International Journal of Pure and Applied Mathematics
Volume 73 No. 4 2011, 405-434

SMART FILTER FOR DYNAMIC SPECT IMAGE RECONSTRUCTION

Joe Qranfal[1] §, Charles Byrne[2]

[1]Department of Mathematics
Simon Fraser University
British Columbia, CANADA
[2]Department of Mathematical Sciences
University of Massachusetts at Lowell
Lowellm, USA

Abstract: We present a new filtering algorithm, the SMART filter (simultaneous multiplicative algebraic reconstruction technique) and provide a convergence result. We test it to solve the inverse problem of reconstructing a dynamic medical image where the signal strength changes substantially over the time required for data acquisition. Our test choice is the time-dependent single photon emission computed tomography (SPECT) which is an ill-posed inverse problem. Based on a linear state-space model of the problem, we provide numerical results to corroborate the effectiveness of our reconstruction method. The SMART filter guarantees a nonnegative and temporally regularized solution, filters out errors from modeling the dynamical system as well as the noise from the data, and outputs an optimal recursive estimate. The SMART filter proves itself to be also computationally time efficient which makes it very suitable for large scale systems such as the ones in medical imaging. In addition, it could be used in any discipline which has used, for instance Kalman filter, or in any one that is interested in time-varying variables such as financial risk assesment/evaluation and forecasting, tracking, or control. Tests in both cases, underdetermined and overdetermined, confirm the convergence result. Getting much better results in the latter case supports the fact that the more information we feed the SMART filter the better it behaves.

AMS Subject Classification: 93E11, 93E10, 34K29, 49N45, 60G35, 62G05, 62M05, 68U10, 94A08, 90C25
Key Words: estimation, stochastic filtering, Kalman filter, optimal filter-

Received: May 31, 2011

§Correspondence author

ing, state estimator, convex optimization, medical image, dynamic SPECT, cross-entropy, nonnegative reconstruction, hidden Markov model, algebraic reconstruction, temporal regularization

1. Introduction

We introduce a new filtering algorithm to find a nonnegative estimate $\hat{x}_k$, $k = 1, \ldots, S$, to the nonnegative unknown x_k of the problem given by the two linear space-state equations,

$$\begin{aligned} x_k &= A_k x_{k-1} + \mu_k \\ z_k &= H_k x_k + \nu_k \end{aligned}$$

μ_k is the error vector, $\mathbb{E}(\mu_k) = 0$ and $\mathbb{E}(\mu_k \mu_k^\top) = Q_k$ is the covariance of the error in modeling the transition from x_{k-1} to x_k. $\mathbb{E}(\nu_k) = 0$ and $\mathbb{E}(\nu_k \nu_k^\top) = R_k$ are the mean and covariance respectively of the noise vector ν_k. Entries of the vector z_k and of both matrices A_k and H_k are nonnegative. We know also that we deal with white error and noise so that,

$$\begin{aligned} Q_k &= \sigma_k^2 I \\ R_k &= diag(z_k) \end{aligned}$$

where I is the identity matrix of order N, $diag(a)$ denotes the square matrix that has the a_i in its main diagonal and 0 otherwise. Our new algorithm that we refer as the *SMART filter* is then numerically tested to solve a reconstruction problem arising in medical imaging, namely dynamic/time-varying SPECT.

Standard SPECT imaging assumes that the distribution of the radioactive tracer is stationary or remains constant during the whole time required for data acquisition. However, nuclear medicine is also interested in investigations of the dynamics of the human body's physiological processes and biochemical function. In this case, the distribution of the radiopharmaceutical (such as for example ^{99m}Tc-Teboroxime which may be used for cardiac imaging) will change over time. In any standard rotating SPECT camera, the projections required for reconstruction of a single image are collected sequentially. But if the concentration of radiotracer changes, then these projections, taken at different times during camera rotation, correspond to different distributions of tracer. Fast changes of activity occurring during SPECT acquisition create so-called "inconsistent" projections which, when processed with standard reconstruction

methods, result in serious image artifacts. Subsequently, different approaches for reconstruction of such non-static images are required.

Time-varying or dynamic SPECT reconstruction is an ill-posed problem that involves a huge number of variables. This ill-posedness of the reconstruction problem is further amplified by physical degradation of the acquired data caused by camera blurring, photon scattering, or attenuation. As a way to diminish sensitivity to noise and other modeling errors, we call on regularization, since it assists in curing an ill-posed problem. Additionally, the reconstructed image has to be a tradeoff between accuracy and damping of the noise within it. Thus arises the need for fast and robust algorithms and regularization can assist to make the solution less sensitive to noise and modeling errors.

Analytical reconstruction techniques such as filtered back projection (FBP) and iterative ones such as ordered subsets expectation maximization (OSEM) can be used in the static case of emission tomography. Classical EM method, works fine for a static image as well, but breaks down to solve a dynamic SPECT problem. So Bauschke et. al. [1] have introduced what they refer to as a "dynamic EM" approach by using the activity dynamics as linear constraints. Some authors assume prior knowledge about the time activity curve (TAC) dynamics [2]. Based on compartmental modeling [3], early works on dynamic SPECT reconstruction use nonlinear least squares techniques to fit the exponential form of the solution; see [4, 5, 6, 7, 8] and references therein. Fitting strategy using the exponential formulation is known as Prony's method [9] and it is less stable because oscillatory solution may occur. Estimation of the radioactive tracer kinetic parameter is very challenging particularly when the number of compartments is greater than two. In this paper we adopt a stochastic hidden Markov model (HMM) to describe the dynamic SPECT imaging problem. That gives rise to a Bayesian filtering problem. Our model does not assume any prior knowledge about the dynamics of the activity and is best suited to treat the general case of two compartments or more.

In 1960, Kalman has proposed in his pioneering work [10] to solve the noise filtering problem using what was subsequently referred to as the Kalman filter (KF). The most powerful feature is that the Kalman filtering technique is an online recursive form in place of an off-line batch form. Therefore, there is no need to store the past measurements in the computer RAM to estimate the present state. The KF behaves extremely well when the object to be reconstructed is constraint free. In medical imaging we require the activity intensity to be nonnegative. Recently, Qranfal et. al. [11, 12] introduced a novel projected Kalman to solve the dynamic SPECT problem using a proximal algorithm

to enforce this nonnegativity constraint into the KF solution. However, it remains computationally time consuming as KF involves many matrix-matrix multiplications and matrix inversions.

The purpose of this paper is to present a new filtering method, the SMART filter, that we apply to solve the problem for dynamic SPECT image reconstruction based on a stochastic model. While keeping the KF temporal regularization feature, our approach remedies mainly to KF drawbacks of time consuming and not embedding the nonnegativity constraint. In addition, SMART filter is an iterative algorithm and only requires matrix-vector multiplication and does not necessitate any matrix inversion.

The remainder of the paper is organized as follows. First we describe in Section 2 the problem and the stochastic model of the state evolution and projection in space that models it. We show also how we could bring a general setting, such as when we do not have white noise or nonnegative system matrices or data, to a desired one. In Section 3, we review the optimal filtering in the linear case. Our proposed algorithm aims to give an alternative to KF and is the SMART algorithm [13] when the activity is static. We review then the BLUE (best linear unbiased estimator) and introduce the KF, as the BLUE, and its drawbacks that we set ourselves to remedy. Our goal is to introduce an alternative to KF that keeps its advantages, such as filtering the noise and errors, while remedying its drawbacks, especially its known computational time consuming. We thus revisit KF in more details. ART, a precursor algorithm to SMART, and SMART algorithms are reviewed as well. Then Section 4 introduces the weighted KL distance. This latter distance is used to derive the SMART filter algorithm with its convergence results. Section 5 covers this in detail as well as an application to dynamic SPECT. Section 6 on numerical experiment, based on simulations of dynamic SPECT, corroborates the effectiveness of our algorithm in terms of convergence and cpu time in both cases, underdetermined and overdetermined. We finally conclude in Section 7 summing up our findings.

2. Problem Formulation

We start off by stating the problem and how we choose to model it in Section 2.1. White noise is a very important assumption in our approach, to use the weighted KL described in Section 4.1, but is not a restrictive one. Section 2.2 shows what should be done first when we have a colored noise instead. System matrices, data, and variables must be nonnegative to apply our method. However, these

conditions are not restrictive neither. Section 2.3 shows what should be done first if any of these three conditions is not met.

2.1. Stochastic Modeling of Dynamic SPECT

We consider a physiological process where the distribution of the radioactive tracer in an organ or a specific region is time dependent. This region is divided into small parts called dynamic voxels in 3D or doxels and dynamic pixels in 2D or dixels. A SPECT camera, that could have one, two or three heads, is used to register the number of photons emitted by the patient. Let $t_k,\ k = 1, \ldots, S$ be an index of a sequence of acquisition times, N the total number of voxels and M the total number of camera heads' bins, we denote by $x_k \in \mathbb{R}^N$ and $z_k \in \mathbb{R}^M$ the spatial distribution of the activity and the measured data during the Δt_k time. The time frames Δt_k may not be equally spaced. The entry $(z_k)_i$ holds the number of photons registered during the time Δt_k at bin i. The observations $z_1, z_2 \ldots, z_S$ are random vectors. Furthermore, each observation z_k depends on x_k only. The nonnegative activities sequence $x_1, x_2 \ldots, x_S$ satisfy Markov property with unknown time varying transition matrix $A_k \in \mathbb{R}^{N\times N}$. That is

$$x_k = A_k x_{k-1} + \mu_k \tag{1}$$

where, μ_k is the error vector, $\mathbb{E}(\mu_k) = 0$ and $\mathbb{E}(\mu_k \mu_k^\top) = Q_k$ is the covariance of the error in modeling the transition from x_{k-1} to x_k. The random variable μ_k does not have to be a Gaussian or Poisson distributed. In many applications the unknown transition matrix A_k is approximated by a random walk or a discrete diffusion-transport operator. The authors of [14, 15] show convincingly that the first-order Markov model covers a wide range of dynamic models that are applicable for modeling tracer kinetics including the diffusion and one-compartment model used in [8].

Let $(h_k)_{ij}$ be the conditional probability that an emission from dixel/doxel j during the acquisition time Δt_k will be detected in bin i. We call projection or observation matrix the time varying matrix defined by $H_k = [(h_k)_{ij}]$. It is assumed to be known from the geometry of the detector array and may include attenuation correction. It is also organ/region dependent, thus time-dependent as well, since the camera "sees" different views of the organ/region at each acquisition time Δt_k. We shall assume throughout that the matrix $H_k \in \mathbb{R}^{M\times N}$ has been constructed so that H_k, as well as any submatrix O obtained from H_k by deleting columns, has full rank. In particular, if O is $M \times L$ and $L \geq M$, then O has rank M. This is not an unrealistic assumption, which is presumed

in KF as well, to obtain the convergence theorem 3. The columns of H_k are vectors in the nonnegative orthant of M-dimensional space. When attenuation, detector response, and scattering are omitted from the design of the projection matrix H_k, it can sometimes happen that H_k or some submatrices O can fail to be full rank. However, the slightest perturbation of the entries of such a H_k will almost surely produce a new H_k having the desired full-rank properties. The observation and activity vectors are related by the following

$$z_k = H_k x_k + \nu_k \tag{2}$$

where $\mathbb{E}(\nu_k) = 0$ and $\mathbb{E}(\nu_k \nu_k^\top) = R_k$ are the mean and covariance respectively of the noise vector ν_k. The random variable ν_k does not have to be a Gaussian or Poisson distributed. We only need to know its mean and covariance matrix. The observation noise is not additive to the measurements in a strict physical sense. However, feasible solutions can also be obtained using this approximate noise model. Multiplicative noise is generally more difficult to remove than additive noise [16], because the intensity of the noise varies with the signal intensity, thus violating the linearity of the observation model. The linear model we choose requires then the noise to be additive. Otherwise we would not have, for instance, an unbiased estimator.

Filtering is an operation that involves the extraction of information about a quantity of interest at time k by using data measured up to and including k. More precisely the determination of the activity x_k from the measurement data z_k is a filtering problem. Stochastic filtering is an inverse problem. Given collected z_k at discrete time steps and provided A_k and H_k are known, one needs to find the optimal $\hat{x}_k$. Equations (1) and (2) are the state-space form of a particular case of a more general filtering problem [17, 18]. The actual model is a linear dynamic system for which the analytic filtering solution is given by the KF [10]. This can be seen as a temporal regularization technique for solving dynamic inverse problems.

In dynamic SPECT, we assume that covariance matrices Q_k and R_k are diagonal; that is we deal with white error, μ_k, and noise, ν_k, respectively. In addition, it is usually assumed that we have at hand systems with nonnegative entries. Diagonal covariance matrices, white noise and error, and nonnegative system matrices are three more assumptions that we shall adopt throughout this paper. In case one or more of these three is violated, the next two subsections show how we could bring a general setting to an equivalent one presented in this paper.

2.2. Pre-whitening Process

White light contains all frequencies. In a similar manner, a random noise signal or process is called *white* when it is composed of a flat power spectral density of all frequencies. In mathematical terms, a random vector v is a white random vector if and only if its mean vector is zero and its covariance matrix is a multiple of the identity; that is

$$\begin{aligned} \mathbb{E}(v) &= 0 \\ \mathbb{E}(vv^T) &= \sigma^2 I \end{aligned}$$

where I is the identity matrix. There are times when the noise or error vector is not white; we say it is *colored*. We then whiten it by a simple linear change of variable. For instance, assume that we have an error or noise vector w that is colored. It means that $\mathbb{E}(w) = \mu_w \neq 0$ or $\mathbb{E}\left[(w-\mu_w)(w-\mu_w)^T\right] = \Sigma_{ww} \neq \sigma^2 I$. Let

$$v = \Lambda^{-1/2} E^\top (w - \mu_w)$$

where E is the orthonormal matrix of eigenvectors and Λ is the diagonal matrix of positive eigenvalues of the spectral decomposition of the definite positive covariance matrix Σ_{ww}. It follows that v is a random white noise vector because,

$$\mathbb{E}(v) = \Lambda^{-1/2} E^\top (\mathbb{E}(w) - \mu_w) = \Lambda^{-1/2} E^\top (\mu_w - \mu_w) = 0$$

and

$$\begin{aligned} \mathbb{E}(vv^\top) &= \mathbb{E}\big(\Lambda^{-1/2} E^\top (w-\mu_w)(w-\mu_w)^\top E \Lambda^{-1/2}\big) \\ &= \Lambda^{-1/2} E^\top \mathbb{E}\big((w-\mu_w)(w-\mu_w)^\top\big) E \Lambda^{-1/2} \\ &= \Lambda^{-1/2} E^\top \Sigma_{ww} E \Lambda^{-1/2} \\ &= \Lambda^{-1/2}\, E^\top E \Lambda E^\top E \Lambda^{-1/2} \\ &= \Lambda^{-1/2} I \Lambda I \Lambda^{-1/2} \\ &= I \end{aligned} \tag{3}$$

Thus even though when we might have a colored random noise or error w, we remedy to it by a simple change of variable to obtain v as a white noise or error. Our algorithm is not restricted to only nonnegative matrices or vectors as it is the case here of its application to dynamic SPECT. Our approach is applicable to any optimal filtering problem, especially where KF had been applied before, even when the system matrices or the data vectors are not necessary nonnegative. We show in the next subsection how we convert general linear systems to equivalent systems having the desired form in order to use our algorithm.

2.3. From General Systems to Nonnegative Systems

On one hand, we assume that there is no column with 0 in all the entries of the system matrix G. If that is the case, it suffices to delete this column, say column j, work with the remaining ones; then set its corresponding $(\hat{x}_k)_j$ to 0, delete it, and work with the remaining unknown variables. On the other hand, we also assume that $d \in \mathbb{R}^M$ has nonnegative entries. In case an entry $d_i < 0$, it suffices to multiply d_i by -1 as well as the row entries G_{ij} $\forall j = 1, \cdots, N$.

We follow four steps to convert a general system to a nonnegative one [13]. Suppose that $Gc = d$ is an arbitrary system of linear equations, such that $G \in \mathbb{R}^{M \times N}$.

1. If a column j has its sum equal zero, we rescale the equations to make the sum different than zero. If by rescaling one equation of a particular column makes another column sum turn to zero, we just choose a different rescaling. The number of columns is finite so we can always reach a system with nonzero column sums in finite steps.

2. Redefine B and y as follows; replace g_{ij} with $b_{ij} = \dfrac{g_{ij}}{\sum_{i'=1} g_{i'j}}$ and c_j with $y_j = c_j \sum\limits_i g_{ij}$. Observe that the new matrix B has column sums equal to one and that the product Gc is equal to By; so that we retain the same system $By = d$. Note also that $\sum\limits_i d_i = d_+ = \sum\limits_j y_j = y_+ > 0$.

3. If U is the matrix whose entries are all 1, we let $t \geq 0$ be large enough so that $P = B + tU$ has all nonnegative entries. If $\mathbf{1}$ is the vector whose entries are all one, then $Py = By + (ty_+)\mathbf{1}$. Consequently the new system of equations to solve is $Py = d + (td_+)\mathbf{1} = z$. The entries of the "new" data z are still nonnegative as it is the case with the original data d. We introduce an algorithm that assumes the column sums of the system are all one, the system is said to be normalized. To achieve this goal, we make one additional renormalization. So

4. Substitute p_{ij} with $h_{ij} = \frac{p_{ij}}{\sum_{i'} p_{i'j}}$ and y_j with $x_j = y_j \sum_{i'} p_{i'j}$. We have $Hx = Py = z$ and the new matrix H and vector z are nonnegative and all the matrix H columns sums are one.

The assumption of the normality of H_k, that is $\sum_i (h_k)_{ij} = 1$, is for convenience. In emission tomography not all emitted particles are detected, so some

rescaling of the original probabilities $(h_k)_{ij}$ and redefinition of what is meant by x_k is required to achieve this simplification.

To sum up, diagonal covariance matrices, white noise and error, nonnegative z_k, and normalized and full rank system matrices H_k $(k = 1, \cdots, S)$ with nonnegative entries are assumptions that we shall adopt throughout this paper. The solution we seek belongs to the optimal filtering topic; this is covered next.

3. Optimal Filtering

In this section, we review the optimal filtering in the linear case. Our proposed algorithm aims to give an alternative to KF and is the SMART algorithm [13] when the activity is static. Section 3.1 reviews then the BLUE and section 3.2 introduces the KF, as the BLUE, and its drawbacks that we set ourselves to remedy. Our goal is to introduce an alternative to KF that keeps its advantages, such as filtering the noise and errors, while remedying its drawbacks, mainly its known computational time consuming in addition to insuring a nonnegative solution. We revisit KF in section 4. The ART (algebraic reconstruction technique), a precursor algorithm to SMART, and SMART are reviewed in section 3.3.

3.1. Best Linear Unbiased Estimation

Consider the problem of finding an estimator $\hat{x}$ as a linear function of the data vector $z \in \mathbb{R}^M$, such that $z = Hx + \nu$, where $H \in \mathbb{R}^{M \times N}$ is known and $\nu \in \mathbb{R}^M$ represents zero-mean noise with known covariance matrix R. The BLUE of x from z is the vector $\hat{x}$, which minimizes

$$J(x) = \|z - Hx\|^2_{R^{-1}}$$

where $\|v\|^2_B$ denotes the weighted Euclidian norm $v^\top Bv$. If H has full rank, then $\hat{x} = V^\top z$, where $V = R^{-1}H(H^\top R^{-1}H)^{-1}$ [13]. Now suppose that, in addition to the data vector z, we have $y = x + \mu$, a prior estimate of x, where μ is the zero-mean error in this estimate, and the known covariance matrix of μ is Q. We want to estimate x as a linear function of both z and y. Applying the BLUE to the augmented system of equations, that is minimizing the cost function

$$F(x) = \|z - Hx\|^2_{R^{-1}} + \|y - x\|^2_{Q^{-1}}$$

we find the solution to be

$$\hat{x} = y + W(z - Hy)$$

where

$$W = QH^\top (R + HQH^\top)^{-1}$$

We see that to obtain the estimate $\hat{x}$ of x, we first check to see how well y, the prior estimate of x, performs as a potential solution of the system $z = Hx$ and correct the estimate y, using the error $z - Hy$, to get the new estimate $\hat{x}$. If $z = Hy$, then $\hat{x} = y$. The KF involves the repeated application of this extension of the BLUE [13].

3.2. Kalman Filter

Both equations (1) and (2) form the state-space model that are suited to be solved using KF, based on the HMM. The KF solution is the BLUE [13, 18]. The HMM is a statistical model where the activity distributions are assumed to be a Markov process with unknown parameters. Based on this assumption, the challenge is then to determine these hidden parameters from the observable projections. However, KF might produce a negative activity (an activity vector where at least one of its components is negative); this is meaningless in medical imaging. Recall the KF approach. Given an unbiased estimate $\hat{x}_{k-1}$ of the state vector x_{k-1}, our prior estimate of x_k based solely on the physics is

$$y_k = A_k \hat{x}_{k-1} \tag{4}$$

The KF is a recursive algorithm to estimate the state vector x_k during the time Δt_k as a linear combination of the vectors z_k and y_k. The estimate $\hat{x}_k$ will have the form, refer for instance to [11]

$$\hat{x}_k = y_k + K_k(z_k - H_k y_k) \tag{5}$$

where

$$P_k = A_k P_{k-1} A_k^\top + Q_k \tag{6}$$

$$K_k = P_k H_k^\top (H_k P_k H_k^\top + R_k)^{-1} \tag{7}$$

P_k and P_{k-1} in (6) are the covariances of the estimated activity $\hat{x}$ at time k and $k-1$ respectively. On one hand, entries of these two matrices are not guaranteed to be positive; the ones of K_k in (7) are neither since K_k involves an inversion of

a matrix that has positive entries. On the other hand, even though (5) involves y_k, z_k, and H_k which have all nonnegative values, the fact that K_k has some nonnegative entries and that we necessitate a subtraction to update $\hat{x}_k$, we will most likely end up with some negative entries in the vector solution $\hat{x}_k$. This solution has no physical meaning in medical imaging. Setting negative values of the reconstructed activity $\hat{x}$ to zero or taking their absolute value did not give an acceptable solution. This has been remedied, for instance, in [11, 12]. The authors use KF to solve for the unknown activity in dynamic SPECT. Since they end up with negative activity, they use a proximal based minimization approach to project the KF output solution into the positive orthant in order to render the activity feasible.

Since KF does not ensure the nonnegativity of the solution, we like to produce a substitute to KF that does so. Another drawback of KF is the matrix-matrix multiplications involved in (6) and (7) and the matrix inversion required in (7). Attempts have been made to rectify these two shortcomings; please refer to [17, 18] for more details. Furthermore, KF needs to calculate, update, and store covariance matrices. Our goal is manyfold. We intent to find a substitute algorithm to KF that

1. filters out errors from modeling the dynamical system,
2. filters out the noise from the data,
3. insures temporal regularization,
4. is an optimal recursive estimate,
5. does not require the storage of past measurement data in computer RAM,
6. guarantees nonnegativity of the solution,
7. does not use matrix-matrix multiplications
8. does not necessitate any matrix inversion, and
9. does not need to calculate, update, or store any covariance matrix.

We aim then to keep the same first five properties of KF while improving it by requesting four more. Each recursive step in the new approach is an iterative reconstruction that involves only matrix-vector multiplication. These should then handle the problems of huge number of variables, such is the case in medical imaging, and would guarantee positive solutions. But first, let us review the algebraic reconstruction techniques that were applied, for instance, in medical imaging; this is the subject of the next section.

3.3. Algebraic Reconstruction Algorithms

The static emission tomography problem amounts to finding $x \in \mathbb{R}^N$ solution of the linear equation

$$z = Hx + \nu$$

where, $z \in \mathbb{R}^M$, $H \in \mathbb{R}^{M \times N}$ are the observation data vector and the observation matrix respectively. The vector $\nu \in \mathbb{R}^M$ represents the additive noise in recording z. We assume the entries of the vector z and of the matrix H are nonnegative and the columns of H each sums to one. We denote by support(x) the set of indexes j of the vector x for which $x_j > 0$.

The ART [19], an instance of the Kaczmarz method, was the first iterative algorithm used in Computerized Tomography. The ART algorithm goes as follow. Begin with an arbitrary vector x^0. Having found x^ℓ, for each nonnegative integer ℓ, let $i = i(\ell) = (\ell \bmod M) + 1; x^{\ell+1}$ is then obtained as

$$x_j^{\ell+1} = x_j^\ell + h_{ij} \frac{z_i - \sum_{n=1}^N h_{in} x_n^\ell}{\sum_{n=1}^N h_{in}^2} \tag{8}$$

In observing (8), we notice that the new estimate $x_j^{\ell+1}$ is determined by adding a correction term to the current estimate and then it is compared by subtracting the estimated projections from the measured ones. This subtraction operation might induce negative $x_j^{\ell+1}$; which is not desirable in application fields such as medical imaging. Closely related to the ART is the multiplicative ART (MART) [19]. The MART, which can be applied only to nonnegative systems, starts with a positive vector x^0. Having found x^ℓ for nonnegative integer ℓ, we let $i = i(\ell) = (\ell \bmod M) + 1$ and define $x^{\ell+1}$ by

$$x_j^{\ell+1} = x_j^\ell \left(\frac{z_i}{(Hx^\ell)_i} \right)^{H_{ij}} \tag{9}$$

The advantage of MART over ART is that the former guarantees a nonnegative solution over the latter. Byrne [13] showed that by minimizing the Kullback-Leibler [20] distance, we obtain the simultaneous version SMART of MART considered earlier by Gordon et. al. [19] and others. The SMART begins with a strictly positive vector x^0 and has the iterative step

$$x_j^{\ell+1} = x_j^\ell \prod_{i=1}^M \left(\frac{z_i}{(Hx^\ell)_i} \right)^{H_{ij}}, \quad j = 1,\ 2,\ \ldots N \tag{10}$$

The SMART is but a particular case of our algorithm 2. Recall that the Kullback-Leibler/cross-entropy distance between nonnegative numbers α and β is

$$KL(\alpha, \beta) = \alpha \log \frac{\alpha}{\beta} + \beta - \alpha$$

We also define $KL(\alpha, 0) = +\infty$, $KL(0, \beta) = \beta$, and $KL(0, 0) = 0$. Extending to nonnegative vectors $a = (a_1, \cdots, a_N)^\top$ and $b = (b_1, \cdots, b_N)^\top$, we have

$$KL(a, b) = \sum_{j=1}^{N} KL(a_j, b_j) = \sum_{j=1}^{N} (a_j \log \frac{a_j}{b_j} + b_j - a_j)$$

We have $KL(a, b) = \infty$ unless support(a) is contained in support(b). Note how the KL distance is not symmetric; we have in general $KL(a, b) \neq KL(b, a)$.

In the consistent case, that is when there are vectors $x \geq 0$ with $z = Hx$, then both MART and SMART converge to the non-negative solution that minimizes $KL(x, x^0)$. When there are no such nonnegative vectors, the SMART converges to the unique nonnegative minimizer of $KL(Hx, z)$ for which $KL(x, x^0)$ is minimized. We are now ready to derive our algorithm; this is the topic of the next two sections.

4. Towards a Nonnegative Constrained Filter

In Kalman filtering we estimate the state x_k based on all the measurements taken up to the time k. The required estimate is obtained by minimizing the following cost function

$$F(x_k) = \frac{1}{2}\|z_k - H_k x_k\|^2_{R_k^{-1}} + \frac{1}{2}\|y_k - x_k\|^2_{Q_k^{-1}} \tag{11}$$

with respect to x_k. Qranfal et. al. [12] implemented a proximal based algorithm to find a nonnegative solution of $F(x_k)$. A nonnegative constraint approach, applicable to nonnegative vectors and matrices, might need to minimize a distance that applies only to nonnegative quantities; the Kullback-Leibler (KL) does the trick. The cost function becomes

$$F(x_k) = KL(H_k x_k, z_k) + KL(x_k, y_k)$$

It is clear that if the prior estimate y_k of x_k satisfies $z_k = H_k y_k$, then the new estimate is y_k again, just as in the classical Kalman filtering. Then the nonnegative filter would use repeated application of the solution to the minimization

problem. However, the covariances do not seem to play a role now, since this is not a least-squares or Gaussian theory. Nevertheless, the two covariance matrices R_k and Q_k play crucial roles in KF to filter out the noise from the data and the errors from our modeling of the dynamic system. We would like then to keep this filtering property by utilizing these two matrices. We introduce then a weighted KL distance that will handle the filtering part; this is covered in the next section.

4.1. Weighted Kullback-Leibler Approach

In this section, the matrix H is any of the observation matrices H_k and the vector z is any of the observation vectors z_k. Similarly, R and Q are any of the covariance matrices R_k and Q_k respectively. Since R and Q are symmetric positive definite, so are their inverses R^{-1} and Q^{-1}. Therefore, there exist nonnegative diagonal matrices L_1, L_2, U_1, and U_2 such that

$$R^{-1} = L_1 L_1^\top = U_1^\top U_1 \tag{12}$$
$$Q^{-1} = L_2 L_2^\top = U_2^\top U_2 \tag{13}$$

with these decompositions, we have $\|x\|^2_{R^{-1}} = \|U_1 x\|^2$ and $\|x\|^2_{Q^{-1}} = \|U_2 x\|^2$. By the same token, we define the weighted KL distance w.r.t., for instance R^{-1}, as follows,

$$KL_{R^{-1}}(a, b) = KL(U_1 a, U_1 b) \tag{14}$$

The cost function is now a sum of weighted Kullback-Leibler distances

$$\widetilde{F}(x) = KL_{R^{-1}}(Hx, z) + KL_{Q^{-1}}(x, y) \tag{15}$$

It goes without saying that first we should make sure that the four vectors $U_1 Hx$, $U_1 z$, $U_2 x$, and $U_2 y$ have nonnegative entries. Recall that we assume that we have diagonal covariance matrices, white noise and error, nonnegative z, and normalized system matrix H with nonnegative entries. On one hand, the matrix H and vectors z, y, and x have nonnegative values. On the other hand, matrices U_1 in (12) and U_2 in (13) will have, in the general case, negative entries to the extend that it is not ensured that the four vectors are all nonnegative coordinate-wise. Nonetheless, when we deal with white noise and error, U_1 and U_2 will be diagonal matrices with nonnegative values. This is exactly the case of most applications including dynamic SPECT where these four vectors are nonnegative. Otherwise, we should first convert a general system to a nonnegative one and then pre-whiten it, as described earlier in sections 2.3

and 2.2. The cost function (15) is used in a recursion to derive our filtering approach; this follows next.

5. SMART Filter Algorithm

In [21] Byrne considers the (possibly inconsistent) linear system of equations $d = Px$; where the entries of d, P, and x are nonnegative. He solves it by considering the regularization problem

$$\min_{x \geq 0} G(x) = \alpha KL(Px, d) + (1 - \alpha) KL(x, y) \tag{16}$$

where $0 \leq \alpha \leq 1$ is a regularization parameter. Recall the alternating projections algorithm [21]. Let x^0 be a starting nonnegative point. Then having got the ℓ^{th} iterate x^ℓ, we obtain for all $j = 1, 2, \ldots, N$

$$r_{ij}^{\ell+1} = P_{ij} \log \frac{d_i}{(Px^\ell)_i} \tag{17}$$

$$x_j^{\ell+1} = (x_j^\ell)^\alpha (y_j)^{1-\alpha} \exp\left(\alpha \sum_{i=1}^{M} r_{ij}^{\ell+1}\right) \tag{18}$$

In this article, we give also an alternative to Byrne's steps (17) and (18) by gathering both in a convex combination compact form as,

$$x_j^{\ell+1} = (y_j)^{1-\alpha} \left[x_j^\ell \prod_{i=1}^{M} \left(\frac{d_i}{(Px^\ell)_i} \right)^{P_{ij}} \right]^\alpha \tag{19}$$

For $\alpha = 1$, we obtain the SMART iteration (10). The following convergence lemma can be found in [21].

Lemma 1. *The sequence $\{x^\ell\}$ converges to a limit x^∞ for all M and N, for all staring $x^0 > 0$, for all $y > 0$, and for all $0 \leq \alpha \leq 1$. For $0 \leq \alpha < 1$, x^∞ is the unique minimizer of $G(x)$. For $\alpha = 1$, x^∞ is the unique nonnegative minimizer of $KL(Px, d)$ if there is no nonnegative solution of $d = Px$. If there are nonnegative solutions, then the limit may depend on the starting value; we have $d = Px^\infty$, x^∞ is the unique solution minimizing $KL(x, x^0)$, and support(x) is contained within support(x^∞) for all $x \geq 0$ with $d = Px$.*

Not only it is not obvious how problem (16) can be used to ensure temporal regularization in dynamic SPECT, but it does not bear any resemblance to a filtering process. In addition, problem (16) was not thought of to solve a dynamic problem neither. How it is done for the first time is detailed next, illustrated through an example.

5.1. Application to Dynamic SPECT

The error and noise covariance matrices in the dynamic SPECT problem are usually modeled as diagonal matrices with nonnegative entries, refer to section 2.1.

$$Q_k = \sigma_k^2 I \tag{20}$$
$$R_k = diag(z_k) \tag{21}$$

where I is the identity matrix of order N, $diag(a)$ denotes the square matrix that has the a_i in its main diagonal and 0 otherwise. Our measurements z_k are modeled as Poisson random variables, thus the choice of $diag(z_k)$ as their covariance matrices. Let

$$U_1 = diag(\frac{1}{\sqrt{z_k}}) \tag{22}$$
$$U_2 = \frac{1}{\sigma_k} I \tag{23}$$

where the vector $\sqrt{z_k}$ is the one having $\sqrt{(z_k)_i}$ in its entry i. In this case, we are sure that $U_1 z \geq 0$, $U_1 Hx \geq 0$, $U_2 y \geq 0$, and $U_2 x \geq 0$. For ease of notation, we drop for a while the subscript k, the weighted KL cost function is then

$$KL((\frac{1}{\sqrt{z}}).Hx, \sqrt{z}) + KL(\frac{1}{\sigma}x, \frac{1}{\sigma}y) \tag{24}$$

or

$$\widetilde{G}(x) = \frac{\sigma - 1}{\sigma} KL(\frac{\sigma}{(\sigma - 1)\sqrt{z}}.Hx, \frac{\sigma}{\sigma - 1}\sqrt{z}) + \frac{1}{\sigma} KL(x, y) \tag{25}$$

where $a.b$ and a/b designate the multiplication and division respectively of the vectors a and b component wise. We can obtain the same functional (25) by using a pre-whitening. Recall

$$z = Hx + \nu \tag{26}$$

where $\mathbb{E}(\nu) = 0$ and $\mathbb{E}(\nu\nu^\top) = R = diag(z)$. Now multiply both sides with $R^{-1/2}$ in order to pre-whiten ν,

$$R^{-1/2} z = (R^{-1/2} H)x + (R^{-1/2}\nu) \tag{27}$$

Take $z_1 = R^{-1/2} z$, $H_1 = R^{-1/2} H$, and $\nu_1 = R^{-1/2}\nu$, we then have

$$z_1 = H_1 x + \nu_1 \tag{28}$$

Notice that $\mathbb{E}(\nu_1) = 0$ but $\mathbb{E}(\nu\nu^\top) = I$. We still have z_1 and H_1 being nonnegative; we do have nonnegative "observations" and "projection" matrices. Hence we do not need a weighted distance in the first KL since the covariance matrix is just I. As per the second KL portion, we do not need to do the same trick with the evolution equation because the covariance matrix Q is just $\sigma^2 I$. Recall

$$x = y + \mu \tag{29}$$

where the predicted state $y = A_k x_{k-1}$, $\mathbb{E}(\mu) = 0$, and $\mathbb{E}(\mu\mu^\top) = \sigma^2 I$.

Recall that we aim to find a nonnegative estimate $\hat{x}_k$, $k = 1, \ldots, S$, to the nonnegative unknown x_k of the problem given by the two linear space-state equations (1) and (2),

$$\begin{aligned} x_k &= A_k x_{k-1} + \mu_k \\ z_k &= H_k x_k + \nu_k \end{aligned}$$

μ_k is the error vector, $\mathbb{E}(\mu_k) = 0$ and $\mathbb{E}(\mu_k \mu_k^\top) = Q_k$ is the covariance of the error in modeling the transition from x_{k-1} to x_k. $\mathbb{E}(\nu_k) = 0$ and $\mathbb{E}(\nu_k \nu_k^\top) = R_k$ are the mean and covariance respectively of the noise vector ν_k. Entries of the vector z_k and of both matrices A_k and H_k are nonnegative. We know also from equations (20) and (21) that,

$$\begin{aligned} Q_k &= \sigma_k^2 I \\ R_k &= diag(z_k) \end{aligned}$$

Minimizing the functional (25) at each recursion step k is the same as solving problem (16). It suffices to use the change of variables given in step 2 of the SMART filter algorithm 2, that follows, and to recall that the predicted activity state y_k is $A_k x_{k-1}$. We apply, recursively at each time step k, the iterative procedure given by the formulas (17) and (18). The clustering point x^∞ will be the estimate $\hat{x}_k$ we are solving for. Thus we obtain the following algorithm 2 that we refer to as SMART filter. We remind the reader that diagonal covariance matrices, white noise and error, nonnegative z_k, and normalized and full rank system matrices H_k $(k = 1, \cdots, S)$ with nonnegative entries are presumed in order to apply the SMART filter.

Algorithm 2. SMART Filter Algorithm

1. *Start with $\hat{x}_0 > 0$. For $k = 1, \cdots, S$ execute the following steps*

2. *Assume we have done the recursive step up to time $k-1$, do the change of variables*

$$\begin{aligned} \alpha &= \frac{\sigma_k - 1}{\sigma_k} \\ P &= R_k^{-1/2} H_k \\ d &= R_k^{-1/2} z_k = \sqrt{z_k} \end{aligned}$$

3. *To get $\hat{x}_k$, start with $x^0 = \hat{x}_{k-1}$*

4. *Make $y_k = A_k \hat{x}_{k-1}$*

5. *Alternate between the next two sub-steps, $\ell = 0, 1, \ldots$*

$$\begin{aligned} r_{ij}^{\ell+1} &= P_{ij} \log \frac{d_i}{(Px^\ell)_i} \\ x_j^{\ell+1} &= (x_j^\ell)^\alpha ((y_k)_j)^{1-\alpha} \exp\left(\alpha \sum_{i=1}^{M} r_{ij}^{\ell+1}\right) \end{aligned}$$

6. *The update formula for the next estimate is $\hat{x}_k = x^\infty$, where x^∞ is the cluster point of the sequence $(x^\ell)_{\ell \in \mathbb{N}}$.*

Observe that we could combine the two sub-steps of step 5 into one step as we have offered before in (19)

$$x_j^{\ell+1} = \left[x_j^\ell \prod_{i=1}^{M} \left(\frac{\sqrt{(z_k)_i}}{(Px^\ell)_i} \right)^{P_{ij}} \right]^\alpha ((y_k)_j)^{1-\alpha} \tag{30}$$

Note how the next iterate $x_j^{\ell+1}$ in (30) is formed as a product of a convex combination of the calculated SMART iterate, in a similar form as in (10), associated with the same coefficient α as in problem (16) relying only on the observation model and of the predicted $(y_k)_j$, associated with the same coefficient $1-\alpha$ as in problem (16), relying only on the evolution model. In step 2 the matrix R_k is diagonal, thus the SMART filter does not involve any matrix-matrix multiplication or any matrix inversion. It involves only matrix-vector multiplication. This algorithm does not need any matrix update or storage neither. The temporal regularization parameter α is well defined and takes values between 0 and 1 when σ_k varies between 1 and ∞. If $\alpha = 0$, that is $\sigma_k = 1$ and

$\hat{x}_k = A_k \hat{x}_{k-1}$, then we are discarding completely the observations to the extend that we rely only on our evolution model. This defeats the purpose of the experiment. When $\alpha = 1$, the predicted y_k in step 4 is not needed in step 6 and we get $x_j^{\ell+1} = x_j^{\ell} \prod_{i=1}^{M} (d_i/(Px^{\ell})_i)^{P_{ij}}$. Hence we retrieve the SMART iteration as mentioned before, which is only valid for the static case. Indeed, choosing $\alpha = 1$ means that $\dfrac{\sigma_k - 1}{\sigma_k} = 1$ or simply $\sigma_k = \infty$. That is the covariance matrix in the transition equation (1) is very huge; which implies we have no confidence at all in our evolution model. In other words, we discard the evolutionary state of the variable. Only the observations z_k are meaningful in finding the $\hat{x}_k$; which is then a stationary state as it should be.

We have the following convergence result; this is a direct consequence of lemma 1

Theorem 3. *For all $k = 1, \cdots, S$, the sequence $\{x^{\ell}\}$ converges to a limit $\hat{x}_k = x^{\infty}$ for all M and N, for all staring $\hat{x}_0 > 0$, and for all $1 \leq \sigma_k \leq \infty$. For $1 \leq \sigma_k < \infty$, $\hat{x}_k$ is the unique minimizer of the functional $\widetilde{G}(x)$ given by (25). For $\sigma_k = \infty$, $\hat{x}_k$ is the unique static nonnegative minimizer of $KL(H_k x_k, z_k)$ if there is no nonnegative solution of $z_k = Hx_k$. If there are nonnegative solutions, then the limit may depend on the starting value; we have $z_k = H_k \hat{x}_k$, $\hat{x}_k$ is the unique solution minimizing $KL(x_k, x^0)$, and support(x_k) is contained within support$(\hat{x}_k)$ for all $x_k \geq 0$ with $z_k = H_k x_k$.*

Remark 4. In considering these particular mixture of KL distances in (25), we unified approaches commonly taken in the underdetermined and overdetermined cases, and we developed an iterative solution method in a recursion within a single framework of alternating projections as well as establishing a convergence result.

Next, we put the SMART filter to test in both cases, underdetermined and overdetermined.

6. Numerical Experiment

6.1. Simulation

Our phantom is composed of six regions of interest (ROI) or segments. Each ROI has a different TAC, see figure 1. The example investigated in this work is based on the teboroxime dynamics in the body during the first hour post

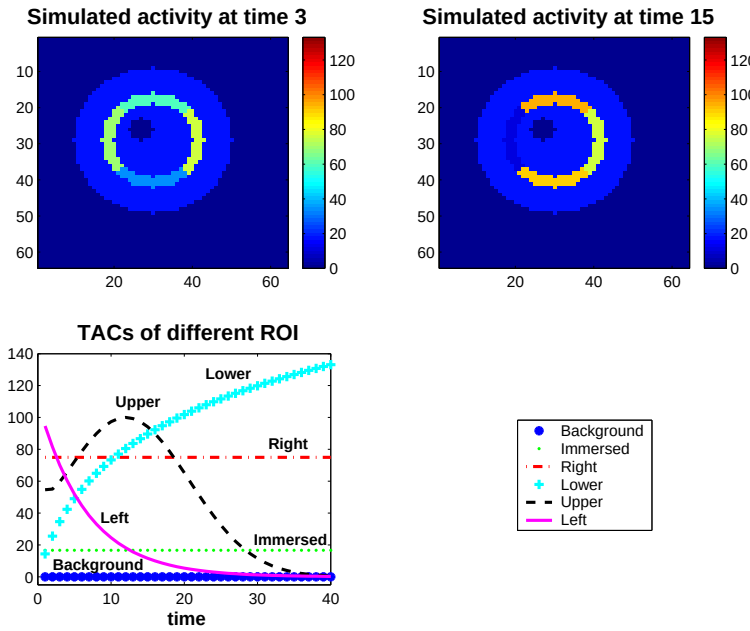

Figure 1: Simulated annulus with its different ROI and their TACs. Upper left: simulated activity at time 3, upper right: simulated activity at time 15, and lower left: TACs of the 6 different ROI.

injection. The choice of the TACs is motivated by the behavior of liver, healthy myocardium, muscles, stenotic myocardium, and lungs. Only one slice is modeled; that is we simulate a 2D object. The star-like shape placed on the left ensures that the phantom is not entirely symmetrical. We simulate 120 projections over 360°, one projection for every 3°, with attenuation and a 2D Gaussian detector response.

There are three camera heads consisting of 64 square bins each measuring 0.625 cm in each side, see figure 2. The distance from the annulus to the camera head rotation axis is 30 cm. We simulate $S = 40$ time instances for three heads; that is we have $3 \times 40 = 120$ projections for a camera rotating clock wise (CW) in a circular orbit. Head 1 starts at $-60°$, head 2 at $60°$, and head 3 at $180°$. A low energy high resolution (LEHR) collimator is used with a full width at half maximum (fwhm). We determine the blurred parallel strip/beam geometry system matrices for all projections with resolution recovery and attenuation correction [2].

We have 64 projection/measurement values for each head, which amounts

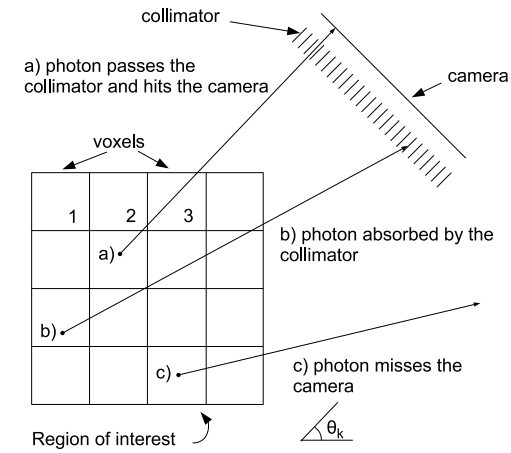

Figure 2: Photon radiating from the region of interest: a) passes the collimator and hits the camera, b) absorbed by the collimator, c) misses the camera.

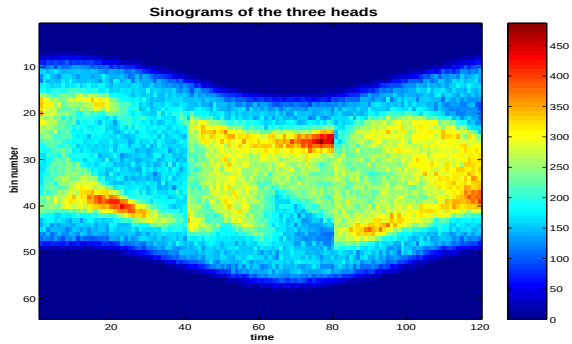

Figure 3: Sinogram or $2D$ projections: y-axis has the bin number and the x-axis has the 40 time instances of the 3 heads. Time instances from 1 to 40 are for head 1, 41 to 80 for head 2, and 81 to 120 for head 3. A color intensity of a pixel is the number of detected photons by a certain bin at a certain time.

to a total of $M = 192$ observations at each time frame, figure 3 shows the data sinogram. Notice how the photons' count in any detector's bin varies between 0 and somewhere around 500; which is a realistic scenario for a 2D slice. The size

of the image we aim to reconstruct at each time frame is $N = 4096 = 64 \times 64$ dixels. We have six kinds of TACs that are very representative for clinical applications. The annulus has four arcs that we name "Left", "Upper", "Right", and "Lower" according to their location. The activity is decreasing in the Left arc, increasing-decreasing in the Upper arc, constant in the Right arc, and increasing in the Lower arc, see figure 1. The star-like shape has zero activity within it and is called the "Star" region; we refer to it as "Background" too. The annulus is immersed within a region that is called "Immersed" and has a constant activity. We have six ROI in total. The SMART filter algorithm should work in both underdetermined and overdetermined settings, refer to remark 4. We aim then to test the algorithm 2 in the underdetermined and overdetermined cases.

The undermined case happens when we reconstruct dixel by dixel; we possess $M = 192$ data for $N = 4096$ unknowns or a ratio of about 1:21 data to unknowns. It is an ill-posed problem. Maltz [23] mentioned that Reutter et. al. method [8] is effective in providing the desired estimates; however, the amount of computation required is large for studies involving many dynamic regions/compartments. Our approach deals with the general case regardless of the number of compartments. It does not assume uniformity of the pixels and should work if someone desires to use multiresolution, as done in [23], since the system matrix H_k captures the information that links the pixel to the data without assuming anything about the shape, size, or location of the pixels. The more noise is introduced, the more Maltz's multiresolution method [23] seems to under-perform as shown in his article. Our approach, as in the case of KF, filters out the noise while reconstructing the images and TACs. This is a real advantage over his. The overdetermined case consists in reconstructing the six ROI, when we assume full knowledge of their locations, so that we have $M = 192$ data for $N = 6$ unknowns or a ratio of 32:1 data to unknowns. We should of course get much better reconstructed images in the latter case than in the former one; this, indeed, will be confirmed shortly.

We provide quantitative analysis of the reconstructed images in order to compare the simulated activity with the reconstructed one. We define the relative deviation error δ of the reconstructed activity v^* from the truth x, refer to formulas (31) through (33). Hence we compare the simulated count $x_{i,k}$ with the corresponding reconstructed one $v^*_{i,k}$ at each time frame k for every location i. We sum over a ROI containing J dixels normalized by the total simulated/true counts in order to diminish the effect of statistical fluctuations. We have a $\delta_{ROI,k}$ for every sector. These indicators allow us to see how the

method performs under different dynamic behaviors. We could compare, for instance, sectors with fast washout with those with slow one. We calculate similar δ_k over the total number of doxels (dynamic voxels) N then we average them over the total number S of time acquisitions; so that we have δ_{avg}. This is an objective comparison of the quality of reconstruction for different sets of parameters such as iteration stopping criteria, noise levels, etc. The closer δ_{avg} is to zero, the better the reconstructed images should be.

$$\delta^2_{ROI,k} = \frac{\sum_{j=1}^{J}(v^*_{j,k} - x_{j,k})^2}{\sum_{j=1}^{J} x^2_{j,k}} \tag{31}$$

$$\delta^2_{k} = \frac{\sum_{j=1}^{N}(v^*_{j,k} - x_{j,k})^2}{\sum_{j=1}^{N} x^2_{j,k}} \tag{32}$$

$$\delta_{avg} = \frac{1}{S}\sum_{k=1}^{S}\delta_k \tag{33}$$

6.2. Results

The SMART filter algorithm by its nature ensures the temporal continuity of the reconstructed TACs since it imposes a temporal regularization in its formulation stated in (25). We could have used, for instance, a diffusion model to model the temporal evolution. The closer the model to reality, the better our algorithm should perform. In our test case, we do not make any assumption about the blood input and we assume that the system dynamics are unknown to us as per (1); therefore we use a pure random walk. In practical terms, we set $A_k = I$, for all $k = 1, \cdots, S$. However, we do not have much confidence in our transition model (1) so we compensate to that by choosing covariance matrices to be pretty high, $10^3 \leq \sigma_k \leq 10^4$. Recall that a random walk is a special first-order autoregressive (AR(1)) process with a unit slope (i.e. unit root) [24]. In its simplest form an AR(1) process is,

$$u_k = a_1 u_{k-1} + \varepsilon_k$$

where the $\{\varepsilon_k\}$ disturbance/error sequence is a white-noise process. A special case is the pure random walk,

$$u_k = u_{k-1} + \varepsilon_k$$

Random walk predicts that the value at time "k" will be equal to the last period value plus a stochastic (non-systematic) component that is a white noise.

For the state transition linear model, we proceeded as follows. We assume that the system dynamics are unknown to us (1); therefore we use a random walk. In practical terms, we set $A_k = I$, for all $k = 1, \cdots, S$. For the state transition linear model, we proceeded as follows. We are not interested in the background and we assume that we know the locations of these zero activities; this is a common practice [8, 14, 22]. We have run experiments without this assumption and results are very comparable to when we have run them with this assumption. One interesting way to deal with this assumption is as this. Set to zero the values of the corresponding positions of the matrix H_k. The updating equations (6) and (7) ensure that the updated activities will remain equal to zero; thus the KF reconstructs perfectly the star/background region(s) [11, 12]. By having these values as zero while eliminating these columns and setting to zero the corresponding entries of the estimated activity [13, 25], we guarantee automatically that those entries stay at the value of zero when using our present algorithm SMART filter.

We experiment with different initial guesses $\hat{x}_0$ such as $(1, \cdots, 1)^\top$ and $(10^{-6}, \cdots, 10^{-6})^\top$. We also start the algorithm with the static image given by OSEM; we call this initial guess *OSEM activtiy*. The average of the deviation error δ_{avg} combined with visual inspection show that there is no pronounced advantage in favor of any. In the underdetermined case, we do not assume the ROI to be known exactly. However, we make use of these segments only to interpret the results. As a consequence, there are some differences in intensity between pixels within the same region. To assess the effectiveness of the method and of the convergence result 3, we show the TACs averaged over the pixels within the same ROI and this is also valid for the overdetermined case. Follow are the results of both reconstruction cases, underdetermined and overdetermined, obtained through SMART filter and their comparison with results obtained using the projected Kalman algorithm [11]. The projected Kalman is the classical KF followed by a projection into the positive octant, using a proximal approach, to ensure the feasibility of the activity. Recall that KF, see equations (6) and (7), necessitates multiplications and inversion of huge matrices which is time consuming and memory hungry.

6.2.1. Underdetermined Case

We are solving the ill-posed inverse problem in reconstructing the dynamic images of the annulus, 192 observations for 4096 unknowns. This is an underdetermined case with a ratio of about 1:21 data to unknowns. We use the

algorithm 2 on a P4 3.00 GHz desktop. It takes about 8 min to run the SMART filter. In contrast to the projected Kalman algorithm [11] which takes more that 2.5 hr, we witness an improvement of more than 18 times faster. We suspect that using the formulation (30), instead of steps 5 and 6 in algorithm 2, could even give us better speed. The δ_{avg} is about 0.52 which is the same as with projected Kalman. Images and TACs look fine and are about of the same quality as with projected Kalman, refer to figure 4.

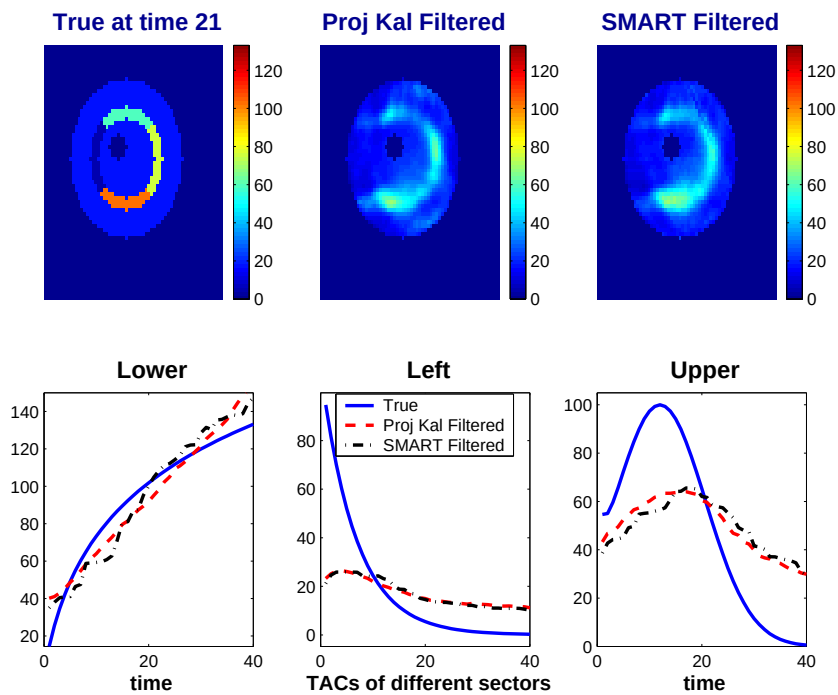

Figure 4: Reconstructed images, pixel by pixel, at time 21 and the averaged out TACs, over their corresponding region, using SMART filter and Projected Kalman. The true phantom at time 21 is on upper left, the reconstructed using the projected Kalman and SMART filter on upper centre and upper right respectively. TACs of three different regions, Lower, Left, and Upper, are shown on bottom left, centre, and right respectively. Blue TACs for simulated, red TACs for reconstructed using projected Kalman, and black TACs for reconstructed using SMART filter.

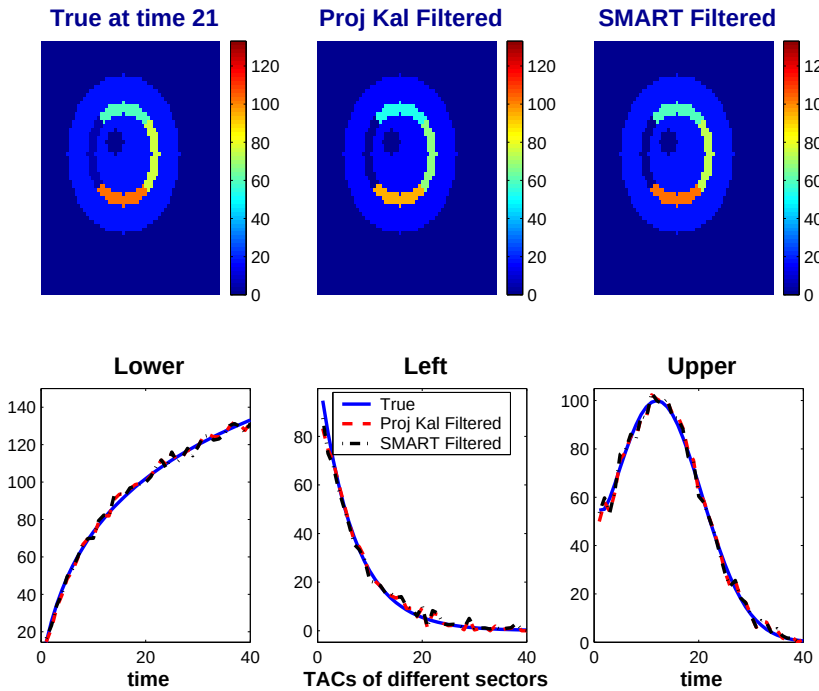

Figure 5: Reconstructed images at time 21 and TACs of 6 ROI, region by region, using SMART filter and Projected Kalman. The true phantom at time 21 is on upper left, the reconstructed using the projected Kalman and SMART filter on upper centre and upper right respectively. TACs of three different regions, Lower, Left, and Upper, are shown on bottom left, centre, and right respectively. Blue TACs for simulated, red TACs for reconstructed using projected Kalman, and black TACs for reconstructed using SMART filter.

6.2.2. Overdetermined Case

In medical imaging, we are sometimes not interested in individual intensities of each and every pixel/voxel but rather on some ROI intensities. We are then more concerned with a segmented reconstruction [8]. A CT scan for instance might give us an idea about the ROI. In case we have this prior knowledge about the selection of ROI before hand, we could include this constraint, reduce the size of our problem, and have by the same token a better image. In our setting, we are then solving the inverse problem in reconstructing the dynamic images of the annulus, 192 observations for 6 unknown ROI. This is an overdetermined

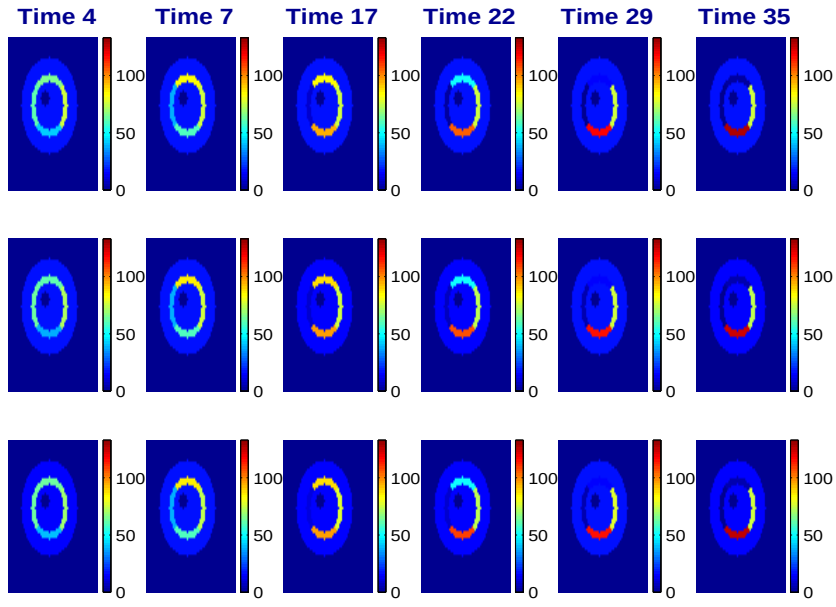

Figure 6: Reconstructed images at various time instances: simulated images in top row, Projected Kalman filter reconstructed images in middle row, and SMART filter reconstructed images in bottom row.

case with a ratio of 32:1 data to unknowns. We test the algorithm 2 on a P4 3.00 GHz desktop. It takes about 15 sec to run the SMART filter. In contrast to the projected Kalman algorithm [11] which takes about 1.7 sec, SMART filter is 9 times slower due probably to the many log and exp functions evaluations in steps 5 and 6 of the algorithm. Using the combined steps 5 and 6, as per equation (30), would most likely accelerate the algorithm. The δ_{avg} of the SMART filter method is about 0.03 which is half of the one with projected Kalman. We witness then a net improvement, convergence wise, with SMART filter than with projected Kalman. As expected with this overdetermined case, we get much better images and TACs than with the underdetermined case, see figures 5 and 6 and compare them to figure 4. Images and TACs of SMART filter method are of the same quality as with projected Kalman. However we get better images with SMART filter; compare in figure 5 the Upper, Right, and Lower arcs color wise of both reconstructions to the simulated ones.

7. Conclusion

We presented here a novel algorithm that we refer to as the SMART filter. It applies to nonnegative normalized full rank systems when a nonnegative solution is desired. Our algorithm guarantees this in addition to a temporal regularization. We also proposed a systematic way in how someone could bring a general system to a normalized nonnegative one in order to use our approach. We tested SMART filter to reconstruct a dynamic image in SPECT while using a pure random walk to model the activity evolution. The SMART filter reveals itself to be about 18 times faster than the projected Kalman in the undetermined case, minutes instead of hours. In the over determined case, SMART filter is about 9 times slower than the projected Kalman, both cpu times in the seconds; however, SMART filter shows better convergence result. We got much better TACs and images in the overdetermined case; this suggests that the more info we feed the algorithm the better it behaves. Thus we suspect that we could improve the quality of the images and TACs, even in the underdetermined case, by using a closer to reality evolution system matrix. The SMART filter algorithm filters out errors from modeling the dynamical system and the noise from the data. It insures temporal regularization and outputs an optimal recursive estimate. It does not need any matrix update or storage. It also does not use any matrix-matrix multiplication and does not necessitate any matrix inversion. These last properties make it very suitable for large scale systems such as the ones in medical imaging, PET (positron emission tomography) for instance, or in electrical impedance tomography. The SMART filter algorithm could be used in any discipline which has used, for instance, KF or in any one that is interested in time-varying variables such as financial risk assesment/evaluation and forecasting or control, especially if they are concerned with nonnegative solutions. Application of our algorithm to time-varying SPECT, a medical imaging modality in nuclear medicine, confirms our convergence theorem. Our results substantiate the efficiency of this novel filtering technique, the SMART filter.

Acknowledgments

This is to thank Dr. Germain Tanoh of Quantimal for his contributions to the present paper.

References

[1] H.H. Bauschke, D. Noll, A. Celler, J. M. Borwein, An EM-algorithm for dynamic SPECT tomography, *IEEE Trans. Med. Imag.*, **18** (1999), 252-261.

[2] G. Tanoh, *Algorithmes du point intérieur pour l'optimisation en tomographie dynamique et en mécanique du contact*, Université Paul Sabatier (2004).

[3] K.R. Godfrey, *Compartmental Models and their Application*, Academic Press, New York, NY (1983).

[4] D.J. Kadrmas, G. T. Gullberg, 4D maximum a posteriori reconstruction in dynamic SPECT using compartemental model-based prior, *Phys. med. Biol.*, **46** (2001), 1553-1574.

[5] M.A. Limber, A. Celler, J. Barney, M.N. Limber, J.M. Borwein, Direct reconstruction of functional parameters for dynamic SPECT, *IEEE Trans. Nuc. Sci.*, **42** (1995), 1249-1256.

[6] J. Maeght, D. Noll, S. Boyd, Dynamic emission tomography regularization and inversion, In: *Bulletin of the Canadian Math. Society*, **27** (2000), 211-234.

[7] J. Maltz, Parsimonious Basis Selection in Exponential Spectral Analysis, *Phys. Med. Biol.*, **47** (2002), 2341-2365.

[8] B.W. Reutter, G.T. Gullberg, R.H. Huesman, Direct least Squares Estimation of Spatiotemporal Distribution from Dynamic SPECT Projections using Spatial Segmentation and Temporal B-splines, *IEEE trans. med. imaging*, **19** (2000), 434-450.

[9] M.R. Osborne, K.G. Smythe, A modified Prony Algorithm for Exponential Function Fitting, *SIAM J. Sci. Computing*, **16** (1995), 119-138.

[10] R.E. Kalman, K.G. Smythe, A new approach to linear filtering and prediction problems, *Trans. of the ASME-Jour. of Basic Eng.*, **82** (1960).

[11] J. Qranfal, *Optimal Recursive Estimation Techniques for Dynamic Medical Image Reconstruction*, Simon Fraser University (2009).

[12] J. Qranfal, G. Tanoh, Regularized Kalman filtering for dynamic SPECT, In: *J. Phys.: Conf. Ser.*, **124** (2008).

[13] C.L. Byrne, *Signal Processing, A Mathematical Approach*, A K Peters, Wellesley, MA (2005).

[14] M. Kervinen, M. Vauhkonen, J.P. Kaipio, P.A. Karjalainen, Time-varying reconstruction in single photon emission computed tomography, *Int. J. of imaging syst. tech*, **14** (2004), 186-197.

[15] N.A. Lassen, W. Perl, *Tracer Kinetic Methods in Medical Physiology*, Raven Press, New York (1979).

[16] M.A. Schulze, An edge-enhancing nonlinear filter for reducing multiplicative noise, In: *Proc. SPIE Vol. 3026*, (1997), 46-56.

[17] B.D.O. Anderson, J.B. Moore, *Optimal Filtering*, Printice-Hall, Englewood, Ciffs, NJ (1979).

[18] D. Simon, *Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches*, Wiley-Interscience, (2006).

[19] R. Gordon, R. Bender, G.T. Herman, Algebraic reconstruction technique (ART) for three-dimentional electron microscopy and X-ray photography, *Ann. Math. Statist*, **29** (1970), 471-481.

[20] S. Kullback, R. Leibler, On information and sufficiency, *J. Theoret. Biol.*, **22** (1951), 79-86.

[21] C.L. Byrne, Iterative image reconstruction algorithms based on cross-entropy minimization, *IEEE Trans. on Image Processing*, **2** (1993), 96-103.

[22] T. Farncombe, *Functional Dynamic SPECT Imaging Using a Single Slow Camera Rotation*, University of British Columbia (2000).

[23] J.S. Maltz, Multiresolution constrained least-squares algorithm for direct estimation of time activity curves from dynamic ECT projection data, In: *Proc. SPIE 3979*, (2000), 586-598.

[24] L. Kónya, Basic Properties of Stationary First-Order Autoregressive Processes and Random Walks, *SSRN eLibrary*, (2000).

[25] J. Qranfal, C. Byrne, EM Filter for Time-Varying SPECT Reconstruction, *Int. J. of Pure and Appli Math*, **to appear** (2011).