Penalized PET/CT Reconstruction Algorithms with Automatic Realignment for Anatomical Priors

Yu-Jung Tsai, Alexandre Bousse, Simon Arridge, Charles W. Stearns, *Fellow, IEEE*, Brian F. Hutton, *Senior Member, IEEE* and Kris Thielemans, *Senior Member, IEEE*

Abstract—Two algorithms for solving misalignment issues in penalized PET/CT reconstruction using anatomical priors are proposed. Both approaches are based on a recently published joint motion estimation and image reconstruction method. The first approach deforms the anatomical image to align it with the functional one while the second approach deforms both images to align them with the measured data. Our current implementation alternates between alignment estimation and image reconstruction. We have chosen Parallel Level Sets (PLS) as a representative anatomical penalty, incorporating a spatially-variant penalty strength. The performance was evaluated using simulated non-TOF data generated with an XCAT phantom in the thorax region. We used the attenuation map in the anatomical prior. The results demonstrated that both methods can estimate the misalignment and deform the anatomical image accordingly. However, the performance of the first approach depends highly on the workflow of the alternating process. The second approach shows a faster convergence rate to the correct alignment and is less sensitive to the workflow. The presence of anatomical information improves the convergence rate of misalignment estimation for the second approach but slow it down for the first approach. Both approaches show improved performance in misalignment estimation as the data noise level decreases.

I. INTRODUCTION

Penalized maximum-likelihood (PML) image reconstruction using penalties derived from anatomical images, such as computed tomography (CT) or magnetic resonance (MR) images, has been shown to be effective in improving object delineation and reducing quantitative error in many studies [1]–[10]. However, in order to utilize the structural information without incurring artifacts, a good alignment between the anatomical and functional images is essential [6], [9], [11]–[13]. This is challenging in practice because these images are most likely obtained separately or sequentially. Even with a multi-modality scanner that performs simultaneous acquisition

Y.-J. Tsai* was with the Institute of Nuclear Medicine, University College London, London NW1 2BU, UK and is now with the Yale PET Center, Yale University, New Haven, CT 06520, USA. (e-mail: yu-jung.tsai@yale.edu).

A. Bousse was with the Institute of Nuclear Medicine, University College London, London, UK and is now with the LaTIM, INSERM, UMR 1101, Université de Bretagne Occidentale, Brest, France.

S. Arridge is with the Department of Computer Science, University College London, London WC1E 6BT, UK.

C. W. Stearns is with MICT Engineering, GE Healthcare, Waukesha, WI 53188 USA.

B. F. Hutton is with the Institute of Nuclear Medicine, University College London, London, UK.

K. Thielemans is with the Institute of Nuclear Medicine, University College London, London NW1 2BU, UK.

(e.g., Siemens mMR system and GE Signa PET/MR), good alignment between the acquisitions is still difficult to achieve due to the different time scales of the scans. For the mis-alignment involving non-rigid deformation, for instance due to respiration, the assumption is even harder to satisfy and accurate image segmentation and co-registration (which are difficult and time consuming) are often required [14], [15].

In many cases, anatomical information can be derived from the attenuation image. Since the literature on solving the misalignment issue between a priori anatomical information and the functional information is quite limited, we instead seek ideas from a similar but previously studied problem in thoracic PET imaging, in which a potentially misaligned CT or MR-derived attenuation map is used for the attenuation correction. As for PML image reconstruction using anatom-ical priors, the misalignment induced by patient respiration degrades resolution of the reconstructed image and introduces artifacts where large movement or deformation of organs is observed [16], [17]. Although these methods are intended to be used for having a better attenuation corrected image from emission data, they offer insights into resolving the misalignment between the anatomical attenuation map and the functional emission image.

This study will concentrate on imaging of the thorax, where respiratory motion is a known problem [18]. One strategy to tackle the respiratory motion is to sort the acquired data from both modalities into several gates where no motion is assumed in each of them. The gated data are then paired up according to their breathing phases estimated from the data themselves or an external tracking system [19]. In addition to reconstructing these gated data pairs individually and then registering them to a reference respiratory phase [20], one can also incorporate the corresponding attenuation information into 4-dimensional (4-D) reconstruction algorithms [21], [22]. However, these methods rely on a relatively consistent breathing pattern during both scans [23]. Moreover, they imply the need of special scans to obtain the gated anatomical information. This can increase patient dose or prolong the overall study time, depending on the applied anatomical imaging modality. To adapt to irregular breathing patterns, another strategy applies individual motion model to deform the input attenuation map [24]. The model can be derived from other imaging data, such as dynamic CT and MR [25], as well as the non-attenuation corrected PET [26]. However, the former approach has the potential problem of propagating the error in the model estimation to the final reconstructed image, while the latter method's performance depends on the tracer distribution

and data statistics. Another way around these issues is to use population-based deformation models [27], [28]. However, these have not been convincingly shown to work in practice [24]. The application of applying an individual model to PML image reconstruction using anatomical priors therefore faces similar challenges. Finally, none of the above methods is able to cope with the residual misalignment caused by other general motion of the patient.

Algorithms that allow simultaneous estimation of the activity distribution and the corresponding attenuation map from the respiratory gated PET data have been proposed [29], [30] in recent years. These methods do not rely on assumptions about the breathing pattern or a pre-estimated motion model. Therefore, they have the potential to be applied to different misalignment problems without suffering from the error propagation issue. However, since the problem is very ill-conditioned, some *a priori* knowledge about the intensity distribution of the attenuation map is required. This can compromise the benefit of using anatomical information during the image reconstruction as the intensity is restricted to several values and most of the anatomical details are lost. Besides, significant cross-talk between the estimated activity and attenuation map is observed in non-time-of-flight (non-TOF) PET. Although the artifacts can be largely eliminated when TOF data are available, the *a priori* knowledge about the intensity distribution is still necessary [30].

In contrast to seeking to align the attenuation map with the emission image, a different joint estimation approach incorporates a warp matrix that deforms both the activity distribution and the attenuation map within the objective function [31]–[33]. By optimizing the objective function using an alternating process between motion estimation and image reconstruction, the motion compensation and attenuation correction are achieved simultaneously. The optimization can be applied to both non-TOF and TOF data albeit with a significantly improved convergence rate when TOF data are available [34]. The *a priori* knowledge of the attenuation distribution is not necessary anymore. This study motivated us to investigate the idea of applying a warp matrix to an anatomical prior. Extending on the existing method, two approaches that account for the misalignment between the functional and anatomical images by incorporating the warp matrix into the penalized objective function are proposed. The proposed methods do not require any image preprocessing, such as segmentation and co-registration. They are also applicable to both non-TOF and TOF datasets and no *a priori* information regarding the misalignment or intensity distribution of both images are needed. To the best of our knowledge, there are no other algorithms that show similar capacities in literature. As a special case of the application, we will only investigate the alignment of one PET position with a single CT derived attenuation map, which is also used to provide anatomical information. This paper is an extension of initial results presented in [35].

II. METHOD

A. Objective function without considering the potential misalignment

In this section, we define the PML objective function with an anatomical prior without misalignment considerations between the activity and anatomical images. Given the emission image $\boldsymbol{f} \in \mathbb{R}^J$, the anatomical image $\boldsymbol{z} \in \mathbb{R}^J$, the attenuation map $\boldsymbol{\mu} \in \mathbb{R}^J$ and the measured data $\boldsymbol{g} \in \mathbb{R}^I$, the objective function can be written as:

$$\Phi(\boldsymbol{f}) = -L(\boldsymbol{f}, \boldsymbol{g}, \boldsymbol{\mu}) + \beta R(\boldsymbol{f}|\boldsymbol{z}), \tag{1}$$

where L is the log-likelihood and R is the penalty function with a parameter β controlling its strength.

As the statistical nature of the measured data $\boldsymbol{g}$ can be described using the Poisson distribution in emission tomography, the log-likelihood function L, omitting terms independent of $\boldsymbol{f}$, can be expressed as:

$$L(\boldsymbol{f}, \boldsymbol{g}, \boldsymbol{\mu}) = \sum_i g_i \log \bar{g}_i(\boldsymbol{f}, \boldsymbol{\mu}) - \bar{g}_i(\boldsymbol{f}, \boldsymbol{\mu}),$$

$$\bar{\boldsymbol{g}}(\boldsymbol{f}, \boldsymbol{\mu}) \triangleq \mathrm{diag}\left\{ e^{-\boldsymbol{A}\boldsymbol{\mu}} \right\} \boldsymbol{A}\boldsymbol{f} + \boldsymbol{n} \tag{2}$$

where $\bar{g}_i$ is the mean measurement in bin i, $\boldsymbol{A} \in \mathbb{R}^{I \times J}$ is the system matrix which characterizes the physical system properties, such as resolution and detector sensitivity, in terms of detection probability and $\boldsymbol{n} \in \mathbb{R}^I$ is the expected background events vector. The attenuation effect is modeled explicitly by the matrix $\mathrm{diag}\left\{ e^{-\boldsymbol{A}\mu} \right\}$, where $\mathrm{diag}\left\{ \cdot \right\}$ is an operator that constructs a diagonal matrix from a vector.

Since it has shown promising results in the literature [12], [36], [37], Parallel Level Sets (PLS) is chosen as a representative anatomical penalty in this study:

$$R(\boldsymbol{f}|\boldsymbol{z}) = \sum_j \sqrt{\epsilon^2 + \| [\nabla \boldsymbol{f}]_j \|_2^2 - \langle [\nabla \boldsymbol{f}]_j, [\boldsymbol{\xi}]_j \rangle^2},$$

$$[\boldsymbol{\xi}]_j := \frac{[\nabla \boldsymbol{z}]_j}{\sqrt{\| [\nabla \boldsymbol{z}]_j \|_2^2 + \eta^2}}, \ \epsilon \text{ and } \eta > 0 \tag{3}$$

where ∇ is the gradient operator, $\langle \cdot, \cdot \rangle$ is the Euclidean scalar product and $\| \cdot \|_2$ denotes the ℓ^2-norm. The edge preserving property of PLS is modulated by the pair of parameters (ϵ, η).

B. Objective function considering the potential misalignment

Two approaches that account for the misalignment between the functional and anatomical images in penalized image reconstruction using anatomical priors are proposed in this section. Both approaches are based on a joint motion estimation and image reconstruction method proposed recently for dealing with the mismatch between the attenuation map and the PET image in respiratory gated PET/CT [31]. Instead of applying a quadratic penalty function to enforce smoothness on the reconstructed activity images as in [31], an anatomical penalty calculated with the attenuation map is employed to improve the image quality. Therefore, artifacts induced by the misalignment between the activity image and the attenuation map will be introduced through both the attenuation correction and the incorporated penalty function. The main

difference between these two approaches is that the first approach Φ_1 (Approach I) deforms the anatomical image (*i.e.*, the attenuation map) to align it with the functional image, while the second approach Φ_2 (Approach II) deforms both images to align them with the measured data. Although the deformation of the attenuation map implies the change of the scatter distribution, the estimated background events are fixed during the optimization process for simplicity. As in [31], we use uniform cubic B-splines for image interpolation and deformation. Despite the use of an anatomical prior, Approach II is quite similar to the method introduced in [31] and Approach I is essentially a simplified algorithm to Approach II. We also modified Approach II to impose positivity on image values as opposed to B-spline coefficients.

Assume a continuous volumetric image $s : \mathbb{R}^3 \to \mathbb{R}$ can be represented as a linear combination of basis functions centered on a voxel grid $\mathsf{C} = \{\boldsymbol{r}_k, \ k = 1, \ldots, \ N\} \subset \mathbb{R}^3$ that coincides with the voxel centers:

$$s(\boldsymbol{r}) = \sum_{k=1}^{N} s'_k B\left(\frac{\boldsymbol{r} - \boldsymbol{r}_k}{\sigma_1}\right), \tag{4}$$

where s'_k is the B-spline coefficient of the basis function centered on voxel k, $\boldsymbol{r} = (x, y, z)$ is the index vector in the 3-D Cartesian coordinate system, $B(\boldsymbol{r}) = b(x)b(y)b(z)$ is an interpolating function based on the cubic B-splines b and σ_1 represents the voxel-spacing. The discretized image can therefore be represented as a collection of the B-spline coefficients $\boldsymbol{s}' = (s'_j)_{j=1}^{N}$. From this section, the prime notation is used to distinguish the B-spline coefficients from the corresponding voxel values for images. Note that the coefficients of the cubic B-splines are not identical to the image values at the grid nodes. Particularly, the B-spline coefficients can be negative. The deformation of the image represented by the coefficients $\boldsymbol{s}'$ is achieved by deforming the continuous image function s followed by a re-sampling on C for every voxel j:

$$[\boldsymbol{W}\boldsymbol{s}']_j = \sum_{k=1}^{N} s'_k B\left(\frac{\nu(\boldsymbol{r}_j) - \boldsymbol{r}_k}{\sigma_1}\right), \tag{5}$$

where $\boldsymbol{W}$ is a square matrix with each element $[\boldsymbol{W}]_{j,k} \triangleq B(\frac{\nu(\boldsymbol{r}_j) - \boldsymbol{r}_k}{\sigma_1})$ and ν is the warping function. Given $\tilde{\mathsf{C}} = \{\tilde{\boldsymbol{r}}_l, \ l = 1, \ldots, \ \tilde{N}\}$ a uniform sub-grid of C with $\tilde{N}$ grid nodes, the function ν_θ can be parametrized by a collection of the deformation coefficients $\boldsymbol{\theta} = (\boldsymbol{\theta}^x, \ \boldsymbol{\theta}^y, \ \boldsymbol{\theta}^z)$:

$$\nu_\theta(\boldsymbol{r}) \triangleq \boldsymbol{r} + \begin{bmatrix} \sum_{l=1}^{\tilde{N}} \theta_l^x B(\frac{\boldsymbol{r} - \tilde{\boldsymbol{r}}_l}{\sigma_2}) \\ \sum_{l=1}^{\tilde{N}} \theta_l^y B(\frac{\boldsymbol{r} - \tilde{\boldsymbol{r}}_l}{\sigma_2}) \\ \sum_{l=1}^{\tilde{N}} \theta_l^z B(\frac{\boldsymbol{r} - \tilde{\boldsymbol{r}}_l}{\sigma_2}) \end{bmatrix} = \begin{bmatrix} \nu^x(\boldsymbol{r}) \\ \nu^y(\boldsymbol{r}) \\ \nu^z(\boldsymbol{r}) \end{bmatrix} \tag{6}$$

where σ_2 is the distance between two grid nodes.

1) Approach I: The first approach optimizes an objective function Φ_1 that considers the deformed anatomical image. Assume the attenuation map $\boldsymbol{\mu}$, represented as a collection of the B-spline coefficients $\boldsymbol{\mu}' = (\mu'_j)_{j=1}^{N}$, is used to provide anatomical information as well, Φ_1 is given by:

$$\Phi_1(\boldsymbol{f}, \boldsymbol{\mu}', \boldsymbol{\theta}) = -L(\boldsymbol{f}, \boldsymbol{g}, \boldsymbol{W}\boldsymbol{\mu}') + \beta R(\boldsymbol{f}|\boldsymbol{W}\boldsymbol{\mu}') \tag{7}$$
$$+ \gamma U(\boldsymbol{\theta}),$$

where $U(\boldsymbol{\theta})$ is a quadratic penalty on the difference between neighboring nodes of the motion grid for reducing the influence of noise and γ is a constant that controls its strength. Note that the misalignment between $\boldsymbol{\mu}$ and $\boldsymbol{f}$ affects the optimization through the attenuation correction in the log-likelihood L and the incorporated anatomical penalty R as both functions use the warped attenuation map $\boldsymbol{W}\boldsymbol{\mu}'$ as inputs. Since this approach does not require deforming the activity image, $\boldsymbol{f}$ in (7) represents a vector of image values. The positivity constraint on $\boldsymbol{f}$ can therefore be achieved by performing constrained image reconstruction. In contrast, the attenuation map is represented as a collection of coefficients $\boldsymbol{\mu}'$ for image warping using B-splines. This could lead to negative values in $\boldsymbol{W}\boldsymbol{\mu}'$ as we are optimizing the B-spline coefficients. However, since small negative values in $\boldsymbol{W}\boldsymbol{\mu}'$ would become attenuation factors very close to one, they have been left unchanged in this study. When the objective function in (7) is optimized, the warped attenuation map should be in the same space as the activity image and $e^{-\boldsymbol{A}\boldsymbol{W}\boldsymbol{\mu}}$ is a coefficient vector that accurately corrects $\boldsymbol{f}$ for attenuation in the projection domain. In other words, $\boldsymbol{W}\boldsymbol{\mu}'$ and $\boldsymbol{f}$ are aligned and adequately projected to fit the data $\boldsymbol{g}$.

2) Approach II: Instead of seeking to align the warped attenuation map with the reconstructed activity image, the second approach deforms both the anatomical and functional images in order to obtain an estimate that optimizes the objective function Φ_2:

$$\Phi_2(\boldsymbol{f}', \boldsymbol{\mu}', \boldsymbol{\theta}) = -L(\boldsymbol{W}\boldsymbol{f}', \boldsymbol{g}, \boldsymbol{W}\boldsymbol{\mu}') + \beta R(\boldsymbol{f}|\boldsymbol{\mu}) \tag{8}$$
$$+ \gamma U(\boldsymbol{\theta}) + \delta E(\boldsymbol{f}'),$$

where

$$E(\boldsymbol{s}) = \sum_{n=1}^{\hat{N}} \min(0, \ s(\nu_\theta(\hat{\boldsymbol{r}}_n)))^2 \tag{9}$$

is a barrier function that penalizes negative values in $\boldsymbol{W}\boldsymbol{f}'$ [38]. Given $\hat{\mathsf{C}} = \{\hat{\boldsymbol{r}}_n, \ n = 1, \ldots, \ \hat{N}\}$ a finer grid that contains a finite number of uniformly spaced locations in each interval of two grid nodes in the voxel grid C, the function computes the spline values (*i.e.*, the image values) centered on the finer grid and penalizes the square of any negatives. The strength of E is determined by the parameter δ. In this study, we defined the distance between two adjacent locations in $\hat{\mathsf{C}}$ equal to one quarter of the grid spacing used for the image.

Although the log-likelihood requires the warped emission and anatomical images as inputs, the anatomical penalty function is calculated with the image value of the non-warped ones. Since the motivation of using an anatomical prior is to encourage edges in the emission image corresponding to those in the anatomical one, finding common edges in the warped or non-warped images is essentially a similar optimization problem. In other words, calculating the anatomical penalty function with $\boldsymbol{W}\boldsymbol{f}'$ and $\boldsymbol{W}\boldsymbol{\mu}'$ should lead to (nearly) the same solution.

In summary, when the objective function in (8) is optimized, the emission and anatomical images are aligned such that the attenuation corrected projections calculated with $\boldsymbol{W}\boldsymbol{f}'$ and $\boldsymbol{W}\boldsymbol{\mu}'$ fit the measured data $\boldsymbol{g}$.

C. Algorithm implementation

In both approaches the optimization was implemented as an alternating process that includes a misalignment estimation subroutine and a penalized image reconstruction subroutine. Pseudo-code that summarizes the implementation can be found in Algorithm 1. The workflow was defined by the number of inner iterations (InnerIter1 and InnerIter2 in Algorithm 1) for these two subroutines and the number of outer iterations (OuterIter in Algorithm 1) that controls the repetition of the alternating process. Since limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) based algorithms show the potential to work with a relatively wide range of penalty functions [39], we applied L-BFGS for unconstrained optimization (misalignment estimation in both approaches and image reconstruction in the second approach) and L-BFGS-B [40] for the positivity constrained image reconstruction in Approach I. The statement $x \leftarrow$ L-BFGS(G, x_0, Niter) in Algorithm 1 means optimizing function G using algorithm L-BFGS with initial point x_0 and Niter iterations, where x can be either activity image or deformation coefficients depending on the context. Ordered subsets expectation maximization (OS-EM) was used to reconstruct the initial activity image $\boldsymbol{f}_0$ of the whole process. The misalignment estimation was then initialized by $\boldsymbol{f}_0$ and the attenuation map $\boldsymbol{\mu}$. The implementation of the misalignment estimation employed in this study was originally proposed in [31]. Every time the misalignment estimation was done, a new initial image $\boldsymbol{f}_{\text{InitInner}}$ for the penalized image reconstruction was recomputed using OS-EM, taking into account the current estimated misalignment. To improve the convergence rate of the penalized image reconstruction, a preconditioner proposed in [41] was incorporated into both approaches. A spatially-variant penalization scheme was also applied to the anatomical penalty function to further achieve uniform local contrast [39]. Both the preconditioner and the spatially-variant penalty strength were calculated with the initial image from OS-EM at every outer iteration as well.

III. EVALUATION

Since both approaches should be able to find a warped attenuation map that helps minimize the objective function, the evaluation is focused on the performance of the misalignment subroutine of each approach.

A. Data

Two XCAT phantoms [42] representing different respiratory phases and the corresponding $\boldsymbol{\mu}$ maps (Figure 1) were produced to simulate different PET positions. Both phantoms were a $128 \times 128 \times 47$ matrix with voxel size of 3.906 mm in all directions. The set of images at end inspiration was forward projected in parallel acquisition geometry to generate

---

Algorithm 1: Pseudo-code for the alternating process

Input: Data $\boldsymbol{g}$, attenuation map $\boldsymbol{\mu}$ and strength of each penalty function (*i.e.*, the set of parameters (β, γ) for Approach I and (β, γ, δ) for Approach II

Output: Estimated tracer distribution $\boldsymbol{f}$ and B-spline deformation coefficient $\boldsymbol{\theta}$

$\boldsymbol{\theta}_0 \leftarrow \boldsymbol{0}$;
$\boldsymbol{f}_0 \leftarrow$ OS-EM$(\boldsymbol{g}, \boldsymbol{\mu}, \boldsymbol{\theta}_0)$;
for $t = 0, \ldots, \text{OuterIter} - 1$ **do**
  misalignment estimation;
  if Approach I **then**
    Define $F \colon \boldsymbol{\theta} \mapsto \Phi_1(\boldsymbol{f}_t, \boldsymbol{\mu}, \boldsymbol{\theta})$;
  else if Approach II **then**
    Define $F \colon \boldsymbol{\theta} \mapsto \Phi_2(\boldsymbol{f}_t, \boldsymbol{\mu}, \boldsymbol{\theta})$;
  end
  $\boldsymbol{\theta}_{t+1} \leftarrow$ L-BFGS $(F, \boldsymbol{\theta}_t, \text{InnerIter1})$;
  image reconstruction;
  $\boldsymbol{f}_{\text{InitInner}} \leftarrow$ OS-EM$(\boldsymbol{g}, \boldsymbol{\mu}, \boldsymbol{\theta}_{t+1})$;
  if Approach I **then**
    Define $H \colon \boldsymbol{f} \mapsto \Phi_1(\boldsymbol{f}, \boldsymbol{\mu}, \boldsymbol{\theta}_{t+1})$;
    $\boldsymbol{f}_{\text{InnerIter2}} \leftarrow$ L-BFGS-B$(H, \boldsymbol{f}_{\text{InitInner}}, \text{InnerIter2})$;
  else if Approach II **then**
    Define $H \colon \boldsymbol{f} \mapsto \Phi_2(\boldsymbol{f}, \boldsymbol{\mu}, \boldsymbol{\theta}_{t+1})$;
    $\boldsymbol{f}_{\text{InnerIter2}} \leftarrow$ L-BFGS$(H, \boldsymbol{f}_{\text{InitInner}}, \text{InnerIter2})$;
  end
  $\boldsymbol{f}_{t+1} \leftarrow \boldsymbol{f}_{\text{InnerIter2}}$;
end

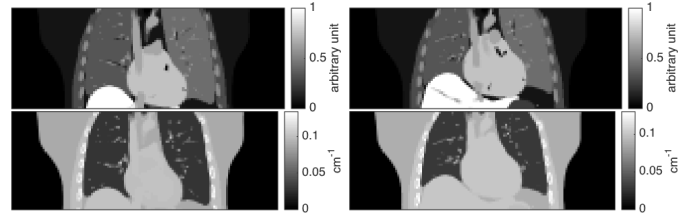Fig. 1. The central coronal view of the XCAT phantoms representing end inspiration (top left) and expiration (top right). The corresponding $\boldsymbol{\mu}$ maps are also provided (bottom).

data with size of 128 bins $\times$ 140 angles $\times$ 47 slices. A uniform projection with a constant intensity was added to the generated data to simulate the background events equivalent to 78.66% of the total prompts. Both approaches were initially evaluated with a noiseless dataset. To further assess performance in the presence of noise, three datasets with total counts of 16 M, 54 M and 161 M were also simulated using a built-in Poisson noise model in MATLAB. The same as for the noiseless dataset, uniform background events were simulated and added to the generated sinograms. The amount of background events was equivalent to 78.66% of the total prompts for each noisy dataset. Note that all simulations took into account the attenuation effect and system blurring using FWHM = 5 mm in all directions. To simulate the misalignment between functional and anatomical images, the attenuation map at end expiration was used as the initial input for both attenuation correction and misalignment estimation.

B. Reconstruction

The selection of the strength of each penalty was based on an initial investigation, where the difference between the

warped and target μ maps at OuterIter = 30 was visually compared with respect to a given set of candidate strengths (results not shown). A simple grid search using relatively large magnitude differences between consecutive candidates (10 fold for β and δ (when applicable), and 100 fold for γ) was performed on the noiseless dataset. The combination of penalty parameters that resulted in the best visual alignment in the diaphragm and heart regions was recorded. As a result, for the noiseless dataset, the set of parameters that determines the strength of each penalty function was $(\beta, \gamma) = (7 \times 10^{-3}, 10^{-4})$ for Φ_1 and $(\beta, \gamma, \delta) = (7 \times 10^{-3}, 10^{-4}, 10^{-1})$ for Φ_2. A stronger $\beta = 2 \times 10^{-1}$ was used when data with noise were considered for either approach. The strength of other penalty function(s) remained the same as for the noiseless dataset. The parameter ϵ in PLS was fixed at 10^{-1} for the noiseless data. To have similar local effect of PLS for data with different noise level, ϵ was chosen according to the scale of the statistics to reconstruct the activity images. Therefore, it was 10^{-2}, 3×10^{-2} and 10^{-1} for data with total counts of 16 M, 54 M and 161 M, respectively. As the scale of the anatomical image, *i.e.*, the attenuation map μ, did not vary with data noise level, $\eta = 10^{-2}$ was used for all datasets. The distance between two grid nodes for the deformation model was 6 voxels. The alternating process as well as the image reconstruction subroutine at every outer iteration were initialized by one full iteration of OS-EM with 14 subsets. This initial image was also used for calculating the preconditioner and the spatially-variant penalty strength for the anatomical prior. Up to 100 outer iterations were used for both approaches. Each reconstructed activity image and the warped attenuation map had the same matrix dimension and voxel size as that of the phantoms.

C. Analysis

The influence of the workflow configuration on the performance of misalignment estimation of each approach was studied with both noiseless and noisy data. For the noiseless data, a two-part study was performed. In the first part of the study, we used 1 inner iteration for the misalignment estimation and explored the minimum iterations required for the image reconstruction subroutine to obtain satisfactory results. The studied inner iteration numbers for the image reconstruction subroutine (InnerIter2) were 1, 5, 10 and 20. We then fixed the iteration number for the image reconstruction to the optimal value found in the first part and increased the number of iterations for the misalignment estimation (InnerIter1) from 1 to 5, 15 or 30 to assess if the performance of the misalignment estimation can be improved by using a higher inner iteration number. Up to 100 outer iterations (controlling the repetition of the alternating process) were performed. A workflow that provided satisfactory visual alignment in the diaphragm and heart regions for both approaches was also noted such that their performance at the same computational cost could be compared. This common workflow was then also used to assess if the use of an anatomical prior is beneficial for either approach. Reconstructions without considering any anatomical information were achieved by substituting the anatomical image (*i.e.*, μ map at end expiration) with a uniform image

when calculating PLS, therefore it is equivalent to using a (smooth) total variation (TV) prior.

The results for the noiseless data formed the basis for the investigation on noisy datasets. We started with seeking for a common workflow for both approaches using the dataset with total counts of 161 M. Recall that a stronger β was used for the anatomical penalty function in the presence of noise. As the convergence rate of the preconditioned L-BFGS-B (L-BFGS-B-PC) varies with the strength of the penalty function and data noise level [41], the alternating process was also performed with a higher InnerIter2 = 20 or 40 for both approaches. The number of outer iterations and the number of inner iterations for the misalignment subroutine were kept the same as for the noiseless dataset. We evaluated the value of InnerIter2 that was able to provide satisfactory results at 100 outer iterations for both approaches and used it to determine the common workflow. This workflow was then applied to all noisy datasets. We did not seek for the optimal common workflow for each noise level; this ensured that the sensitivity of the approaches to noise was explored without interferences from changing of workflow.

The difference image between the warped and target μ maps at a given outer iteration was used to evaluate the performance of misalignment estimation. As the lungs are the main target of the respiratory motion alignment, a mask slightly larger than the lungs was applied to the difference images and the root-mean-square errors (RMSE) was computed in the mask to reflect the misalignment estimation of borders and small structures of the lungs. The yellow contour in the top-left image of Fig. 2 illustrates the edges of the mask in the central coronal view. The RMSE in the mask was then plotted against the outer iteration numbers to assist evaluations of the performance change of each approach in misalignment estimation in response to the change of workflow configuration, availability of anatomical information and data noise level.

IV. RESULTS

A. Influence of workflow configuration (noiseless data)

For the sub-study in which the inner iteration number for the misalignment estimation was fixed at 1, the central coronal view of the difference images between the warped and target μ maps at OuterIter = 100 are shown in Fig. 2. For both approaches, using InnerIter2 = 1 was problematic, resulting in severe distortion of structures in the warped μ map (Fig. 2, top row). Satisfactory results were obtained with Approach I when InnerIter2 = 5 or 10 was chosen (Fig. 2, left column, second and third rows). However, when a higher InnerIter2 = 20 was applied, the misalignment around the diaphragm (Fig. 2, left column, bottom row) was still observed after 100 outer iterations. In contrast to Approach I, the performance of the misalignment estimation of Approach II was improved as InnerIter2 increased (Fig. 2, right column). When InnerIter2 ≥ 10 was used, the algorithm was able to realign the input μ map to the target one at OuterIter = 100.

The RMSE in the mask was consistent with the visual observation from the difference images. As shown in Fig. 3, for both approaches, at OuterIter = 100, the RMSE in the mask
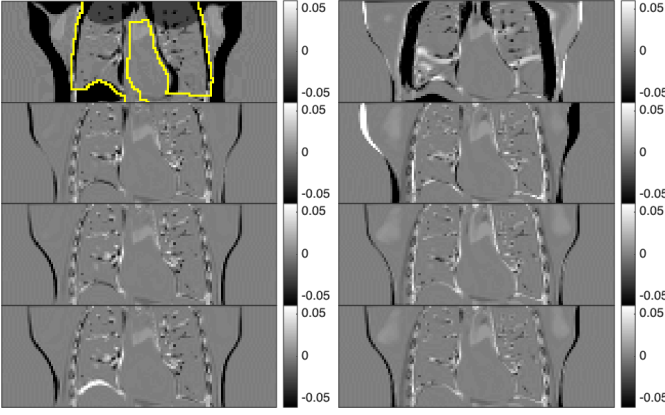
Fig. 2. The central coronal view of the difference images between the target and $W\mu'$ maps for Approach I (left column) and Approach II (right column) at 100 outer iterations. The applied workflows were 1 inner iteration for the misalignment estimation and 1 (top row), 5 (second row), 10 (third row) and 20 (bottom row) inner iterations for the image reconstruction. The noiseless dataset was used. The yellow contour in the top-left image illustrates the edges of the mask in the central coronal view.
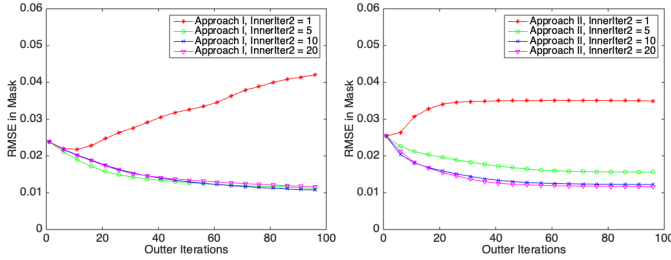
Fig. 3. The RMSE in the mask plotted against the outer iteration numbers for Approach I (left) and II (right). The applied workflows were 1 inner iteration for the misalignment estimation (InnerIter1) and 1, 5, 10 or 20 inner iterations for the image reconstruction (InnerIter2). The noiseless dataset was used.
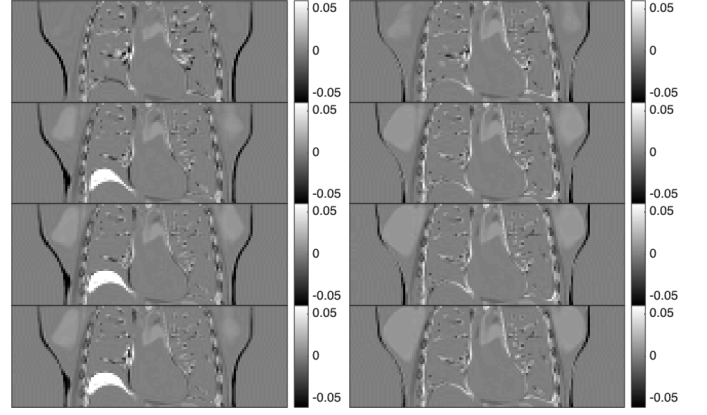

Fig. 4. The central coronal view of the difference images between the target and $W\mu'$ maps for Approach I (left column) and Approach II (right column) after 100 outer iterations. The applied workflows were 10 inner iteration for the image reconstruction and 1 (top row), 5 (second row), 15 (third row) and 30 (bottom row) inner iterations for the misalignment estimation. The noiseless dataset was used.
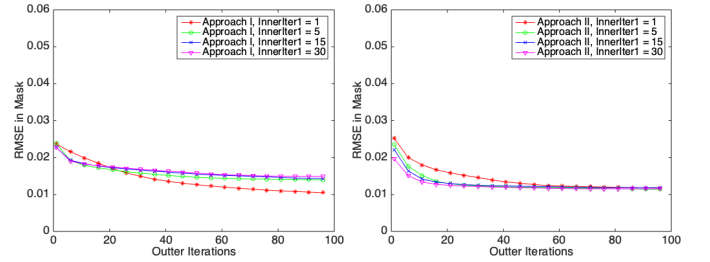

Fig. 5. The RMSE in the mask plotted against the outer iteration numbers for Approach I (left) and II (right). The applied workflows were 10 inner iteration for the image reconstruction (InnerIter2) and 1, 5, 15 or 30 inner iterations for the misalignment estimation (InnerIter1). The noiseless dataset was used.

reached the highest error when InnerIter2 = 1 was used. For Approach I, the optimal InnerIter2 was 10, while the RMSE in the mask decreased as InnerIter2 increased for Approach II. In terms of the convergence rate, Approach II was able to reach a relatively stable RMSE in the mask after 60 outer iterations. In contrast, Approach I did not yet reach a stable RMSE in the lungs at 100 outer iterations for all evaluated workflows.

Since both approaches provided visually and numerically good results when InnerIter2 = 10 was used, we fixed InnerIter2 = 10 and increased the iteration number used in the misalignment estimation subroutine. Fig. 4 shows the central coronal view of the difference image between the warped and target μ maps for both approaches with various InnerIter1 and a fixed InnerIter2 = 10. The misalignment around the diaphragm region became apparent for Approach I at OuterIter = 100 as InnerIter1 > 1 was chosen (Fig. 4, left column). For Approach II, however, the difference images at OuterIter = 100 were visually identical, regardless of the number of the applied InnerIter1 (Fig. 4, right column).

The RMSE in the mask for Approach I and II with different InnerIter1 plotted against the outer iteration numbers are given in Fig. 5. Consistent with the visual comparison, for the first approach, the smallest RMSE in the mask at OuterIter = 100 was achieved by the workflow with InnerIter1 = 1.

When Approach II was adopted, all workflows with different InnerIter1 settings were able to achieve a similar RMSE in the mask at OuterIter = 100. The convergence rate of the misalignment estimation of Approach II was improved as InnerIter1 increased. However, the performance of Approach I in terms of the convergence rate of the RMSE in the mask seemed insensitive to the change of the inner iteration number for the misalignment estimation when the applied InnerIter1 was larger than 1.

Based on the results shown in this section, 1 iteration of misalignment estimation (InnerIter1 = 1), followed by 10 iterations of image reconstruction (InnerIter2 = 10) was defined as the workflow that provides satisfactory results for these two approaches when the noiseless dataset is considered. The corresponding central coronal view of the reconstructed functional images at OuterIter = 100 are provided in Fig. 6 (top row). Note that the anatomical information was utilized throughout this two-part study. As shown in the figure, the selection of each penalty strength based on the performance of misalignment estimation has led to over-regularized functional images. Therefore, an additional reconstruction (using the same framework as the image reconstruction subroutine of Approach I) with a smaller $\beta = 7 \times 10^{-5}$ was also performed following the last outer iteration (Fig. 6, second row). To
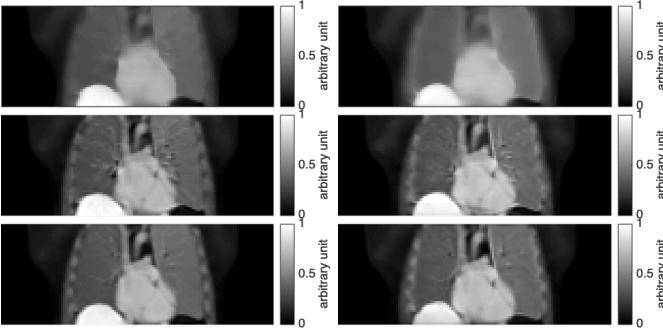
Fig. 6. The central coronal view of the activity images at 100 outer iterations (top row) and after an additional reconstruction with a smaller β (second row) for Approach I (left column) and II (right column). The target activity images for the case where well-aligned μ map and a smaller β are used (bottom row) are also provided.

The applied workflow was 1 inner iteration for the misalignment estimation and 10 inner iterations for the image reconstruction. The noiseless dataset was used.

provide insights into the relevance between the realigned and target activity images when clinically appropriate settings for the image reconstruction are used, the estimated activity images using well-aligned μ map for calculating the attenuation correction factors and anatomical prior are given as well (Fig. 6, bottom row).

B. Influence of availability of anatomical information

As observed in the central coronal view of the difference images at 100 outer iterations (Fig. 7 and 8, bottom row), both Approach I and II were able to estimate the misalignment and warp the input attenuation map accordingly, regardless of the presence of the anatomical information. However, in terms of the convergence rate, these two approaches had different responses to the use of the anatomical information. For the first approach, incorporating the anatomical information degraded the convergence rate of the misalignment estimation. The central coronal view of the difference images at OuterIter = 20 and 60 for the reconstructions without using anatomical information showed less apparent misalignment around the diaphragm region compared to those for the reconstructions considering the anatomical information (Fig. 7, top and second rows). In contrast, Approach II was able to achieve a faster convergence rate when the additional anatomical information was available (Fig. 8, top and second rows). These observations were further demonstrated by the RMSE in the mask plotted against the outer iteration numbers (Fig. 9). When the anatomical information was considered, Approach II reached a stable RMSE in the mask after around 60 outer iterations, while Approach I required more than 80 outer iterations to achieve that.

C. Influence of data noise level

Fig. 10 shows the central coronal view of the difference image between the warped and target μ maps for each reconstruction condition at OuterIter = 100 for the noisy dataset with total counts of 161 M. As observed in the figure, the
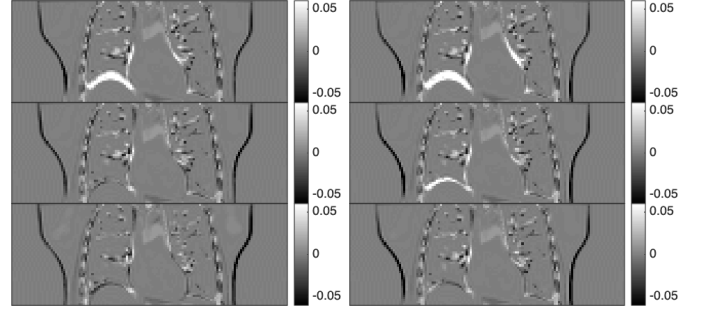
Fig. 7. The central coronal view of the difference images between the target and $W\mu'$ maps for Approach I at 20 (top row), 60 (second row) and 100 (bottom row) outer iterations. The results for the reconstructions without and with considering the anatomical information are shown in the left and right column, respectively. The noiseless dataset was used.
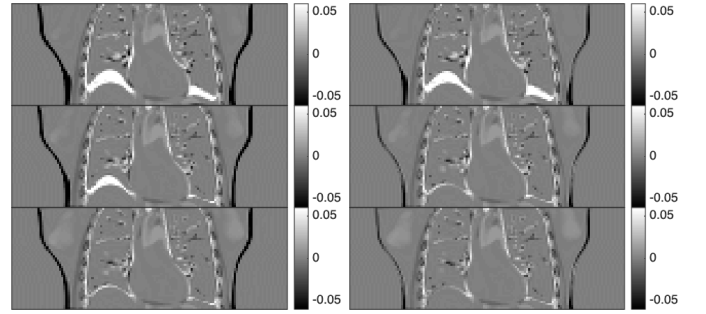
Fig. 8. The central coronal view of the difference images between the target and $W\mu'$ maps for Approach II at 20 (top row), 60 (second row) and 100 (bottom row) outer iterations. The results for the reconstructions without and with considering the anatomical information are shown in the left and right column, respectively. The noiseless dataset was used.

first approach still suffered from the misalignment issue at 100 outer iterations when InnerIter2 = 10 or 20 were applied (Fig. 10, left column, top and second images). Consistent with the results for the noiseless dataset (Fig. 2), the performance of the misalignment estimation of the second approach was less sensitive to the applied number of InnerIter2. We obtained satisfactory results at 100 outer iterations for both approaches using 1 iteration of misalignment estimation, followed by 40 iterations of image reconstruction.

The central coronal view of the difference image between the warped and target μ maps for each noisy dataset at OuterIter = 100 is shown in Fig. 11. The common workflow
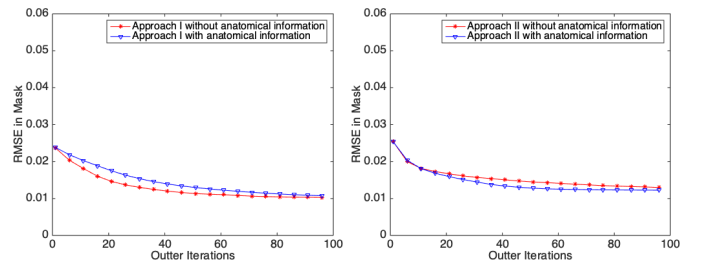
Fig. 9. The RMSE in the mask plotted against the outer iteration numbers for Approach I (left) and II (right) without and with considering the anatomical information. The applied workflow was 1 inner iteration for the misalignment estimation and 10 inner iterations for the image reconstruction. The noiseless dataset was used.
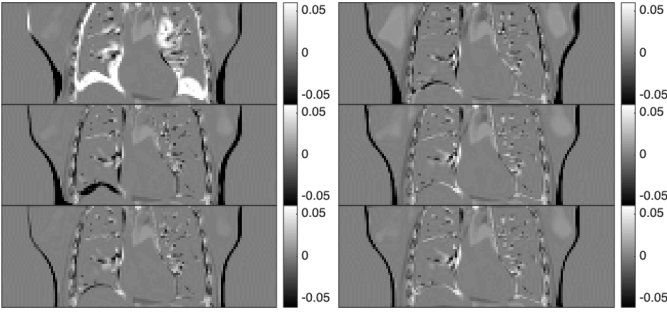
Fig. 10. The central coronal view of the difference images between the target and $\boldsymbol{W}\boldsymbol{\mu}'$ maps for Approach I (left column) and Approach II (right column) at 100 outer iterations. The noisy dataset with total counts of 161 M was used. The applied workflows were 1 inner iteration for the misalignment estimation and 10 (top row), 20 (middle row), 40 (bottom row) inner iterations for the image reconstruction.
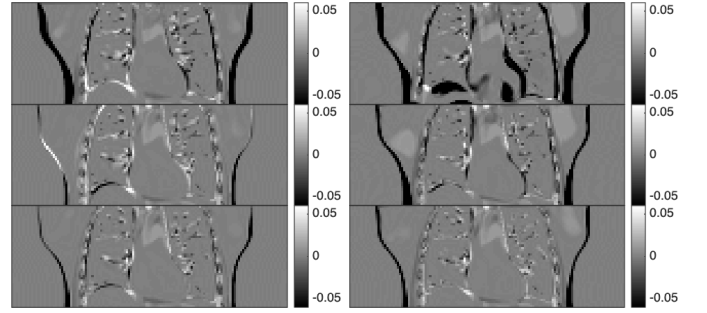

Fig. 11. The central coronal view of the difference images between the target and $\boldsymbol{W}\boldsymbol{\mu}'$ maps for Approach I (left column) and (right column) Approach II at 100 outer iterations. The noisy datasets with total counts of 16 M (top row), 54 M (second row) and 161 M (bottom row) were used. The applied workflows were 1 inner iteration for the misalignment estimation and 40 inner iterations for the image reconstruction.
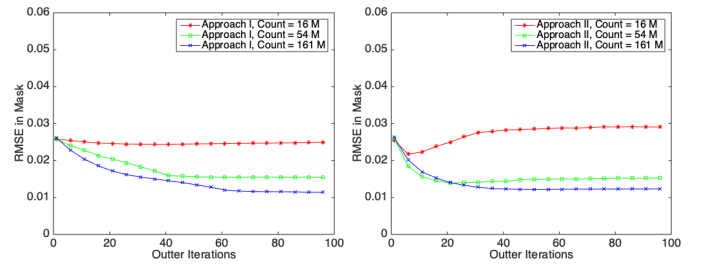

Fig. 12. The RMSE in the mask on noisy data plotted against the outer iteration numbers for Approach I (left) and II (right). The applied workflows were 1 inner iteration for the misalignment estimation and 40 inner iterations for the image reconstruction.

found using the dataset with total counts of 161 M was applied to all noisy datasets. When the data with 16 M total count was considered, significant misalignment was still observed for both approaches at 100 outer iterations (Fig. 11, top row). The second approach seemed to be more sensitive to high noise as the misalignment around the diaphragm and heart regions in the difference image was more apparent for the second approach compared to that for the first approach. The performance of the misalignment estimation was improved as the data noise level decreased for both approaches (Fig. 11, second and bottom row). The RMSE in the mask plotted against the outer iteration numbers support our observations from the difference images (Fig. 12). For data with high noise, Approach I and Approach II stuck at a high RMSE in the mask at 100 outer iterations. When the data noise level was lower, both of them were able to achieve a lower RMSE in the mask. Moreover, Approach II converged faster than Approach I for data with low noise. The corresponding reconstructed functional images are provided in Fig. 13. The misalignment around the diaphragm and heart led to distortions in the activity image for the high noise dataset when Approach II was applied (Fig. 13, top row). Similar to the results for the noise-less dataset, the activity images were over-regularized due to the use of settings optimized for the misalignment estimation. An additional reconstruction (using the same framework as the image reconstruction subroutine of Approach I) with a smaller $\beta = 2 \times 10^{-2}$ was also performed. The final images are given in Fig. 14.

V. DISCUSSION

The potential misalignment between functional and anatomical images is the main concern for incorporating an anatomical prior into image reconstruction. Expanding on the algorithm proposed in [31], two approaches that perform alternating misalignment estimation and penalized image reconstruction using anatomical priors are proposed. In this study, we focused on the performance of the misalignment estimation of each approach for the thorax with one gate of non-TOF PET data in which no motion was assumed. The $\boldsymbol{\mu}$ map was used for attenuation correction and anatomical prior calculation. Therefore, misalignment affects both factors, with subsequent

influence on the optimization. Both approaches have shown the ability to estimate the misalignment and warp the anatomical image accordingly, but with a different convergence rate, depending on the applied workflow, data noise level and if the anatomical information is included as a prior.

A. Performance comparison in terms of the convergence rate of the misalignment estimation

To study how the change of the workflow and the presence of anatomical information influence the performance of the
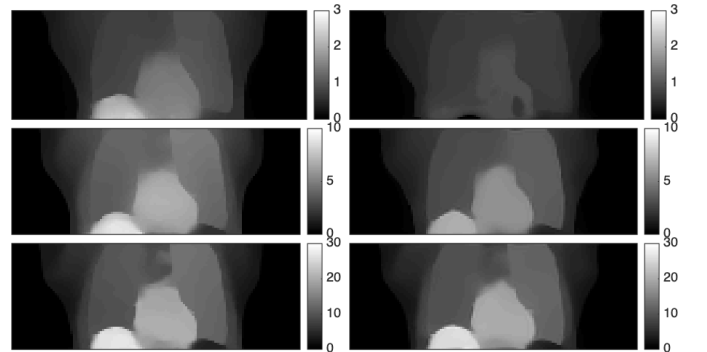

Fig. 13. The central coronal view of the activity images for Approach I (left) and Approach II (right) at 100 outer iterations. The noisy datasets with total counts of 16 M (top row), 54 M (second row) and 161 M (bottom row) were used. The applied workflows were 1 inner iteration for the misalignment estimation and 40 inner iterations for the image reconstruction.
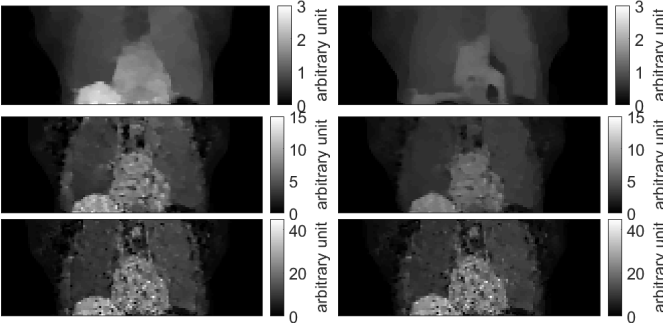
Fig. 14. The central coronal view of the activity images for Approach I (left) and Approach II (right) after the additional image reconstruction. The noisy datasets with total counts of 16 M (top row), 54 M (second row) and 161 M (bottom row) were used. The applied workflows were 1 inner iteration for the misalignment estimation and 40 inner iterations for the image reconstruction.

proposed approaches, initial evaluations on noiseless datasets were conducted. As the maximum outer iteration number that controls the repetition of the alternating process was set to 100, the workflow of both approaches was determined by the number of iterations for the misalignment estimation (InnerIter1) and image reconstruction (InnerIter2) subroutines. Based on the results shown in Section IV-A, when a sufficient number of image reconstruction subroutine was applied, Approach II showed the ability to achieve a good alignment in less outer iterations compared to Approach I. Its performance was also less sensitive to the change of workflow. The use of a larger InnerIter1 or InnerIter2 improved the convergence rate of the misalignment estimation for Approach II but slowed it down for Approach I. For the given workflow with InnerIter1 = 1 and InnerIter2 = 10, Approach II required a larger outer iteration to achieve a good result (Fig. 8) in the case where the anatomical information was absent. However, the convergence rate of misalignment estimation was improved for Approach I when the anatomical information was not available (Fig. 7). For noisy data, we found that an increased number of InnerIter2 was preferable. When the data noise level was reasonably low, both approaches showed the ability to achieve satisfactory misalignment estimation at 100 outer iterations with the given workflow. However, for data with high noise, the proposed approaches tended to converge to a high RMSE in the mask (Fig. 12), indicating that they might suffer from the local minimum problem. Since it is hard to attribute the cause to the high noise level or the applied workflow, further investigation on the need of optimizing the workflow for different data noise level is necessary.

B. Performance comparison in terms of the RMSE value in the mask and computation time

Throughout the study, the RMSE in the mask was used to evaluate how the convergence of each approach depended on different factors. When using this metric to compare both methods, in most cases similar RMSE in the mask could be obtained by both approaches, with a slightly lower RMSE in the mask observed for Approach I in Fig. 9. However, the RMSE in the mask might not be sufficient to evaluate the performance superiority of one approach over the other, and an

application-dependent metric would be advisable. In addition, it is likely that a more thorough optimization of the workflow and parameters would allow improving both methods.

In this study, a workflow that gave good results for both approaches was utilized to be able to compare their performance at a given outer iteration. Since one iteration of motion estimation is computationally more expensive than one iteration of image reconstruction, a measure that reflects the computational demand of each approach is therefore required. Although the computation time might be able to serve the purpose, the current implementation of both approaches was not optimized to have a fair comparison. They were implemented using a combination of C, C++, Fortran and MATLAB without applying acceleration techniques, such as multithreading or graphics processing unit (GPU) computing. Under the current implementation structure, Approach I and Approach II took respectively 56 and 260 seconds to complete one outer iteration of the suggested workflow on a desktop with an Intel i7-8700 CPU. Given that similar number of forward and back projections was involved in both approaches for the applied workflow, the difference between them in computation time could be caused by the non-optimized implementation. Future work includes further investigation and optimization of the performance of the approaches in terms of computation time.

C. Influence of the strength of each penalty function

For both approaches, a spatially-variant penalization scheme proposed in [39] was applied to the anatomical penalty function to achieve uniform local contrast. Based on the bias-variance analysis presented in [39], the use of a stronger β can lead to a lower image variance but higher bias, regardless of the consideration of the spatially-variant penalization. In this study, the strength of γ and δ did not have a direct impact to the image quality as they were used to control the smoothing penalty on the motion field and barrier function on negative image values. However, since the performance of the misalignment estimation of both approaches varies with the strength of each penalty function, the image quality can be influenced indirectly by the residual misalignment caused by insufficient inner and outer iterations for a given set of hyper-parameters. We have chosen each penalty strength based on the performance of the misalignment estimation of the proposed approaches with respect to a limited set of candidate values. As can be seen in Fig. 6 and Fig. 13, optimizing for accurate alignment led to over-regularized functional images. This could be overcome by running a final image reconstruction incorporating warping with settings, especially the strength of β, optimized for the functional image estimation. Example reconstructions can be found in Fig. 6 and Fig. 14.

D. Potential strategies to further improve the performance

Since the deformation of the attenuation map is likely to lead to a different scatter distribution, the estimated background events should be updated accordingly during the optimization. However, the effect was assumed to be small and ignored for simplicity in the current study. To achieve accurate quantification, performing active scatter correction

based on the update of the estimated activity distribution might be necessary. In practice, this could be done by re-estimating the scatter after a number of iterations of the current algorithm.

In the current study, we have adapted the strategy often used in CT or MR derived attenuation correction that down-samples the anatomical image to match the resolution of PET and surrenders the structural details carried by the high-resolution anatomical image. However, it could be beneficial to reconstruct the PET image at the same voxel size as the anatomical image instead [43], [44]. The benefit of using TOF data for misalignment estimation was studied in [34] where a similar algorithm was applied to obtain reconstructed activity images with aligned attenuation correction. As Approach I and II are extensions of that algorithm, practical convergence of misalignment estimation in less outer iterations can be expected when TOF data are available.

E. Future demonstrations in more realistic scenarios

The algorithms proposed in this study have been demon-strated with simulation. A thorough validation with more realistic data is required to demonstrate usefulness in future applications in clinic. In addition, as the algorithm perfor-mance and the quantitative accuracy can be affected by other parameters that determine the edge-preserving property of the penalty (*e.g.*, ϵ and η in PLS), future work should also include parameter optimization with respect to different applications. It remains to be investigated if the introduction of the anatomical prior would introduce errors in the alignment estimation for cases where edges in functional and anatomical images are not consistent. To investigate the effectiveness of applying the proposed approaches in improving quantitative accuracy, evaluations using phantoms with inserted features or patient datasets with pseudo-lesions should also be included. In our current work, we assumed that the attenuation correction and anatomical image used for PLS were the same. However, the algorithms can be generalized to other user cases where the anatomical image is independent of the attenuation image. One particular case of this might be reconstructions of PET/CT data using an anatomical prior derived from MR images.

F. Potential limitations

The methods described here are applicable to any differen-tiable deformation models that can be parameterized using a linear sum of basis-functions. As in [31], we have chosen to use the uniform cubic B-splines to model the image deforma-tion, taking advantage of no *a priori* knowledge regarding the misalignment is required. Although the model should be able to adapt to various motion effects, including respiration and organ contraction and expansion, it might fail to accurately model sliding motion such as occurs between the lower lung and ribcage. Moreover, as the current model can only address smooth motion, it is not able to handle misalignment caused by abrupt motion, for example, objects moving in and out the field of view or sudden repositioning of the arms. To improve the accuracy of the misalignment estimation and make max-imum use of the anatomical information, the investigation of other valid (*i.e.*, differentiable and parametrizable) deformation models specific to different applications and motion effects should be included in future work.

VI. CONCLUSIONS

Two approaches for solving the potential misalignment between the functional and anatomical images in penalized image reconstruction using anatomical priors have been pro-posed in this study. The main difference between them is that the first approach deforms the anatomical image to align it with the functional one, while the second approach deforms both images to align them with the measured data. Both approaches were implemented using alternation between misalignment estimation and image reconstruction. The results demonstrated that both methods are able to estimate the misalignment and deform the anatomical image when the data noise level is reasonably low and a proper workflow for the alternating optimization is applied. Moreover, the second approach shows the ability to converge to the correct alignment faster than the first approach and is less sensitive to variations in the workflow. The use of anatomical information improves the convergence rate of misalignment estimation for the second approach, although slowed it down for the first approach. These encouraging results indicate that it is possible to align functional and anatomical information, overcoming a serious limitation in practical use of anatomical priors.

REFERENCES

[1] K. Vunckx, A. Atre, K. Baete, A. Reilhac, C. M. Deroose, K. V. Laere, and J. Nuyts, "Evaluation of three MRI-based anatomical priors for quantitative PET brain imaging," *IEEE Trans. Med. Imag.*, vol. 31, no. 3, pp. 599–612, 2012.

[2] Y.-J. Tsai, H.-M. Huang, C.-Y. Chou, W. Wang, and I.-T. Hsiao, "Effective anatomical priors for emission tomographic reconstruction," *J Med. Biol. Eng.*, vol. 35, no. 1, pp. 52–61, 2015.

[3] K. Sakaguchi, H. Shinohara, T. Hashimoto, and T. Yokoi, "An iterative reconstruction using median root prior and anatomical prior from the segmented μ-Map for count-limited transmission data in PET imaging," *Ann. Nucl. Med.*, vol. 22, no. 4, pp. 269–79, 2008.

[4] G. Gindi, M. Lee, A. Rangarajan, and I. G. Zubal, "Bayesian recon-struction of functional images using anatomical information as priors," *IEEE Trans. Med. Imag.*, vol. 12, no. 4, pp. 670–80, 1993.

[5] B. Lipinski, H. Herzog, E. R. Kops, W. Oberschelp, and H. W. Müller-Gärtner, "Expectation maximization reconstruction of positron emission tomography images using anatomical magnetic resonance information," *IEEE Trans. Med. Imag.*, vol. 16, no. 2, pp. 129–36, 1997.

[6] C. Comtat, P. E. Kinahan, J. A. Fessler, T. Beyer, D. W. Townsend, M. Defrise, and C. Michel, "Clinically feasible reconstruction of 3D whole-body PET/CT data using blurred anatomical labels," *Phys. Med. Biol.*, vol. 47, no. 1, pp. 1–20, 2002.

[7] Y. Mameuda and H. Kudo, "New anatomical-prior-based image recon-struction method for PET/SPECT," *Proc. IEEE Nucl. Sci. Symp. Med. Imag. Conf. Rec.*, vol. 6, pp. 4142–8, 2007.

[8] J. Nuyts, K. Baete, D. Bequé, and P. Dupont, "Comparison between MAP and postprocessed ML for image reconstruction in emission tomography when anatomical knowledge is available," *IEEE Trans. Med. Imag.*, vol. 24, no. 5, pp. 667–75, 2005.

[9] P. P. Bruyant, H. C. Gifford, G. Gindi, and M. A. King, "Numerical observer study of MAP-OSEM regularization methods with anatomical priors for lesion detection in 67-Ga images," *IEEE Trans. Nucl. Sci.*, vol. 51, no. 1, pp. 193–7, 2004.

[10] J. Browne and A. B. d. Pierro, "A row-action alternative to the EM algorithm for maximizing likelihood in emission tomography," *IEEE Trans. Med. Imag.*, vol. 15, no. 5, pp. 687–99, 1996.

[11] S. Kulkarni, P. Khurd, I. Hsiao, L. Zhou, and G. Gindi, "A channelized Hotelling observer study of lesion detection in SPECT MAP recon-struction using anatomical priors," *Phys. Med. Biol.*, vol. 52, no. 12, pp. 3601–17, 2007.

[12] M. J. Ehrhardt and S. R. Arridge, "Vector-valued image processing by parallel level sets," *IEEE Trans. Image Processing*, vol. 23, no. 1, pp. 9–18, 2014.

[13] D. Deidda, N. A. Karakatsanis, P. M. Robson, N. Efthimiou, Z. A. Fayad, R. G. Aykroyd, and C. Tsoumpas, "Effect of PET-MR inconsistency in the kernel image reconstruction method," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 3, no. 4, pp. 400–9, 2019.

[14] J. Liao and J. Qi, "PET image reconstruction with anatomical prior using multiphase level set method," *Proc. IEEE Nucl. Sci. Symp. Med. Imag. Conf. Rec.*, 2007.

[15] O. Camara, G. Delso, O. Colliot, A. Moreno-Ingelmo, and I. Bloch, "Explicit incorporation of prior anatomical information into a nonrigid registration of thoracic and abdominal CT and 18-FDG whole-body emission PET images," *IEEE Trans. Med. Imag.*, vol. 26, no. 2, pp. 164–78, 2007.

[16] Q. Xu, K. Yuan, and D. Ye, "Respiratory motion blur identification and reduction in ungated thoracic PET imaging," *Phys. Med. Biol.*, vol. 56, no. 14, pp. 4481–98, 2011.

[17] M. J. Nyflot, T. C. Lee, A. M. Alessio, S. D. Wollenweber, C. W. Stearns, S. R. Bowen, and P. E. Kinahan, "Impact of CT attenuation correction method on quantitative respiratory-correlated (4D) PET/CT imaging," *Med. Phys.*, vol. 42, no. 1, pp. 110–20, 2015.

[18] A. Pépin, J. Daouk, P. Bailly, S. Hapdey, and M. Meyer, "Management of respiratory motion in PET/computed tomography: the state of the art," *Nucl. Med. Commun.*, vol. 35, no. 2, pp. 113–22, 2014.

[19] S. A. Nehmeh and Y. E. Erdi, "Respiratory motion in positron emission tomography/computed tomography: a review," *Semin. Nucl. Med.*, vol. 38, no. 3, pp. 167–76, 2008.

[20] S. A. Nehmeh, Y. E. Erdi, T. Pan, A. Pevsner, K. E. Rosenzweig, E. Yorke, G. S. Mageras, H. Schoder, P. Vernon, O. Squire, H. Mostafavi, S. M. Larson, and J. L. Humm, "Four-dimensional (4D) PET/CT imaging of the thorax," *Med. Phys.*, vol. 31, no. 12, pp. 3179–86, 2004.

[21] R. Manjeshwar, X. Tao, E. Asma, and K. Thielemans, "Motion compensated image reconstruction of respiratory gated PET/CT," *IEEE Int. Symp. Biomed. Imag.*, pp. 674–7, 2006.

[22] F. Qiao, T. Pan, J. W Clark Jr, and O. R. Mawlawi, "A motion-incorporated reconstruction method for gated PET studies," *Phys. Med. Biol.*, vol. 51, no. 15, pp. 3769–83, 2006.

[23] B. K. Teo, B. Saboury, R. Munbodh, J. Scheuermann, D. A. Torigian, H. Zaidi, and A. Alavi, "The effect of breathing irregularities on quantitative accuracy of respiratory gated PETCT," *Med. Phys.*, vol. 39, no. 12, pp. 7390–7, 2012.

[24] J. R. McClelland, D. J. Hawkes, T. Schaeffter, and A. P. King, "Respiratory motion models: a review," *Med. Image Anal.*, vol. 17, no. 1, pp. 19–42, 2013.

[25] R. Manber, K. Thielemans, B. F. Hutton, S. Wan, J. McClelland, A. Barnes, S. Arridge, S. Ourselin, and D. Atkinson, "Joint PET-MR respiratory motion models for clinical PET motion correction," *Phys. Med. Biol.*, vol. 61, no. 17, pp. 6515–30, 2016.

[26] F. Kalantari and J. Wang, "Attenuation correction in 4D-PET using a single-phase attenuation map and rigidity-adaptive deformable registration," *Med. Phys.*, vol. 44, no. 2, pp. 522–32, 2017.

[27] S. J. McQuaid, T. Lambrou, and B. F. Hutton, "A novel method for incorporating respiratory-matched attenuation correction in the motion correction of cardiac PET-CT studies," *Phys. Med. Biol.*, vol. 56, no. 10, pp. 2903–15, 2011.

[28] H. Fayad, J. F. Clement, T. Pan, C. Roux, C. C. L. Rest, O. Pradier, and D. Visvikis, "Towards a generic respiratory motion model for 4D CT imaging of the thorax," *Proc. IEEE Nucl. Sci. Symp. Med. Imag. Conf. Rec.*, pp. 3975–9, 2009.

[29] J. Nuyts, P. Dupont, S. Stroobants, R. Benninck, L. Mortelmans, and P. Suetens, "Simultaneous maximum a posteriori reconstruction of attenuation and activity distributions from emission sinograms," *IEEE Trans. Med. Imag.*, vol. 18, no. 5, pp. 393–403, 1999.

[30] A. Rezaei, M. Defrise, G. Bal, C. Michel, M. Conti, C. Watson, and J. Nuyts, "Simultaneous reconstruction of activity and attenuation in time-of-flight PET," *IEEE Trans. Med. Imag.*, vol. 31, no. 12, pp. 2224–33, 2012.

[31] A. Bousse, O. Bertolli, D. Atkinson, S. Arridge, S. Ourselin, B. F. Hutton, and K. Thielemans, "Maximum-likelihood joint image reconstruction/motion estimation in attenuation-corrected respiratory gated PET/CT using a single attenuation map," *IEEE Trans. Med. Imag.*, vol. 35, no. 1, pp. 217–28, 2016.

[32] A. Rezaei, C. Michel, M. E. Casey, and J. Nuyts, "Simultaneous reconstruction of the activity image and registration of the CT image in TOF-PET," *Phys. Med. Biol.*, vol. 61, no. 4, pp. 1852–74, 2016.

[33] A. Bousse, R. Manber, B. F. Holman, D. Atkinson, S. Arridge, S. Ourselin, B. F. Hutton, and K. Thielemans, "Evaluation of a direct motion estimation/correction method in respiratory-gated PET/MRI with motion-adjusted attenuation," *Med. Phys.*, vol. 44, no. 6, pp. 2379–2390, 2017.

[34] A. Bousse, O. Bertolli, D. Atkinson, S. Arridge, S. Ourselin, B. F. Hutton, and K. Thielemans, "Maximum-likelihood joint image reconstruction and motion estimation with misaligned attenuation in TOF-PET/CT," *Phys. Med. Biol.*, vol. 61, no. 3, pp. L11–19, 2016.

[35] Y.-J. Tsai, A. Bousse, S. Ahn, C. W. Stearns, S. Arridge, and K. Thielemans, "Algorithms for solving misalignment issues in penalized PET/CT reconstruction using anatomical priors," *Proc. IEEE Nucl. Sci. Symp. Med. Imag. Conf. Rec.*, 2018.

[36] M. J. Ehrhardt, K. Thielemans, L. Pizarro, D. Atkinson, S. Ourselin, B. F. Hutton, and S. R. Arridge, "Joint reconstruction of PET-MRI by exploiting structural similarity," *Inverse Problems*, vol. 31, no. 1, pp. 1–23, 2015.

[37] M. J. Ehrhardt, P. Markiewicz, M. Liljeroth, A. Barnes, V. Kolehmainen, J. S. Duncan, L. Pizarro, D. Atkinson, B. F. Hutton, S. Ourselin, K. Thielemans, and S. R. Arridge, "PET reconstruction with an anatomical MRI prior using parallel level sets," *IEEE Trans. Med. Imag.*, vol. 35, no. 9, pp. 2189–99, 2016.

[38] T. E. Nichols, J. Qi, E. Asma, and R. M. Leahy, "Spatiotemporal reconstruction of list-mode PET data," *IEEE Trans. Med. Imag.*, vol. 21, no. 4, pp. 396–404, 2002.

[39] Y.-J. Tsai, G. Schramm, S. Ahn, A. Bousse, S. Arridge, J. Nuyts, B. F. Hutton, C. W. Stearns, and K. Thielemans, "Benefits of using a spatially-variant penalty strength with anatomical priors in PET reconstruction," *Early Access, IEEE Trans. Med. Imag.*, 2019.

[40] R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, "A limited memory algorithm for bound constrained optimization," *SIAM J Sci. Comput.*, vol. 16, no. 5, pp. 1190–208, 1995.

[41] Y.-J. Tsai, A. Bousse, M. J. Ehrhardt, C. W. Stearns, S. Ahn, B. F. Hutton, S. Arridge, and T. K., "Fast quasi-Newton algorithms for penalized reconstruction in emission tomography and further improvements via preconditioning," *IEEE Trans. Med. Imag.*, vol. 37, no. 4, pp. 1000–10, 2018.

[42] W. P. Segars, G. Sturgeon, S. Mendonca, J. Grimes, and B. M. W. Tsui, "4D XCAT phantom for multimodality imaging research," *Med. Phys.*, vol. 37, no. 9, pp. 4902–15, 2010.

[43] M. A. Belzunce, A. Mehranian, J. Bland, and A. J. Reader, "ML and MAP PET reconstruction with MR-voxel sizes for simultaneous PET-MR," *Proc. IEEE Nucl. Sci. Symp. Med. Imag. Conf. Rec.*, 2017.

[44] M. A. Belzunce, A. Mehranian, and A. J. Reader, "Enhancement of partial volume correction in MR-guided PET image reconstruction by using MRI voxel sizes," *IEEE Trans. Radiat. Plasma Med. Sci.*, vol. 3, no. 3, pp. 315–26, 2019.