac{\beta_{\xi_0}}{v_{\xi_i}} \right] + \ln(p(\boldsymbol{\theta}|\mathcal{M})). \end{aligned} \quad (37)$$

According to Bayes's rule, each step of the algorithm is equivalent to maximizing the conditional distribution of each unknown given the others. In the following of this section, we detail these steps for the general case with any prior model  $\mathcal{M}$  on the volume.

1) *Conditional distribution of "true" projections  $\mathbf{g}_0$* : According to Bayes's rule, the conditional distribution of the "true" projections given the other unknowns reads

$$\begin{aligned} p(\mathbf{g}_0|\mathbf{v}_\epsilon, \mathbf{f}, \mathbf{v}_\xi, \boldsymbol{\theta}; \mathbf{g}, \mathcal{S}, \mathcal{M}) &\propto p(\mathbf{g}|\mathbf{g}_0, \mathbf{v}_\epsilon)p(\mathbf{g}_0|\mathbf{f}, \mathbf{v}_\xi) \\ &\propto \exp \left[ -\frac{1}{2}\|\mathbf{g} - \mathbf{g}_0\|_{\mathbf{V}_\epsilon}^2 - \frac{1}{2}\|\mathbf{g}_0 - \mathbf{H}\mathbf{f}\|_{\mathbf{V}_\xi}^2 \right] \end{aligned} \quad (38)$$

where  $\mathbf{V}_\epsilon = \text{diag}[v_{\epsilon_i}]$  and  $\mathbf{V}_\xi = \text{diag}[v_{\xi_i}]$ . Hence, this conditional distribution is a Gaussian

$$p(\mathbf{g}_0|\mathbf{v}_\epsilon, \mathbf{f}, \mathbf{v}_\xi, \boldsymbol{\theta}; \mathbf{g}, \mathcal{S}, \mathcal{M}) = \mathcal{N}(\mathbf{g}_0|\mathbf{m}_{g_0}, \mathbf{V}_{g_0}) \quad (39)$$

where

$$\begin{cases} \mathbf{V}_{g_0} = (\mathbf{V}_\epsilon^{-1} + \mathbf{V}_\xi^{-1})^{-1} \\ \mathbf{m}_{g_0} = \mathbf{V}_{g_0} (\mathbf{V}_\epsilon^{-1}\mathbf{g} + \mathbf{V}_\xi^{-1}\mathbf{H}\mathbf{f}) \end{cases} \quad (40)$$

Step (36a) sets

$$\hat{\mathbf{g}}_0 = \mathbf{m}_{g_0} = (\mathbf{V}_\epsilon^{-1} + \mathbf{V}_\xi^{-1})^{-1} (\mathbf{V}_\epsilon^{-1}\mathbf{g} + \mathbf{V}_\xi^{-1}\mathbf{H}\mathbf{f}) \quad (41)$$

which is simple to compute since  $\mathbf{V}_{g_0}$  is diagonal.

2) *Conditional distribution of variances  $\mathbf{v}_\epsilon$  of the measurement uncertainties*: By applying Bayes's rule, we have

$$p(v_{\epsilon_i}|\mathbf{g}_0, \mathbf{f}, \mathbf{v}_\xi, \boldsymbol{\theta}; \mathbf{g}, \mathcal{S}, \mathcal{M}) = \mathcal{IG}(v_{\epsilon_i}|\hat{\alpha}_{\epsilon_0_i}, \hat{\beta}_{\epsilon_0_i}) \quad (42)$$

where

$$\begin{cases} \hat{\alpha}_{\epsilon_0_i} = \alpha_{\epsilon_0} + \frac{1}{2} \\ \hat{\beta}_{\epsilon_0_i} = \beta_{\epsilon_0} + \frac{1}{2}(g_i - g_{0_i})^2, \forall i. \end{cases} \quad (43)$$

So, step (36b) sets

$$\hat{v}_{\epsilon_i} = \frac{\hat{\beta}_{\epsilon_0_i}}{\hat{\alpha}_{\epsilon_0_i} + 1}, \forall i. \quad (44)$$

3) *Conditional distribution of the volume*: The conditional distribution of the volume given the other unknowns depends on its prior model

$$\begin{aligned} p(\mathbf{f}|\mathbf{g}_0, \mathbf{v}_\epsilon, \mathbf{f}, \mathbf{v}_\xi, \boldsymbol{\theta}; \mathbf{g}, \mathcal{S}, \mathcal{M}) &\propto p(\mathbf{g}_0|\mathbf{f}, \mathbf{v}_\xi)p(\mathbf{f}|\boldsymbol{\theta}; \mathcal{M}) \\ &\propto \exp \left[ -\frac{1}{2}\|\mathbf{g}_0 - \mathbf{H}\mathbf{f}\|_{\mathbf{V}_\xi}^2 \right] p(\mathbf{f}|\boldsymbol{\theta}; \mathcal{M}) \end{aligned} \quad (45)$$

So, step (36c) consists in minimizing

$$J_f(\mathbf{f}) = \frac{1}{2}\|\mathbf{g}_0 - \mathbf{H}\mathbf{f}\|_{\mathbf{V}_\xi}^2 - \ln(p(\mathbf{f}|\boldsymbol{\theta}; \mathcal{M})) \quad (46)$$

by an optimization algorithm which is dependent on the prior model on the volume.

TABLE I

JMAP ALGORITHM USING THE ERROR-SPLITTING FORWARD MODEL

1: initialize $\mathbf{v}_\epsilon, \mathbf{f}, \mathbf{v}_\xi$ and $\boldsymbol{\theta}$
2: $t := 1$
3: <b>while</b> $t \leq t^{(max)}$
4:   update $\mathbf{g}_0$ by (41)
5:   update $\mathbf{v}_\epsilon$ by (44)
6:   update $\mathbf{f}$ by minimizing criterion (46)
7:   update $\mathbf{v}_\xi$ by (49)
8:   update $\boldsymbol{\theta}$ by maximizing (50)
9: $t := t + 1$

4) *Conditional distribution of variances  $\mathbf{v}_\xi$  of the linear model uncertainties*: Thanks to the use of conjugate prior, we have :

$$p(v_{\xi_i}|\mathbf{g}_0, \mathbf{v}_\epsilon, \mathbf{f}, \boldsymbol{\theta}; \mathbf{g}, \mathcal{S}, \mathcal{M}) = \mathcal{IG}(v_{\xi_i}|\hat{\alpha}_{\xi_0_i}, \hat{\beta}_{\xi_0_i}) \quad (47)$$

where

$$\begin{cases} \hat{\alpha}_{\xi_0_i} = \alpha_{\xi_0} + \frac{1}{2} \\ \hat{\beta}_{\xi_0_i} = \beta_{\xi_0} + \frac{1}{2}(g_{0_i} - [\mathbf{H}\mathbf{f}]_i)^2, \forall i. \end{cases} \quad (48)$$

So, step (36d) sets

$$\hat{v}_{\xi_i} = \frac{\hat{\beta}_{\xi_0_i}}{\hat{\alpha}_{\xi_0_i} + 1}, \forall i. \quad (49)$$

5) *Conditional distribution of hyperparameters  $\boldsymbol{\theta}$* : The conditional distribution of hyperparameters  $\boldsymbol{\theta}$

$$p(\boldsymbol{\theta}|\mathbf{g}_0, \mathbf{v}_\epsilon, \mathbf{f}, \mathbf{v}_\xi, \boldsymbol{\theta}; \mathbf{g}, \mathcal{S}, \mathcal{M}) \propto p(\mathbf{f}|\boldsymbol{\theta}; \mathcal{M})p(\boldsymbol{\theta}|\mathcal{M}) \quad (50)$$

only depends on the prior model on the volume and not on the used forward model.

The overall algorithm is summarized in table I. As we already pointed, the algorithm is applicable with any prior model on the volume. In section IV, we combine the error-splitting forward model with several prior models on the volume to complete the algorithm with explicit updates of volume  $\mathbf{f}$  and hyperparameters  $\boldsymbol{\theta}$ . Before that, we compare the cost of the use of the error-splitting forward model, with respect to the usual one.

#### B. Comparison with the usual forward model

It is interesting to compare the use of the error-splitting forward model (32) with the usual one (1). With the usual forward model, total uncertainties  $\zeta$  are Gaussian, with variances  $v_\zeta$ , as presented in equation (5). A conjugate prior can be assigned to  $v_\zeta$ , as done in [14], [16]. We denote this model for total uncertainties by  $\mathcal{U}$ .

With the usual forward model (1), true projections  $\mathbf{g}_0$  do not appear since uncertainties are not splitted. To estimate volume  $\mathbf{f}$ , variances  $v_\zeta$  and hyperparameters  $\boldsymbol{\theta}$ , their joint posterior distribution

$$\begin{aligned} p(\mathbf{f}, v_\zeta, \boldsymbol{\theta}|\mathbf{g}; \mathcal{U}, \mathcal{M}) &\propto p(\mathbf{g}|\mathbf{f}, v_\zeta)p(v_\zeta|\mathcal{U}) \\ &p(\mathbf{f}|\boldsymbol{\theta}; \mathcal{M})p(\boldsymbol{\theta}|\mathcal{M}) \end{aligned} \quad (51)$$

is maximized

$$(\hat{\mathbf{f}}, \hat{v}_\zeta, \hat{\boldsymbol{\theta}}) = \arg \max_{(\mathbf{f}, v_\zeta, \boldsymbol{\theta})} \{p(\mathbf{f}, v_\zeta, \boldsymbol{\theta}|\mathbf{g}; \mathcal{U}, \mathcal{M})\} \quad (52)$$

by the alternate optimization [14], [16]

$$\begin{cases} \mathbf{f} \leftarrow \arg \max_{\mathbf{f}} \{p(\mathbf{f}, \mathbf{v}_\zeta, \boldsymbol{\theta} | \mathbf{g}; \mathcal{U}, \mathcal{M})\} & (53a) \\ \mathbf{v}_\zeta \leftarrow \arg \max_{\mathbf{v}_\zeta} \{p(\mathbf{f}, \mathbf{v}_\zeta, \boldsymbol{\theta} | \mathbf{g}; \mathcal{U}, \mathcal{M})\} & (53b) \\ \boldsymbol{\theta} \leftarrow \arg \max_{\boldsymbol{\theta}} \{p(\mathbf{f}, \mathbf{v}_\zeta, \boldsymbol{\theta} | \mathbf{g}; \mathcal{U}, \mathcal{M})\} & (53c) \end{cases}$$

In section III-A, using the error-splitting forward model, step (36c) reads

$$\mathbf{f} \leftarrow \arg \min_{\mathbf{f}} \left\{ \frac{1}{2} \|\mathbf{g}_0 - \mathbf{H}\mathbf{f}\|_{V_\zeta}^2 - \ln(p(\mathbf{f} | \boldsymbol{\theta}; \mathcal{M})) \right\} \quad (54)$$

while step (53a) with usual forward model (1) has the form

$$\mathbf{f} \leftarrow \arg \min_{\mathbf{f}} \left\{ \frac{1}{2} \|\mathbf{g} - \mathbf{H}\mathbf{f}\|_{V_\zeta}^2 - \ln(p(\mathbf{f} | \boldsymbol{\theta}; \mathcal{M})) \right\}, \quad (55)$$

so minimization (54) can be performed in the same way as (55) by replacing  $\mathbf{g}$  by  $\mathbf{g}_0$ , and  $\mathbf{v}_\zeta$  by  $\mathbf{v}_\xi$ . The update of hyperparameters  $\boldsymbol{\theta}$  is done as in section III-A. Hence, alternate optimizations (35) or (52) only differ by the estimation of  $\mathbf{g}_0$ ,  $\mathbf{v}_\epsilon$  and  $\mathbf{v}_\xi$  with the error-splitting forward model, and by the estimation of  $\mathbf{v}_\zeta$  with the usual forward model. As we have seen in section III-A, estimations of  $\mathbf{g}_0$ ,  $\mathbf{v}_\epsilon$  and  $\mathbf{v}_\xi$  have a very little computational cost since they are performed by analytical formulae thanks to the use of conjugate priors. Therefore, especially when a conjugate prior is assigned to  $\mathbf{v}_\zeta$  (such as Inverse-Gamma prior [14], [16]), the cost of one global iteration of JMAP when using the error-splitting forward model is approximately the same as the one with usual forward model (1), for any prior model on the volume.

One drawback of our error-splitting approach is the memory cost. The error-splitting forward model introduces variables  $\mathbf{g}_0$ ,  $\mathbf{v}_\epsilon$  and  $\mathbf{v}_\xi$  in place of variances  $\mathbf{v}_\zeta$  in the usual forward model. These variables are all the size of projections  $\mathbf{g}$ , which can be very large. In order to save memory usage, one way is to assign only one variance  $v_{cell_c}$  to all measurement uncertainties related to cell  $c$  of the detector :

$$v_{\epsilon_i} = v_{cell_c}, \quad \text{for all ray } i \text{ hitting cell } c, \quad (56)$$

so variances of measurement uncertainties are only size of an image with  $N_u \times N_v$  pixels. This is what we do in our experiments in section V. Compared to the case where we have one  $v_{\epsilon_i}$  per  $\epsilon_i$ , we did not notice troubles due to this memory saving.

#### IV. COMBINATION WITH GAUSS-MARKOV-POTTS PRIOR MODEL ON THE VOLUME

In order to test its efficiency in full MBIR method, we now combine our error-splitting forward model with a specific prior model  $\mathcal{M}$  on the volume. In this paper, the aim is to reconstruct industrial parts for NDT. The industrial parts we want to control are quasi-piecewise-constant volumes. The piecewise-constant assumption holds in many CT imaging problems such as industrial inspection [33], [34] or assessment of bone microstructure in medicine [35], [36]. Many different priors are possible for such volumes, such as minimum  $\mathcal{L}_0$ -norm [13], [37], sparse Haar transform [16], or Gauss-Markov-Potts prior [14].

Gauss-Markov-Potts prior model enables to connect the attenuation coefficient  $f_j$  of voxel  $j$  to the material to which it belongs. This material is identified by a label  $z_j$  assigned to voxel  $j$  :  $z_j = k$  when voxel  $j$  is in material  $k$ ,  $k \in \{1, \dots, K\}$ , where  $K$  is the number of materials in the volume. The compacity of a region composed of material  $k$  is enforced thanks to a Potts model for labels  $z$ . Using Hammersley-Clifford theorem [38], this Potts model is

$$p(\mathbf{z} | \boldsymbol{\alpha}, \gamma_0) \propto \exp \left[ \sum_{j=1}^N \left( \sum_{k=1}^K \alpha_k \delta(z_j - k) + \gamma_0 \sum_{i \in \mathcal{V}(j)} \delta(z_j - z_i) \right) \right] \quad (57)$$

where  $\mathcal{V}(j)$  is the first-order neighbourhood of voxel  $j$ . Parameter  $\gamma_0$  is called Potts coefficient and weights the dependence of label  $z_j$  on the neighbours. Parameter  $\alpha_k, \forall k$ , enables to insert a prior knowledge about the probability for voxel  $j$  to be in class  $k$ . In a region exclusively composed of material  $k$ , the distribution of attenuation coefficients is modeled as Gaussian, with mean  $m_k$  and variance  $v_k$  depending on the material :

$$p(f_j | z_j = k, m_k, v_k) = \mathcal{N}(f_j | m_k, v_k). \quad (58)$$

Means  $\mathbf{m}$  and variances  $\mathbf{v}$  of the classes are considered as unknowns and assigned conjugate priors :

$$\begin{cases} p(m_k | m_0, v_0) = \mathcal{N}(m_k | m_0, v_0), \forall k, \\ p(v_k | \alpha_0, \beta_0) = \mathcal{IG}(v_k | \alpha_0, \beta_0) \end{cases}, \quad (59)$$

where  $m_0, v_0, \alpha_0$  and  $\beta_0$  are parameters.

Gauss-Markov-Potts prior model has been used for instance in microwave imaging [39] and diffraction imaging [40]. In [14], we used this model to propose a joint reconstruction and segmentation algorithm in 3D CT. Like DART [33], the proposed algorithm in [14] mainly consists in alternating reconstruction (estimation of  $\mathbf{f}$ ) and segmentation (estimation of  $\mathbf{z}$ ) steps. The main advantage with respect to DART is that the levels of the classes (which correspond to means  $\mathbf{m}$ ) do not need to be known [41]. Moreover, Gauss-Markov-Potts prior allows variability in the classes thanks to the introduction of variances  $\mathbf{v}$ . Another Potts model, which promotes minimum  $\mathcal{L}_0$ -norm for the gradient of the volume, is used in [13] for CT. The reconstruction algorithm in [13] is an instance of ADMM [42], and performs a variable-splitting in each direction used to compute the gradient : as a result, it requires more memory than the reconstruction algorithm proposed in [14], and has been only applied to small volumes in 3D [37].

Sparse Haar transform prior is also suited for modelling piecewise-constant volumes, as proposed in [16]. Nevertheless, the algorithm proposed in [16] is memory-costly due to its high number of variables having the size of the volume. For Gauss-Markov-Potts prior model, only labels  $\mathbf{z}$  have the same size as the volume, but they requires much less memory to be stored, since they are integers. Moreover, because of intermediate results which have to be stored for multi-level transformation [43], Haar transform is more difficult to parallelize on GPU than the segmentation step proposed in [14]. At last, the



TABLE II

JMAP ALGORITHM COMBINING THE ERROR-SPLITTING FORWARD MODEL AND GAUSS-MARKOV-POTTS PRIOR

---

```

1: initialize  $v_\epsilon, \mathbf{f}, v_\xi, \mathbf{z}, \mathbf{m}$  and  $\mathbf{v}$ 
2:  $t := 1$ 
3: while  $t \leq t^{(max)}$ 
4:   update  $\mathbf{g}_0$  by (41)
5:   update  $v_\epsilon$  by (44)
6:   update  $\mathbf{f}$  by minimizing criterion (60)
7:   update  $v_\xi$  by (49)
8.1:  update  $\mathbf{z}$  by ICM algorithm [14], [44]
8.2:  update  $\mathbf{m}$  by (61)
8.3:  update  $\mathbf{v}$  by (62)
9:    $t := t + 1$ 

```

---

parameters in [16] are more difficult to tune than the ones for Gauss-Markov-Potts prior [14].

For all the reasons explained above, we choose to combine the error-splitting forward model with Gauss-Markov-Potts prior in order to reconstruct piecewise-constant volumes. The joint maximization a posteriori with the error-splitting forward model is performed as described in section III. The criterion minimized to update the volume, corresponding to criterion (46) in section III, reads

$$J_f(\mathbf{f}) = \frac{1}{2} \|\mathbf{g}_0 - \mathbf{H}\mathbf{f}\|_{V_\epsilon}^2 + \frac{1}{2} \|\mathbf{f} - \mathbf{m}_z\|_{V_z}^2 \quad (60)$$

where  $m_{z_j} = m_k$  and  $v_{z_j} = v_k$  if  $z_j = k$ , and  $V_z = \text{diag}[v_z]$ . We perform this minimization by a simple gradient descent with adaptive stepsize, which is very similar to the one we presented in [14] with usual forward model (1), except that  $\mathbf{g}$  is replaced by  $\mathbf{g}_0$  and  $v_\zeta$  by  $v_\xi$ .

For Gauss-Markov-Potts prior model, hyperparameters are  $\theta = (\mathbf{z}, \mathbf{m}, \mathbf{v})$ . As we have seen, their update do not depend on the used forward model. Consequently,  $\mathbf{z}$ ,  $\mathbf{m}$  and  $\mathbf{v}$  are updated as done in [14]. The labels are updated by running few iterations of Iterated Conditional Mode (ICM) algorithm [44], which is highly-parallelizable [14]. Next, the updating formulae for  $\mathbf{m}$  and  $\mathbf{v}$  are [14] :

$$\hat{m}_k = \frac{\frac{m_0}{v_0} + \frac{1}{v_k} \sum_{j=1}^N f_j \delta(z_j - k)}{\frac{1}{v_0} + \frac{N_k}{v_k}}, \forall k \quad (61)$$

and

$$\hat{v}_k = \frac{\beta_0 + \frac{1}{2} \sum_{j=1}^N (f_j - m_k)^2 \delta(z_j - k)}{\alpha_0 + \frac{N_k}{2} + 1}, \forall k \quad (62)$$

where  $N_k$  is the number of voxels in class  $k$ . The full MBIR algorithm combining Gauss-Markov-Potts prior and the error-splitting forward model is given in table II and is simply an instance of the one given in table I. Initial labels  $\mathbf{z}^{(0)}$  are obtained by a segmentation of initial volume  $\mathbf{f}^{(0)}$ , on which initial means  $\mathbf{m}^{(0)}$  and variances  $\mathbf{v}^{(0)}$  are computed. Concerning the fixation of parameters of Gauss-Markov-Potts prior :  $K, \gamma_0, (\alpha_k)_k, m_0, v_0, \alpha_0$  and  $\beta_0$ , the applied strategy is the same explained in [14].

## V. RESULTS ON REAL DATA

We evaluate the error-splitting forward model combined with Gauss-Markov-Potts prior by applying the algorithm

given in table II to real data. We compare the results with those obtained with usual forward model (1). Through the iterations of each reconstruction algorithm, we observe the evolution of *Root Mean Square Difference* (RMSD) :

$$RMSD = \sqrt{\frac{1}{N} \sum_{j=1}^N (f_j - f_j^*)^2} \quad (63)$$

between estimation  $\mathbf{f}$  and a reference volume  $\mathbf{f}^*$ . We run each algorithm during 50 iterations. We also compare our results with TV-regularized least-squares, by applying memory-efficient Primal-Dual Frank-Wolfe (PDFW) algorithm [45].

The used projector and backprojector are parallelized on GPU and are ray-driven and voxel-driven respectively [46]. This pair is unmatched to reduce the computational cost. An unmatched pair of projector and backprojector can be used under conditions on the eigenvalues of  $\mathbf{H}^T \mathbf{H}$  [47]. We did not notice troubles due to this mismatching, and all the tested algorithms have converged. The segmentation step, corresponding to step 8.1 of the algorithm in table II, is performed on CPU.

### A. Framework for the experiments

In order to use usual forward model (1), we need to assign a prior to variances  $v_\zeta$  of total uncertainties  $\zeta$  : to do simple calculations, we consider a conjugate prior, which is Inverse-Gamma :

$$p(v_{\zeta_i} | \alpha_{\zeta_0}, \beta_{\zeta_0}) = \mathcal{IG}(v_{\zeta_i} | \alpha_{\zeta_0}, \beta_{\zeta_0}) \quad (64)$$

with parameters  $(\alpha_{\zeta_0}, \beta_{\zeta_0})$ , as done in [14]. With this prior, step (53b) presented in section III-B is done similarly to updates (36b) and (36d). The way to fix parameters  $\alpha_{\zeta_0}$  and  $\beta_{\zeta_0}$  is important, since, according to

$$\begin{cases} p(\zeta_i | v_{\zeta_i}) = \mathcal{N}(\zeta_i | 0, v_{\zeta_i}) \\ p(v_{\zeta_i} | \alpha_{\zeta_0}, \beta_{\zeta_0}) = \mathcal{IG}(v_{\zeta_i} | \alpha_{\zeta_0}, \beta_{\zeta_0}) \end{cases} \quad (65)$$

the distribution of  $\zeta_i$  marginalized with respect to  $v_{\zeta_i}$  is a generalized Student- $t$ . From this point, two ways are possible. One is to model total uncertainties as Gaussian, and so parameters  $\alpha_{\zeta_0}$  and  $\beta_{\zeta_0}$  are fixed similarly to  $\alpha_{\epsilon_0}$  and  $\beta_{\epsilon_0}$  in section II-D for measurement uncertainties  $\epsilon$ . Another way is to enforce robustness with respect to outliers in projections  $\mathbf{g}$ , so a heavy-tailed prior is more appropriate for  $\zeta$ . In this case,  $\alpha_{\zeta_0}$  and  $\beta_{\zeta_0}$  are fixed similarly to  $\alpha_{\xi_0}$  and  $\beta_{\xi_0}$  in section II-D for linear model uncertainties  $\xi$ .

### B. Evaluation of the accuracy : Image Quality Indicator (IQI) volume

We evaluate the accuracy of the error-splitting forward model on an Image Quality Indicator (IQI) volume [48]. The field-of-view is sampled into  $512 \times 512 \times 256$  voxels. To reconstruct this volume by JMAP, 300 projections uniformly distributed over  $[0, 2\pi]$  are used. The detector is sampled into  $512 \times 256$  pixels. The source-to-rotation-center distance is 975 mm, and the source-to-detector distance is 1300 mm. The IQI volume is a cylindrical plate in which holes are drilled [48]. It is obliquely positioned in order to minimize

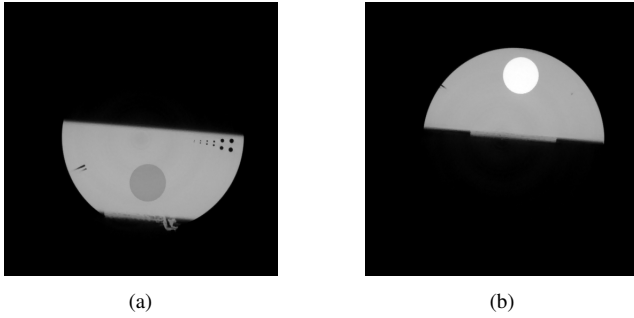


Fig. 4: Reference for IQI, reconstructed from 2400 projections [14] (bottom (a) and top (b))

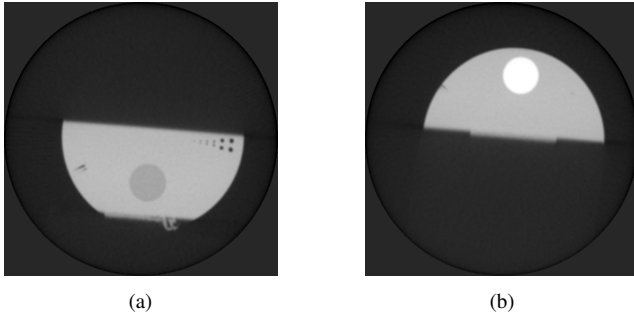


Fig. 5: Reconstruction from 300 projections of IQI volume by filtered backprojection [4] (bottom (a) and top (b))

the attenuation of X-rays inside it. As a result, the signal-to-noise ratio is improved. Reference volume is reconstructed from 2400 projections [14] and is shown in figure 4. From 300 projections, the filtered backprojection [4], which initializes JMAP, is shown in figure 5. TV-reconstruction is shown in figure 6. To tune JMAP, we consider that the SNR is 20 db, which is a realistic value. We set  $\beta_{\epsilon_0} = 1$  to ensure that constraint (28) is fulfilled. The parameters of the error-splitting forward model for IQI reconstruction are summarized in table III. They are fixed according to the strategy explained in section II-D.

Parameters	SNR	$\alpha_{\epsilon_0}$	$\beta_{\epsilon_0}$	$\alpha_{\xi_0}$	$\beta_{\xi_0}$
Values	20	Use (31)	1	0.01	0.0001

TABLE III: Parameters of the error-splitting forward model for IQI reconstruction

In order to compare tested forward models, the values of

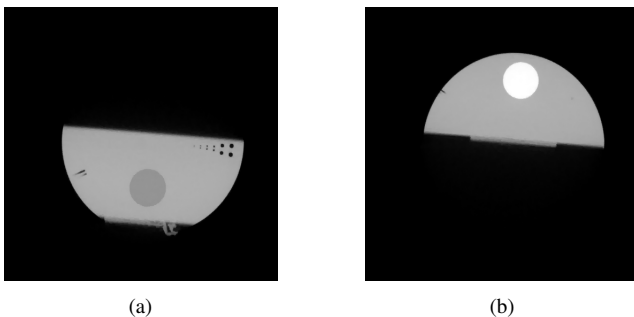


Fig. 6: Reconstruction from 300 projections of IQI volume by TV-regularized least-squares [45] (bottom (a) and top (b))

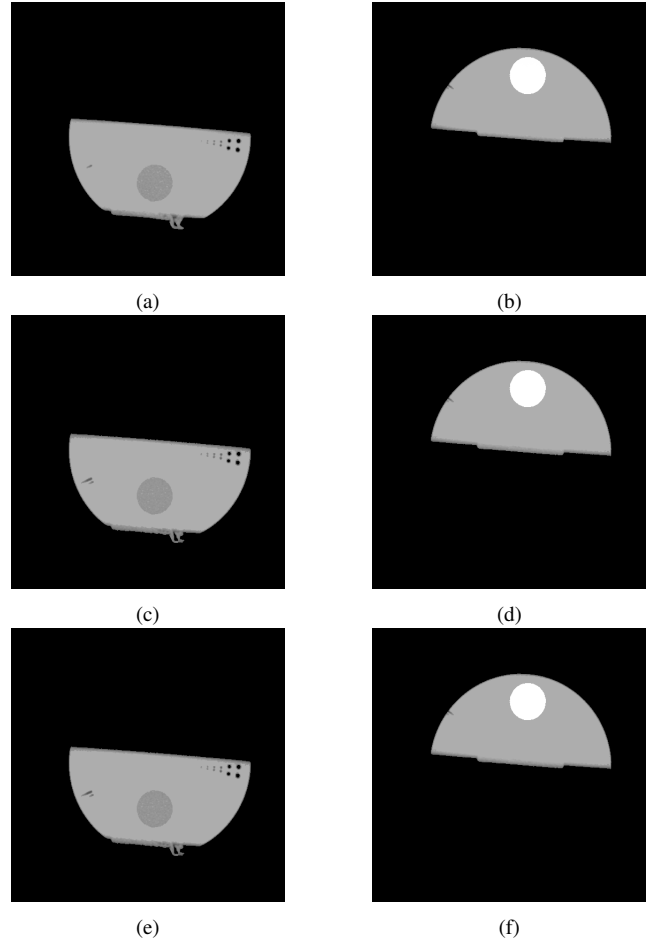


Fig. 7: Reconstructions from 300 projections of IQI volume by JMAP with Gauss-Markov-Potts prior and : the usual forward model modelling  $\zeta$  as heavy-tailed (bottom (a) and top (b)), the usual forward model and modelling  $\zeta$  as Gaussian (bottom (c) and top (d)), and the error-splitting forward model (bottom (e) and top (f))

the parameters of the Gauss-Markov-Potts prior on the volume remain the same whatever the forward model is. These values are summarized in table IV. Since the volume is composed of nylon, polytetrafluoroethylene (PTFE), polyamide and air, we fix  $K = 4$ . Parameters  $m_0$  and  $(\alpha_k)_k$  are automatically set as done in [14]. The other parameters in table IV are fixed following the strategies explained in [14].

Parameters	$K$	$\gamma_0$	$v_0$	$\alpha_0$	$\beta_0$
Fixed values	4	3	1	5	0.01

TABLE IV: Parameters for Gauss-Markov-Potts prior model on IQI volume

Reconstructions of IQI volume by JMAP with Gauss-Markov-Potts prior are shown in figure 7. As we see, the reconstructions obtained with the different forward models look the same : SSIM indices [49] computed between these reconstructions are approximately 1. We see they are smoother than FDK reconstruction. Like TV-reconstruction in figure 6, the reconstructions by JMAP have compact and homogeneous regions thanks to the use of Gauss-Markov-Potts prior model. For each forward model, the segmentations obtained jointly with the reconstruction are shown in figure 8 : as we see, they

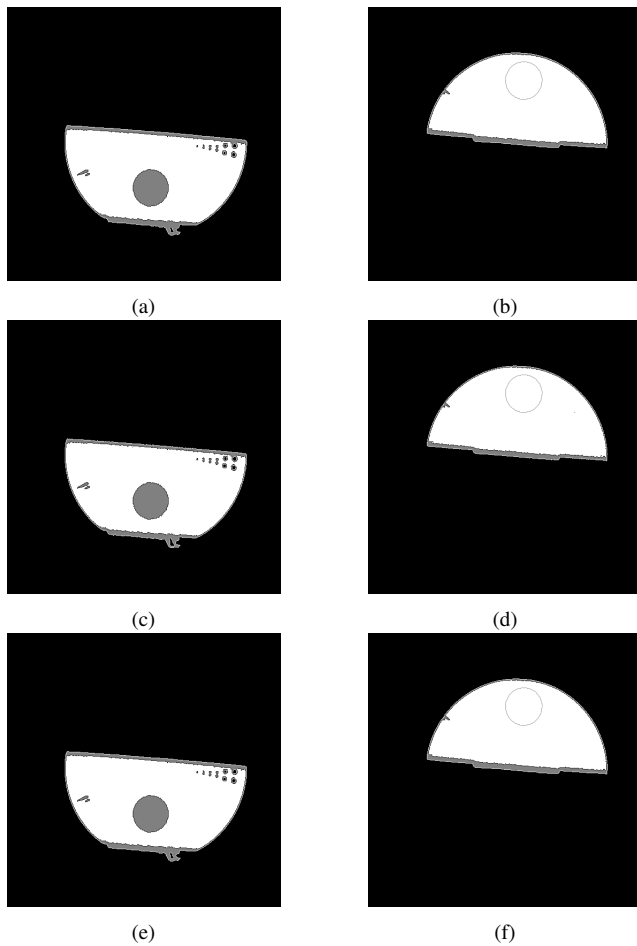


Fig. 8: Segmentations obtained jointly with the reconstruction from 300 projections of IQI volume by JMAP with Gauss-Markov-Potts prior and : the usual forward model modelling  $\zeta$  as heavy-tailed (bottom (a) and top (b)), the usual forward model and modelling  $\zeta$  as Gaussian (bottom (c) and top (d)), and the error-splitting forward model (bottom (e) and top (f))

are the same. To compare the accuracy of each reconstruction, we focus on the little holes of the volume visible in subfigures (a), (c) and (e) of figure 7. We show their profiles in figure 9, with the ones of reference, FDK and TV reconstructions. We observe that JMAP with Gauss-Markov-Potts prior enhances the contrast, making transitions between holes much sharper than FDK and TV. For the two biggest holes, whatever the used forward model is, JMAP reaches the lowest value and is better than FDK and TV. Nevertheless, this value is slightly under the theoretical value of 0 due to the fact that no non-negativity constraint is enforced in our algorithm. For the two intermediate holes, TV-reconstruction's profile goes lower than JMAP. For these holes, the profile is the same for JMAP reconstructions whatever the forward model is. For the two most little holes, the accuracy is quite similar for JMAP with the error-splitting forward model, JMAP with the usual forward model modelling  $\zeta$  as Gaussian, and TV-reconstruction. It is slightly lower for JMAP with the usual forward model modelling  $\zeta$  as heavy-tailed. Hence, we see that the error-splitting forward model does not imply a loss of accuracy, and is even better than the usual forward model for a certain set of parameters.

We analyze the convergence of JMAP with each forward model. For the error-splitting forward model, the criterion (37) of algorithm in table II converges to a maximum value as shown in figure 10. Computation times are shown in table V : because the error-splitting forward model adds very few operations, the reconstruction times for JMAP with each forward model are similar and are approximately 30 minutes. For volume update step at each iteration of JMAP, 10 sub-iterations are performed. Hence, because PDFW does not have sub-iterations and because we run JMAP during 50 iterations, TV-reconstruction is obtained in 500 iterations. The reconstruction time for TV is approximately 20 minutes. The computation time for JMAP is larger than the one for TV due to the segmentation step done on CPU. We show the evolution of RMSD with respect to reference volume in figure 11 for TV and in figure 12 for JMAP. Compared to TV, in table V, we see that a lower value of RMSD is reached for JMAP with Gauss-Markov-Potts prior, whatever the forward model is. Nevertheless, in figure 12, we see that the RMSD reaches its lowest value with the error-splitting forward model. Thus, for this reconstruction with Gauss-Markov-Potts prior, JMAP with the error-splitting forward model is more accurate than JMAP with usual forward model (1).

### C. Evaluation of robustness : composite material with metallic edges

We now evaluate the robustness of our algorithm on challenging data with metal : the controlled volume is made of composite material with titanium on the edges. We denote it by CMTE (Composite Material with Titanium on the Edges). The field-of-view is sampled into  $512^3$  voxels. The reconstruction is performed with 300 projections uniformly distributed over  $[0, 2\pi]$ . The detector is sampled into  $512^2$  pixels. The source-to-rotation-center distance is 432 mm, and the source-to-detector distance is 807 mm. Reference volume is reconstructed from 1500 projections [14] and is shown in figure 13. As we see, even with a lot of projections, scattering remains around the titanium edges. This scattering is hard to remove because the values for scattering are near the values for composite material. The filtered backprojection from 300 projections is shown in figure 14. The parameters of the error-splitting forward model are summarized in table VI. Because we have a lot of absorption and scattering due to the presence of titanium, we fix  $\alpha_{\xi_0} < \beta_{\xi_0} < 1$ , as explained in section II-D. The values of the parameters for Gauss-Markov-Potts prior model on CMTE volume are presented in table VII. CMTE is composed of composite material and titanium. Furthermore, we have to take the air in the field-of-view into account. Hence, we fix  $K = 3$ .

Because of high beam-hardening, the number of sub-iterations for volume update in JMAP has to be fixed carefully in order to avoid local optima : at each iteration of JMAP, 2 sub-iterations are performed for volume estimation. We run JMAP during 50 iterations. Hence, TV-reconstruction is obtained in 100 iterations. Joint reconstructions and segmentations of CMTE volume by JMAP with Gauss-Markov-Potts prior are shown in figure 16. Figure 17 shows the convergence

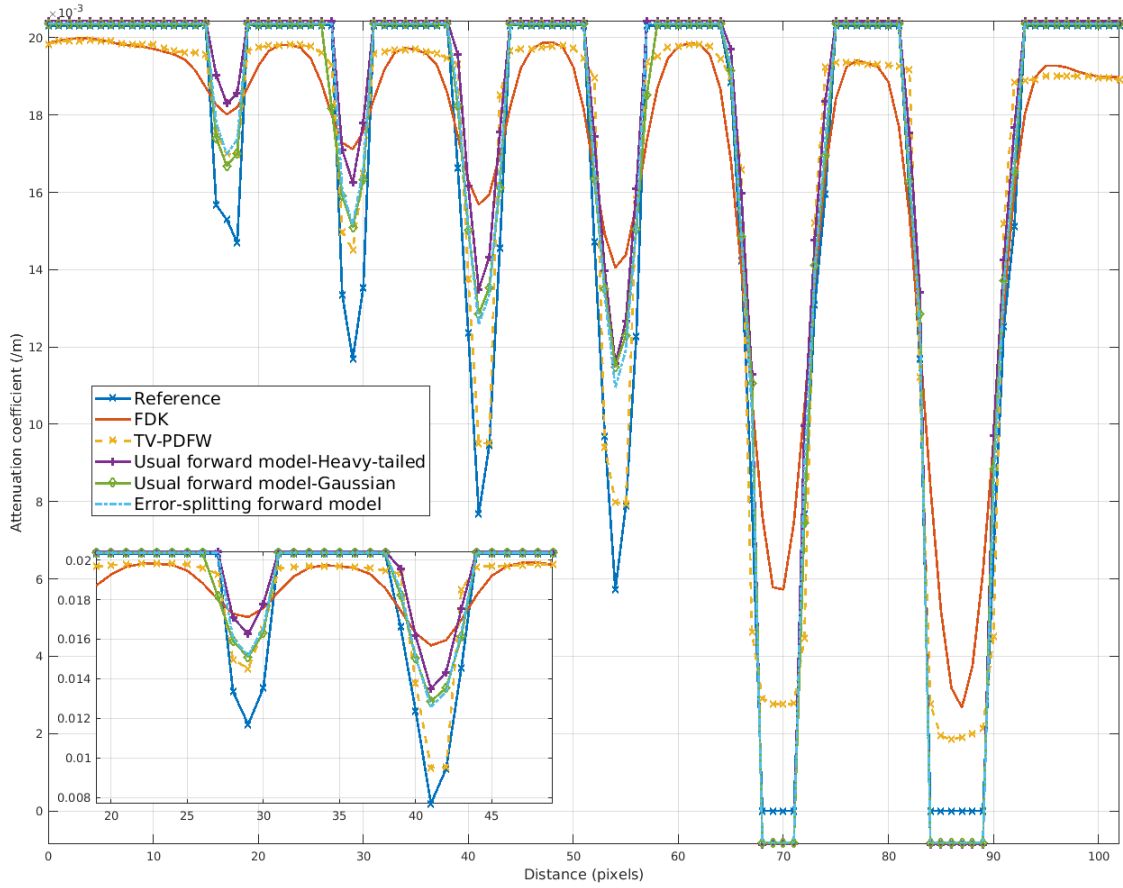


Fig. 9: Accuracy : profiles of the holes of IQI volume for each reconstruction, with Gauss-Markov-Potts prior for JMAP

Forward model	Usual	Usual	Usual	Error-splitting
Prior model	TV	Gauss-Markov-Potts	Gauss-Markov-Potts	Gauss-Markov-Potts
Algorithm	PDFW	JMAP-Heavy-tailed	JMAP-Gaussian	JMAP-Error-splitting
Computation time	<b>1180 s</b>	1876 s	1777 s	1854 s
Final value of RMSD (/m)	$22.10 \times 10^{-4}$	$9.09 \times 10^{-4}$	$8.86 \times 10^{-4}$	<b><math>8.78 \times 10^{-4}</math></b>

TABLE V: Performance comparison on IQI reconstruction from 300 projections, of TV-PDFW and JMAP with Gauss-Markov-Potts prior model and : the usual forward model modelling  $\zeta$  as heavy-tailed, the usual forward model modelling  $\zeta$  as Gaussian, and the error-splitting forward model

Parameters	SNR	$\alpha_{\epsilon_0}$	$\beta_{\epsilon_0}$	$\alpha_{\xi_0}$	$\beta_{\xi_0}$
Values	20	Use (31)	1	0.01	0.1

TABLE VI: Parameters of the error-splitting forward model for CMTE reconstruction

Parameters	$K$	$\gamma_0$	$v_0$	$\alpha_0$	$\beta_0$
Fixed values	3	3	1	5	0.01

TABLE VII: Parameters for Gauss-Markov-Potts prior model on CMTE volume

of the criterion of JMAP with the error-splitting forward model. In figure 16, we see that reconstructions by JMAP still have scattering around the titanium. Because the values for scattering are very near the values for composite material, scattering is misclassified in the segmentation and is in the

same class as composite. With the error-splitting forward model and with the usual forward model modelling  $\zeta$  as heavy-tailed, JMAP is very close to the reference. This is not the case with the usual forward model modelling  $\zeta$  as Gaussian : the reconstruction has very strong beam-hardening artifacts at the extremities of the titanium. We see in figure 19 that JMAP with the usual forward model modelling  $\zeta$  as Gaussian diverges far from the reference volume. Indeed, because the prior on  $\zeta$  is Gaussian, i.e. short-tailed, the algorithm tries to be too close to the data, which is not a good point, since high absorption is present in the data due to the titanium edges. The same phenomena occurs for TV-reconstruction in figure 15. As shown in table VIII, TV-reconstruction reaches a higher value of RMSD than JMAP.

On the opposite, JMAP with the error-splitting forward

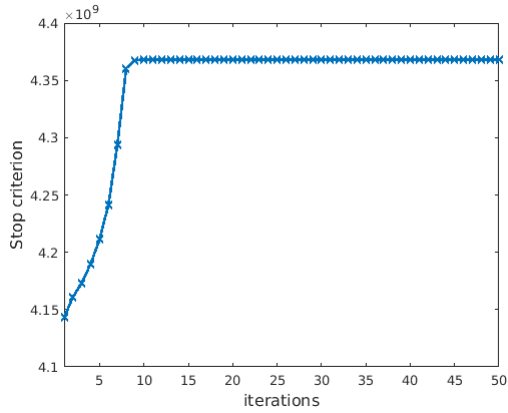


Fig. 10: Convergence for JMAP with the error-splitting forward model and Gauss-Markov-Potts prior for IQI reconstruction from 300 projections

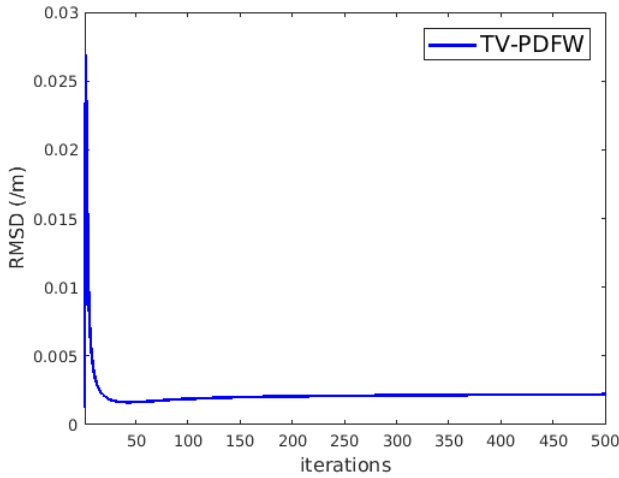


Fig. 11: Convergence to reference volume for IQI reconstruction by TV-regularized least-squares

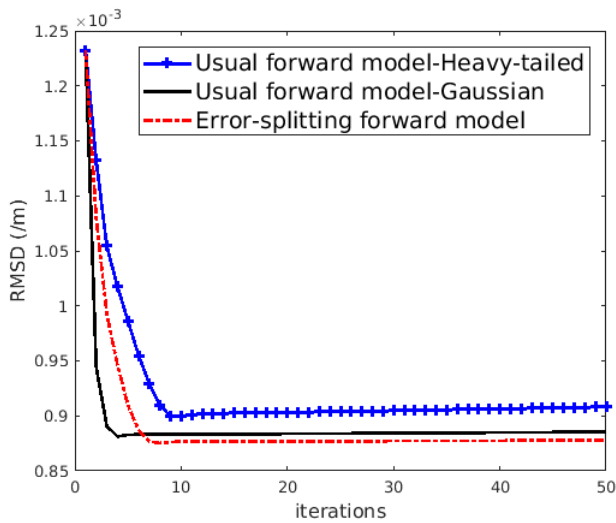


Fig. 12: Convergence to reference volume for IQI reconstruction by JMAP from 300 projections with Gauss-Markov-Potts prior and : the usual forward model modelling  $\zeta$  as heavy-tailed, the usual forward model modelling  $\zeta$  as Gaussian, and the error-splitting forward model

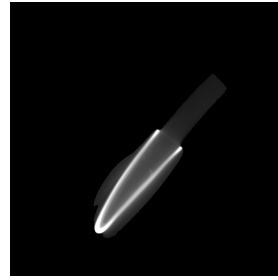


Fig. 13: Reference for CMTE, reconstructed from 1500 projections [14] (middle slice)

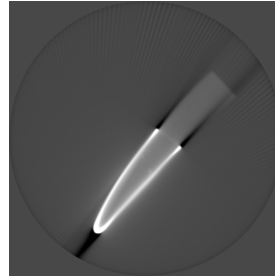


Fig. 14: Reconstruction from 300 projections of CMTE volume by filtered backprojection [4] (middle slice)

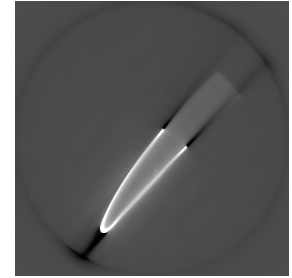


Fig. 15: Reconstruction from 300 projections of CMTE volume by TV-regularized least-squares [45] (middle slice)

model and JMAP with the usual forward model modelling  $\zeta$  as heavy-tailed manage to remove beam-hardening artifacts, and, as a result, converge to the reference volume, as shown by the evolution of RMSD in figure 19. In table VIII, we see that final RMSD is practically the same for these reconstructions.

Hence, for this reconstruction with Gauss-Markov-Potts prior, JMAP with the error-splitting forward model is as robust as JMAP with the usual forward model modelling  $\zeta$  as heavy-tailed. The interesting point is that, when evaluating the accuracy on IQI volume in the previous section, JMAP with the error-splitting forward model obtains the same accuracy than JMAP with the usual forward model modelling  $\zeta$  as Gaussian, while the reconstruction is less accurate with heavy-tailed prior on  $\zeta$  in the usual forward model. Indeed, because there are not high uncertainties on the projections for IQI volume, modelling  $\zeta$  as short-tailed for JMAP with the usual forward model enables to be close to the data, and then to be more accurate. But, with CMTE volume, being too close to the data leads to a failure in removing strong beam-hardening artifacts. For the error-splitting forward model, our better physical modelling leads to the best accuracy for IQI reconstruction, and, thanks to the heavy-tailed prior on linear model uncertainties, to robustness for CMTE reconstruction. Hence, in our experiments with Gauss-Markov-Potts prior, JMAP with the error-splitting forward model combines accuracy and robustness, while, with the usual forward model, accuracy is only achieved by modelling  $\zeta$  as Gaussian, and robustness only by modelling  $\zeta$  as heavy-tailed.

#### D. Discussion

Our experiments have shown that the error-splitting forward model enables to get robustness with heavy-tailed prior on

Forward model	Usual	Usual	Usual	Error-splitting
Prior model	TV	Gauss-Markov-Potts	Gauss-Markov-Potts	Gauss-Markov-Potts
Algorithm	PDFW	JMAP-Heavy-tailed	JMAP-Gaussian	JMAP-Error-splitting
Computation time	<b>381 s</b>	5074 s	5359 s	5496 s
Final value of RMSD (/m)	$11.883 \times 10^{-3}$	<b><math>1.765 \times 10^{-3}</math></b>	$8.167 \times 10^{-3}$	$1.785 \times 10^{-3}$

TABLE VIII: Performance comparison on CMTE reconstruction from 300 projections, of TV-PDFW and JMAP with Gauss-Markov-Potts prior model and : the usual forward model modelling  $\zeta$  as heavy-tailed, the usual forward model modelling  $\zeta$  as Gaussian, and the error-splitting forward model

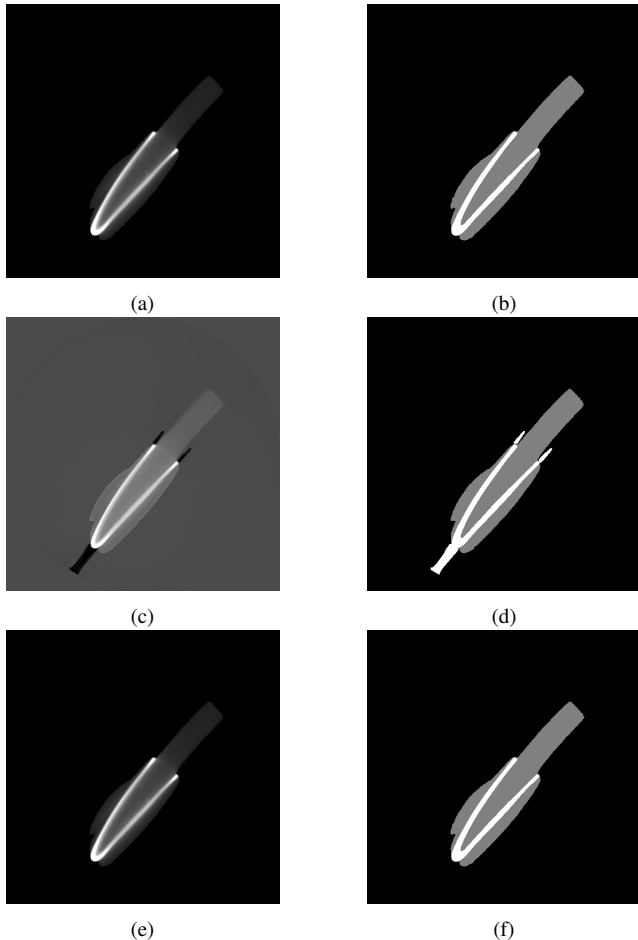


Fig. 16: Joint reconstructions and segmentations (middle slices) from 300 projections of CMTE volume by JMAP with Gauss-Markov-Potts prior and : the usual forward model modelling  $\zeta$  as heavy-tailed (reconstruction (a) and segmentation (b)), the usual forward model and modelling  $\zeta$  as Gaussian (reconstruction (c) and segmentation (d)), and the error-splitting forward model (reconstruction (e) and segmentation (f))

linear model uncertainties  $\xi$ , and accuracy with short-tailed prior on measurement uncertainties  $\epsilon$ , while the usual forward model only enables to get one of these features. These results have been obtained with Gauss-Markov-Potts prior for quasi-piecewise-constant volumes, which is useful for many CT imaging problems in industry and medicine [33]–[36]. Nevertheless, Gauss-Markov-Potts prior fails in reconstructing texture-rich details such as in woven composite materials. Since our derivation of the error-splitting forward model in section II-A is valid for any X-ray CT imaging task, future works will mainly focus on combining the error-splitting forward model with appropriate priors for textured, or more generally no longer piecewise-constant, volumes.

## VI. CONCLUSION

Taking into account the error in the monochromatic model with respect to the polychromatic one, we have included a noise term in the mean of Poisson statistics of photon-count in 3D X-ray CT. From this addition, we have derived a new forward model for Model-Based Iterative Reconstruction (MBIR) methods, where two terms of uncertainties are present instead of one. This error-splitting forward model distinguishes what we called measurement uncertainties and linear model uncertainties. Measurement uncertainties appeared as Gaussian in our second order Taylor series expansion of the Poisson log-likelihood, while linear model uncertainties appeared as an unknown bias between measured and theoretical projections  $g$  and  $Hf$ . Thanks to the use of generalized Student- $t$  distribution, a heavy-tailed prior has been assigned to this bias in order to ensure robustness in the reconstruction process. We have given details about how to fix each parameter of the error-splitting forward model, so each uncertainty was modeled as we wished.

Compared to the usual forward model, the error-splitting forward model adds very few calculations thanks to the simplicity of generalized Student- $t$  distribution. We have combined this new forward model and Gauss-Markov-Potts prior in order to reconstruct piecewise-constant volumes. In our experiments with real data, we have compared our new forward model with the usual one. Thanks to our better physical modelling, we have shown that, with Gauss-Markov-Potts prior, using the error-splitting forward model enables to obtain more accurate and more robust reconstructions, while only one of these two features is achievable with the usual forward model. Since our experiments have been done with Gauss-Markov-Potts prior model on the volume, in future works, the emphasis will be on the impact of the error-splitting forward model combined with other priors. Efforts will also be made to see how to improve the error-splitting forward model to tackle strong scattering around titanium edges of the second volume used in our experiments. Concerning Gauss-Markov-Potts prior, the acceleration of the segmentation step on GPU will also be done.

## APPENDIX A

### QUADRATIC APPROXIMATION FOR THE LOG-LIKELIHOOD TO DERIVE THE ERROR-SPLITTING FORWARD MODEL

From equation (4) and by performing a second order Taylor series expansion, we get

$$I_0 e^{-[Hf]_i} + n_i + r_i = I_i + n_i + (I_i - r_i)(g_i - [Hf]_i) + \frac{1}{2}(I_i - r_i)(g_i - [Hf]_i)^2 + o((g_i - [Hf]_i)^2). \quad (66)$$

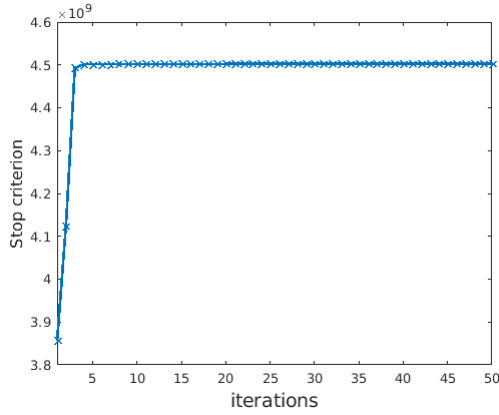


Fig. 17: Convergence for JMAP with the error-splitting forward model and Gauss-Markov-Potts prior for CMTE reconstruction from 300 projections

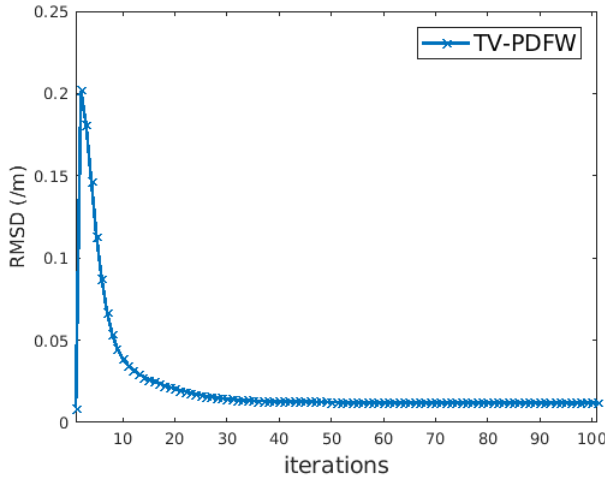


Fig. 18: Convergence to reference volume for CMTE reconstruction by TV-regularized least-squares

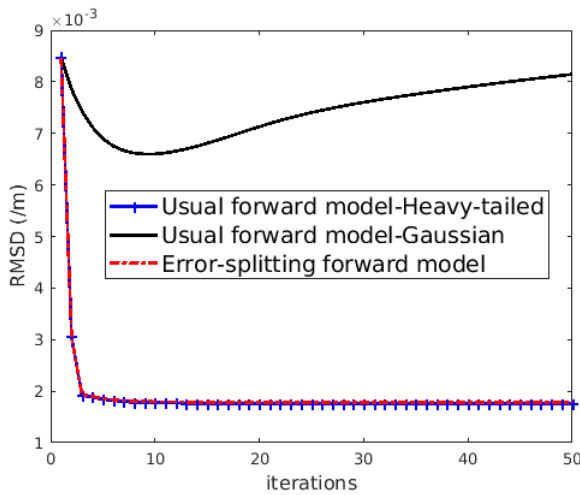


Fig. 19: Convergence to reference volume for CMTE reconstruction by JMAP from 300 projections with Gauss-Markov-Potts prior and : the usual forward model modelling  $\zeta$  as heavy-tailed, the usual forward model modelling  $\zeta$  as Gaussian, and the error-splitting forward model

Hence, we have

$$\begin{aligned} I_i \ln \left( I_0 e^{-[\mathbf{H}\mathbf{f}]_i} + n_i + r_i \right) &= I_i \ln(I_i + n_i) \\ &+ \frac{I_i(I_i - r_i)}{I_i + n_i} (g_i - [\mathbf{H}\mathbf{f}]_i) \\ &+ \frac{I_i(I_i - r_i)(r_i + n_i)}{2(I_i + n_i)^2} (g_i - [\mathbf{H}\mathbf{f}]_i)^2 \\ &+ o\left((g_i - [\mathbf{H}\mathbf{f}]_i)^2\right). \end{aligned} \quad (67)$$

From (9), and after reparametrization (4), it leads to

$$\begin{aligned} -\ln(p(g_i|\mathbf{f}, n_i, r_i)) &= \ln(I_i!) + I_i + n_i - I_i \ln(I_i + n_i) \\ &+ b_i(g_i - [\mathbf{H}\mathbf{f}]_i) + \frac{a_i}{2}(g_i - [\mathbf{H}\mathbf{f}]_i)^2 \\ &+ o\left((g_i - [\mathbf{H}\mathbf{f}]_i)^2\right) \end{aligned} \quad (68)$$

which leads to approximation (14), with  $a_i$  and  $b_i$  given by (11).

## APPENDIX B

### CALCULATIONS FOR FIXING THE PARAMETERS OF THE ERROR-SPLITTING FORWARD MODEL

We suppose  $\beta_{\epsilon_0}$  is correctly fixed to comply with constraint (28), so measurement uncertainties are actually Gaussian. In order to do simple calculations, we make the assumption that true projections and measurement uncertainties are independent, and remove the expectations :

$$\|\mathbf{g}\|_2^2 = \|\mathbf{g}_0\|_2^2 + \|\boldsymbol{\epsilon}\|_2^2. \quad (69)$$

We define parameter  $SNR$  by :

$$SNR = 10 \log \left( \frac{\|\mathbf{g}_0\|_2^2}{\|\boldsymbol{\epsilon}\|_2^2} \right) = 10 \log \left( \frac{\|\mathbf{g}\|_2^2}{\|\boldsymbol{\epsilon}\|_2^2} - 1 \right). \quad (70)$$

Since we approximate  $\|\boldsymbol{\epsilon}\|_2^2$  by its expectation, we have

$$\|\boldsymbol{\epsilon}\|_2^2 \approx \mathbb{E}(\|\boldsymbol{\epsilon}\|_2^2 | \alpha_{\epsilon_0}, \beta_{\epsilon_0}) = \sum_{i=1}^M \mathbb{E}(\epsilon_i^2 | \alpha_{\epsilon_0}, \beta_{\epsilon_0}) \quad (71)$$

and

$$\mathbb{E}(\epsilon_i^2 | \alpha_{\epsilon_0}, \beta_{\epsilon_0}) = \frac{\beta_{\epsilon_0}}{\alpha_{\epsilon_0} + \frac{1}{2}} \quad (72)$$

according to (30). Combining equations (70) and (71), we obtain formula (31) to automatically fix  $\alpha_{\epsilon_0}$  with respect to the data and  $SNR$ .

## ACKNOWLEDGEMENTS

The authors are grateful to Lionel Gay and Nicolas Cochenec at Safran Composites for having provided the data used for the tests.

## REFERENCES

- [1] K. Lange, R. Carson *et al.*, "EM reconstruction algorithms for emission and transmission tomography," *J Comput Assist Tomogr.*, vol. 8, no. 2, pp. 306–16, 1984.
- [2] K. T. Smith and F. Keinert, "Mathematical foundations of computed tomography," *Applied Optics*, vol. 24, no. 23, pp. 3950–3957, 1985.
- [3] J. Idier, *Bayesian approach to inverse problems*. John Wiley & Sons, 2008.
- [4] L. Feldkamp, L. Davis, and J. Kress, "Practical cone-beam algorithm," *JOSA A*, vol. 1, no. 6, pp. 612–619, 1984.

- [5] R. L. Siddon, "Fast calculation of the exact radiological path for a three-dimensional CT array," *Medical physics*, vol. 12, no. 2, pp. 252–255, 1985.
- [6] R. Gordon, R. Bender, and G. T. Herman, "Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and X-ray photography," *Journal of theoretical Biology*, vol. 29, no. 3, pp. 471–481, 1970.
- [7] P. Gilbert, "Iterative methods for the three-dimensional reconstruction of an object from projections," *Journal of theoretical biology*, vol. 36, no. 1, pp. 105–117, 1972.
- [8] A. H. Andersen and A. C. Kak, "Simultaneous algebraic reconstruction technique (SART): a superior implementation of the ART algorithm," *Ultrasonic imaging*, vol. 6, no. 1, pp. 81–94, 1984.
- [9] J. A. Fessler, "Statistical image reconstruction methods for transmission tomography," *Handbook of medical imaging*, vol. 2, pp. 1–70, 2000.
- [10] E. Y. Sidky, H. Jakob, X. Pan *et al.*, "Convex optimization problem prototyping for image reconstruction in computed tomography with the Chambolle & Pock algorithm," *Physics in medicine and biology*, vol. 57, no. 10, p. 3065, 2012.
- [11] Y. Boykov, O. Veksler, and R. Zabih, "Fast approximate energy minimization via graph cuts," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 23, no. 11, pp. 1222–1239, 2001.
- [12] R. Ramlau and W. Ring, "A Mumford–Shah level-set approach for the inversion and segmentation of X-ray tomography data," *Journal of Computational Physics*, vol. 221, no. 2, pp. 539–557, 2007.
- [13] M. Storath, A. Weinmann, J. Frikkel, and M. Unser, "Joint image reconstruction and segmentation using the Potts model," *Inverse Problems*, vol. 31, no. 2, p. 025003, 2015.
- [14] C. Chapdelaine, A. Mohammad-Djafari, N. Gac, and E. Parra, "A 3D Bayesian Computed Tomography Reconstruction Algorithm with Gauss-Markov-Potts Prior Model and its Application to Real Data," *Fundamenta Informaticae*, vol. 155, no. 4, pp. 373–405, 2017.
- [15] B. Vandeghinste, B. Goossens, R. Van Holen, C. Vanhove, A. Pižurica, S. Vandenberghe, and S. Staels, "Iterative CT reconstruction using shearlet-based regularization," *IEEE Transactions on Nuclear Science*, vol. 60, no. 5, pp. 3305–3317, 2013.
- [16] L. Wang, A. Mohammad-Djafari, and N. Gac, "X-ray computed tomography using a sparsity enforcing prior model based on haar transformation in a bayesian framework," *Fundamenta Informaticae*, vol. 155, no. 4, pp. 449–480, 2017.
- [17] Q. Xu, H. Yu, X. Mou, L. Zhang, J. Hsieh, and G. Wang, "Low-dose X-ray CT reconstruction via dictionary learning," *IEEE Transactions on Medical Imaging*, vol. 31, no. 9, pp. 1682–1697, 2012.
- [18] X. Zheng, S. Ravishankar, Y. Long, and J. A. Fessler, "PWLS-ULTRA: An efficient clustering and learning-based approach for low-dose 3D CT image reconstruction," *IEEE transactions on medical imaging*, vol. 37, no. 6, pp. 1498–1510, 2018.
- [19] J. Sanders and E. Kandrot, *CUDA by Example: An Introduction to General-Purpose GPU Programming, Portable Documents*. Addison-Wesley Professional, 2010.
- [20] D. Kim, D. Pal, J.-B. Thibault, and J. A. Fessler, "Accelerating ordered subsets image reconstruction for X-ray CT using spatially nonuniform optimization transfer," *IEEE transactions on medical imaging*, vol. 32, no. 11, pp. 1965–1978, 2013.
- [21] D. Kim, S. Ramani, and J. A. Fessler, "Combining ordered subsets and momentum for accelerated X-ray CT image reconstruction," *IEEE transactions on medical imaging*, vol. 34, no. 1, pp. 167–178, 2015.
- [22] B. De Man and S. Basu, "Distance-driven projection and backprojection in three dimensions," *Physics in medicine and biology*, vol. 49, no. 11, p. 2463, 2004.
- [23] Y. Long, J. A. Fessler, and J. M. Balter, "3D forward and back-projection for X-ray CT using separable footprints," *IEEE transactions on medical imaging*, vol. 29, no. 11, pp. 1839–1850, 2010.
- [24] J. Nuyts, B. De Man, J. A. Fessler, W. Zbijewski, and F. J. Beekman, "Modelling the physics in the iterative reconstruction for transmission computed tomography," *Physics in medicine and biology*, vol. 58, no. 12, p. R63, 2013.
- [25] K. Sauer and C. Bouman, "A local update strategy for iterative reconstruction from projections," *IEEE Transactions on Signal Processing*, vol. 41, no. 2, pp. 534–548, 1993.
- [26] J.-B. Thibault, C. A. Bouman, K. D. Sauer, and J. Hsieh, "A recursive filter for noise reduction in statistical iterative tomographic imaging," in *Electronic Imaging 2006*. International Society for Optics and Photonics, 2006, pp. 60650X–60650X.
- [27] J.-B. Thibault, K. D. Sauer, C. A. Bouman, and J. Hsieh, "A three-dimensional statistical approach to improved image quality for multislice helical CT," *Medical physics*, vol. 34, no. 11, pp. 4526–4544, 2007.
- [28] L. Gjestebj, B. De Man, Y. Jin, H. Paganetti, J. Verburg, D. Giantsoudi, and G. Wang, "Metal artifact reduction in CT: Where are we after four decades?" *IEEE Access*, vol. 4, pp. 5826–5849, 2016.
- [29] P. Jin, C. A. Bouman, and K. D. Sauer, "A Model-Based Image Reconstruction Algorithm With Simultaneous Beam Hardening Correction for X-Ray CT," *IEEE Trans. Computational Imaging*, vol. 1, no. 3, pp. 200–216, 2015.
- [30] C. Martinez, J. A. Fessler, M. Desco, and M. Abella, "Statistical Image Reconstruction with Sample-Based Beam-Hardening compensation for X-ray CT," in *The Fifth International Conference on Image Formation in X-Ray Computed Tomography*, 2018.
- [31] M. Dumitru, "A Bayesian Approach for Periodic Components Estimation for Chronobiological Signals," Ph.D. dissertation, Université Paris-Saclay, 2016.
- [32] M. Dumitru, N. Gac, L. Wang, and A. Mohammad-Djafari, "Unsupervised sparsity enforcing iterative algorithms for 3D image reconstruction in X-ray Computed Tomography," in *The 14th International Meeting on Fully Three-Dimensional Image Reconstruction in Radiology and Nuclear Medicine*, 2017.
- [33] K. J. Batenburg and J. Sijbers, "DART: a practical reconstruction algorithm for discrete tomography," *Image Processing, IEEE Transactions on*, vol. 20, no. 9, pp. 2542–2553, 2011.
- [34] M. Zisler, J. H. Kappes, C. Schnörr, S. Petra, and C. Schnörr, "Non-binary discrete tomography by continuous non-convex optimization," *IEEE Transactions on Computational Imaging*, vol. 2, no. 3, pp. 335–347, 2016.
- [35] M. L. Bouxsein, S. K. Boyd, B. A. Christiansen, R. E. Guldberg, K. J. Jepsen, and R. Müller, "Guidelines for assessment of bone microstructure in rodents using micro-computed tomography," *Journal of bone and mineral research*, vol. 25, no. 7, pp. 1468–1486, 2010.
- [36] L. Wang, B. Sixou, S. Rit, and F. Peyrin, "Binary tomography reconstruction from few projections with Total Variation regularization for bone microstructure studies," *Journal of X-ray science and technology*, vol. 24, no. 2, pp. 177–189, 2016.
- [37] M. Storath, D. Rickert, M. Unser, and A. Weinmann, "Fast segmentation from blurred data in 3D fluorescence microscopy," *IEEE Transactions on Image Processing*, vol. 26, no. 10, pp. 4856–4870, 2017.
- [38] J. Besag, "Spatial interaction and the statistical analysis of lattice systems," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 192–236, 1974.
- [39] O. Féron, B. Duchêne, and A. Mohammad-Djafari, "Microwave imaging of inhomogeneous objects made of a finite number of dielectric and conductive materials from experimental data," *Inverse Problems*, vol. 21, no. 6, p. S95, 2005.
- [40] H. Ayasso, B. Duchêne, and A. Mohammad-Djafari, "Bayesian inversion for optical diffraction tomography," *Journal of Modern Optics*, vol. 57, no. 9, pp. 765–776, 2010.
- [41] X. Zhuge, W. J. Palenstijn, and K. J. Batenburg, "TVR-DART: a more robust algorithm for discrete tomography from limited projection data with automated gray value estimation," *IEEE Transactions on Image Processing*, vol. 25, no. 1, pp. 455–468, 2016.
- [42] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.
- [43] T. M. Quan and W.-K. Jeong, "A fast discrete wavelet transform using hybrid parallelism on GPUs," *IEEE Transactions on Parallel and Distributed Systems*, vol. 27, no. 11, pp. 3088–3100, 2016.
- [44] J. Besag, "On the statistical analysis of dirty pictures," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 259–302, 1986.
- [45] G. Ongie, N. Murthy, L. Balzano, and J. A. Fessler, "A Memory-Efficient Algorithm for Large-Scale Sparsity Regularized Image Reconstruction," in *The Fifth International Conference on Image Formation in X-Ray Computed Tomography*, 2018.
- [46] N. Gac, A. Vabre, A. Mohammad-Djafari, A. Rabanal, and F. Buyens, "GPU implementation of a 3D Bayesian CT algorithm and its application on real foam reconstruction," in *The First International Conference on Image Formation in X-Ray Computed Tomography*, 2010, pp. 151–155.
- [47] G. L. Zeng and G. T. Gullberg, "Unmatched projector/backprojector pairs in an iterative reconstruction algorithm," *IEEE transactions on medical imaging*, vol. 19, no. 5, pp. 548–555, 2000.
- [48] L. Gay, P. Arslan, Y. Rambourg, and A. Chandelle, "Fantôme destiné à être utilisé pour le contrôle de la qualité d'images tomographiques," France Patent 3 030 753, June 24, 2016.
- [49] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.