Simultaneous Estimation of Attenuation and Activity Images Using Optimization Transfer

Matthew Jacobson and Jeffrey A. Fessler

Department of Electrical Engineering and Computer Science
University of Michigan, Ann Arbor, MI 48109-2122

Abstract— **This paper addresses the application of optimization transfer to simultaneous statistical estimation of attenuation and activity images in tomographic image reconstruction. Although the technique we propose has wider applicability, we focus on the problem of reconstructing from data acquired via a post-injection transmission scan protocol. In this protocol, emission scan data is supplemented with transmission scan data that is acquired after the patient has received the injection of radio-tracer. The negative loglikelihood function for this data is a complicated function of the activity and attenuation images, leading to an objective function for the model that is difficult to minimize for the purpose of estimation.**

Previous work on this problem showed that when either the attenuation or activity image was held fixed, a paraboloidal surrogate could be found for the negative loglikelihood as a function of the remaining variables. This led to an algorithm in which the model's objective function is alternately minimized as a function of the attenuation and activity, using the optimization transfer technique. In the work we present here, however, we develop bivariate surrogates for the loglikelihood, i.e., functions that serve as surrogates with respect to both the attenuation and activity variables. Hence, simultaneous minimization in all variables can be carried out, potentially leading to convergence in fewer surrogate minimizations.

Keywords—**optimization transfer, surrogate, post-injection, bivariate**

I. Introduction

The material presented here revisits work done in [3] on the application of optimization transfer to simultaneous statistical estimation of attenuation and activity images in tomographic image reconstruction. As in [3], we treat this problem in the specific context of the post-injection transmission data acquisition protocol (although our technique could be applied to other reconstruction problems of a similar form). In this protocol, emission scan data is supplemented by transmission data acquired after the patient has received the injection of radio-tracer. This protocol has advantages such as increased patient throughput and reduced misregistration due to patient motion. Since radio-tracer is present in the patient during the transmission scan, it emits photons that contaminate the transmission data. Using both sets of data, one wishes to estimate simultaneously the a priori unknown attenuation and activity images μ and λ.

Statistical estimation usually involves minimizing a cost function, one term of which is the negative loglikelihood. For the post-injection problem, the negative loglikelihood, and hence also the total cost function, is a non-convex function of μ and λ, making global minimization difficult. One way to address this is to start with reasonable estimates of μ and λ (obtained, say, from an analytical reconstruction), and apply a monotonic iterative algorithm to descend to a local minimum. In taking this approach, one hopes that the initial estimates lie sufficiently close to the global minimizers so that local and global minimization coincide.

To implement the technique, one desires a monotonic minimization algorithm. The method of optimization transfer is an attractive candidate, since it requires only convex optimization techniques. With this method, one minimizes a sequence of convex *surrogate* functions which are tangent to the true cost function. The sequence of minimizers monotonically reduce the cost function as desired.

To implement the optimization transfer method, one must find a surrogate for the given cost function. Some tools for doing this for the post-injection problem cost function were developed in [2] and [3]. There, paraboloidal surrogates were identified for "plain" transmission and emission tomography. Then, recognizing that when either μ or λ was held fixed, the post-injection model's negative logliklihood resembled a plain transmission or emission model, it was suggested to use these paraboloidal surrogates to alternately minimize the post-injection cost function with respect to μ and λ.

A possible limitation of this approach is that it constrains the minimization to take steps in μ and λ separately rather than simultaneously. This could, conceivably, slow convergence. As an alternative to this, we indicate several ways in which bivariate surrogates can be derived, i.e., functions which are surrogates with respect to *both* μ and λ. Consequently, both variables can be updated simultaneously in the minimizations, possibly leading to convergence using fewer optimization transfers.

In what follows, we present these methods and test their performance. For simplicity, we consider unregularized maximum likelihood estimation, although the concepts extend readily to penalized likelihood estimation.

II. Preliminaries

Let μ and λ respectively denote vectors of attenuation and activity image values for a tomographically scanned object. In addition, let y denote the vector of measured projection data with components y_i, and let $\bar{y}$ be the statistical mean of y. In general, $\bar{y}$ depends on μ and/or λ in a manner specific to the system model.

For plain transmission and emission tomography, commonly considered models are,

$$\bar{y}_i(\mu) = b_i e^{-l_i(\mu)} + r_i \quad (1)$$

$$\bar{y}_i(\lambda) = s_i p_i(\lambda) + r_i. \quad (2)$$

In the above, $l_i(\mu)$ and $p_i(\lambda)$ denote discrete geometric forward projections of images into bin i. Equation (1) models transmission tomography data with blank scan data b_i and mean random counts r_i. Equation (2) is a model for emission tomography data with a priori known attenuation effects in s_i.

In [5], the following generalization of (1) was covered,

$$\bar{y}_i(\mu) = \sum_{m \in M_i} [b_{im} \exp(-l_i^m(\mu))] + r_i. \quad (3)$$

It models the case where the counts in bin i result from a set M_i of overlapping transmission source beams.

In the post-injection transmission scan protocol, one acquires emission data y^E and transmission data y^T, both while radio-tracer is present in the patient. These data have statistical means given by

$$\begin{aligned} \bar{y}_i^E(\mu,\lambda) &= \tau^E d_i^E(\epsilon_i e^{-l_i(\mu)} p_i(\lambda) + r_i^E) \quad (4) \\ \bar{y}_i^T(\mu,\lambda) &= \tau^T d_i^T(b_i e^{-l_i(\mu)} + r_i^T \\ &\quad + k_i \epsilon_i e^{-l_i(\mu)} p_i(\lambda)). \quad (5) \end{aligned}$$

Here τ^E, d_i^E, and r_i^E respectively denote emission scan acquisition time, dead time factors, and mean random counts (and similarly for the transmission scan). In addition b_i are blank scan data, ϵ_i are detector efficiency factors, and k_i are contamination factors determining the contribution to y^T of photons emitted from the radio-tracer.

For a given model of $\bar{y}$, the Poisson negative loglikelihood function has the form

$$L(\bar{y}_i) = \sum_{i=1}^{N} h_i(\bar{y}_i) \quad (6)$$

where $h_i(t) \stackrel{\triangle}{=} t - y_i \log t$. Maximum likelihood estimation corresponds to treating L as a cost function and minimimizing it as a function of μ and/or λ, depending on which of these variables are unknown.

For the post-injection problem, (6) can be written

$$\begin{aligned} L(\mu,\lambda) &= \sum_i [h_i^T(\bar{y}_i^T(\mu,\lambda)) + h_i^E(\bar{y}_i^E(\mu,\lambda))] \quad (7) \\ &= L^T(\mu,\lambda) + L^E(\mu,\lambda) \quad (8) \end{aligned}$$

where we have let

$$\begin{aligned} h_i^E(t) &\stackrel{\triangle}{=} t - y^E \log t \\ h_i^T(t) &\stackrel{\triangle}{=} t - y_i^T \log t \\ L^E(\mu,\lambda) &\stackrel{\triangle}{=} \sum_i h_i^E(\bar{y}_i^E(\mu,\lambda)) \\ L^T(\mu,\lambda) &\stackrel{\triangle}{=} \sum_i h_i^T(\bar{y}_i^T(\mu,\lambda)) \end{aligned}$$

on the right hand side of (7) and (8).

III. Reparametrization Technique and Optimization Transfer

The technique for minimizing the cost function that we consider here (conceptually illustrated in Fig. 1) is that of optimization transfer. Given a generic cost function $F(x)$ and a point x_n, we shall say that a differentiable, convex function $Q(x; x_n)$ is a *surrogate* for F generated at x_n if $Q(x; x_n) \geq F(x)$ with equality at x_n. Minimization of Q results in a new point x_{n+1} satisfying $F(x_n) \geq F(x_{n+1})$. Repeating this iteratively results in a sequence $\{x_n\}$ which monotonically reduces F. When the cost function is differentiable (which will always be the case here), it is also true that $\nabla Q(x_n) = \nabla F(x_n)$, i.e., $Q(x_n)$ is tangent to $F(x_n)$. When the iterations are judiciously initialized, one may hope that $\{x_n\}$ converges to a global minimizer.

In [2], [3], and [4], it was shown how paraboloidal surrogates could be generated for the transmission and emission tomography cost functions (when the means had the forms in (1) and (2)). In [5], an extension of these results was made for the overlapping beams transmission tomography problem (corresponding to equation (3)). In [3], the post-injection problem was also considered. There, an approach was proposed based on the observation that both (4) and (5) have the form of (1) when λ is held fixed, and the form of (2) when μ is held fixed. Since paraboloidal surrogates for these simpler models had been previously identified in [3], the optimization transfer technique could be applied to alternately minimize (7) with respect to μ and λ.

Our work proposes an alternative to this by observing that (4) and (5) can be converted into (3). To do so, we first let $p_i(\lambda) = \sum_{m \in M_i} g_{im} \lambda_m$ where m is an index for image voxels and $M_i = \{m \,|\, g_{im} > 0\}$.

Making the change of variables $\lambda_m = \gamma^{(1)} e^{-\gamma^{(2)} \tilde{\mu}_m}$ (where the $\gamma^{(1)}, \gamma^{(2)} > 0$ are constant scale factors), we substitute into (4) yielding,

$$\begin{aligned} \bar{y}_i^E(\mu,\tilde{\mu}) &= \sum_{m \in M_i} [\underbrace{(\tau^E d_i^E \epsilon_i g_{im} \gamma^{(1)})}_{b_{im}} e^{-\overbrace{(l_i(\mu) + \gamma^{(2)} \tilde{\mu}_m)}^{l_i^m(\mu,\tilde{\mu})}}] \\ &\quad + \underbrace{\tau^E d_i^E r_i^E}_{r_i} \quad (9) \end{aligned}$$

Equation (9) can now be likened to (3) where μ is replaced by the augmented vector $(\mu, \tilde{\mu})$ and r_i, b_{im}, and $l_i^m(\mu, \tilde{\mu})$ are as shown.

A similar transformation can be applied to (5). The overall result is that the functions $L^E(\mu, \tilde{\mu})$, $L^T(\mu, \tilde{\mu})$, and hence also $L(\mu, \tilde{\mu})$ are negative loglikelihoods for overlapping beam models. We shall see that it is possible to develop surrogates for this transformed problem in several ways.

IV. A Bivariate Paraboloidal Surrogate

Using the transform technique of section III, $L^E(\mu, \tilde{\mu})$ and $L^T(\mu, \tilde{\mu})$ become cost functions for overlapping beam models. Paraboloidal surrogates were developed for such functions in [5], and can be used here to get bivariate paraboloidal surrogates for $L^E(\mu, \tilde{\mu})$ and $L^T(\mu, \tilde{\mu})$. We briefly outline the construction of these surrogates.

For the L^E component of the cost function, the transform was accomplished, as we have shown, by letting,

$$\begin{aligned} b_{im} &= \tau^E d_i^E \epsilon_i g_{im} \gamma^{(1)} \\ r_i &= \tau^E d_i^E r_i^E \\ l_i^m(\mu,\tilde{\mu}) &= l_i(\mu) + \gamma^{(2)} \tilde{\mu}_m. \quad (10) \end{aligned}$$

Proceeding as in [5], we make the following definitions,

$$l_{im}^n = l_i^m(\mu_n, \tilde{\mu}_n)$$

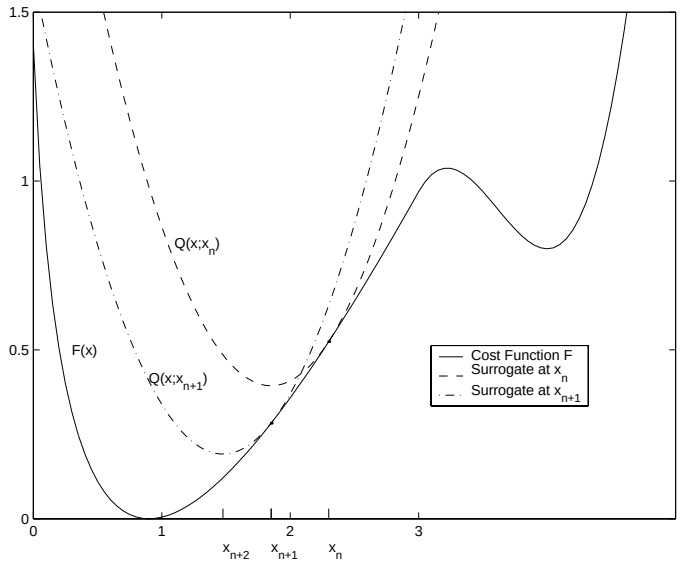

Fig. 1. Conceptual 1D illustration of the optimization transfer method for a hypothetical cost function $L(x)$.

$$
\begin{aligned}
u_{im}^n &= b_{im}\exp(-l_{im}^n) + r_i/|M_i| \\
\bar{y}_{im}^n &= \sum_{m\in M_i} u_{im}^n \\
b_{im}^n &= \frac{\bar{y}_{im}^n}{u_{im}^n} b_{im} \\
r_{im}^n &= \frac{\bar{y}_{im}^n}{u_{im}^n} r_i \\
g_{im}^n(l) &= (b_{im}^n e^{-l} + r_{im}^n) - y_i^E \log(b_{im}^n e^{-l} + r_{im}^n) \\
q_{im}^n(l) &= g_{im}^n(l_{im}^n) + \dot{g}_{im}^n(l_{im}^n)(l - l_{im}^n) \\
&\quad + \frac{1}{2} c_{im}^n (l - l_{im}^n)^2.
\end{aligned}
$$

where the optimal choice for the curvatures c_{im}^n is given in [5], equation (30). With these definitions, one can show that

$$Q(\mu,\tilde{\mu};\mu_n,\tilde{\mu}_n) = \sum_i \sum_{m\in M_i} \frac{u_{im}^n}{\bar{y}_{im}^n} q_{im}^n(l_i^m(\mu,\tilde{\mu})) \qquad (11)$$

is a bivariate paraboloidal surrogate for $L^E(\mu,\tilde{\mu})$ generated at $(\mu_n,\tilde{\mu}_n)$. An analogous surrogate can be developed for $L^T(\mu,\tilde{\mu})$ and the sum of the two surrogates is a surrogate for $L(\mu,\tilde{\mu})$.

An immediately apparent drawback to this approach is that the array c_{im}^n of curvature parameters is the same size as the sparse matrix of projection weights g_{im}. It is clear that it will involve time consuming operations to compute, a difficulty exacerbated by the fact that it must be computed on the fly for each new surrogate.

V. A Bivariate Non-Paraboloidal Surrogate

Observe that the transformed cost function can be written

$$L(\mu,\tilde{\mu}) = \Gamma(\mu,\tilde{\mu}) - \Upsilon(\mu,\tilde{\mu}) \qquad (12)$$

where

$$\Gamma(\mu,\tilde{\mu}) = \sum_i [\bar{y}_i^E(\mu,\tilde{\mu}) + \bar{y}_i^T(\mu,\tilde{\mu})] \qquad (13)$$

$$\Upsilon(\mu,\tilde{\mu}) = \sum_i [y_i^E \log \bar{y}_i^E(\mu,\tilde{\mu}) + y_i^T \log \bar{y}_i^T(\mu,\tilde{\mu})] \qquad (14)$$

are both differentiable, convex functions. By linearizing $\Upsilon(\mu,\tilde{\mu})$ at $(\mu_n,\tilde{\mu}_n)$, we have from the gradient inequality

$$
\begin{aligned}
L(\mu,\tilde{\mu}) \le Q(\mu,\tilde{\mu};\mu_n,\tilde{\mu}_n) \;\stackrel{\triangle}{=}\; & \Gamma(\mu,\tilde{\mu}) - \Upsilon(\mu_n,\tilde{\mu}_n) \\
& - < \nabla_\mu \Upsilon(\mu_n,\tilde{\mu}_n),\, \mu - \mu_n > \\
& - < \nabla_{\tilde{\mu}} \Upsilon(\mu_n,\tilde{\mu}_n),\, \tilde{\mu} - \tilde{\mu}_n >
\end{aligned}
$$

and clearly $Q(\mu,\tilde{\mu};\mu_n,\tilde{\mu}_n)$ is a bivariate, but non-paraboloidal surrogate for $L(\mu,\tilde{\mu})$ generated at $(\mu_n,\tilde{\mu}_n)$.

To minimize Q, we will typically require the gradients $\nabla_\mu Q$ and $\nabla_{\tilde{\mu}} Q$, with respect to μ and $\tilde{\mu}$. Let us define system matrices A and G such that $l_i(\lambda) = [A\lambda]_i$ and $p_i(\lambda) = [G\lambda]_i$ and let ∇_l and ∇_p denote gradients with respect to projection arrays l_i and p_i. Noting that Γ and Υ depend on μ only through l_i, it then follows from the chain rule that

$$\nabla_\mu Q(\mu,\lambda;\mu_n,\lambda_n) = A^T[\nabla_l \Gamma(\mu,\lambda) - \nabla_l \Upsilon(\mu_n,\lambda_n)]. \quad (15)$$

Furthermore, noting from the chain rule that $\nabla_{\tilde{\mu}} = -\gamma^{(2)}\lambda \odot \nabla_\lambda$ (here, $\odot$ denotes component-wise multiplication) then we also have

$$
\begin{aligned}
&\nabla_{\tilde{\mu}} Q(\mu,\lambda;\mu_n,\lambda_n) = \\
&\quad \gamma^{(2)}[\lambda_n \odot \nabla_\lambda \Upsilon(\mu_n,\lambda_n) - \lambda \odot \nabla_\lambda \Gamma(\mu,\lambda)]. \quad (16)
\end{aligned}
$$

Equations (15) and (16) show that the surrogate gradients can be computed straightforwardly in the original (μ,λ) space.

The surrogate parameters to be computed and stored are $\nabla_l \Upsilon(\mu_n,\lambda_n)$ and $\lambda_n \odot \nabla_\lambda \Upsilon(\mu_n,\lambda_n)$. Observe that,

$$
\begin{aligned}
[\nabla_l \Upsilon(\mu_n,\lambda_n)]_i &= \frac{y_i^E}{\bar{y}_i^E(\mu_n,\lambda_n)} \frac{\partial \bar{y}_i^E(\mu_n,\lambda_n)}{\partial l_i} \\
&\quad + \frac{y_i^T}{\bar{y}_i^T(\mu_n,\lambda_n)} \frac{\partial \bar{y}_i^T(\mu_n,\lambda_n)}{\partial l_i} \quad (17) \\
[\nabla_p \Upsilon(\mu_n,\lambda_n)]_i &= \frac{y_i^E}{\bar{y}_i^E(\mu_n,\lambda_n)} \frac{\partial \bar{y}_i^E(\mu_n,\lambda_n)}{\partial p_i} \\
&\quad + \frac{y_i^T}{\bar{y}_i^T(\mu_n,\lambda_n)} \frac{\partial \bar{y}_i^T(\mu_n,\lambda_n)}{\partial p_i} \quad (18) \\
\nabla_\lambda \Upsilon(\mu_n,\lambda_n) &= G^T \nabla_p \Upsilon(\mu_n,\lambda_n). \quad (19)
\end{aligned}
$$

We see in (17) and (18) that once the means and their derivatives are computed (a seemingly inevitable step in any loglikelihood based algorithm) then only 3 multiplications and 1 addition per projection data element are required to find $\nabla_l \Upsilon$ and $\nabla_p \Upsilon$. Computing $\nabla_\lambda \Upsilon$ via (19) requires a backprojection. However, we will see that a reconstruction typically requires two surrogate minimizations at most, making this additional computational burden marginal. The burden can also be offset by parallel computation.

VI. Results

In this section, we compare the performance of optimization transfer using the bivariate surrogates proposed in sections IV and V, and the alternating descent scheme proposed in [3] on

simulated data. Two sets of data were considered, which we label D1 and D2. Both sets were generated by first forward projecting simulated attenuation and activity images of a torso phantom, then using (4) and (5) to generate mean data, and finally adding Poisson noise. Images were 128×128 while projections were discretized into 60 equi-spaced angles and 185 radial bins. For both the $l_i(\mu)$ and $p_i(\lambda)$ forward projectors, discrete approximations to line integrals were used.

For D1, the mean total counts for both the emission and transmission data was 0.5 million. In D2, the noisier set, the mean total counts was 0.05 million. In all cases, dead time was ignored, and uniform detector efficiency and blank scan data were assumed. Mean randoms count rates for all data were 30% and a 3% contamination factor was used.

Iterative reconstructions were initialized using analytical reconstructions, obtained by substituting the measured projection data in place of $\bar{y}^E$ and $\bar{y}^T$ into (4) and (5). These equations were then solved for $l_i(\mu)$ and $p_i(\lambda)$ (incorporating smoothing and thresholding where appropriate) whereupon filtered back projection was used to reconstruct an initial point (μ_0, λ_0). Surrogate minimization was implemented using the Conjugate Barrier iterative algorithm [1]. Many sub-iterations were run to ensure that the surrogates were approximately minimized.

For the purposes of transformation between (μ, λ) and $(\mu, \tilde{\mu})$ space, $\gamma^{(1)}$ was set to an upper bound on the activity (assumed known). Some preliminary trial and error was used to determine a good working value for $\gamma^{(2)}$. When bivariate surrogates were used, the activity image λ was constrained to have a small, strictly positive lower bound, so as to prevent situations where $\tilde{\mu}$ might approach infinity.

Plots comparing convergence rates appear in Figs. 2 and 3. For each of the three optimization transfer techniques considered, the descent of the cost function as a function of outer iterations is shown. For the bivariate surrogate methods an outer iteration refers to one surrogate minimization. For the alternating descent method, an outer iteration refers to two surrogate minimizations, one to update the attenuation image and one to update the activity image.

We see that the non-paraboloidal surrogate approach minimizes the cost function in essentially one outer iteration for both data sets. For D1, this approach only slightly outperforms alternating descent (see Fig. 2). This is seemingly because the data is non-noisy. Hence, the initial analytically reconstructed point is sufficiently close to the maximum likelihood estimate that Alternating Descent can also achieve the minimum in essentially one outer iteration. For the noisier data D2 (see Fig. 3), the initial point is farther away and convergence is slower compared to the non-paraboloidal surrogate approach.

The performance of the bivariate paraboloidal surrogate is significantly poorer than the two other alternatives in terms of convergence rate. It is also slow in terms of computation time as discussed in section IV, making it all the more unattractive.

The above findings are corroborated by trends in Figs. 4 and 5, which show sample 1D profiles of the cost function for D1 together with profiles of the two bivariate surrogates. For Fig. 4, surrogates were generated at the initial analytically reconstructed point. For Fig. 5, this point was distanced from the minimum by applying a scale factor of 0.5 and surrogates were

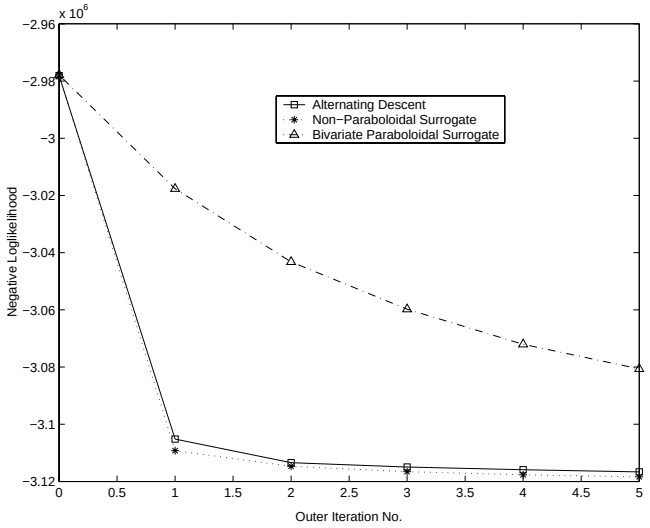

Fig. 2. Comparison of convergence rates of the different optimization transfer approaches for D1. The plot shows cost function descent with outer iterations.

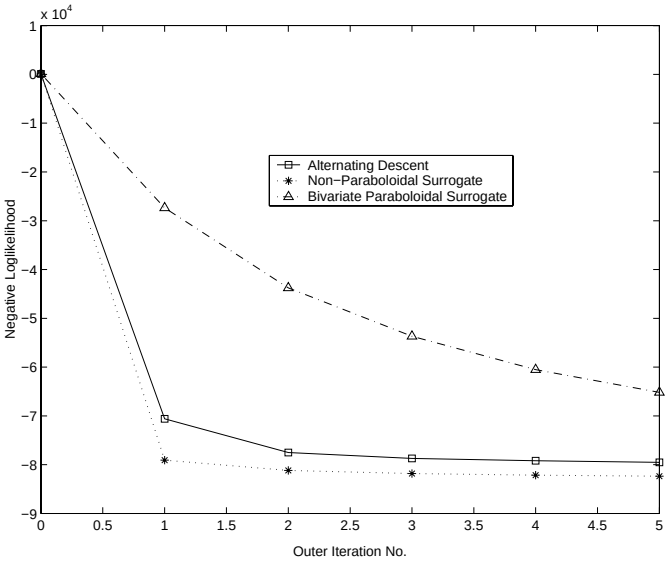

Fig. 3. Comparison of convergence rates of the different optimization transfer approaches for D2. The plot shows cost function descent with outer iterations.

generated at the resulting point. In both cases, we see that the non-paraboloidal surrogate closely approximates the cost function whereas the paraboloidal surrogate provides a poor, high curvature approximation. This accounts for their differing convergence rates. It also shows that the effectiveness of the non-paraboloidal surrogate persists as the generating point is distanced from the cost function minimum, possibly explaining its robust performance under increased noise in Fig. 3.

VII. Conclusions

For statistical tomographic image reconstruction problems where both the attentuation and activity images are unknown, we have proposed a transform methodology which effectively turns the model into a familiar form from transmission tomogra-

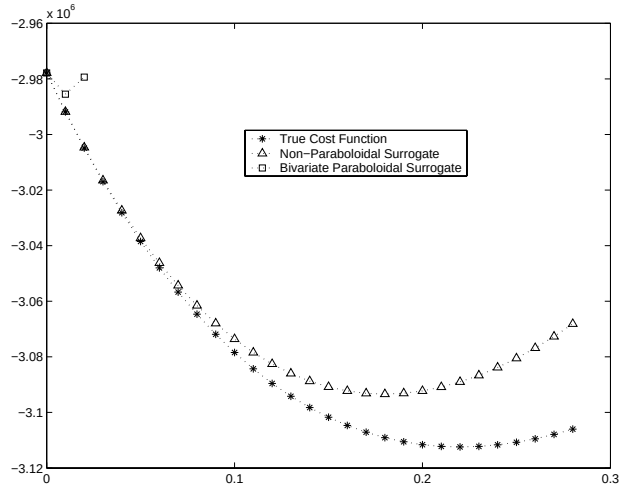

Fig. 4. Comparitive profiles of the cost function for D1 and the bivariate surrogates generated at an analytically reconstructed point in image space.

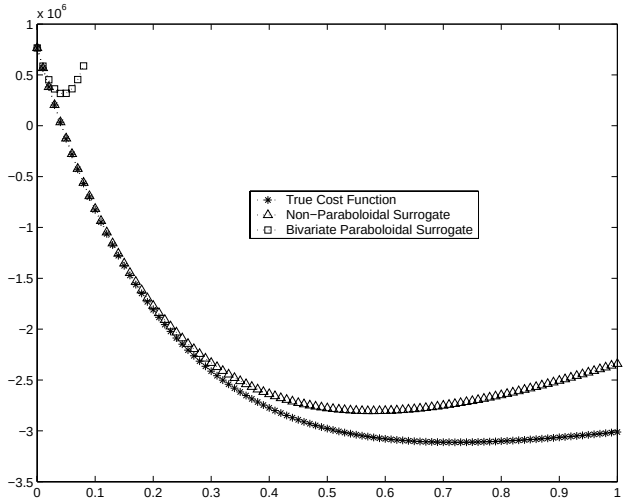

Fig. 5. Comparitive profiles of the cost function for D1 and the bivariate surrogates. In generating the surrogates, a point in image space was analytically reconstructed and then perturbed to a more distant location.

phy. This allowed us to obtain bivariate paraboloidal surrogates for the post-injection cost function based on previous work, as well as non-paraboloidal ones.

We compared the performance of these surrogates to an alternating descent method previously proposed in [3]. The bivariate paraboloidal surrogate proved unattractive in terms of both computation and convergence rate. The convergence rate of the non-paraboloidal surrogate approach proved competitive with the alternating descent method when convergence rate was measured in outer iterations (number of surrogate minimizations). Also, the computational effort required to generate surrogates was similar in both.

Thus far, we have not made any effort to optimize the computational effort involved in minimizing the surrogates, so it remains to be seen whether the non-paraboloidal surrogates approach can be exploited to reduce overall CPU time. However, various ordered subsets algorithms are now available which, hopefully, will allow one to quickly minimize the surrogates and, hence, exploit their full potential.

Future work could also include extending the techniques proposed here from unregularized maximum likelihood to penalized likelihood reconstruction. This may require that one define roughness penalties on the transformed activity image $\tilde{\mu}$, since the transformation of variables that we have proposed will not generally preserve the convexity of penalty functions defined on λ.

Finally, the findings in this work, as well as numerous other tests not shown, indicate that the non-paraboloidal surrogate approach typically requires only the first surrogate to approximately achieve the minimum. This might mean that the first surrogate calculation amounts to an analytical procedure for turning the non-convex post-injection problem into an equivalent convex one. This implication might follow for pure transmission tomography models (e.g. (1) and (3)), as well, since the same surrogate technique applies.

REFERENCES

[1] A. Ben-Tal and A. Nemirovski, "The conjugate barrier method for non-smooth convex optimization," Minerva Optimization Center, Technion - Israel Institute of Technology, TR #5/99, 1999.

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

[3] H. Erdogan, "Statistical image reconstruction using paraboloidal surrogates," *PhD Thesis*, University of Michigan, Ann Arbor, MI, 48109-2122, Jul. 1999.

[4] J. Fessler and H. Erdogan, "A paraboloidal surrogates algorithm for convergent penalized-likelihood emission image reconstruction," in *Proc. IEEE Nuc. Sci. Symp. Med. Im. Conf.*, 1998, vol. 2, pp. 1132–5.

[5] D. Yu, J. Fessler, and E. Ficaro, "Maximum likelihood transmission image reconstruction for overlapping transmission beams" *IEEE Trans. Med. Imaging*, vol. 19, no. 11, pp. 1094–1105, Nov. 2000.