Fast Hybrid Algorithms for PET Image Reconstruction

Quanzheng Li, *Student Member, IEEE,* Sangtae Ahn, *Member, IEEE,* and Richard Leahy, *Fellow, IEEE,*

***Abstract*— We describe a hybrid approach to iterative PET image reconstruction in which we combine three algorithms: preconditioned conjugate gradient (PCG), ordered subsets separable paraboloidal surrogate (OSSPS), and emission reconstruction incremental optimization transfer (ERIOT). These algorithms exhibit quite different convergence behavior, e.g. the initial convergence of OSSPS is fast but it soon enters a limit cycle. Conversely, initial convergence of PCG is slow compared to OSSPS and ERIOT but its asymptotic behavior is the fastest. The hybrid approach estimates convergence behavior for each method by fitting an exponential to the objective function as a function of iteration number and switches between the algorithms to optimize the convergence behavior throughout the iterations. This hybrid approach is compared to each of the component algorithms in application to simulated PET data demonstrating a reduction of at least 50% in iterations required for effective convergence compared to use of PCG alone.**

***Index Terms*— PET Image Reconstruction, Incremental Optimization Transfer, PCG.**

I. Introduction

STATISTICALLY based PET image reconstruction algorithms can achieve superior image quality when compared with conventional analytic reconstruction methods. However, they require iterative optimization algorithms and computational times for fully 3D reconstruction can limit the practical use of these algorithms. Consequently, in recent years many algorithms have been proposed to reduce the complexity and number of iterations required for effective convergence. These algorithms range from general purpose gradient based optimization methods to the use of incremental gradient or subset based methods and optimization transfer techniques.

Conjugate gradient methods have been widely used for iterative reconstruction of PET images [1], [2], [5]. The diagonal EM-based preconditioner [4] and nondiagonal Fourier based preconditioners [5] can significantly improve convergence rates for conjugate gradient methods, however initial convergence tends to be slow so that several 10's of iterations are often required for effective convergence.

Ordered Subset (OS) methods use subsets of the data rather than the entire data set to update the image at each iteration. Since computation cost is dominated by forward and backprojection, the use of subsets methods reduces cost per iteration by a factor equal to the number of subsets used. The ordered subsets expectation maximization (OSEM) algorithm applies this subset approach to the EM algorithm [8] and produces significant acceleration in the initial convergence behavior relative to the original EM algorithm. However, OSEM eventually reaches a limit cycle rather than converging to a maximum of the likelihood function [8], [21]. While the EM and OSEM algorithms were originally developed and applied to maximum likelihood estimation for PET image reconstruction, these approaches are readily extended to allow maximization of a penalized likelihood function, or equivalently to compute a maximum a posteriori (MAP) image estimate, which is the problem we address here.

To avoid the limit cycle and ensure convergence, the OSEM algorithm can be modified to use a reduced step size at each iteration. Methods that employ a relaxed update scheme include block sequential regularized expectation maximization (BSREM) [6], modified BSREM [21], relaxed Ordered Subset Separable Paraboloidal Surrogate (OS-SPS) [21], and row-action maximum likelihood algorithm (RAMLA) [7]. An alternative way to achieve convergence is to gradually reduce the number of subsets during reconstruction [9].

The convergent relaxed OS algorithms require the user to specify relaxation parameters, which can have a significant effect on convergence rate; however there is no theoretical basis for selection of the optimal parameter. The closely related class of incremental EM algorithms [10], [11], [12] and incremental aggregated gradient algorithms [13] use the OS concept without requiring user specified parameters. Ahn and Fessler generalized incremental EM algorithms to an incremental optimization transfer method [16] by using the optimization transfer principles [14] and successfully applied them to transmission PET reconstruction [16]. Here we apply this incremental optimization transfer method to emission reconstruction using a quadratic surrogate function. Using the same framework, Hsiao *et al.* developed Complete-data OSEM (COSEM) and applied it to emission tomography using EM (rather than quadratic) surrogates [11], [12].

The early convergence behavior of the above algorithms often differs from their asymptotic behaviour: while incremental gradient methods show rapid initial convergence, their asymptotic behavior can be slower than that of full gradient methods. Here we examine the relative behavior of four fast algorithms: (i) the PCG algorithms as described in [2]; (ii) the OSSPS described in [20]; (iii) Relaxed OSSPS proposed in [21]; and (iiii) Emission Reconstruction Incremental Optimization Transfer (ERIOT) described in Section IV. Based on a comparison of their respective convergence rates we propose a

This work was supported by Grant R01 EB000363 from the National Institute of Biomedical Imaging and Bioengineering.

fast and convergent hybrid algorithm that takes advantage of the properties of different algorithms to minimize the total number of iterations.

In Section II we describe penalized maximum likelihood PET image reconstruction; we then review PCG and (relaxed) OSSPS in Section III, and describe the new ERIOT algorithm in Section IV. In Section V we compare the convergence rate of ERIOT with the other three methods. Based on the comparison, we propose a hybrid algorithm that combines OSSPS, ERIOT and PCG in Section VI; we also describe an automated procedure to decide at which point to switch between algorithms.

II. Maximum a Posteriori Image Reconstruction

Given data $\mathbf{y}$ and an image $\mathbf{x}$, penalized maximum likelihood or MAP image reconstruction estimates the image by maximizing the log posterior density:

$$\mathbf{x}^* = \arg\max_{x\in\Omega} \sum_{i=1}^{N_s} L_i(y_i|\mathbf{x}) - \beta U(\mathbf{x}) \tag{1}$$

where Ω is the nonnegativity constraint set defined as $\Omega = \{\mathbf{x}|x_j \geq 0, \forall j\}$, and N_s is the number of detector pairs. The constant β is the hyper-parameter for a Gibbs prior. Here we use a Gaussian form for the Gibbs energy:

$$U(\mathbf{x}) = \sum_j \sum_{k>j, k\in N_j} w_{jk}(x_j - x_k)^2; \tag{2}$$

where N_j denotes the set of neighbors of voxel j, and w_{jk} is a weight for voxel j and its neighbor voxel k which varies as a function of the Euclidean distance between them. Assuming Poisson data, the log likelihood for the data y_i is given by

$$L_i(y_i|\mathbf{x}) = y_i \log \bar{y}_i(\mathbf{x}) - \bar{y}_i(\mathbf{x}) \tag{3}$$

with $\bar{y}_i(\mathbf{x}) = \sum_j p_{ij}x_j + r_i$ where r_i includes the randoms and scatter contribution; p_{ij} is the ij^{th} element of the projection matrix that represents the probability of a positron emission from voxel j being detected at detector pair i.

III. Optimization algorithms

The Preconditioned Conjugate Gradient (PCG) algorithm with a Newton-Raphson bent line search is used in this study [2]. We use the Polak-Ribiere format of updating equation with an expectation maximization (EM) type preconditoner described in [2].

Separable Paraboloidal Surrogate (SPS) methods maximize a quadratic surrogate function at each iteration. The surrogate function is chosen in such a way as to guarantee monotonicity in the MAP objective function [18], [19]. Its ordered subsets version, OS-SPS (called OSTR in the context of transmission tomography [20]), achieves an order-of-magnitude acceleration relative to the ordinary SPS algorithm [20]. However, OS-SPS does not converge to a maximizer of the objective function but to a limit cycle in the same way that ordinary OS methods do. However, by introducing suitable relaxation parameters, one can make OS-SPS converge to a maximizer [21]. The relaxed OS-SPS algorithm is written as follows:

$$\mathbf{x}^{n,m+1} = \left[\mathbf{x}^{n,m} + \alpha_n \boldsymbol{C}^{-1}\nabla\Phi_m(\mathbf{x}^{n,m})\right]_+ \tag{4}$$

where n and m denote the iteration and the subiteration (subset) index respectively, Φ_m denotes the mth subobjective function corresponding to the mth subset, α_n represents the relaxation parameter, and $[\mathbf{x}]_+$ denotes the orthogonal projection onto the nonnegative orthant and is computed componentwise as $([\mathbf{x}]_+)_j = \max\{x_j, 0\}$. $\boldsymbol{C}$ is a diagonal matrix whose jth diagonal element is given by

$$c_j = \frac{1}{M}\left(\sum_{i=1}^{N} p_{ij}p_i\check{c}_i + 8\beta \sum_{k>j,\ k\in N_j} w_{jk}\right) \tag{5}$$

where $p_i = \sum_j p_{ij}$ and

$$\check{c}_i = \begin{cases} 1/\max\{y_i, 1\}, & \text{if } [y_i - r_i]_+ > 0 \\ y_i/r_i^2, & \text{if } [y_i - r_i]_+ = 0. \end{cases} \tag{6}$$

The $\check{c}_i$ is called "precomputed curvature" and is an approximate Newton curvature for the marginal log-likelihood function in the sense that $\check{c}_i = -\ddot{h}_i(\hat{l})$ where $h_i(l) = y_i \log(l+r_i) - (l+r_i)$ and $\hat{l} = \arg\max_{l\geq 0} h_i(l)$ (see [20] for details).

In this study, we use $\alpha_n = \frac{\alpha_0}{n\times\gamma+1}$ as the step-size function, where α_0 and γ are relaxation parameters. Note that if we choose $\alpha_0 = 0$, then relaxed OS-SPS becomes OS-SPS.

IV. Emission Reconstruction Incremental Optimization Transfer (ERIOT)

Incremental optimization transfer is a convergent ordered-subset type algorithm, which was developed and applied to transmission tomography in [22]. By construction, it converges to a maximum of the objective function and its convergence rate is usually faster than its nonincremental counterpart. The incremental optimization transfer method effectively achieves fast convergence rates particularly when an ordered subset algorithm gets stuck or close to a limit cycle [22]. We adapt the incremental optimization transfer method to emission reconstruction following the development in [22] for the transmission case to yield the following algorithm, which refer to as Emission Reconstruction Incremental Optimization Transfer (ERIOT):

For initial variables $\bar{\boldsymbol{x}}_1, \ldots, \bar{\boldsymbol{x}}_M$, where M is the number of subsets:

for $n = 0, \ldots$

for $m = 1, \ldots, M$

$$\bar{\boldsymbol{x}}_m := \left[\left[\sum_{k=1}^{M} \boldsymbol{C}_k(\bar{\boldsymbol{x}}_k)\right]^{-1} \cdot \sum_{k=1}^{M} [\boldsymbol{C}_k(\bar{\boldsymbol{x}}_k)\bar{\boldsymbol{x}}_k + \nabla\Phi_k(\bar{\boldsymbol{x}}_k)]\right]^+ \tag{7}$$

end

$$\mathbf{x}^{n+1} = \bar{\boldsymbol{x}}_M$$

end

where $\boldsymbol{C}_m$ is a diagonal matrix whose jth diagonal element c_{mj} is given by

$$c_{mj} = \max\left\{ \sum_{i \in S_m} p_{ij} p_i \check{c}_i + \frac{8\beta}{M} \sum_{k>j,\ k \in N_j} w_{jk},\ \epsilon \right\} \qquad (8)$$

for some small $\epsilon > 0$ to ensure $c_{mj} > 0$, where $\check{c}_i$ is defined in (5). Note the similarity between (5) and (8) (for only one subset $(M = 1)$ they are equivalent).

OS-SPS uses only the partial gradient $\nabla\Phi_m$ for each update in (4); this greedy approach leads to fast initial convergence rates but when unrelaxed it usually does not achieve global convergence, which requires information about the whole gradient $\nabla\Phi = \sum_m \nabla\Phi_m$ in view of the first-order necessary optimality condition. In contrast, ERIOT uses all $\nabla\Phi_k$'s for each update in (7), although the gradients are not computed at the current iterate. The initial convergence rate of ERIOT is usually slower than OS-SPS due to the conservative approach of ERIOT. However, it was found in the context of transmission tomography in [22] that if OS-SPS is switched to the incremental optimization transfer algorithm when OS-SPS is stuck at the limit cycle, the incremental optimization transfer algorithm increases the objective function significantly, partly due to its built-in averaging of the limit cycle points [see the first term on the right hand side in (7)]. Therefore, ERIOT is particularly effective if one switches to it from OS algorithms when the OS algorithm approaches a limit cycle.

V. The Comparison of Convergence Rates

For OS methods, there are many parameters that can affect the convergence rate. We compared the convergence rates for reconstruction with different parameter combinations by simulation, found the best parameter combination and used these in the comparisons that follow. For the PCG algorithm, the step-size was selected using a Newton-Raphson bent line search. The ERIOT algorithm requires no user specified relaxation parameter.

We simulated a 4 ring small animal PET scanner with the transaxial physical parameters of a Concorde F220 scanner. A 128x128x4 subset of the 3D Hoffman brain phantom with voxel size 0.4mm was used to generate simulated sinogram data of total size 288x252x13 with an average of 30 counts/LOR. One section of the Hoffman brain phantom is shown in figure 1. All results shown below were initialized with two iterations of 3D OSEM.

We define the normalized objective function as: $(\Phi(\mathbf{x}^*) - \Phi(\hat{\mathbf{x}}^n))/(\Phi(\mathbf{x}^*) - \Phi(\hat{\mathbf{x}}^0))$ where $\mathbf{x}^0$ is the initial image and the optimal image $\mathbf{x}^*$ is approximated by 500 iterations of the PCG algorithm. We compared the normalized objective function vs. iteration for the OSSPS, relaxed OSSPS, ERIOT and PCG algorithm, and the results are shown in figure 2. Fig. 2 reflects our general observations: the OSSPS exhibits the fastest initial convergence rate; ERIOT exhibits a fast convergence rate for the first few iterations after the switch from OSSPS at iteration 7, and PCG exhibits faster asymptotic convergence behavior. Based on this observation, we developed the hybrid algorithm described below.

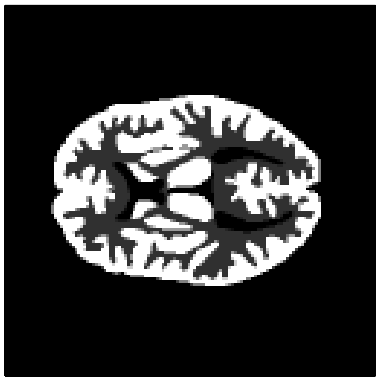

Fig. 1. One section of the Hoffman brain phantom.

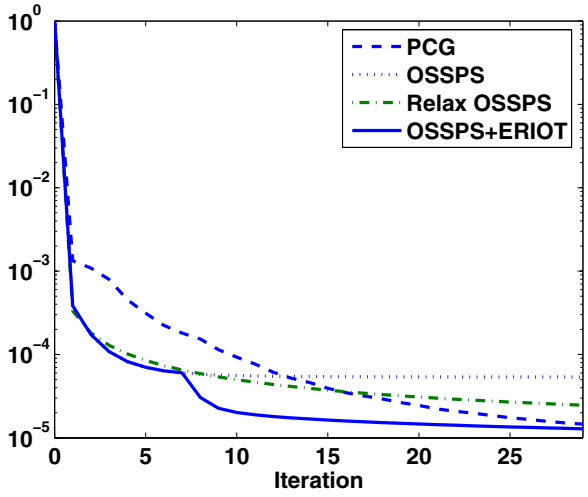

Fig. 2. Normalized objective function vs. iteration number for OSSPS, relaxed OSSPS, ERIOT and PCG following initialization with two iterations of 3D OSEM. All methods other than PCG used M=63 subsets. One iteration represents a single pass through the complete data.

VI. The Hybrid Algorithm

The OSSPS, ERIOT and PCG algorithms have different advantages and disadvantages at different stages of the reconstruction. Our hybrid algorithm attempts to combine the advantages of these algorithms. We begin with OSSPS, once the OSSPS reaches its limit cycle we switch to ERIOT. Finally, as the convergence rate of ERIOT reduces, we switch to PCG. An example of the convergence behavior of this approach is shown in Fig. 3. We achieve similar behavior in about 12-15 iterations of the hybrid algorithm as with 30 iterations of PCG.

To optimize the performance of this hybrid algorithm it is important to switch between algorithms at the appropriate

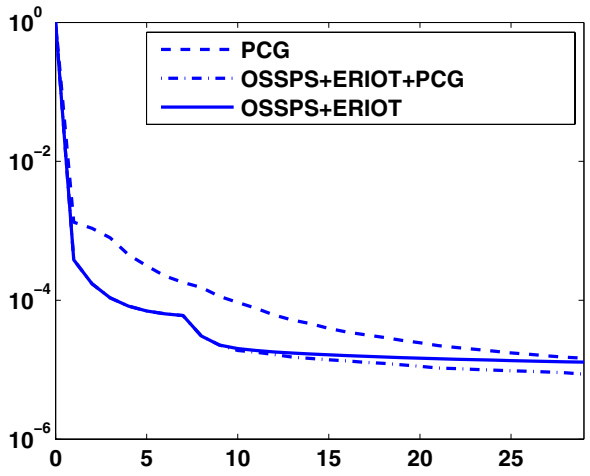

Fig. 3. Normalized objective function vs. iteration number for PCG, OSSPS+ERIOT and OSSPS+ERIOT+PCG following initialization with two iterations of 3D OSEM. For OSSPS, ERIOT: M=63 subsets.

iteration. To automatically select this point we fit the following exponential function to the objective values:

$$f(n) = a * (1 - e^{-(n-1)*b}) + c. \tag{9}$$

and estimate the parameters a, b, and c by:

$$\min_{a,b,c} \sum_{n=N_0}^{N} (\Phi(\mathbf{x}^n) - f(n))^2 \tag{10}$$

where $\Phi(\mathbf{x}^n)$ is the objective value at iteration n, N_0 is the iteration at which the current algorithm started and N is the current iteration number. Observing that $\max \Phi(\mathbf{x}^n)$ is $(a+c)$, we move to the next algorithm when the objective value for the current iteration is greater than 98% of $(a+c)$.

Because in the OSSPS and ERIOT algorithms the calculation of $\Phi(\mathbf{x}^n)$ requires an additional full forward projection, which is time consuming, we use the following approximation:

$$\Phi(\mathbf{x}) \approx M \times \sum_{y_i \in S_m} L_i(y_i|\mathbf{x}) - \beta U(\mathbf{x}) \tag{11}$$

where M is the subset number and S_m is the m^{th} subset. Since $\sum_{y_i \in S_m} L_i(y_i|\mathbf{x})$ is computed as part of the image update procedure for OSSPS and ERIOT, there is little additional cost in performing these computations. Fig. 4 shows an example of this approximation resulting in the same switching iterations as the exact values for different thresholds.

VII. Conclusions

We describe a new convergent incremental optimization transfer method, ERIOT, for PET image reconstruction which exhibits guaranteed convergence and has no user selected parameters. We have also proposed a hybrid method for combining different iterative algorithms to minimize the total iterations required for effective convergence. This hybrid approach was based on the a comparison of the convergence rates of

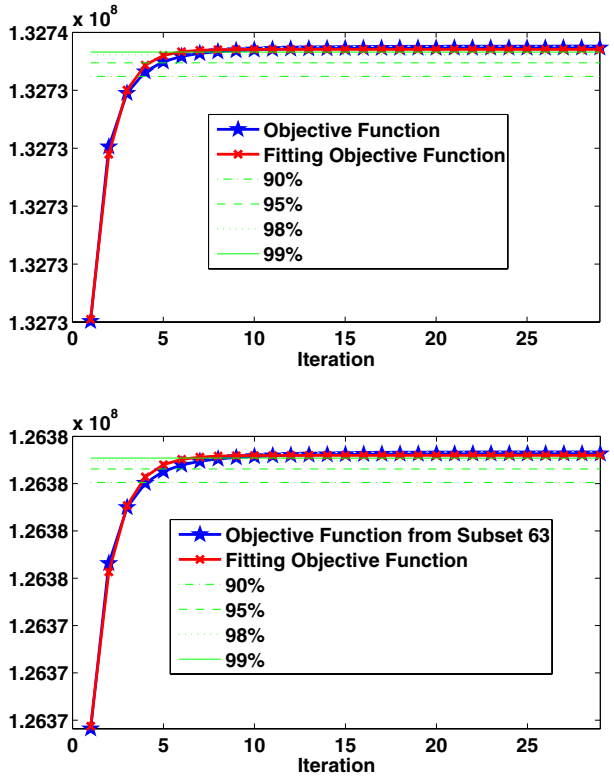

Fig. 4. Top figure: $\Phi(\mathbf{x}^n)$ and its corresponding $f(n)$; Bottom figure: approximation of $\Phi(\mathbf{x}^n)$ using data of subset 63 and its corresponding $f(n)$.

the OSSPS, ERIOT and PCG algorithms. Through automatic selection of the point at which we switch between these three algorithms we achieve superior performance at each iteration compared to the use of any one of these algorithms alone.

References

[1] E. Mumcuoglu, R.M. Leahy and S. Cherry, "Bayesian reconstruction of pet images: quantitative methodology and performance analysis," *Physics in medicine and Biology*, 41(9):1777–1807, Sept. 1996.

[2] J. Qi, R. Leahy, C. Hsu, T. Farquhar, S. Cherry, "Fully 3D Bayesian image reconstruction for the ECAT EXACT HR+," *IEEE Trans. Nucl. Sci.*, 45(6):1096–1103, June 1998.

[3] J. Qi, R. M. Leahy, S. R. Cherry, A. Chatziioannou and T. H. Farquhar, "High resolution 3d bayesian image reconstruction using the micropet small-animal scanner," *Physics in Medicine and Biology*, 43:1001–1013, 1998.

[4] L. Kaufman, Maximum likelihood, least squares, and penalized least squares for PET, *IEEE Trans. Med. Imag.,* vol. 12, pp. 200C214, June 1993.

[5] J.A. Fessler, S. Booth, "Conjugate-gradient preconditioning methods for shift-variant pet image reconstruction," *Image Processing, IEEE Transactions on*, 8(5):688 – 699, May 1999.

[6] A.R. De Pierro and M.E.B. Yamagishi, "Fast EM-like Methods for Maximum a posteriori Estimates in Emission Tomography," *IEEE Trans. Medical Imaging,* vol.20, no.4, pp.280-88, Apr. 2001.

[7] J.A. Browne and A.R. De Pierro, "A Row-action Alternative to the EM Algorithm for Maximizing Likelihoods in Emission Tomography," *IEEE Trans. Medical Imaging,* vol.15, no.5, pp.687-699, Oct. 1996.

[8] H. M. Hudson, R. S. Larkin, "Accelerated Image Reconstruction using Ordered Subsets of Projection Data," *IEEE Trans. Medical Imaging,* **13**, 1994, 100-108

[9] D. P. Bertsekas, "A new class of incremental gradient methods for least squares problems," *SIAM J. Optim.,* **7**, 1997, 913-926

[10] R. Neal and G. E. Hinton, "A view of the EM algorithm that justifies incremental, sparse and other variants", In M. I. Jordan, editor, *Learning in Graphical Models,* pages 255.268. Kluwer, Dordrencht, 1998.

[11] I. Hsiao, A. Rangarajan, and G. Gindi, "A new convergent MAP reconstruction algorithm for emission tomography using ordered subsets and separable surrogates," In *Proc. IEEE Intl. Symp. Biomedical Imaging,* pages 409.412, 2002.

[12] I. T. Hsiao, A. Rangarajan, P. Khurd, and G. Gindi, "An accelerated convergent ordered subsets algorithm for emission tomography," *Phys. Med. Biol.*, vol. 49, no. 11, pp. 2145–2156, June 2004.

[13] D. Blatt, A. Hero, and H. Gauchman, "An incremental gradient method that converges with a constant step size," *SIAM J. Optim.,* 2004, Submitted.

[14] K. Lange, D. R. Hunter, and I. Yang, "Optimization transfer using surrogate objective functions," *J. Computational and Graphical Stat.,* 9(1):1.20, March 2000.

[15] Bai B, Smith A M, Newport D F and Leahy R M, "On the effect of modeling block detector structure in small animal PET scanners," in *Proc. IEEE Nuclear Science Symp. Medical Imaging Conf*, 2004

[16] S. Ahn, J. A. Fessler, D. Blatt and A. O. Hero "Incremental optimization transfer algorithms: Application to transmission tomography." in *Proc. IEEE Nuclear Science Symp. Medical Imaging Conf.*, 2004.

[17] Q. Li, E. Asma and R. M. Leahy, "A Fast Fully 4D Incremental Gradient Reconstruction Algorithm for List Mode PET Dat," in *Proc. 2004 IEEE International Symposium on Biomedical Imaging,* Arlington, VA, 2004

[18] H. Erdogan and J. A. Fessler, "Monotonic algorithms for transmission tomography," *IEEE Trans. Med. Imag.*, vol. 18, no. 9, pp. 801–814, Sept. 1999.

[19] J. A. Fessler and H. Erdogan, "A paraboloidal surrogates algorithm for convergent penalized-likelihood emission image reconstruction," in *Proc. IEEE Nuclear Science Symp. Medical Imaging Conf.*, 1998, vol. 2, pp. 1132–1135.

[20] H. Erdogan and J. A. Fessler, "Ordered subsets algorithms for transmission tomography," *Phys. Med. Biol.*, vol. 44, no. 11, pp. 2835–2851, Nov. 1999.

[21] S. Ahn and J. A. Fessler, "Globally convergent image reconstruction for emission tomography using relaxed ordered subsets algorithms," *IEEE Trans. Med. Imag.*, vol. 22, no. 3, pp. 613–626, May 2003.

[22] S. Ahn, J. A. Fessler, D. Blatt, and A. O. Hero, "Convergent incremental optimization transfer algorithms: Application to tomography," *IEEE Trans. Med. Imag.*, 2005, To appear.