

# Performance Improvement and Validation of a New MAP Reconstruction Algorithm

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

**Abstract**—We previously proposed a fast maximum a posteriori (MAP) algorithm, LBFSG-B-PC, combining LBFSG-B with diagonal preconditioning. Previous results have shown in simulations that it converges using around 40 projections independent of many factors. The aim of this study is to improve the algorithm further by using a better initial image and a modified preconditioner that is less sensitive to noise and data scale. By initializing the algorithm with the best initial image (one full iteration of OSEM with 35 subsets), ROI values can converge almost twice as fast for the same computation time. Moreover, the new preconditioner makes the performance more consistent between high and low count data sets. In addition, we have found a means to choose the stopping criteria to reach a desired quantitative accuracy level in the reconstructed image. Based on the results with patient data, the optimized LBFSG-B-PC shows promise for clinical imaging.

## I. INTRODUCTION

We have previously proposed a MAP algorithm LBFSG-B-PC which combines preconditioning with the use of a quasi-Newton optimization method based on the constrained LBFSG-B approach. Previous results showed that the algorithm can converge several times faster than relaxed SPS and LBFSG-B. Moreover, its performance is independent of many factors, such as penalty types, penalty strengths, noise levels and object geometries. The aim of this study is to make further improvements and practical demonstrations to increase the algorithm’s robustness.

Manuscript received November 28, 2016. This project is supported by GE Healthcare, and by the National Institute for Health Research, University College London Hospitals Biomedical Research Centre.

Yu-Jung Tsai is with the Division of Medicine, University College London, London, UK (e-mail: yu-jung.tsai.14@ucl.ac.uk).

Alexandre Bousse is with the Institute of Nuclear Medicine, University College London Hospital, London, UK (e-mail: a.bousse@ucl.ac.uk).

Charles W. Stearns is with the MICT Engineering, GE Healthcare, Waukesha, WI 53188 USA (e-mail: Charles.Stearns@med.ge.com).

Sangtae Ahn is with the Image Reconstruction Lab, GE Global Research, Niskayuna, NY 12309 USA (e-mail: ahns@ge.com).

Brian F. Hutton is with the Institute of Nuclear Medicine, University College London Hospital, London, UK. He is also with the Centre for Medical Radiation Physics, University of Wollongong, Wollongong, Australia (e-mail: b.hutton@ucl.ac.uk).

Simon Arridge is with the Department of Computer Science, University College London, London, UK (e-mail: S.Arridge@cs.ucl.ac.uk).

Kris Thielemans is with the Institute of Nuclear Medicine, University College London Hospital, London, UK (e-mail: k.thielemans@ucl.ac.uk).

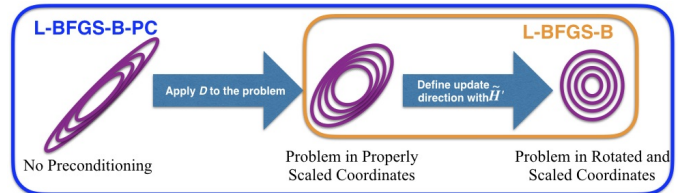


Fig. 1: Illustration of the transformation and optimization process of LBFSG-B-PC using contour-plots of the objective function.

## II. METHOD

### A. Objective function

MAP reconstruction is performed by maximizing a function  $\Phi$  consisting of 2 parts, the un-penalized likelihood  $L$  and the penalty function  $R$  with a parameter  $\beta$  which controls the strength of the penalty.

### B. LBFSG-B-PC

We proposed to use LBFSG-B in a transformed coordinate system to circumvent its potential slow convergence and sensitivity to global scale factors [1]. The transformation was achieved as follows, given  $\mathbf{x}$  the estimated image and  $\mathbf{g}$  the measured data:

$$\mathbf{x}' = \mathbf{D}^{-1}\mathbf{x}, \quad (1)$$

$$\mathbf{D} = \sqrt{\text{diag}\{\mathbf{P}^t \text{diag}\{\frac{1}{g+1}\}\mathbf{P} \cdot \mathbf{1} + \beta \nabla^2 R(\mathbf{x})\}}$$

where  $\mathbf{D}$  represents the preconditioner proposed in [1] and  $\mathbf{P}$  is the system matrix. The LBFSG-B optimization is then performed by maximizing the function  $\Phi$  at each iteration along a search line  $\mathbf{p}$  with the transformed image:

$$\mathbf{x}'^{\text{new}} = \mathbf{x}'^{\text{old}} + \delta^* \mathbf{p}, \quad \delta^* = \arg \max_{\delta \geq 0} \Phi'(\mathbf{x}'^{\text{old}} + \delta \mathbf{p}), \quad (2)$$

where  $\Phi'(\mathbf{x}') = \Phi(\mathbf{x})$  and  $\mathbf{p} = \tilde{H}'^{-1} \nabla \Phi'(\mathbf{x}'^{\text{old}})$  with  $\tilde{H}'$  the approximation of the Hessian of  $\Phi$  at the current iteration. The overall process is explained in Fig. 1.

### C. New preconditioner

Although superior performance of the proposed algorithm has been shown in our previous study [1], we notice that the “+1” term in the denominator of  $\mathbf{D}$  in (1) can be problematic if the data scale is very small. For adapting to

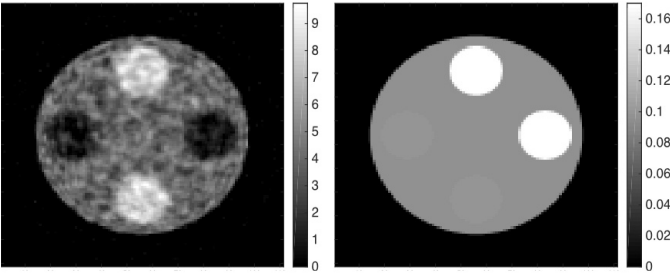


Fig. 2: Reconstructed image from LFBGS-B-PC2 at the 41<sup>st</sup> iteration and the corresponding attenuation map for the cylindrical phantom.

noisy data, we propose a new preconditioner precomputed by the initial image ( $\mathbf{x}_{ini}$ ):

$$\tilde{\mathbf{D}} = \sqrt{\text{diag}\{\mathbf{P}^t \text{diag}\{\frac{\mathbf{g}}{(\mathbf{P}\mathbf{x}_{ini} + \mathbf{r})^2}\}\mathbf{P} \cdot \mathbf{1} + \beta \nabla^2 R(\mathbf{x})\}} \quad (3)$$

where  $r$  is the background events vector. This preconditioner is no longer data scale dependent, however, the performance could be affected by  $\mathbf{x}_{ini}$ . The preconditioners (1) and (3) are referred to as PC1 and PC2, respectively.

#### D. Data

1) *Digital cylindrical phantom*: The projection data were generated to simulate the GE Discovery STE in 2D. A cylindrical phantom containing 4 ROIs, 2 hot spots and 2 cold spots was used. The radius was 26.367 mm for each of them. The activity ratios of the cold and hot regions to the background were 0.25 and 2, respectively. For each group of activity levels, different attenuation materials simulating the effects of bone and soft tissue were applied to each spot. The phantom and the attenuation map can be found in Fig. 2. Two data sets with total counts of 594 K and 29.3 K were generated.

2) *Patient data*: Three FDG PET/CT data sets were acquired on a GE Discovery STE. The PET data were collected in 3D mode with total counts of 255M, 355M and 181M, all in the thorax region.

#### E. Algorithm optimization

Choosing an initial image can improve convergence rate by starting closer to the solution but also by improving the preconditioner  $\tilde{\mathbf{D}}$ . In our previous study we used 1 iteration of MLEM [2] for initialisation. Here we will investigate the use of OSEM [3].

To simplify the problem of finding the best initial image, a two-part study was conducted. The first part tried to speed up the convergence rate by increasing the number of ordered-subsets. 8 different subsets (1, 2, 5, 7, 10, 14, 35 and 70) were employed for one full iteration to keep the same computational burden as one iteration of MLEM. We then fixed the subsets to the limit found in the first part and increased the number of full iterations from one to two to assess if the performance can be improved even

further in the second part of the study. The reconstruction was then performed by LFBGS-B-PC2 initialized with various images described above. After finding the best initial image, we evaluated how the preconditioners change performance with noisy data. Finally, an example reconstruction with patient data is shown.

We used the quadratic prior with penalty factor 0.1 for the simulated data and 4 for the patient data. Since we have proved that the performance of LFBGS-B-PC is independent of the penalty type and strength in [1], we did not include other penalty functions or study the performance differences with various penalty strengths. We used the acquisition model of STIR [4] for the simulated data, and extracted raw and calibration clinical data from the Discovery STE raw data files. The implementation was based on the Fortran LFBGS-B proposed in [5].

#### F. Analysis

Since the true image is known for simulated data, the convergence rate of each algorithm was evaluated in terms of normalized total regional recovery ratio ( $\text{RR}_{\text{total}}$ ):

$$\text{RR}_{\text{total}} = \sqrt{\frac{1}{n} \sum_{i=1}^{i=n} \left( \frac{\text{RBR}_i(\mathbf{f})}{\text{RBR}_i(\mathbf{f}^{\text{true}})} \right)^2} \quad (4)$$

where  $\text{RBR}_i$  is the  $i_{\text{th}}$ -ROI to background ratio,  $n$  is the number of ROIs and  $\mathbf{f}^{\text{true}}$  is the true image. The ROIs were drawn in the center of the spots with size of 26.37 mm<sup>2</sup>. For the patient data, we evaluated the performance by calculating the recovery ratio of a selected hot spot and a background region with respect to the converged values ( $\text{RR}_{\text{conv}}$ ). Here we considered the algorithm was converged when no intended update that satisfies Wolfe conditions could be found. The criteria were defined as follows:

$$\Phi^{(k+1)} \leq \Phi^{(k)} - \lambda_1 \alpha_k \nabla \Phi^{(k)T} \mathbf{p}_k \quad (5)$$

$$|\nabla \Phi^{(k+1)T} \mathbf{p}_k| \leq \lambda_2 |\nabla \Phi^{(k)T} \mathbf{p}_k|$$

where  $\alpha_k$  is the tested step length during the line search,  $\mathbf{p}_k$  is the search direction at  $k_{\text{th}}$  iteration and  $\{\lambda_1, \lambda_2\}$  is a set of fixed parameters controlling the strength of these conditions. In this study, the values were set to  $10^{-4}$  and 0.9, respectively.

#### G. Stopping criteria

A robust stopping criterion requires a stable relationship between the relative amount of image change per iteration (something we can measure) and the closeness of the image to its convergence value (what we want to know, but cannot), in order to stop iterations. Using the simulated data with high counts, we defined metrics  $F(k)$  and  $M(k)$  to measure the image change between two iterations and the distance from the convergence, respectively:

$$F(k) = \frac{(\Phi^{(k)} - \Phi^{(k+1)})}{\max(|\Phi^{(k)}|, |\Phi^{(k+1)}|)} \quad (6)$$

$$M(k) = \frac{\frac{1}{N} \sum_{j=1}^{j=N} (x_j^{(k)} - x_j^{\text{conv}})^2}{(\bar{\mathbf{x}}^{\text{conv}})^2}$$

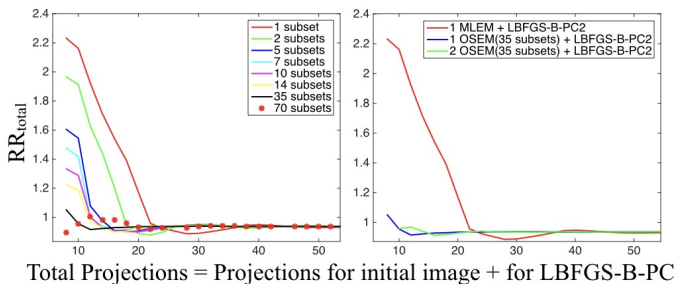


Fig. 3: The  $RR_{total}$  values plotted against projection numbers for LBFGS-B-PC2 initialized by one full iteration of OSEM with various subsets (left) and various full iterations of OSEM with 35 subsets (right). The reconstruction was performed on the simulated high counts data.

where  $N$  is the number of pixels in the image. Instead of normalizing the objective function  $\Phi$  with the number of total counts  $T$ , we took into account the data scale while calculating  $M(k)$ . Therefore, the comparison of the convergence rate in terms of  $M(k)$  would be normalized to the same scale.  $\bar{x}^{conv}$  is the mean value of all the pixels in the converged image. We used SPS at 1000<sub>th</sub> iteration as the reference since its convergence has been well-proved.

### III. RESULTS

As shown in Fig. 3 (left), the convergence rate of the  $RR_{total}$  is improved as the number of ordered subsets is increased. The convergence trend for 70 subsets (there were only 4 projections in one subset) is quite different from others. Therefore, we chose 35 as the highest number of subsets and increased the number of full iterations. Based on the results in Fig. 3 (right), no significant improvement was observed after one full iteration. The performance comparison between the two preconditioners can be found in Fig. 4. Similar convergence rates were observed for the high counts data. However, when the total counts was reduced to a very low level, the preconditioner PC1 compromises the convergence rate. The linear relationship between the  $\log(F)$  and  $\log(M)$  implies a power law relationship between  $M$  and  $F$  (Fig. 5). As a faster algorithm will make larger improvements in the objective function per iteration, this plot confirms that the preconditioners speed-up LBFGS-B. Similar results for XCAT phantom were obtained (not shown). A slice from one of the patient data sets and the selected ROIs on it are shown in Fig. 6 (left) as an example. The recovery ratio of the selected hot spot and the background region with respect to the converged values ( $RR_{conv}$ ) for all patients are given in Fig. 6 (right). Comparable convergence rate in both regions are observed among different patients.

### IV. CONCLUSION

The performance and robustness of the previously proposed preconditioned algorithm, LBFGS-B-PC, have been further improved. The optimized algorithm shows promise for use in a clinical context.

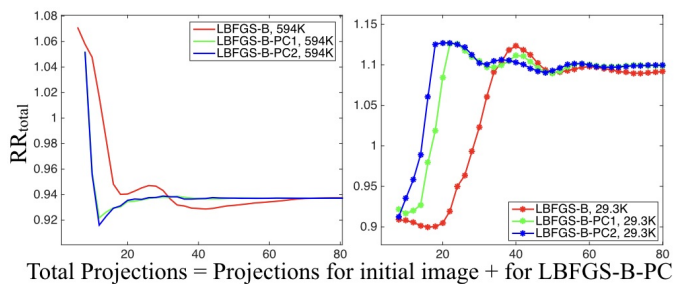


Fig. 4: The  $RR_{total}$  values plotted against projection numbers for all algorithms with data of 594 K counts (left) and 29.3 K counts (right).

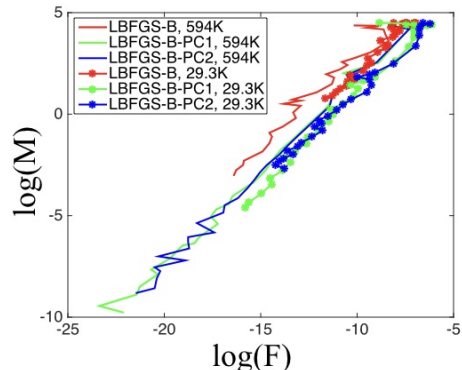


Fig. 5: The  $M(k)$  values plotted against the  $F(k)$  for all algorithms. Both x and y-axis of the plot are in log scale.

### REFERENCES

- [1] Y. J. Tsai *et al.*, *IEEE MIC Conf. Proc.*, 2015.
- [2] L. Shepp and Y. Vardi, *IEEE TMI*, vol. 1, 1982.
- [3] H. M. Hudson and R. S. Larkin, *IEEE TMI*, vol. 13, 1994.
- [4] K. Thielemans *et al.*, *Phys. Med. Biol.*, vol. 57, 2012.
- [5] C. Zhu *et al.*, *ACM Trans. Math. Softw.*, vol. 23, no. 4, 1997.

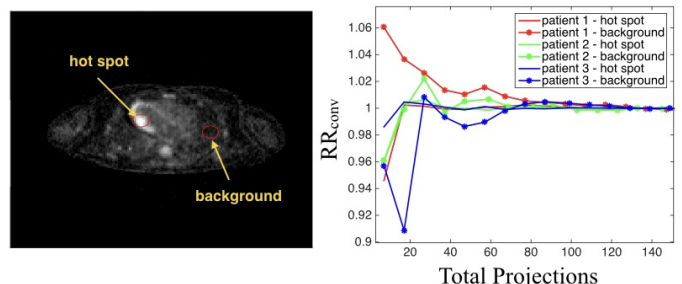


Fig. 6: A slice of the patient data reconstructed from the optimized LBFGS-B-PC2 (left) and the recovery ratio of the selected hot spot and the background region with respect to the converged values (right).