Canadian Mathematical Society
Conference Proceedings

Dynamic Emission Tomography – Regularization and Inversion

Jean Maeght, Dominikus Noll, and Stephen Boyd

To Jonathan Borwein, on the occasion of his award of a Dhc

Abstract. The problem of emission tomography, inverting the attenuated Radon transform, is moderately ill-posed if the unknown emission source is static. Here we consider the case where the emission source is dynamic due to the movement of molecules after metabolism. Inversion now requires a model of the underlying dynamics and presents a seriously ill-posed inverse problem. We present a regularization approach and a numerical experiment to validate our method.

1. Problem Setting

Single Photon Emission Computed Tomography (SPECT) is a non-invasive diagnostic technology which is used to show blood flow in the heart muscle, extent of damage in stroke patients, presence and degree of malignancy of tumors, and much else. SPECT images the function of the body through a *tracer*, a biochemical molecule labelled with radioactivity. The radioactive material is incorporated by the patient and metabolized by the organ under investigation. The emissions are recorded by a SPECT camera rotating around the patient, and a 3D visualization is created from the 2D projection data via an inverse mathematical method.

The traditional inverse method in tomography, filtered back-projection (FBP), may no longer be used if the unknown radiating source is dynamic, i.e., changes significantly during the scan. However, representing *dynamic* phenomena is the truly interesting case, since it allows to assess the movements of molecules during biodistribution, metabolism and washout, and consequently gives a much deeper insight into the function of the body. Recently, a first step towards *dynamic* SPECT reconstruction has been achieved through a mathematical approach replacing FBP [**22, 19, 23, 13, 1**]. This approach will be further developed here.

Key words and phrases. Dynamic emission tomography, dynamic attenuated Radon transform, compartment model, exponential fit, augmented Lagrangian SQP method, photon transport equation.

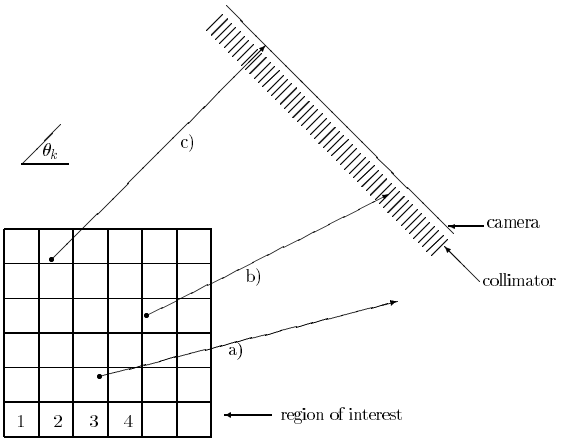


Figure 1. Photons radiating from the region of interest. a) misses the camera, b) is absorbed by the collimator, c) passes the collimator and hits the camera.

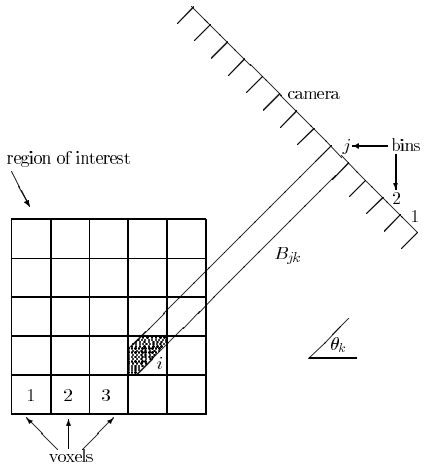


Figure 2. The coefficient c_{ijk} is the relative volume of voxel i lying within the beam B_{jk} connecting i and the receptor bin j during the camera position θ_k.

Similar to the static case, the reconstruction of the unknown 3D dynamic source may be decoupled into a series of 2D slices if Compton scatter is ignored in the transport model. The details of this process are recalled in Appendix 1. In each slice, the unknown dynamic source $f(x,t)$, representing the radio activity at position x and time t, is then related to the projection data $d(t,s,\omega)$ through the *dynamic attenuated Radon transform*

$$(1)\quad \mathcal{R}[\mu, f(\cdot,t)](s,\omega) = \int_{-\infty}^{\infty} f(s\omega^{\perp}+\tau\omega,t)\, e^{-\int_{\tau}^{\infty}\mu(s\omega^{\perp}+\tau'\omega)\,d\tau'}\, d\tau = d(t,s,\omega).$$

Here $\omega = (\cos\theta, \sin\theta)$ and $\omega^{\perp} = (-\sin\theta, \cos\theta)$, and the integrals in (1) are line integrals along the lines $L(s,\omega) = \{s\omega^{\perp}+\tau\omega : \tau \in R\}$. The term $d(t,s,\omega)$ represents the number of counts recorded at time t by a sensor positioned on the line $L(s,\omega)$. Figure 1 illustrates the situation in the case of a parallel hole collimator which at a given angular position admits only photons traveling perpendicular to the camera surface. It is the effect of collimation which allows attributing a direction ω of propagation to a count detected at time t and position s on the camera.

Notice that $\mu(x)$, the unknown attenuation map, which essentially represents the density of the tissue at position x, is not the primordial goal of the SPECT reconstruction. Knowledge of $\mu(x)$ is required, however, in order to reconstruct the unknown emission source, $f(x,t)$, and the basic problem in SPECT, therefore, is to solve (1) for the unknown source and attenuation map. In practice, two types of approaches are employed, leading to different mathematical problems.

A first possibility is to do a transmission scan (CT) in parallel with the SPECT acquisition, which allows to calculate μ beforehand. This reduces the complexity of the problem, but requires technology not always available.

The second line of attack is to estimate μ entirely by mathematical means, using the emission data $d(t,s,\omega)$ only. This may be done either by simultaneously solving (1) for both unknowns, [**12**], a difficult task even in the static case, or by estimating μ beforehand via Helgason's consistency formula, an approach followed in [**26, 28**], and more recently in [**37**]. More elementary methods assume for instance a constant

attenuation over the contour of the body if the latter can be recovered, as is often the case, from the emission data. This is usually satisfactory in brain or renal SPECT, but less satisfactory in myocardiac imaging.

We will eventually assume that the attenuation map μ is known, estimated e.g. by any of the above techniques. This renders the integral equation (1) linear in f. In the static case, inverting (1) is a well-studied moderately ill-posed linear inverse problem. The state of the art is to attack it by iterative techniques based on versions of the EM-algorithm [**36, 6**].

In the dynamic case, the problem remains intrinsically more difficult even with known μ, as we proceed to indicate. Considering exemplarily the case of a rotating single head camera, schematically displayed in Figure 1, it is crucial to observe that the camera head, rotating e.g. over 180^0, will acquire *inconsistent projection data*. By that we mean that since a considerable time (usually of the order of 20 minutes) will have elapsed between the initial and final positions of the camera, the unknown source will have changed significantly. The acquired projections will therefore correspond to entirely different dynamic states of the source, which means that the integral equation (1) is largely under determined. Formally, this may be seen by the fact that we have to reconstruct a 3D function $f(x,t)$ on the basis of data given on a the 2D submanifold

$$\mathcal{M} = \{(t,s,\omega) : \omega = (\cos\theta, \sin\theta),\, \theta = 2\pi t/T,\, 0 \le t \le T\},$$

of the full 3D data manifold $[0,T] \times R \times S^1$ (T the time of the scan). In order to address this fact, we have to provide additional information to reconstruct the unknown dynamic image $f(x,t)$.

2. Tracer Dynamics

The missing information is based on a suitable model of the chemical reaction the tracer molecule undergoes during metabolism. As this is highly dependent on the situation, we proceed to give a rather general description as it applies to our situation.

Based on theoretical knowledge of the underlying reaction, it is usually possible to identify a number of different chemical states, called *compartments*, the tracer molecule will assume. The compartment model generally represents a simplified version of the complex reaction. With $c_\nu(t)$ the concentration of tracer in the νth compartment, the reactive process is represented by a dynamical system

$$c'(t) = Kc(t) + c_B(t) \tag{2}$$

with a system matrix $K = (k_{ij})$ involving exchange rates k_{ij} between compartments, and a driving term c_B representing the supply from the blood pool. As radio active events originating from any one of the chemical states c_ν, $\nu = 1, \ldots, P$, could not possibly be distinguished from each others, the emission source is of the form $f(t) = \sum_{\nu=1}^{P} c_\nu(t)$, which means that we cannot assess the tracer concentrations c_ν directly.

So far we have represented the traditional point of view of compartmental modeling which seeks a *single* dynamic model $f(t)$ for the organ as a whole. This point of view changes with tomography, where we wish to account for spatially varying behavior. Assuming that the chemical reaction remains in principal the same over the entire organ, we allow the exchange rates to vary locally. This

leads to a *localized compartment model* of the form $c'(x,t) = K(x)c(x,t) + c_B(x,t)$ involving localized exchange rates $k_{ij}(x)$ and leading to a source term of the form

$$f(x,t) = \sum_{\nu=1}^{P} c_\nu(x,t).$$

The inverse problem of dynamic emission tomography, in a rather general form, is then the following:

Find $f(x,t)$, $\mu(x)$ and exchange rates $k_{ij}(x)$ such that

$$(\mathcal{P}) \quad \begin{aligned} &(i) && \mathcal{R}[\mu, f(\cdot,t)](s,\omega) = d(t,s,\omega) \text{ for } (t,s,\omega) \in \mathcal{M} \\ &(ii) && c'(x,t) = K(x)c(x,t) + c_B(x,t) \\ &(iii) && f(x,t) = \sum_{\nu=1}^{P} c_\nu(x,t) \end{aligned} \tag{3}$$

are satisfied.

This model may in practice be completed by prior information about the exchange rates such as bounds or physiologically meaningful ranges.

The shape of the input function, an impulse, is qualitatively known and may conveniently be modeled as $c_B(t) = Ae^{-\mu t} - Ae^{-\nu t}$ with $A > 0$ and $\mu > \nu > 0$, if $t = 0$ corresponds to the moment of injection of the pharmaceutical. If a considerable time between the injection of the drug and the initialization of the scan has elapsed, the source term may generally have a simpler form and is often even assumed constant. Assuming that the system matrices $K(x)$ are diagonalizable, which in practice is always the case, the source term will have the functional form

$$f(x,t) = \sum_{\nu=1}^{P+2} a_\nu(x)\, e^{-\lambda_\nu(x)t}. \tag{4}$$

Recovering the exchange rates $k_{ij}(x)$ from (4) is in general impossible, but the parameters $\lambda_i(x)$, $a_j(x)$ still carry useful diagnostic information.

Dynamics of existing radio tracers for SPECT use scarcely more than two compartments (cf. [**16**]), even though more complex models with up to five compartments have been proposed for some positron emitters (cf. [**34**]). During the following, we shall therefore restrict our analysis to a two compartment model with constant input function c_B. This is already of sufficient generality for a number of applications, including for instance a myocardiac viability study presented by Limber et al. [**19**], or a renal dynamic SPECT scan discussed in [**13**]. The functional form of the unknown emission source, accordingly, is a biexponential

$$f(x,t) = a_1 e^{-\lambda_1(x)t} + a_2 e^{-\lambda_2(x)t} + a_3(x), \tag{5}$$

and we may substitute (5) for the equations (ii), (iii) of $(\mathcal{P})$. The dynamic SPECT inversion may now be transformed into a nonlinear optimization problem:

$$(NL) \quad \begin{aligned} &\text{minimize} && ||\mathcal{R}[\mu]f - d|| \\ &\text{subject to} && f(x,t) = a_1(x)e^{-\lambda_1(x)t} + a_2(x)e^{-\lambda_2(x)t} + a_3(x). \\ & && \lambda_1(x) \geq 0,\ \lambda_2(x) \geq 0. \end{aligned}$$

This approach is the starting point for our algorithmic discussion. Before continuing this line, we discuss a possible discretization of $(\mathcal{P})$.

3. Discretization

We assume that the camera cross-section has been divided into M bins $j = 1, \dots, M$ of size Δs, the jth bin corresponding to an interval $[s_j - \Delta s/2, s_j + \Delta s/2]$. Suppose the camera takes S angular positions $\omega_k = (\cos\theta_k, \sin\theta_k)$, $k = 1, \dots, S$, with $\theta_k = (k-1)\Delta\theta$, and remains in the kth position during the time interval $[t_k - \Delta t/2, t_k + \Delta t/2]$. Then the datum acquired in the jth camera bin during the kth stop is

$$y_{jk} = \int_{t_k-\Delta t/2}^{t_k+\Delta t/2} \int_{s_j-\Delta s/2}^{s_j+\Delta s/2} d(t, s, \omega_k)\, ds\, dt.$$

Integrating (1) accordingly gives the condensed form of equation (1):

$$\int_{t_k-\Delta t/2}^{t_k+\Delta t/2} \int_{s_j-\Delta s/2}^{s_j+\Delta s/2} \int_{L(s,\omega_k)} f(x,t)\, e^{-\int_\tau^\infty \mu(s\omega^\perp + \tau'\omega)\, d\tau'} d\tau\, ds\, dt = y_{jk}.$$

We proceed to discretize the unknown image into N pixels P_i, $i = 1, \dots, N$ of size $\Delta x \times \Delta x$, and assume the activity $f(x,t)$ constant on a pixel at every instant t, i.e., $f(x,t) = f(x_i, t)$ for the barycenter x_i of the ith pixel. Then the unknown variables are the activities originating from these pixels and cumulated over the stops

$$f_{ik} := \int_{t_k-\Delta t/2}^{t_k+\Delta t/2} \int_{P_i} f(x,t)\, dx\, dt = |P_i| \int_{t_k-\Delta t/2}^{t_k+\Delta t/2} f(x_i, t)\, dt. \tag{6}$$

Assuming the attenuation map $\mu(x)$ constant on each pixel, and using (6), the discretized form of (1) becomes

$$\sum_{i=1}^{N} R_{ijk}[\mu]\, f_{ik} = y_{jk} \tag{7}$$

where the coefficients $R_{ijk}[\mu]$ are

$$R_{ijk}[\mu] = \int_{t_k-\Delta t/2}^{t_k+\Delta t/2} \int_{s_j-\Delta s/2}^{s_j+\Delta s/2} \int_{L(s,\omega_k)\cap P_i} e^{-\int_\tau^\infty \mu(s\omega^\perp + \tau'\omega)\, d\tau'} d\tau\, ds\, dt. \tag{8}$$

During the sequel, (7) will be abbreviated as $R[\mu]f = y$, the discretized version of (1). It is instructive to see the meaning of the coefficients in the unattenuated case $\mu = 0$. Here $R_{ijk}[0]$ is the surface of that part of the ith pixel lying in the strip connecting to the jth camera bin during stop k (see Figure 2).

Observe that, unless $\mu = 0$, (8) has to be calculated numerically, which by itself requires a discretization. We mention that our approach is not the only possible, and various other discretizations for $\mathcal{R}$ have been proposed, including wavelet bases [**40**], or C^0-elements [**12**].

Let us proceed to discretize (5). Using (6) we obtain the estimate

$$f_{ik} = \big(a_1(x_i)\, e^{-\lambda_1(x_i)t_k} + a_2(x_i)\, e^{-\lambda_2(x_i)t_k} + a_3(x_i)\big)(\Delta t \Delta x^2 + o(\Delta t \Delta x^2))$$

if x_i is the barycenter of the ith pixel. Setting $\lambda_{1i} = \lambda_1(x_i)\Delta t$, $\lambda_{2i} = \lambda_2(x_i)\Delta t$ and $a_{\nu i} = a_\nu(x_i)\Delta t \Delta x^2$, we are led to replace the constraint (5) by its discretized version

$$f_{ik} = a_{1i} e^{-\lambda_{1i} k} + a_{2i} e^{-\lambda_{2i} k} + a_{3i}. \tag{9}$$

4. Prony's Method Extended

The projection data are in practice always corrupted by noise, and it is necessary to solve the discretized problem (7), (9) in the least squares sense. Moreover, since the problem is ill-posed, it is wise to include a regularizing term $I[f]$ into the objective function. Assuming that μ is known suggests the following approach via nonlinear programming:

$$(NL)\qquad \begin{array}{ll} \text{minimize} & \ell(f) = \|R[\mu]f - y\|^2 + I[f] \\ \text{subject to} & f_{ik} = a_{1i}e^{-\lambda_{1i}k} + a_{2i}e^{-\lambda_{2i}k} + a_{3i} \\ & \lambda_{1i} \geq 0,, \lambda_{2i} \geq 0 \end{array}$$

Possible choices for $I[f]$ will be discussed in Section 8. The problem has been analyzed by Borwein and Sun [**3**] in the case $I[f] = 0$ and found well-behaved if the matrix $R[\mu]$ has maximal rank and MS is larger than $5N$, the number of unknown variables in (NL). While the rank condition is often satisfied, in particular if every camera bin can see at least one pixel at each stop, the second condition is in practice only satisfied for SPECT devices with a large number of receptors, and does not hold for the usual rotating camera systems. As an example, consider the case of a camera cross section divided into $M = 64$ bins, the camera taking $S = 64$ stops, and divide the region of interest into 64×64 pixels. Then $5N > MS$ for a single head camera, and even $5N > 3MS$ when a triple head camera is used.

And in fact, our implementation, based on the limited memory BFGS approach of Nocedal et al. [**21, 5**] shows that (NL) is a difficult problem subject to instabilities in particular when the eigenvalues tend to coalesce, $\lambda_{1i} \approx \lambda_{2i}$. Reconstructing a *typical slice* above may take between 1 and 2 hours CPU, which is too slow for convenient clinical applications.

In order to circumvent the numerical difficulties of (NL), we consider a generalization of Prony's method for fitting exponentials (cf. [**31**]). The key idea is the simple fact that a curve f_k of biexponential form (9) satisfies a difference equation

$$f_k = \alpha_1 f_{k-2} + \alpha_2 f_{k-1} + \alpha_3, \tag{10}$$

where the five parameters $(a_1, a_2, a_3, \lambda_1, \lambda_2)$ in (9) correspond to the five pieces of information $\alpha_1, \alpha_2, \alpha_3$ along with the initial values f_0, f_1 in (10). The links are the following. If f_k has the form (9), then

$$\alpha_1 = -e^{-\lambda_1+\lambda_2},\ \alpha_2 = e^{-\lambda_1} + e^{-\lambda_2},\ \alpha_3 = (1 - \alpha_1 - \alpha_2)a_3, \tag{11}$$

along with the obvious expressions for f_0, f_1. Conversely, if f_k satisfies a difference equation (10) and if we impose conditions to avoid the possible oscillatory solutions, we retrieve the form (9) through

$$a_1 + a_2 = f_0 - \frac{\alpha_3}{1 - \alpha_1 - \alpha_2},\ a_1u_1 + a_2u_2 = f_1 - \frac{\alpha_3}{1 - \alpha_1 - \alpha_2} \tag{12}$$

where u_i are the (real) roots of $x^2 - \alpha_2 x - \alpha_1 = 0$.

In dynamic emission tomography we are mainly interested in two cases: (i) washout only, where the dynamic profiles are decaying, and (ii) uptake and washout, where activity in the organ initially increases and after reaching the peak activity, decays (cf. [**4, 10, 15, 18, 19, 30, 34**]). The corresponding information may be expressed in terms of the difference equation:

Lemma 1. *The following statements are equivalent:*

1. $f_k = a_1 e^{-\lambda_1 k} + a_2 e^{-\lambda_2 k} + a_3$ *with* $\lambda_1 \neq \lambda_2$ *is monotonically decreasing and nonnegative on* $k = 0, 1, \ldots$
2. f_k *satisfies the difference equation* $f_k = \alpha_1 f_{k-2} + \alpha_2 f_{k-1} + \alpha_3$, *along with the constraints*

$$-1 < \alpha_1 < 0,\ 0 < \alpha_2 < 2,\ \alpha_2^2 + 4\alpha_1 > 0,\ \alpha_1 + \alpha_2 < 1 \text{ and } a_1 \geq 0,\ a_2 \geq 0,$$

where a_i *are expressed in terms of the parameters* α_i, f_j *via (12).* □

Similarly, the increasing-decreasing case is covered by the following

Lemma 2. *The following are equivalent:*

1. f_k *of the form (9) is increasing-decreasing and nonnegative on the range* $k = 0, 1, \ldots$ *with* $\lambda_1 \neq \lambda_2$.
2. f_k *satisfies the difference equation (10) along with the constraints*

$$-1 < \alpha_1 < 0,\ 0 < \alpha_2 < 2,\ \alpha_2^2 + 4\alpha_1 > 0,\ \alpha_1 + \alpha_2 < 1 \text{ and } a_1 + a_2 \geq 0,\ a_2 \geq 0,$$

where the a_i *are expressed in terms of* α_i, f_j *via (12).* □

Let us now show in which way the idea of representing exponentials by difference equations may be exploited algorithmically. We start with the case (i) where we expect decaying activity. We replace program (NL) by the following

$$(NL_1)\qquad \begin{array}{ll} \text{minimize} & \ell(f,\alpha) = \|R[\mu]f - y\|^2 + I[f,\alpha] \\ \text{subject to} & f_{i1} \geq f_{i2} \geq \cdots \geq f_{iS} \geq 0, \\ & f_{ik} - \alpha_{1i} f_{i,k-2} - \alpha_{2i} f_{i,k-1} - \alpha_{3i} = 0, \\ & \text{for} \quad i = 1, \ldots, N, \quad k = 3, \ldots, S. \end{array}$$

with unknown variables $f = (f_{ik})$ and $\alpha = (\alpha_{\nu i})$, and where the regularizing term $I[f,\alpha]$ may involve both variables. The linear inequality constraints, which in tandem with the difference equation will yield a decaying positive exponential (9) will be referred to as *shape constraints*. These constraints form a cocktail of the conditions we have isolated in Lemma 1, which are easier to handle than the nonlinear constraints on α in condition 2 of Lemma 1. As we shall see, the shape constraints turn out extremely useful during the initial period of the minimization, but become more and more redundant as the iteration approaches a local minimum.

Before discussing this model any further, let us catch up with more general dynamic profiles. The most interesting case is where we expect increasing-decreasing curves. Here the shape constraints are more delicate to calibrate, since we do not, in general, know the exact position of the peak activity in each dixel (dixel: short for dynamic pixel). We propose the program

$$(NL_2)\qquad \begin{array}{ll} \text{minimize} & \ell(f,\alpha) = \|R(\mu)f - y\|^2 + I[f,\alpha] \\ \text{subject to} & 0 \leq f_{i1} \leq f_{i2} \leq \cdots \leq f_{i,m_i-\nu_i}, \\ & f_{i,m_i+\nu_i} \geq \cdots \geq f_{iS} \geq 0, \\ & f_{ik} - \alpha_{1i} f_{i,k-2} - \alpha_{2i} f_{i,k-1} - \alpha_{3i} = 0, \\ & \text{for} \quad i = 1, \ldots, N, \quad k = 3, \ldots, S. \end{array}$$

which assumes a peak activity in dixel i among the positions $k \in [m_i - \nu_i, m_i + \nu_i]$, with m_i and ν_i determined beforehand, say from prior knowledge or by inspecting the projection data. A second possibility, to constrain an increasing-decreasing profile with the peak position assumed among four previously assigned positions, was used in [**22**], and is considered again in the experimental section.

5. Augmented Lagrangian Method

Our first algorithmic approach to (NL) is an augmented Lagrangian method. We consider the partially augmented objective function

$$\ell_c(f,\alpha) := \ell(f,\alpha) + \frac{c}{2}\sum_{i=1}^{N}\sum_{k=3}^{S} e_{ik}(f,\alpha)^2 = \ell(f,\alpha) + \frac{c}{2}\,e(f,\alpha)^T e(f,\alpha), \tag{13}$$

where the e_{ik} are the equality constraints in (NL),

$$e_{ik}(f,\alpha) = f_{ik} - \alpha_{1i} f_{i,k-2} - \alpha_{i2} f_{i,k-1} - \alpha_{i3} = 0, \tag{14}$$

and e stands for the vector of e_{ik}. We do *not* include the shape constraints into (13). Due to their linear structure, we prefer to carry them explicitly through the approach. The augmented Lagrangian associated with (13) is then

$$\mathcal{L}_c(f,\alpha;\lambda,\mu) = \ell_c(f,\alpha) + \lambda^T e(f,\alpha) + \mu^T(Af - b), \tag{15}$$

where $Af \le b$ represents the shape constraints. This suggests the following algorithm:

Algorithm 1

1. Initialize the multiplier estimates λ and $\mu \ge 0$, and the penalty parameter $c > 0$.
2. Given multiplier estimates λ and $\mu \ge 0$, and the penalty parameter c, solve

$$(P) \quad \begin{array}{ll} \text{minimize} & \mathcal{L}_c(f,\alpha;\lambda,\mu) \\ \text{subject to} & Af \le b \end{array}$$

   The solution is (f^+,α^+). Let μ^+ the Lagrange multiplier belonging to the inequality constraint in (P).
3. Update the multiplier estimate through $\lambda^+ = \lambda + c\,e(f^+,\alpha^+)$.
4. Update $c^+ \ge c$.
5. Return to step 2 until convergence.

This scheme is known to have local superlinear convergence if the penalty parameter sequence c_k it generates tends to infinity (cf. [**2**]), which is in practice prohibitive due to ill-conditioning.

Fixing c at a sufficiently large value $\bar{c}$ still ensures local linear convergence (cf. [**2**]). In practice, the correct value $\bar{c}$ is not known, and in order to guarantee numerical stability of the algorithm, after some updates we usually fix c at a moderate size. This means that the solution so obtained, although practically useful, will *not* fully match the difference equation (10). We may therefore consider Algorithm 1 as a relaxation approach, where the regularizer $I[f,\alpha]$ includes, possibly among other terms, the term $(c/2)\,e(f,\alpha)^T e(f,\alpha)$.

We consider a variation of the augmented Lagrangian approach which divides step 2 of the above Algorithm 1 into two steps, a minimization over f, followed by a minimization over α. Given the current estimates f, α, λ and $\mu \ge 0$ as well as the constant $c > 0$, we obtain the update f^+ and the new multiplier $\mu^+ \ge 0$ by minimizing $\mathcal{L}_c(\cdot,\alpha;\lambda,\mu)$ subject to the shape constraints. In a second step, we obtain the new estimate α^+ by solving the problem

$$\begin{array}{ll} \text{minimize} & \displaystyle\sum_{i,k}(f^+_{ik} - \alpha_{i1} f^+_{i,k-2} - \alpha_{i2} f^+_{i,k-1} - \alpha_{i3})^2 \\ \text{subject to} & -1 \le \alpha_{i1} \le 0,\ 0 \le \alpha_{i2} \le 2,\ \alpha_{i2}^2 + 4\alpha_{i1} \ge 0,\ \alpha_{i1} + \alpha_{i2} \le 1. \end{array}$$

The multiplier λ^+ is updated as before, as is the constant c^+.

This method is algorithmically easier than optimizing simultaneously in (f, α). Using the constraints from condition 2 in Lemma 1 resp. Lemma 2 is possible here since a posteriori fitting, that is finding α^+ with f^+ given, leads to N problems of size S, and is therefore algorithmically simpler. The drawback of this method is that its convergence properties are not clear. Alternating schemes converge at best with a linear rate, but worse, may even fail to converge. We therefore use this scheme during the initial stages of the optimization only.

6. Lagrangian SQP Method

The canonical way to deal with the equality constraints is the Lagrangian SQP method, which consists in applying the sequential quadratic programming (SQP) method to the augmented objective function (13), cf. [**2**].

Given a current primal estimate (f, α) and Lagrange multiplier estimates λ, and $\mu \geq 0$, the SQP method generates new iterates $(f^+, \alpha^+) = (f, \alpha) + (\Delta f, \Delta\alpha)$ by solving the quadratic program

$$(QP) \qquad \begin{array}{ll} \text{minimize} & q_{f,\alpha}(\Delta f, \Delta\alpha) \\ \text{subject to} & A(f + \Delta f) \leq b \\ & M\Delta f + N\Delta\alpha + d = 0 \end{array}$$

choosing as the new multiplier estimates λ^+ and $\mu^+ \geq 0$ the Lagrange multipliers belonging to the linearized constraints in (QP). Here $q_{f,\alpha}$ is the quadratic functional

$$q_{f,\alpha}(\Delta f, \Delta\alpha) = \nabla_f \ell_c(f,\alpha)^T \Delta f + \nabla_\alpha \ell_c(f,\alpha)^T \Delta\alpha + \frac{1}{2}(\Delta f, \Delta\alpha)\nabla^2 \mathcal{L}_c(f, \alpha; \lambda, \mu)(\Delta f, \Delta\alpha)^T$$

and $\mathcal{L}_c(f, \alpha; \lambda, \mu)$ is the augmented Lagrangian (15). Moreover, the (i,k)th line of the linearized equality constraint $M\Delta f + N\Delta\alpha + d = 0$ is the linearization of (14) about the current (f, α), that is:

$$\begin{aligned} f_{ik} + \Delta f_{ik} - \alpha_{1i} f_{i,k-2} - \alpha_{1i}\Delta f_{i,k-2} - \Delta\alpha_{1,i} f_{i,k-2} - \alpha_{i2} f_{i,k-1} \\ -\alpha_{i2}\Delta f_{i,k-1} - \Delta\alpha_{i2} f_{i,k-1} - \alpha_{i3} - \Delta\alpha_{i3} = 0. \end{aligned}$$

While having fast local convergence properties (cf. [**2**]), the SQP scheme is certainly demanding in large size applications, as the number of equality constraints is important. We therefore propose a relaxation which bypasses this difficulty.

7. Relaxation

We observe a particular block structure in the equality constraints, which makes it possible to explicitly calculate the least squares solution

$$\Delta\alpha = -(N^T N)^{-1} N^T (M\Delta f + d) =: W\Delta f + v \tag{16}$$

of the linear equality constraints, and to substitute (16) into the objective $q_{f,\alpha}$. This suggests the reduced quadratic program

$$(QP') \qquad \begin{array}{ll} \text{minimize} & q_f(\Delta f) := q_{f,\alpha}(\Delta f, W\Delta f + v) + \frac{k}{2}||M\Delta f + N(W\Delta f + v) + d||^2 \\ \text{subject to} & A(f + \Delta f) \leq b \end{array}$$

where the new penalty term $(k/2)||M\Delta f + N(W\Delta f + v) + d||^2$ is intended to keep the feasible variables Δf in (QP') close to producing feasible pairs $(\Delta f, W\Delta f + v)$ for (QP), which we aim at since (QP) presents the correct Newton step ensuring fast local convergence. Since (QP'), nonetheless, allows elements Δf which are *not*

admitted in (QP), the problems are not equivalent, and we may at the end expect solutions which do not fully match the equality constraints. Notice that

$$M\Delta f + N(W\Delta f + v) + d = (I - N(N^TN)^{-1}N^T)(M\Delta f + d),$$

with $\Pi := I - N(N^TN)^{-1}N^T$ the projection onto $\ker(N^T)$, so $\ker(\Pi) = R(N)$. Consequently, the penalty term in q_f attributes a high cost to a mismatch in $M\Delta f + N\Delta\alpha + d = 0$, and (QP') may be expected to be close to the original quadratic program (QP).

While the substitution (16) gives a more convenient quadratic subproblem (QP'), no longer involving equality constraints, we recall that ironically we still have to maintain a Lagrange multiplier estimate λ for the linearized equality constraints we have dispensed with. In fact, the terms in $q_{f,\alpha}$ including $L_{f\alpha} = \nabla^2_{f\alpha}\mathcal{L}_c$ require the estimate λ. Due to linearity, this is not the case for the multiplier estimate μ, so maintaining an update of μ is not mandatory here.

The Kuhn-Tucker conditions to (QP') guarantee a multiplier $\nu \geq 0$ such that with the notations $L_{ff} := \nabla^2_{ff}\mathcal{L}(f, \alpha; \lambda, \mu)$, etc.,

$$\begin{array}{ll} (a) & \nabla_f \ell_c + W^T\nabla_\alpha \ell_c + L_{ff}\Delta f + L_{f\alpha}W\Delta f + L_{f\alpha}v + W^T L^T_{f\alpha}\Delta f \\ & \quad + W^T L_{\alpha\alpha}(W\Delta f + v) + A^T\nu + k\,\Pi(\Delta f + Nv + d) \;=\; 0 \\ (b) & A(f + \Delta f) \;\leq\; b, \quad \nu^T\big(A(f + \Delta f) - b\big) = 0 \end{array} \tag{17}$$

are satisfied. These have to be matched with the Kuhn-Tucker conditions for the original program (QP), with (16) substituted for $\Delta\alpha$:

$$\begin{array}{ll} (i) & \nabla_f \ell_c + L_{ff}\Delta f + L_{f\alpha}(W\Delta f + v) + A^T\mu^+ + M^T\lambda^+ = 0 \\ (ii) & \nabla_\alpha \ell_c + L^T_{f\alpha}\Delta f + L_{\alpha\alpha}\Delta\alpha + N^T\lambda^+ = 0 \\ (iii) & A(f + \Delta f) \leq b, \quad \mu^+(A(f + \Delta f) - b) = 0, \quad \mu^+ \geq 0 \\ (iv) & M\Delta f + N(W\Delta f + v) + d = 0 \end{array} \tag{18}$$

Matching (a) of (17) with (i) of (18) implies

$$\begin{array}{l} M^T\lambda^+ + A^T\mu^+ = \\ A^T\nu + W^T\big(\nabla_\alpha \ell_c + L^T_{f\alpha}\Delta f + L_{\alpha\alpha}(W\Delta f + v)\big) + k\Pi(\Delta f + Nv + d), \end{array} \tag{19}$$

while matching (a) of (17) with (ii) of (18), the latter pre-multiplied by W^T, leads to

$$\begin{array}{l} W^TN^T\lambda^+ = \\ A^T\nu + \nabla_f \ell_c + L_{ff}\Delta f + L_{f\alpha}(W\Delta f + v) + k\,W^T\Pi(\Delta f + Nv + d). \end{array} \tag{20}$$

While (19), (20) as they stand would seem difficult to satisfy due to the condition $\mu^+ \geq 0$, $\nu \geq 0$, we may heuristically assume that $\mu^+ = \nu$, as this would be the case if the solution Δf to (QP') matched the constraint $M\Delta f + N\Delta\alpha + d = 0$ exactly, producing a pair $(\Delta f, W\Delta f + v)$ admissible in (QP). With $\mu^+ = \nu$, (19), (20) are conveniently solved for λ^+, possibly in the least squares sense using the same block structure of the constraints. An alternative would be to estimate λ^+ only on the basis of (19), as this is even an explicit update not involving any inversion.

Resuming the above observations, we propose the following algorithm:

Algorithm 2

1. Initialize f, α, λ, and c. Fix $k > 0$.
2. Given the current f, α, λ and c, calculate a step Δf by solving the quadratic program

$$(QP) \qquad \begin{array}{ll} \text{minimize} & q_f(\Delta f) \\ \text{subject to} & A(f + \Delta f) \leq b \end{array}$$

Obtain $\Delta\alpha$ through (16).
3. Calculate a step $\Delta\lambda$ by estimating λ^+ through the formulae (19) and (20).
4. Do a line search with an appropriate merit function to find $t > 0$ and $f^+ = f + t\Delta f$, $\alpha^+ = \alpha + t\Delta\alpha$ and $\lambda^+ = \lambda + t\Delta\lambda$.
5. Update $c^+ \geq c$.
6. Return to step 2 until convergence.

We recall that if the correct SQP method using (QP) was used, the second order sufficient optimality conditions at a local minimum would imply fast local convergence with the penalty parameter c fixed at a sufficiently large value (cf. [**2**]). This suggests the same strategy for the relaxed version and fixing c at a moderate value. Since the correct c is not known, we propose to initially use an update rule like in [**2**] until a satisfactory value of c has been found.

We mention that a plausible idea is to apply the same least squares relaxation to the original program (NL) directly, that is, eliminate the unknown variable α in (NL) by calculating for every i the least squares solution $(\alpha_{1i}, \alpha_{2i}, \alpha_{3i})$ to the set of equations $e_{i3} = 0, \ldots, e_{iS} = 0$. This is, however, less convenient than the corresponding substitution in the linearized setting, since the functional relationship $\alpha = \alpha(f)$ so obtained is more complicated and renders the gradient and Hessian calculation for the reduced problem difficult.

8. Practical Aspects

Designing a correct compartmental model for a specific situation is usually difficult due to the high complexity of metabolism. As a consequence, reconstructing a dynamic tomographic image with a particular tracer dynamic model imposed may bear the risk to miss or suppress dynamic phenomena apparent in the unknown source, but not included in the chosen compartment model. This risk would be less severe if we were able to reconstruct the dynamic image *before* doing the exponential fit. In Positron Emission Tomography (PET) this strategy *is* valid, since the movie may be reconstructed frame by frame. As we have indicated, this is not possible in SPECT if rotating camera systems are used, because the number of views we could possibly scan during a fixed state of the unknown dynamic source is too small. Nonetheless, with specially equipped triple head cameras one can mimic the situation in PET by doing *fast rotations* which collect sufficiently many views at a single dynamic state of the source, cf. [**15, 25**]. This method suffers from poor data statistics, and we claim that with our present techniques we can improve, doing *slow* rotations.

The approach discussed in [**22**] - similar to the situation in PET - reconstructs a movie prior to exponential fitting by imposing shape constraints only. This seems a crude procedure, but in many cases produces good results, in particular for double and triple head cameras (cf. [**13**]). What we presented here is a more sophisticated approach, which again reconstructs a dynamic image f prior to the definitive exponential fitting, now aiming at reconstructions which are already close to the exponential form (5). Using shape constraints in tandem with the augmented Lagrangian type penalty terms encourages doubly exponential fits.

A second practical aspect is CPU time, which should not be exceedingly large if the method is to be clinically useful. This means that we may perform only a limited number of SQP or augmented Lagrangian steps, probably never reaching

any neighborhood of fast local convergence of the method. The approximate solution so obtained may then still be improved by doing an a posteriori fit of the chosen exponential model (5), a procedure which may be repeated several times, as it is fast in comparison to the reconstruction process.

These practical aspects emphasize the idea of regularization versus exact fitting, and one may ask about other possible regularizers for the dynamic emission tomography problem. Reconstructing a movie requires regularization in space and in time. For simplicity we consider regularizers of the form $I[f] = I_s[f] + I_t[f]$, where I_s is a spatial regularizer, I_t a regularizer in time.

The classical choice among the spatial regularizers is of course the Tychonoff functional $I_s[f] = (k/2)||\nabla f||^2$, and its offsprings, which we may apply frame by frame. Following [**24**], we prefer a Fourier domain regularizer whose philosophy is based on the Fourier slice theorem for the unattenuated Radon transform: $\hat{f}(\sigma\omega) = (\mathcal{R}f)\hat{}(\sigma,\omega)$. This means that the expected signal f could not possibly have a better spatial resolution than each of the projections. Having decided on a spatial bandwidth b in the s-variable, and after choosing a high-pass filter $\mathcal{H}_b$ with the corresponding cutoff frequency, the spatially regularizing term should be chosen as

$$I_s[f] = (k/2)||\mathcal{H}_b f||^2.$$

This filter could be applied either to each frame, or to a subset of frames. The approach should still be qualitatively correct if the transform is attenuated, even though the Fourier slice theorem is no longer true in this case.

Regularizing in the time domain could be done in several ways. A Fourier regularizer could be designed in much the same way as for spatial smoothing. The corresponding dynamic cutoff frequency could be obtained by a prior Fourier analysis of the expected exponential curves (5), as soon as realistic bounds on the eigenvalues λ_i are known.

However, a particular difficulty with Fourier filtering in the time domain, already analyzed in [**23**], is the fact that most practical time profiles are *not* continuous in the sense of the space $C_{\text{per}}[0,T]$, and Fourier filtering would therefore introduce unwanted smoothing at the early and late views. Ways to overcome this have been proposed in [**23**] and could be adapted to build an appropriate temporal high pass filter.

We conclude by presenting a dynamic regularizer based on a different philosophy, motivated by the dynamic heart-in-thorax phantom built at the Vancouver Hospital (cf. [**7**]). The phantom heart consists of several cylindrical containers with a drain and an inflow. Each container is equipped with a mixing propeller to guarantee a homogeneous flow. Tuning the motor which steers the pump could now, at least in principle, produce any possible activity profile in the containers. Consider for instance the situation of washout, where the container is initially filled with activity $f(0)$. Pumping fresh water at flow rate $R(t)$ means that the remaining activity in the container of volume V at time t is

$$f(t) \;=\; f(0)\,\exp\Big\{-\int_0^t \frac{R(\tau)}{V}\,d\tau\Big\}.$$

The idea is now to avoid highly irregular curves $f(t)$ by penalizing irregular behavior of the motor. And in a sense the most natural profile is a flow at constant rate, that is, $R'(t) = 0$. In terms of the curve f, this leads to $R'/V = -(f'/f)' =$

$-(\log f)'' = 0$, so in the continuous model the dynamic regularizer could be chosen as

$$I_t[f] = \int_0^T \left((\log f(t))''\right)^2 dt.$$

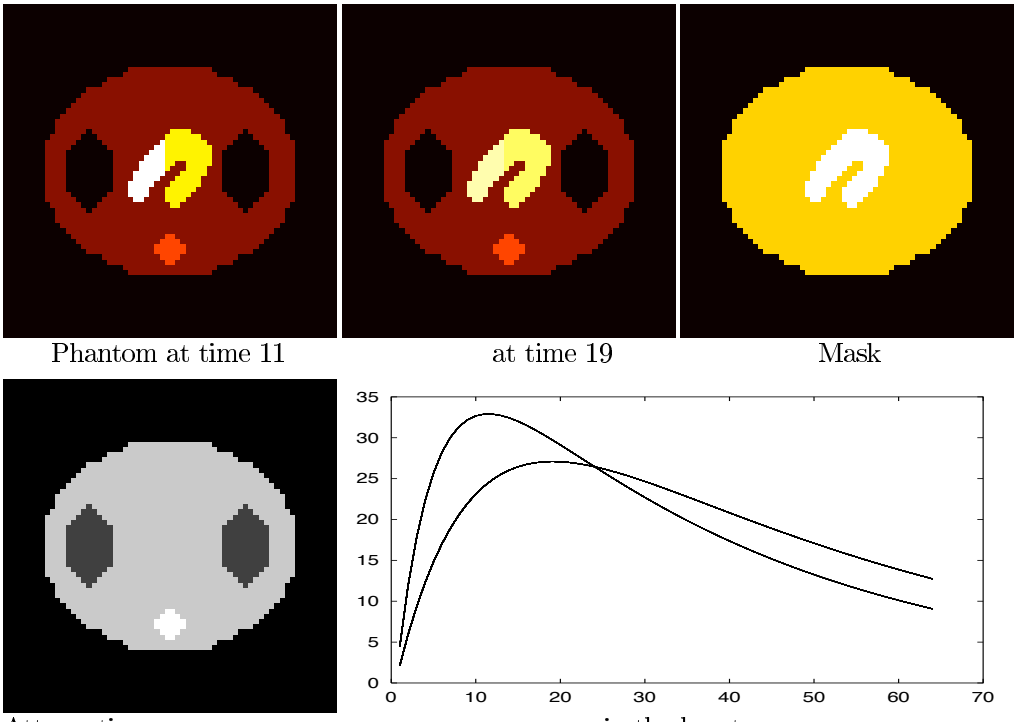
Discretizing the time step and using finite differences to replace time derivatives leads to

$$I_t[f] = \sum_{k=2}^{S-1} (\log f_{k+1} + \log f_{k-1} - 2\log f_k)^2 = \sum_{k=2}^{S-1} \log^2 \left(\frac{f_{k+1} f_{k-1}}{f_k^2}\right).$$

The ultimate link with tomography is the hypothesis that each cell, in a very oversimplified way, acts similar to the macroscopic containers of the phantom heart.

9. Simulation

The following simulation uses a dynamic heart-in-thorax phantom, displayed in Figure 3a and 3b. The scan assumes 64 time steps, and in line 1 the phantom is displayed at times 11 and 19. The heart has been segmented into two dynamic zones exhibiting different up-take and washout profiles, simulating a stenosis in the left part of the muscle. The other sections of the body contour are assumed static. The ideal time profiles for the two zones are shown in Figure 3a, line 2 (right). From an experimental standpoint, the question is that of knowing whether the different dynamic modes are recognized by our inverse method for ideal and noisy data.

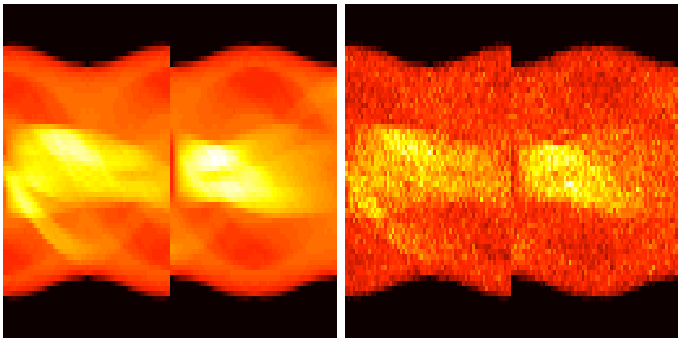


Phantom at time 11 at time 19 Mask

Attenuation map curves in the heart

Figure 3a. Simulated phantom and attenuation map.

The data acquisition simulates a rotating double head camera with two heads at 90^0, each head rotating over a 90^0 sector, with a total of 180^0 scanned. Each camera head takes 32 angular positions. The ideal sinogram is displayed in Figure 3b (left), while the right hand image shows the data rendered Poisson distributed. Notice that in a static scan, the sinograms of the two heads could usually be assembled to a single sinogram of twice as many steps for one head. Due to changing activity, this is no longer possible in dSPECT, and we emphasis this fact by presenting the two sinograms in reversed order. We assume a realistic tissue attenuation map, displayed in Figure 3a, line 2 (left).



Sinogram Noised sinogram

Figure 3b. Sinograms. The two camera heads are placed at 90^0.

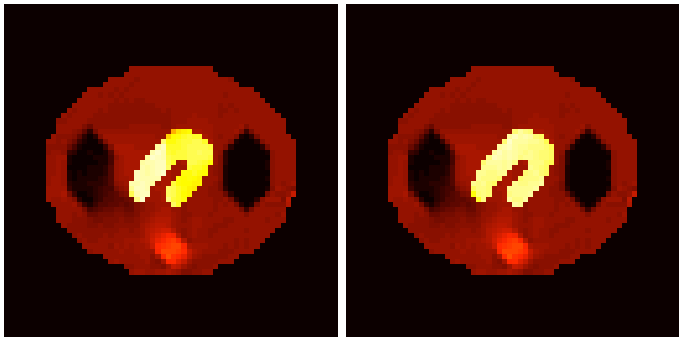
The reconstruction process now starts with a preliminary phase which serves three different purposes. A first reconstruction (a) is obtained by imposing decreasing activities only, a second one (b) imposing increasing activities only. While (a) will show a flat part at the early views, with decay at the later views, (b) will have the opposite form. Matching these reconstructions will provide a first guess for the peak activity position in each pixel.

A third reconstruction (c) is now produced with the assumed peak positions in each pixel, and the reconstruction is used to identify the body contour. Reconstruction (c) also serve the purpose of identifying pixels which are apparently static, and rendering these dixels static reduces the total number of unknown parameters to fit. The resulting mask is shown in Figure 3a, line 1 (right), assuming pixels outside the phantom heart as static. Obtaining the mask may be more difficult in a clinical scan, and some practical issues of this phase, in particular, its automatization, will be discussed elsewhere.

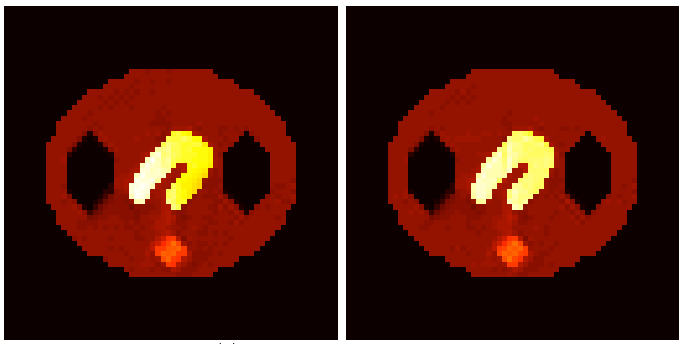
A fourth reconstruction (d) is now obtained on the basis of the previously obtained mask, and is used to initialize the Prony type algorithm. This phase of the reconstruction process uses version (NL_2), either without regularizing term, or with some of the regularizers discussed in section 8. A definitive conclusion on what regularizer behaves best at step (d) has not been drawn as yet, and its choice may depend on the situation.

Reconstruction (d), displayed at times 11 and 19 in Figure 4, line 1, is already of satisfactory visual quality, but the dynamic profiles shown in Figure 5, line 1, are not sufficiently closed to the sought for biexponential form. Step (e) of the procedure now does an a posteriori curve fitting of a biexponential on a pixel-by-pixel basis, translating the model (9) to the form (10), or directly uses the method

from section 5. This step is again critical, since some of the reconstructed profiles in (d) may be far from biexponentials and may lead to aberrant parameter estimates. We address this problem by switching to a local fit which uses the weighted sum of 9 neighboring pixels to a given pixel to stabilize.



Reconstruction (d): before Prony at times 11 and 19



Reconstruction (f): after 4 Prony iterations at times 11 and 19

Figure 4. Noise-free reconstruction. The total CPU was 3 minutes.

Estimate (e), translated into the corresponding $(\alpha_{i1}, \alpha_{i2}, \alpha_{i3})$ information is then used as the initial guess to the Prony type procedure, which in our study is the augmented Lagrangian approach in Section 5 with initial $\lambda = 0$. The result is the reconstruction (f), shown in Figure 4, line 2. As we can see, in the noise-free case a few iterations suffice to produce satisfactory curves which even allow separating the ideal time profiles visually. However, the initial stages of the procedure, shown in Figure 5, seem to indicate oscillatory behavior of the method.

The reconstruction may be terminated by a last step (g) doing a second a posteriori fit, eventually after drawing regions of interest for specific purposes. Notice here that in contrast with other approaches, segmenting is *not* used in the reconstruction process itself, but serves for the interpretation of the results. This is an advantage of our method, since segmentation for reconstruction, usually employed in order to reduce the number of unknown variables, is critical as it may oversimplify or introduce unnecessary bias in the reconstructions.

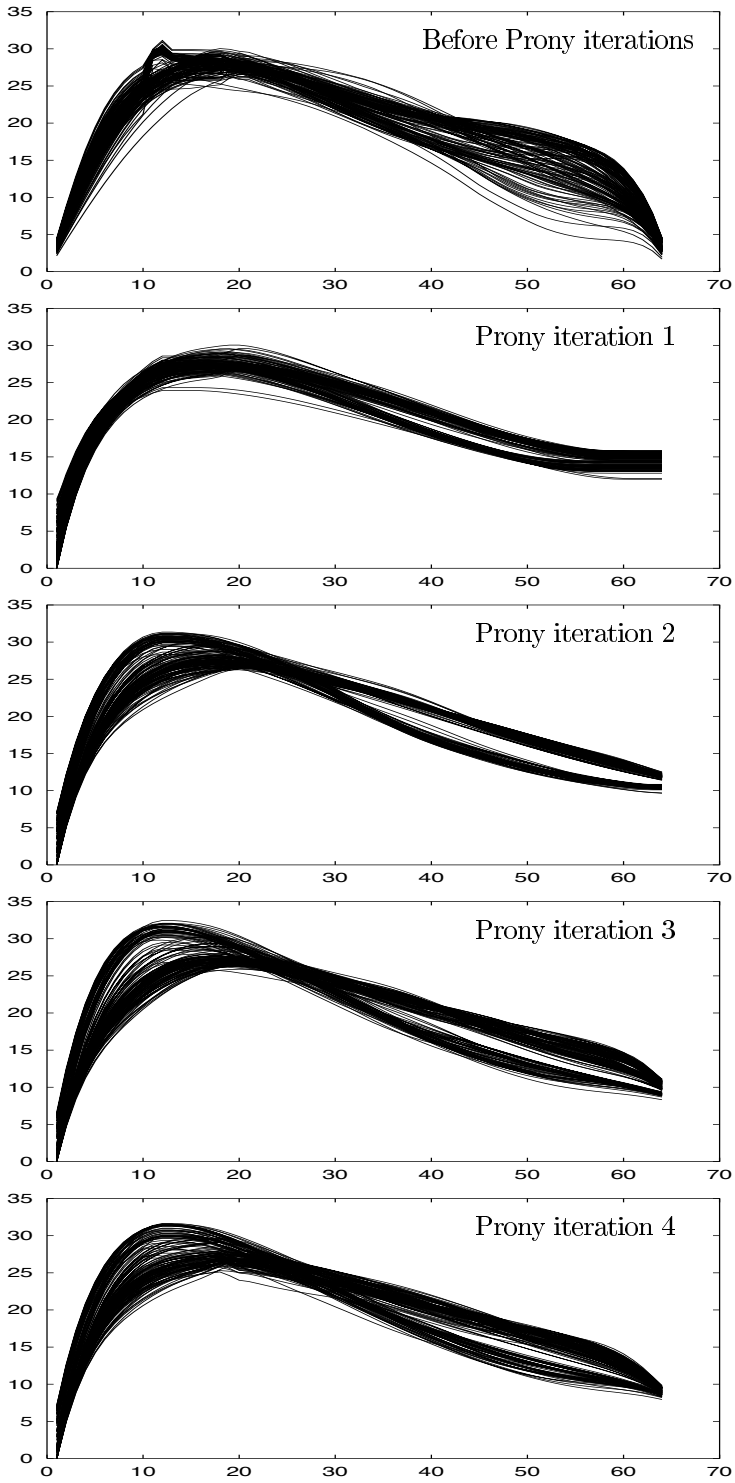
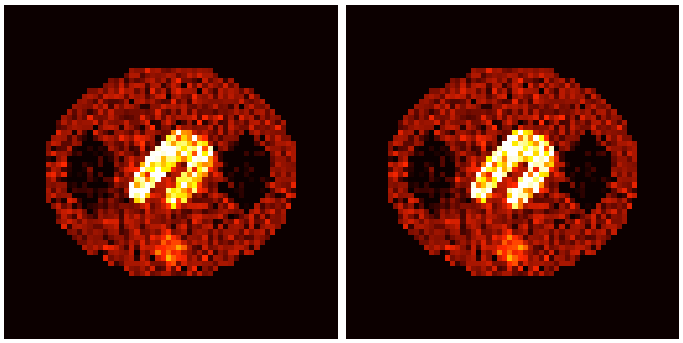
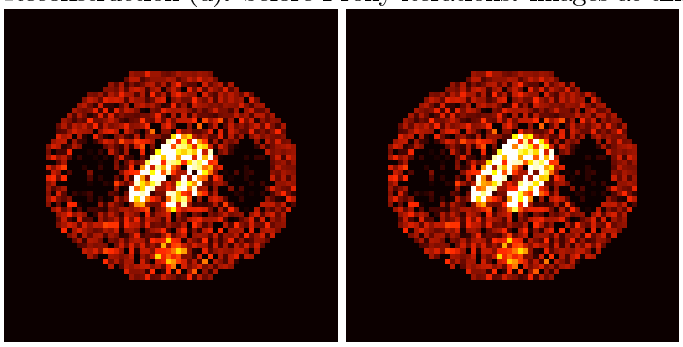


Figure 5. Prony iterates in the heart for the noise free case.

The results in Figure 6 show the same phantom experiment with the noisy sinogram obtained by rendering the ideal data Poisson distributed (Figure 3b, right). Figure 6, lines 1 and 2 show the reconstructed slice, with little to no visual difference between the raw reconstruction and the result obtained after 4 Prony type iterations.



Reconstruction (d): before Prony iterations: images at times 11 and 19



Reconstruction (f): after 4 Prony iterations: images at times 11 and 19

Figure 6. Reconstruction for the noisy case. The total computation time was 3 minutes 30 seconds. Neither the data nor the reconstruction have been filtered.

A visual check of the reconstructed curves as in the noise free case (Figure 5) is no longer possible here. We therefore draw masks for the two parts of the heart we wish to discriminate, shown in Figure 7, line 1. We represent the result for each region by taking the averaged curves. In the left region, this means averaging 46 curves, in the right hand regions 91 curves. The result is displayed in line 3. The correct input is re-displayed in line 2.

As we can see, the reconstruction is capable of representing the cross-over of the two curves apparent in the true image. Equally, the peak positions in each profile are rather accurately represented. The reconstruction is off by about 10% in absolute values. This may not be blamed on attenuation artifacts, as the correct attenuation map has been used in the reconstructions. It is more likely that this

is caused by the imperfections of the reconstruction process itself. A refined estimation disregarding outliers in the reconstructed profiles from each region would probably help to improve the accuracy in absolute values.

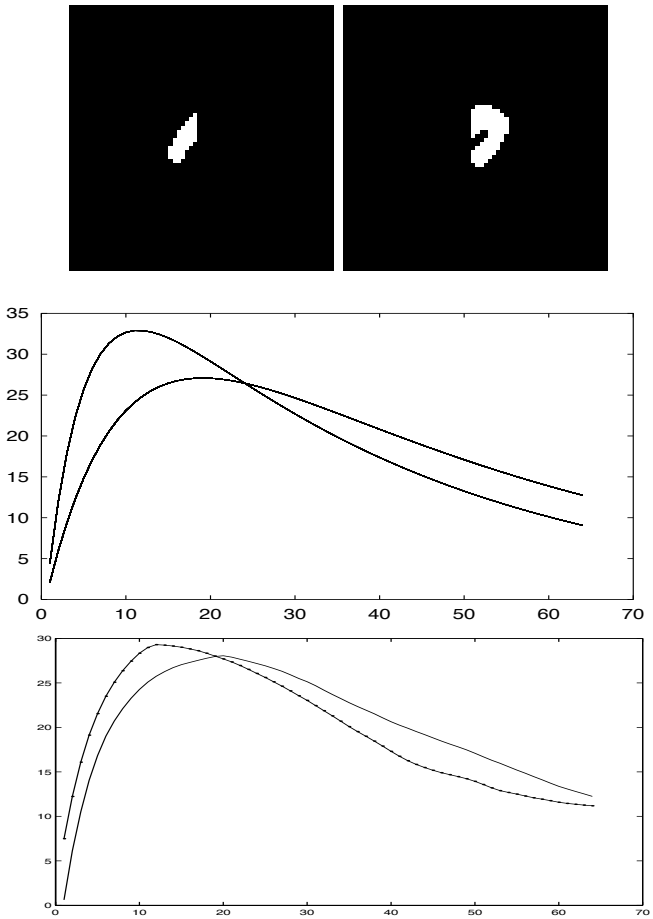


Figure 7. Evaluation of the noisy case. Line 1 shows the masks drawn to distinguish the different dynamic profiles. Line 3 shows the result of the reconstruction after 4 Prony type iterations. We plot the averages of the curves in each region of the heart against each others. Line 2 shows the true curves. The reconstruction is able to represent the cross-over of the curves and the peak activities, but absolute heights around the peaks are off by approximately 10%.

10. Conclusion

We have presented a way of estimating kinetic parameters for compartmental models directly from a dSPECT reconstruction. The approach is an extension of Prony's method and the optimization uses an augmented Lagrangian type procedure. Due to breaking the iteration into two separate steps (section 5), a procedure which accelerates, the method tends to oscillate after some rapid improvements during the early steps. The approach has been validated through a simulated case

study. A more refined scheme (described in section 6 and 7) avoiding the oscillatory behavior has been presented but needs to be validated by means of numerical simulations. Several dynamic regularizers have been successfully used, but a qualitative comparison offering a preference for a specific regularizers has not been obtained as yet.

11. Appendix - Transport Model

Dynamic emission tomography is described by the dynamic photon transport equation (compare [**11**, Vol. I, Section 5]):

$$\begin{aligned} &\frac{1}{c}\frac{\partial}{\partial t}u(x,\omega,E,t)+\omega\cdot\nabla u(x,\omega,E,t)+\mu(x,E)u(x,\omega,E,t)\\ &\qquad=\int_{\mathbb{S}^2}b(x,E',E,\omega',\omega)u(x,\omega',E',t)\,d\omega'+f(x,E,t) \end{aligned} \tag{21}$$

Here $u(x,\omega,E,t)$ is the specific intensity or density of the angular transport of photons at position x, time t, in direction $\omega\in\mathbb{S}^2$, and with energy E; c is the speed of light. $f(x,E,t)$ is the unknown radio active source we wish to estimate, and $\mu(x,E)$ is the linear attenuation coefficient which decomposes as $\mu(x,E)=\rho(x)\,\mu_m(E)$ with $\rho(x)$ the unknown density at position x, and $\mu_m(E)$ a known term representing all physical processes which end flights of photons at energy E. For the energies relevant to nuclear medecine, these comprise photo effect, Rayleigh scattering, and Compton scattering, while pair production becomes relevant at higher photon energies encountered in radio therapy. For energies E between 100keV and 1000keV, photo effect and Rayleigh scatter are almost negligeable, and we may model the scattering kernel as $b(x,E,E',\omega,\omega')=\rho(x)\,\beta(E,\theta)$, where $\theta=\widehat{(\omega,\omega')}$ is the angle between ω and ω', and β is the known Klein-Nishina scattering cross-section (cf. [**12**]). Notice here that due to conservation of energy and momentum in the Compton scatter process, the energy E' of the incident photon scattered from direction ω' into the direction ω with energy E is fixed through the formula

$$E=\frac{E'}{1+(E'/mc^2)(1-\cos\theta)} \tag{22}$$

with $\theta=\widehat{(\omega,\omega')}$ and m the electron mass (cf. [**39**]).

Before discussing (21), it is instructive to see the general form of the dynamic emission tomography problem:

Find an input function $c_B(x,t)$, exchange constants $k_{ij}(x)$, and a mass distribution $\rho(x)$ such that with a dynamic source $f(x,E,t)$ satisfying (ii) and (iii) of $(\mathcal{P})$, the photon transport $u(x,\omega,E,t)$ solving (21) matches the projection data $d(t,x,\omega,E)$ acquired at time t, sensor position x, angular direction ω, and energy E.

This cast of the problem, although not very helpful from the algorithmic point of view, reveals it as a control problem with control vector (c_B,K,ρ). In this context, our regularization approach may be understood as an optimal control technique.

Let us continue our investigation of the transport model. Notice that the energy of photons used in nuclear medicine is often monochromatic, so ideally $f(x,E,t)=f(x,t)\,\delta(E-E_0)$ for a fixed energy level E_0. (For example, $E_0=140$keV for ^{99}Tc used in SPECT, and $E_0=511$keV in PET). Assuming that our camera has a perfect energy resolution, we are capable of recognizing photons which have Compton scattered with angle $\theta\neq 0$ due to the loss of energy (22). Therefore,

for monochromatic sources, the scattering kernel in (21) becomes irrelevant, and on suppressing the dependence of the variables on E_0, we obtain the simplified transport equation

$$\frac{1}{c}\frac{\partial}{\partial t}u(x,\omega,t)+\omega\cdot\nabla u(x,\omega,t)+\mu(x)\,u(x,\omega,t)=f(x,t), \tag{23}$$

with $\omega\in\mathbb{S}^2$, $x\in\mathbb{R}^3$, $t\geq 0$. Clearly this equation decouples into a series of equations on 2D slices.

In cases where the emission source is not monochromatic, the scattering kernel is required. Even for monochromatic sources, scattering is of practical important due to the finite camera resolution which implies that we collect data over an energy window $E_0\pm\Delta E$.

Despite these observations it is customary to ignore the scatter contribution or to interpret it as a source of noise in the data. Introducing the cumulated angular transport

$$u(x,\omega,t)=\int_{E_0-\Delta E}^{E_0+\Delta E}u(x,\omega,E,t)\,dE$$

we obtain, on assuming $\mu(x,E)=\mu(x,E_0)=\mu(x)$ energy independent over $E_0\pm\Delta E$, the same simplified equation (23) in $u(x,\omega,t)$.

Discussing (23) in a slice we parameterize lines by $s\in\mathbb{R}$, and $\omega=(\cos\theta,\sin\theta)\in\mathbb{S}^1$, letting $L(s,\omega)=\{s\omega^{\perp}+t\omega\,:\,t\in\mathbb{R}\}$. Suppose the unknown source and attenuation coefficient are zero outside the unit disk Ω. Then the measured data may be understood as boundary information on Ω. Indeed,

$$u(x_0,\omega,t)=0\qquad\text{for }\ \nu_{x_0}\cdot\omega\leq 0,\quad x_0\in\partial\Omega \tag{24}$$

($\nu_x=$ outer unit normal to Ω), i.e., there is no incoming radiation. And further, the outgoing radioactive transport is known:

$$u(x_1,\omega,t)=:d(t,s,\omega)\quad\text{for }\nu_{x_1}\cdot\omega\geq 0,\quad x_1\in\partial\Omega, \tag{25}$$

where

$$d(t,s,\omega)\;=\;\int_{E_0-\Delta E}^{E_0+\Delta E}d(t,s,\omega,E)\,dE$$

represents the counts collected over the camera energy window $E_0\pm\Delta E$ about the expected energy E_0 of the isotope. Here $L(s,\omega)\cap\Omega=[x_0,x_1]$, and we call $x_0=x_0(s,\omega),x_1=x_1(s,\omega)$ the entry and exit points.

Assuming that the source $f(x,t)$ and data $d(t,s,\omega)$ decay sufficiently fast as $t\to\infty$ we may Fourier transform equation (23). This leads to the family of equations

$$\omega\cdot\nabla U(x,\omega,\sigma)+(\mu(x)+i\sigma/c)U(x,\omega,\sigma)=F(x,\sigma) \tag{26}$$

with boundary information

$$U(x_0,\omega,\sigma)=0,\qquad U(x_1,\omega,\sigma)=D(\sigma,s,\omega). \tag{27}$$

Here capital letters refer to the Fourier transforms of the corresponding lower case functions, and σ stands for the Fourier domain variable. Integrating equation (26)

using (27) on each segment $L(s,\omega)\cap\Omega=[x_0,x_1]$ separately leads to the formula

$$\int_{x_0}^{x_1} F(x,\sigma)\cdot e^{-i\sigma|x_1-x|/c}\cdot e^{-\int_x^{x_1}\mu(y)\,dy}\,dx \;=\; D(s,\omega,\sigma), \tag{28}$$

which is a Fourier domain counterpart of the known integral equation obtained in the case of stationary SPECT based on the attenuated Radon transform (cf. [**27, 26, 28**]). Multiplying (28) by $e^{i\sigma t}$, Fourier inversion then gives the family of integral equations

$$\int_{x_0}^{x_1} f(x,t-|x_1-x|/c)\cdot e^{-\int_x^{x_1}\mu(y)\,dy}\,dx \;=\; d(t,s,\omega), \tag{29}$$

exhibiting a delay between the data acquired at time t and the source terms emitting at times $t-|x_1-x|/c$. The method of characteristics (cf. [**11**]) would of course have produced the same result without using the Fourier transform. Since photons travel with the speed of light, we may in practice assume

$$t-|x_1-x|/c\approx t,$$

which leads to the approximate equation involving the *dynamic attenuated Radon transform* $\mathcal{R}[\mu,f]$:

$$\mathcal{R}[\mu,f(\cdot,t)](s,\omega):=\int_{x_0}^{x_1} f(x,t)\cdot e^{-\int_x^{x_1}\mu(y)\,dy}\,dx \;=\; d(t,s,\omega). \tag{30}$$

Notice that for monochromatic sources, (30) represents the *correct* transport model, while in more general situations it stands for the simplification obtained by setting $b=0$.

For completeness, let us consider the situation in PET. Assuming again that the relevant information is captured by the disk Ω, we are led to use the PET-geometry which parameterizes the lines in the plane via the entry and exit points x_0,x_1 on $\partial\Omega$. The boundary information being symmetric

$$d(x_0,x_1,t)=d(x_1,x_0,t),$$

and with the incoming radiation assumed zero as before, the formula replacing (28) is

$$e^{-\int_{x_0}^{x_1}\mu(y)\,dy}\cdot e^{-is|x_1-x_0|/c}\cdot\int_{x_0}^{x_1} F(x,s)\,dx \;=\; D(x_0,x_1,s),$$

and Fourier inversion provides the following family of integral equations

$$e^{-\int_{x_0}^{x_1}\mu(y)\,dy}\cdot\int_{x_0}^{x_1} f(x,t-|x_1-x_0|/c)\,dx \;=\; d(x_0,x_1,t),$$

showing a similar type of delay between the source and data terms. In the case of PET a simple change of variables leads to the compact formula

$$\exp\{-\mathcal{R}[\mu](x_0,x_1)\}\cdot\mathcal{R}[f(\cdot,t)](x_0,x_1) \;=\; d(x_0,x_1,t+|x_1-x_0|/c) \tag{31}$$

involving the classical Radon transform parametrized through the lines $L(x_0, x_1)$ spanned by points $x_0, x_1 \in \partial\Omega$. Again we may in practice assume $t \pm |x_1 - x_0|/c \approx t$ in (31), leading to

$$\exp\{-\mathcal{R}[\mu](x_0, x_1)\} \cdot \mathcal{R}[f(\cdot, t)](x_0, x_1) = d(x_0, x_1, t) \tag{32}$$

The interpretation of (32) is that it is justified to do dynamics in PET simply by reconstructing a series of static images, as proposed already in the basic reference for probabilistic approaches to PET [**36**].

References

[1] Bauschke HH, Noll D, Celler A, Borwein JM, An EM-algorithm for dynamic SPEC tomography, IEEE Trans. Med. Imag. 18, no. 3, 1999, 252 – 261.

[2] Bertsekas DP, Constrained Optimization and Lagrange multiplier methods, Academic Press, New York, 1982.

[3] Borwein JM, Sun W, The stability analysis of dynamic SPECT systems, Numerische Mathematik 77, 1997, 283 – 298.

[4] Buck A, Westera G, vonSchulthess GK, Burger C, Modeling alternatives for cerebral carbon-11-iomazenil kinetics, J. Nuclear Med. 37, 1996, 699 – 705.

[5] Byrd RH, Lu P, Nocedal J, Zhu C, A limited memory algorithm for bound constrained optimization, SIAM J. on Scientific Computing 16, 1996, 1190 – 1208.

[6] Carson RE, Lange K, The EM parametric image reconstruction algorithm, J. Amer. Stat. Assoc. 80, 1985, 20 – 22.

[7] Celler A, Farncombe T, Harrop R, Lyster D, Dynamic heart-in-thorax phantom for functional SPECT, IEEE Transactions on Nuclear Science 44, 1997, 1600 – 1605.

[8] Celler A, Sitek A, Harrop R, Reconstruction of multiple line source attenuation maps. In: 1996 IEEE Nuclear Science Symposium Conference Record, pp. 1420 – 1424, November 1996.

[9] Celler A, Sitek A, Stoub E, Lyster D, Dykstra C, Worsley D, Fung A, Development of a multiple line source attenuation array for SPECT transmission scans. J. Nuclear Medicine 38, 1997, 215ff.

[10] Chen K, Huang SC, Feng D, New estimation methods that directly use the time accumulated counts in the input function in quantitative dynamic PET studies, Phys. Med. Biol. 39, 1994, 2073 – 2090.

[11] Dautrey R, Lions JL, Analyse mathématique et calcul numérique, tome 9, Masson, Paris 1984.

[12] Dicken V, Simultaneous activity and attenuation reconstruction in emission tomography, to appear in Inverse Problems.

[13] T. Farncombe, A. Celler, D. Noll, J. Maeght, Dynamic SPECT Imaging Using a Single Camera Rotation (dSPECT), IEEE Transactions on Nuclear Science,vol. 46, no. 4, pp.1055-1061, 1999.

[14] Gould NIM, Hribar ME, Nocedal J, On the solution of equality constrained quadratic programming problems arising in optimization. Preprint.

[15] Gullberg GT, Huesman RH, Ross SG, Di Bella EVR, Zeng GL, Reutter BW, Christian PE, Foresti SA, Dynamic cardiac single photon emission computed tomography, In: Nuclear Cardiology: State of the Art and Future Directions. Mosby-Year Book Inc., Philadelphia, PA. In press.

[16] Gjedde A, Compartmental Analysis. In: Principles of Nuclear Medicine, 2nd edition, eds Wagner HNJr, Szabo Z, Buchanan JW. Saunders, Philadelphia, pp. 451 – 461.

[17] Kammler DW, Least squares approximation of completely monotonic functions by sums of exponentials, SIAM J. Numerical Anal. 16, 1979, 801 – 818.

[18] Kao CM, Yap JT, Mukherjee J, Cooper M, Chen CT, An image reconstruction method for dynamic PET, IEEE NSS/MIC 1995 Conference Record.

[19] Limber MA, Limber MN, Celler A, Barney JS, Borwein JM, Direct reconstruction of functional parameters for dynamic SPECT, IEEE Trans. Nuclear Sciences 42, 1995, 1249 – 1256.

[20] Links JM, Frank TL, Becker LC, Effect of differential tracer washout during SPECT acquisition, Journal of Nuclear Medicine vol 32(12) pp 2253-2257, 1991.

[21] Liu DC, Nocedal J, On the limited memory BFGS method for large scale optimization problems, Mathematical Programming 45, 1989, 503 – 528.

[22] Maeght J, Noll D, Methods in dynamic emission tomography,
Technical Report, http://mip.ups-tlse.fr/publi/rapp99/99.26.html
[23] Maeght J, Noll D, Resolution in dynamic emission tomography. To appear in SIAM Journal on Mathematical Analysis.
[24] Maréchal P, Togane D and Celler A, A new reconstruction methodology for computerized tomography: FRECT (Fourier Regularized Computed Tomography), Conference Records of the IEEE Medical Imaging Conference (Seattle, USA), 1999.
[25] Nakajima K, Taki J, Bunko H, Matsudaira M, Muramori A, Matsunari I, Hisada K, Ichihara T, Dynamic acquisition with a three-headed SPECT system: application of Technetium 99m-SQ30217 myocardial imaging, J. Nuclear Medicine 32, 1991, 1273 – 1277.
[26] Natterer F, Computerized tomography with unknown sources, SIAM J. Appl. Math. 43, 1983, 1201 – 1212.
[27] Natterer F, The Mathematics of Computerized Tomography, Teubner Verlag, Stuttgart, 1986.
[28] Natterer F, Determination of tissue attenuation in emission tomography of optically dense media, Inverse Problems 9, 1993, 731 – 736.
[29] O'Connor MK, Cho DS, Rapid radiotracer washout from the heart: Effect on image quality in SPECT performed with a single-headed gamma camera system, J. Nuc. Med., 33(6), 1992, 1146 – 1151.
[30] Oppenheim BE, Krepshaw JD, Dynamic hepatobiliary SPECT: a method for tomography of a changing radioactivity distribution, J. Nuclear Medicine 29, 1988, 98 – 102.
[31] Osborne MR, Smythe GK, A modified prony algorithm for exponential function fitting, SIAM J. Sci. Comput. 16, 1995, 119 – 138.
[32] Ruhe A, Fitting empirical data by positive sums of exponentials, SIAM J. Sci. Stat. Comp. 1, 1980, 481 – 498.
[33] Seber GAF, Wild, CJ, Nonlinear Regression, Wiley and Sons, 1988.
[34] van den Hoff J, Burchert W, Wolpers HG, Meyer GJ, Hundshagen H, A kinetic model for cardiac PET with [1-Carbon-11]-acetate, J. Nuclear Medicine 37, 1996, 521 – 529.
[35] Varah, JM, On fitting exponentials by nonlinear least squares, SIAM J. Sci. Stat. Comput. 6, 1985, 30 – 44.
[36] Vardi Y, Shepp LA, Kaufman L, A statistical model for Positron Emission Tomography, J. Amer. Stat. Assoc. 80, 1985, 8 – 20.
[37] Welch A, Clark R, Natterer F, Gullberg GT, Toward accurate attenuation correction in SPECT without transmission measurements, IEEE Trans. Med. Imag., 16(5), 1997, 532 – 541.
[38] Wells RG, Celler A, Harrop R, Analytic calculation of photon distribution in SPECT projections, IEEE Trans. Nuclear Sci.
[39] Young T, Attenuation correction in SPECT by simultanious reconstruction of emission and intrinsic attenuation information using an ART algorithm. PhD thesis, Graduate School of Syracuse University, 1995.
[40] Zhao S, Welland G, Wang G, Wavelet sampling and localization schemes for the Radon transform in two dimensions, SIAM J. Appl. Math. 57, no. 6, 1997, 1749 – 1762.

Université Paul Sabatier, Mathématiques pour l'Industrie et la Physique, 118 route de Narbonne, 31062 Toulouse, France
E-mail address: maeght@mip.ups-tlse.fr (Jean Maeght)

Université Paul Sabatier, Mathématiques pour l'Industrie et la Physique, 118 route de Narbonne, 31062 Toulouse, France
E-mail address: noll@dumbo.ups-tlse.fr (Dominikus Noll)

Stanford University, Department of Electrical Engineering, Information Systems Laboratory, Stanford, CA, USA
E-mail address: boyd@ISL.Stanford.EDU (Stephen Boyd)