System Identification using Nonlinear Filtering Methods with Applications to Medical Imaging[1]

J.Yin, V.L.Syrmos and D.Y.Y.Yun
Department of Electrical Engineering
University of Hawaii at Manoa
Honolulu, HI 96822
hellas@euclid.eng.hawaii.edu

Abstract

In this paper, we first review the concept of computational tomography (CT) and laser technique using the photon diffusion equation. The forward and the inverse problem are two key problems concerned with the diffusion equation, while the solution to the later one is the goal of research in optical CT. The inverse problem can be stated as follows: given the photon density measured from the detectors outside the tissue, we need to find the anomalies (benign or malignant) inside the tissue. We model the forward and the inverse problem using state-space equations and pose the inverse problem as a system identification problem. The nonlinear filtering techniques, namely the extended Kalman filter and the second order filter are proposed to solve the inverse problem. Comparisons are made through an example of a medical imaging problem.

1 Introduction

Computational tomography for the photon diffusion equation has been the subject of extensive research recently [3, 4, 9]. Photon laser applications in medical imaging are being investigated because of their advantages compared to X-ray radiation. An application where photon laser techniques have found a vast number of applications is the detection of malignant and/or benign tumors in the human tissue.

There are two main tasks concerned with the photon diffusion equation, namely the forward and the inverse problem. In the forward problem, we are given all the tissue cells' scattering and absorption coefficients, we compute the photon density at every position in the medium and at all times. Conversely, given the photon density that is measured by the detectors, we reconstruct the medium structure by estimating the scattering and absorption coefficients at every location inside the medium; this is called the inverse problem.

To solve the inverse problem, we cast it as a parameter identification problem. We use as the "nominal" system the homogeneous system. That is, we use as a baseline the system data obtained from a healthy tissue. Due to the modeling of the system we can represent the heterogeneous system (tissue with anomalies) as a "perturbation" of the homogeneous one. Furthermore the system "perturbation" appears only in the diagonal of the system matrix. If we define the perturbation as an unknown parameter of the heterogeneous system, our task is to determine the value of the parameter using measurements from the detectors. Therefore, the inverse problem of calculating the optical coefficients becomes a parameter identification problem. In the proposed approach, we use the nonlinear filtering techniques to proceed the parameter identification.

2 Preliminaries

We assume that the photons radiate using an extremely fast pulse laser, and the photon propagation through such media is described by the Boltzmann transport equation. The simplest approximation to the equation, i.e. the P_1-approximation, known as the diffusion equation is used to simulate the photon propagation which is given by

$$\begin{aligned}\frac{\partial}{\partial t}\Phi(r,t) &= \mathrm{div}(D(r)\mathrm{grad}\Phi(r,t)) - c\mu_a(r)\Phi(r,t) \\ &\quad +q(r,t), \qquad (2.1)\end{aligned}$$

where $\Phi(r,t)$ is the light density at position r (3D) and time t, c is the speed of light in the medium, $\mu_a(r)$ is the absorption coefficient and $q(r,t)$ is the light density of the source. Furthermore, the optical diffusion coefficient $D(r)$ is given by $D(r) = c/3(\mu_a(r) + \mu_s(r))$, where $\mu_s(r)$ is the scattering coefficient. In this paper we assume that the media is scatter-dominated with $\mu_s(r) \gg \mu_a(r)$. Therefore the optical coefficient $D(r) \approx c/3\mu_s(r)$. In [9], four conditions are given which guarantee that the diffusion equation represents a good approximation of light propagation. Usually, the human tissue, which we study, satisfies those four conditions.

To solve the partial differential equation (2.1), the most commonly used numerical method is the finite element method (FEM), which is based on Galerkin approximation [3]. Suppose that the medium size is $x \times y \times z$, then we divide the x direction by M_x, divide the y direction by M_y, divide the z direction M_z, where M_x, M_y, M_z are the numbers of grids in each direction. As a result, there are $n = M_x M_y M_z$ cubes in the medium,

---

[1] This research was supported by the Central Research Laboratory, Hamamatsu Photonics K.K., 5000 Hirakuchi, Hamakita 434, Japan.

which corresponds to the dimension of the basis functions. In [11], the PDE of equation (2.1) is transformed to a state-space system as follows

$$\begin{aligned} \dot{x}(t) &= -E^{-1}(SG + cP)x(t) + E^{-1}Fu(t) \quad (2.2)\\ &= Ax(t) + bu(t) \\ y(t) &= Cx(t). \quad (2.3) \end{aligned}$$

with $E = h^3 I$, where h is the grid size, while $S = c/3\mu_s I$, $P = \mathrm{diag}(p_1, p_2, \ldots, p_n)$ and $p_i, i = 1, 2, \ldots, n$ denote the space distribution of μ_a, $Fu(t)$ relates to the system input, and matrix $G \in \mathbb{R}^{n\times n}$ is a block tridiagonal sparse matrix, representing the geometric relationship of the cubes. Matrix $C \in \mathbb{R}^{p\times n}$ represents the position of detectors, where p is the number of detectors. $x_i(t)$ is the light density $\Phi(r, t)$ at each cube of the medium and $y_i(t)$ is the light density $\Phi(r, t)$ at each detectors' position.

According to equations (2.2) and (2.3), the inverse problem can be defined as: given the measurements $y(t)$ from the detectors and the scattering coefficients (matrix S), we are asked to calculate the absorption coefficients at every position (matrix P). From system's point of view, the inverse problem can be treated as a parameter identification problem, which can be solved using nonlinear filtering techniques [1, 2, 5, 6, 7, 8, 10]. The extended Kalman filtering method is a standard approach to nonlinear filtering problem and has particular applications in parameter estimations [7]. To enhance the convergence properties of the extended Kalman filter, higher order terms of the power series are also considered in the model linearization and the second order filter is then derived.

3 Nonlinear Filtering Approach

To simplify the model in equations (2.2) and (2.3) to better fit the inverse problem, we propose the following model

$$\begin{aligned} \dot{x}(t) &= (A + D)x(t) + bu(t) \quad (3.1)\\ y(t) &= Cx(t), \quad (3.2) \end{aligned}$$

where, matrices A, b, C are known and matrix D is an unknown diagonal matrix to be identified. The diagonal matrix D represents the difference $\delta\mu_a$ in absorption coefficients between the homogeneous and heterogeneous case, at every position in the medium. The system is called homogeneous system if $D = 0$ and heterogeneous system if $D \neq 0$. Furthermore, we incorporate the sensor noise $v(t)$ into the state-space equation (3.2) to get the stochastic model:

$$\begin{aligned} \dot{x}(t) &= (A + D)x(t) + bu(t) \quad (3.3)\\ y(t) &= Cx(t) + v(t), \quad (3.4) \end{aligned}$$

where

$$\begin{aligned} Ex(0) &= 0 \qquad Ex(0)x^T(0) = P_{x(0)}, \\ Ev(t) &= 0 \qquad Ev(t)v^T(s) = R\delta_{ts}. \end{aligned}$$

$v(t)$ describe the measurement uncertainties introduced by the detectors. $x(0)$ are the initial states of the system, which are assumed to be random. When certain assumptions on the noise signal and the initial states are satisfied, the Kalman filter can be applied to this linear system (3.3) and (3.4) to achieve an optimal state estimate. The extended Kalman filter is a nonlinear filtering technique. For a nonlinear system, we linearize at each point along the estimated state trajectory to get an approximated linear model. A standard Kalman filter is applied to the approximated linear model to get a state estimate. The estimation is not guaranteed optimal for the original nonlinear model, but this approach is a viable candidate for the nonlinear filtering problem. One particular application of extended Kalman filter is the joint state and parameter estimation of a linear system. The unknown parameters can be treated as additional state variables resulting in a nonlinear model, since these new states are multiplied by the original state variables in the model. Apply the extended Kalman filter to the augmented nonlinear model, we obtain a simultaneous state and parameter estimation.

To simplify notation and computations we note that in (3.3) capital letter D represents a diagonal matrix, and we define the operation d=diag$\{D\}$ where d is a vector that contains the diagonal elements of the square matrix D. Also D=diag$\{d\}$, denotes the diagonal matrix D with vector d being its diagonal line. The same notation is used for the state vector x, with X=diag$\{x\}$ and x=diag$\{X\}$.

By treating the parameter d as additional states of the original model, and define the augmented state vector z as:

$$z(t) = \begin{bmatrix} x(t) \\ d(t) \end{bmatrix},$$

we can get an augmented nonlinear model as follows:

$$\begin{aligned} \dot{z}(t) &= f_a(z(t), u(t)) \quad (3.5)\\ y(t) &= h(z(t)) + v(t), \quad (3.6)\\ z(0) &\sim (\begin{bmatrix} 0 \\ d \end{bmatrix}, \begin{bmatrix} P_{x(0)} & 0 \\ 0 & 0 \end{bmatrix}), \quad v(t) \sim (0, R) \end{aligned}$$

where

$$\begin{aligned} f_a(z(t), u(t)) &= \begin{bmatrix} (A + D)x(t) + bu(t) \\ 0 \end{bmatrix}, \\ h(z(t)) &= Cx(t). \end{aligned}$$

The bottom part of $f_a(z(t), u(t))$ is zero, which corresponds to the derivative of $d(t)$. In the sequel, we can apply the continuous extended Kalman filtering algorithm on this model to get the estimation of the vector d. An alternative is to discretize model (3.5) and (3.6), then use the discrete extended Kalman filter to estimate d. We chose the discrete version instead of the continuous for computer implementation purpose.

The forward Euler's method is used to discretize the system. If T is the sample time, then the discretized version of (3.5) and (3.6) is given by:

$$\begin{aligned} z_{k+1} &= f_d(z_k, u_k) && (3.7)\\ y_k &= h(z_k) + v_k && (3.8)\\ z_0 &\sim \left(\begin{bmatrix} 0 \\ d \end{bmatrix}, \begin{bmatrix} P_{x(0)} & 0 \\ 0 & 0 \end{bmatrix}\right), \quad v_k \sim (0, R) \end{aligned}$$

where

$$\begin{aligned} f_d(z_k, u_k) &= T f_a(x_k, u_k) + z_k \\ &= \begin{bmatrix} (I + TA + TD_k)x_k + Tbu_k \\ d_k \end{bmatrix} \end{aligned}$$

Equation (3.7) and (3.8) describe the discrete nonlinear model. Applying the extended Kalman filtering algorithm to system (3.7) and (3.8), we get a joint state and parameter estimator as follows.
The signal model is given by:

$$\begin{aligned} z_{k+1} &= f_d(z_k, u_k) \\ y_k &= h(z_k) + v_k \\ z_0 &\sim \left(\begin{bmatrix} 0 \\ d \end{bmatrix}, \begin{bmatrix} P_{x(0)} & 0 \\ 0 & 0 \end{bmatrix}\right), \quad v_k \sim (0, R). \end{aligned}$$

The filter equations are given by:

- Initialization (*a priori* information)

$$P_0^- = \begin{bmatrix} P_{x(0)} & 0 \\ 0 & P_d \end{bmatrix}, \qquad \hat{z}_0^- = \begin{bmatrix} 0 \\ \bar{d} \end{bmatrix}$$

- Measurement Update

$$\begin{aligned} K_k &= P_k^- H^T (H P_k^- H^T + R)^{-1} \\ P_k &= P_k^- - K_k H P_k^- \\ \hat{z}_k &= \hat{z}_k^- + K_k[y_k - h(\hat{z}_k^-)] \end{aligned}$$

- Time Update

$$\begin{aligned} P_{k+1}^- &= J_k P_k J_k^T && (3.9)\\ \hat{z}_{k+1}^- &= f_d(\hat{z}_k, u_k) && (3.10) \end{aligned}$$

- Jacobians

$$J_k = \left.\frac{\partial f_d(z, u)}{\partial z}\right|_{z=\hat{z}_k} = \begin{bmatrix} (A + \hat{D}_k)T + I & \hat{X}_k T \\ 0 & I \end{bmatrix}$$

$$H = \left.\frac{\partial h(z)}{\partial z}\right|_{z=\hat{z}_k^-} = [C \quad 0],$$

where

$$\hat{D}_k = \mathrm{diag}\{\hat{d}(k)\}, \qquad \hat{X}_k = \mathrm{diag}\{\hat{x}(k)\}$$

$\hat{z}_k$ denotes the estimated value of z_k. The solution of the inverse problem, i.e. the estimation of vector d, is contained in the bottom part of the vector $\hat{z}_k$. The above estimator equations contain a time update and measurement update. In the simulation, these two updates do not occur with the same rate. We assume that the time update is faster than the measurement update. It is useful to rewrite the estimator equations in terms of the original state and parameter vectors.
Signal Model:

$$\begin{aligned} x_{k+1} &= (I + TA + TD)x_k + Tbu_k \\ y_k &= Cx_k + v_k \\ x_0 &\sim (0, P_{x(0)}), \quad v_k \sim (0, R). \end{aligned}$$

Filter Equations:

- Initialization

$$P_0^- = \begin{bmatrix} P_{x(0)} & 0 \\ 0 & P_d \end{bmatrix}, \qquad \begin{bmatrix} \hat{x}_0^- \\ \hat{d}_0^- \end{bmatrix} = \begin{bmatrix} 0 \\ \bar{d} \end{bmatrix}$$

- Measurement Update

$$K_k = P_k^- H^T (H P_k^- H^T + R)^{-1} = \begin{bmatrix} K_k^1 \\ K_k^2 \end{bmatrix}$$

$$\begin{aligned} P_k &= P_k^- - K_k H P_k^- \\ \hat{x}_k &= \hat{x}_k^- + K_k^1 (y_k - C\hat{x}_k^-) \\ \hat{d}_k &= \hat{d}_k^- + K_k^2 (y_k - C\hat{x}_k^-) \end{aligned}$$

- Time Update

$$\begin{aligned} P_{k+1}^- &= J_k P_k J_k^T \\ \hat{x}_{k+1}^- &= (I + TA + T\hat{D}_k)\hat{x}_k + Tbu_k \\ \hat{d}_{k+1}^- &= \hat{d}_k \end{aligned}$$

Preliminary investigation shows that the convergence of the extended Kalman filtering algorithm is not very good. In [7] a detailed analysis of the asymptotic behavior of the extended Kalman filter as a parameter estimator is given. It is shown that in general, the estimates may be biased or divergent and the reasons for this behavior are studied in [7]. However, for the case of *deterministic model*, this process is guaranteed to converge. According to *Theorem 6.1* in [7], if the extended Kalman filter is complemented with a projection facility to keep $\hat{d}_k$ in the stable region of the system, then the estimates $\hat{d}_k$ will converge with probability 1 to the point $\hat{d}$ with $\hat{d}$ satisfying the transfer function equivalence,

$$C(sI - (A + D))^{-1} b = C(sI - (A + \hat{D}))^{-1} b.$$

At this point, we have to assume enough input output data to guarantee the identifiability of the system because when we can only get information from the detectors, transfer function match is the best we can do. The projection algorithm suggested in [6] turns out to be very simple. At each step of $\hat{d}_k$ update, do the following,

$$\begin{aligned} \bar{d}_k &= \hat{d}_k^- + K_k^2 (y_k - C\hat{x}_k^-) \\ \hat{d}_k &= \begin{cases} \bar{d}_k, & |\mathrm{eig}(I + TA + T\bar{D}_k)| < 1 \\ \hat{d}_k^-, & |\mathrm{eig}(I + TA + T\bar{D}_k)| \geq 1 \end{cases} \end{aligned}$$

where eig$(\cdot)$ are the eigenvalues of a matrix. The idea of the projection algorithm is: if the update of $\hat{d}_k$ goes out of the stability region, we just ignore this measurement. The projection facility guarantees the stability of the extended Kalman filter and improves its convergence behavior. However, when the dimension of the problem gets higher, the convergence rate is very slow. In order to accelerate the convergence speed, we examine the second order filter. In the model linearization of the extended Kalman filter, the power series are truncated after the first order terms so the extended Kalman filter can be viewed as a first order approximated filter. Obviously, considering higher order terms in the Taylor series can lead to a more accurate filter. This approach is specifically attractive in our case, because our state-space model of the inverse problem is a bilinear system (3.1). It is clear that if we consider second order terms in the Taylor series of the model linearization, we will get an exact relationship instead of the approximate.

The second order filter we derived here is called second order conditional mean filter in [10]. All the computations involved are only the first and second order statistics. In order to get a close form solution, it is necessary to assume the magnitude of the estimation error is Gaussian distributed.

We define the conditional mean of z_{k+1} given the previous measurements $y_1, y_2, ..., y_k$ as

$$\begin{aligned} \hat{z}_{k+1}^- &= E\{z_{k+1}|y_1, y_2, ..., y_k\} \\ \tilde{z}_{k+1}^- &= z_{k+1} - \hat{z}_{k+1}^-. \end{aligned}$$

$\tilde{z}_{k+1}^-$ is the estimation error so the conditional error covariance matrix P_{k+1}^- is

$$P_{k+1}^- = E\{\tilde{z}_{k+1}^- \tilde{z}_{k+1}^{-T}|y_1, y_2, ..., y_k\}.$$

For simplicity, we use capital Y_k to denote all the previous measurements $y_1, y_2, ..., y_k$ up to time k. We define the first and second order moments of z_{k+1} conditioned on all the previous measurements Y_k and y_{k+1} as,

$$\begin{aligned} \hat{z}_{k+1} &= E\{z_{k+1}|Y_{k+1}\} \\ \tilde{z}_{k+1} &= z_{k+1} - \hat{z}_{k+1} \\ P_{k+1} &= E\{\tilde{z}_{k+1}\tilde{z}_{k+1}^T|Y_{k+1}\}. \end{aligned}$$

The assumption we need to derive the second order conditional mean filter is that the estimate error process $\tilde{z}_k$ conditioned on Y_k are assumed to be zero mean white Gaussian process.

For the nonlinear model (3.7), we have

$$\hat{z}_{k+1}^- = E\{z_{k+1}|Y_k\} = E\{f_d(z_k, u_k)|Y_k\} \qquad (3.11)$$

To proceed with (3.11), we expand the nonlinear function $f_d(z_k, u_k)$ into power series around the estimated value $\hat{z}_k$. Because the second derivative of $f_d(z_k, u_k)$ with respect to z_k is a three dimensional tensor, the rest of the derivation is fairly complicated. We present here a straight forward derivation which does not need the tensor operation but provides the same result. If we consider the second order terms in the Taylor series, we will have an exact relationship,

$$f_d(z_k, u_k) = f_d(\hat{z}_k, u_k) + J_k\tilde{z}_k + \begin{bmatrix} T\tilde{D}_k\tilde{x}_k \\ 0 \end{bmatrix}, \qquad (3.12)$$

where

$$\tilde{D}_k \;=\; D_k - \hat{D}_k, \qquad \tilde{x}_k = x_k - \hat{x}_k.$$

Taking expectations on both sides of (3.12) conditioned on Y_k, we have,

$$\hat{z}_{k+1}^- = f_d(\hat{z}_k, u_k) + \begin{bmatrix} TE\{\tilde{D}_k\tilde{x}_k|Y_k\} \\ 0 \end{bmatrix}. \qquad (3.13)$$

The quantity $E\{\tilde{D}_k\tilde{x}_k|Y_k\}$ is by definition part of the covariance matrix P_k, since

$$P_k = \begin{bmatrix} E\{\tilde{x}_k\tilde{x}_k^T|Y_k\} & E\{\tilde{x}_k\tilde{d}_k^T|Y_k\} \\ E\{\tilde{d}_k\tilde{x}_k^T|Y_k\} & E\{\tilde{d}_k\tilde{d}_k^T|Y_k\} \end{bmatrix} = \begin{bmatrix} P_k^{xx} & P_k^{xd} \\ P_k^{dx} & P_k^{dd} \end{bmatrix}.$$

Obviously, we have,

$$E\{\tilde{D}_k\tilde{x}_k|Y_k\} = \mathrm{diag}\{P_k^{xd}\}.$$

Therefore, by substituting into (3.13), we get

$$\hat{z}_{k+1}^- = f_d(\hat{z}_k, u_k) + \begin{bmatrix} T\mathrm{diag}\{P_k^{xd}\} \\ 0 \end{bmatrix}. \qquad (3.14)$$

Equation (3.14) is the time update of the conditional mean. Comparing with the extended Kalman filter time update (3.10), the difference is the additional term of the second order moment, namely diag$\{P_k^{xd}\}$. To simplify the notations in the following derivation, we denote $f_d(z_k, u_k) = f_d(k)$ and $f_d(\hat{z}_k, u_k) = \hat{f}_d(k)$. The time update of the conditional covariance matrix P_{k+1}^- can be derived as follows,

$$\begin{aligned} P_{k+1}^- &= E\{(z_{k+1} - \hat{z}_{k+1}^-)(z_{k+1} - \hat{z}_{k+1}^-)^T|Y_k\} \\ &= E\{[f_d(k) - \hat{z}_{k+1}^-][f_d(k) - \hat{z}_{k+1}^-]^T|Y_k\} \\ &= E\{f_d(k)f_d^T(k)|Y_k\} - \hat{z}_{k+1}^-\hat{z}_{k+1}^{-T}. \qquad (3.15) \end{aligned}$$

To proceed with $f_d(k)f_d^T(k)$ in (3.15), we expand $f_d(k)$ as in (3.12) to get

$$\begin{aligned} f_d(k)f_d^T(k) = {} & \hat{f}_d(k)\hat{f}_d^T(k) + J_k\tilde{z}_k\hat{f}_d^T(k) + \hat{f}_d(k)\tilde{z}_k^TJ_k^T \\ & + \begin{bmatrix} T\tilde{D}_k\tilde{x}_k \\ 0 \end{bmatrix}\hat{f}_d^T(k) + \hat{f}_d(k)\begin{bmatrix} T\tilde{D}_k\tilde{x}_k \\ 0 \end{bmatrix}^T \\ & + J_k\tilde{z}_k\begin{bmatrix} T\tilde{D}_k\tilde{x}_k \\ 0 \end{bmatrix}^T + \begin{bmatrix} T\tilde{D}_k\tilde{x}_k \\ 0 \end{bmatrix}\tilde{z}_k^TJ_k^T \\ & + J_k\tilde{z}_k\tilde{z}_k^TJ_k^T + \begin{bmatrix} T\tilde{D}_k\tilde{x}_k \\ 0 \end{bmatrix}\begin{bmatrix} T\tilde{D}_k\tilde{x}_k \\ 0 \end{bmatrix}^T. \end{aligned}$$

By taking expectations on both sides of the above equation conditioned on Y_k, and noticing that all the odd order moments go to zero due to the Gaussian assumptions of the density function, we left with five terms, namely,

$$\begin{aligned} E\{f_d(k)f_d^T(k)|Y_k\} = {} & \hat{f}_d(k)\hat{f}_d^T(k) + J_k E\{\tilde{z}_k\tilde{z}_k^T|Y_k\}J_k^T \\ & + \begin{bmatrix} T^2E\{\tilde{D}_k\tilde{x}_k\tilde{x}_k^T\tilde{D}_k|Y_k\} & 0 \\ 0 & 0 \end{bmatrix} \\ & + \begin{bmatrix} TE\{\tilde{D}_k\tilde{x}_k|Y_k\} \\ 0 \end{bmatrix} \hat{f}_d^T(k) \\ & + \hat{f}_d(k) \begin{bmatrix} TE\{\tilde{D}_k\tilde{x}_k|Y_k\} \\ 0 \end{bmatrix}^T \end{aligned} \quad (3.16)$$

To proceed with $\hat{z}_{k+1}^-\hat{z}_{k+1}^{-T}$ in (3.15), we use equation (3.13) to express this quantity as

$$\begin{aligned} \hat{z}_{k+1}^-\hat{z}_{k+1}^{-T} = {} & \hat{f}_d(k)\hat{f}_d^T(k) + \begin{bmatrix} TE\{\tilde{D}_k\tilde{x}_k|Y_k\} \\ 0 \end{bmatrix} \hat{f}_d^T(k) \\ & + \begin{bmatrix} T^2E\{\tilde{D}_k\tilde{x}_k|Y_k\}E\{\tilde{D}_k\tilde{x}_k|Y_k\} & 0 \\ 0 & 0 \end{bmatrix} \\ & + \hat{f}_d(k) \begin{bmatrix} TE\{\tilde{D}_k\tilde{x}_k|Y_k\} \\ 0 \end{bmatrix}^T . \end{aligned} \quad (3.17)$$

Subtracting (3.17) from (3.16), we get an expression for the covariance matrix P_{k+1}^-,

$$\begin{aligned} P_{k+1}^- = {} & J_k E\{\tilde{z}_k\tilde{z}_k^T|Y_k\}J_k^T + \begin{bmatrix} T^2E\{\tilde{D}_k\tilde{x}_k\tilde{x}_k^T\tilde{D}_k|Y_k\} & 0 \\ 0 & 0 \end{bmatrix} \\ & - \begin{bmatrix} T^2E\{\tilde{D}_k\tilde{x}_k|Y_k\}E\{\tilde{D}_k\tilde{x}_k|Y_k\} & 0 \\ 0 & 0 \end{bmatrix} . \end{aligned} \quad (3.18)$$

In equation (3.18), if we use S_k to denote matrix

$$T^2E\{\tilde{D}_k\tilde{x}_k\tilde{x}_k^T\tilde{D}_k|Y_k\} - T^2E\{\tilde{D}_k\tilde{x}_k|Y_k\}E\{\tilde{D}_k\tilde{x}_k|Y_k\}$$

and notice that the quantity $E\{\tilde{z}_k\tilde{z}_k^T|Y_k\}$ is just P_k by definition, then we get the time update of the conditional covariance matrix,

$$P_{k+1}^- = J_kP_kJ_k^T + \begin{bmatrix} S_k & 0 \\ 0 & 0 \end{bmatrix} . \quad (3.19)$$

Comparing (3.19) with (3.9), the difference is the additional term of the fourth order moment matrix S_k. To evaluate this matrix, we need the formula of the fourth-product moment of the zero mean Gaussian random variable. In general, suppose x_1, x_2, x_3, x_4 are jointly Gaussian random variables with zero mean, we have the following result,

$$\begin{aligned} E\{x_1x_2x_3x_4\} = {} & E\{x_1x_2\}E\{x_3x_4\} + E\{x_1x_3\}E\{x_2x_4\} \\ & + E\{x_1x_4\}E\{x_2x_3\}. \end{aligned}$$

Notice that in this formula, x_1, x_2, x_3, x_4 do not need to be a different variable. In particular, we have

$$\begin{aligned} E\{x_1^4\} &= 3E\{x_1^2\} \\ E\{x_1^2x_2^2\} &= E\{x_1^2\}E\{x_2^2\} + 2E\{x_1x_2\}E\{x_1x_2\}. \end{aligned}$$

With the help of the formula of the fourth-product moment, matrix S_k can be expressed by the elements of the matrix P_k as follows,

$$S_k(i,j) = P_k^{dd}(i,j)P_k^{xx}(i,j) + P_k^{xd}(i,j)P_k^{dx}(i,j), \quad (3.20)$$

where $A(i,j)$ is the (i,j)th element of matrix A. Equation (3.14) and (3.19) are the time update equations of the first and second order moments of z_{k+1} conditioned on measurements Y_k. The measurement update equations of the second order filter are the same as the extended Kalman filter. This is because the measurement model in our case is linear $h(z_k) = Cx_k$, the second order or the first order linearization of $h(z)$ makes no difference.

The advantage of this second order filter is that it achieves a much better result than the extended Kalman filter while the computational cost is slightly increased. We summarize the second order conditional mean filter as follows,

Signal Model:

$$\begin{aligned} x_{k+1} &= (I + TA + TD)x_k + Tbu_k \\ y_k &= Cx_k + v_k \\ x_0 &\sim (0, P_{x(0)}), \quad v_k \sim (0, R). \end{aligned}$$

Filter Equations:

- Initialization

$$P_0^- = \begin{bmatrix} P_{x(0)} & 0 \\ 0 & P_d \end{bmatrix}, \qquad \begin{bmatrix} \hat{x}_0^- \\ \hat{d}_0^- \end{bmatrix} = \begin{bmatrix} 0 \\ \bar{d} \end{bmatrix}$$

- Measurement Update

$$\begin{aligned} K_k &= P_k^-H^T(HP_k^-H^T + R)^{-1} \\ P_k &= P_k^- - K_kHP_k^- \\ \hat{x}_k &= \hat{x}_k^- + K_k^1(y_k - C\hat{x}_k^-) \\ \bar{d}_k &= \hat{d}_k^- + K_k^2(y_k - C\hat{x}_k^-) \\ \hat{d}_k &= \begin{cases} \bar{d}_k, & |\mathrm{eig}(I + TA + T\bar{D}_k)| < 1 \\ \hat{d}_k^-, & |\mathrm{eig}(I + TA + T\bar{D}_k)| \geq 1 \end{cases} \end{aligned}$$

- Time Update

$$P_{k+1}^- = J_kP_kJ_k^T + \begin{bmatrix} S_k & 0 \\ 0 & 0 \end{bmatrix}$$

(S_k is given by equation (3.20))

$$\hat{x}_{k+1}^- = (I + TA + T\hat{D}_k)\hat{x}_k + T\mathrm{diag}\{P_k^{xd}\} + Tbu_k$$

4 Simulation Results

In this paper we have chosen a 80mm×80mm 2D square medium as the test object, shown in Figure 4.1. The scattering and absorption coefficients of the homogeneous medium are 1.0 and 0.01 mm^{-1}, respectively. We put three absorbers (A,B,C) inside the medium. Their positions, sizes and absorption coefficients are:

A: (30,40), 15mm×15mm, $0.02\mathrm{mm}^{-1}$
B: (50,30), 10mm×10mm, $0.015\mathrm{mm}^{-1}$
C: (25,15), 5mm×5mm, $0.02\mathrm{mm}^{-1}$.

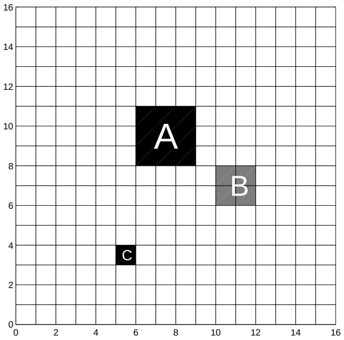

Figure 4.1: The Test Medium: 80mm×80mm

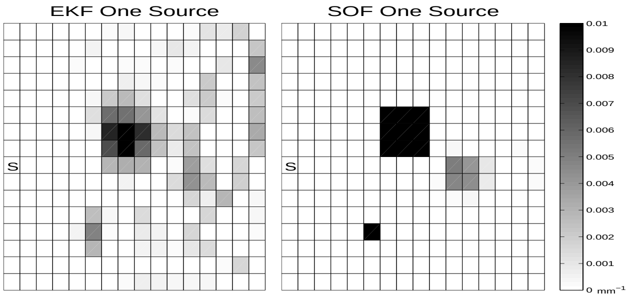

Figure 4.2: Comparision of Estimated $\delta\mu_a$ between EKF and SOF with One Source "S"

In Figure 4.1, the value of the difference $\delta\mu_a$ in the absorption coefficient is represented by the gray scale with white representing 0 and black representing 0.01mm^{-1}. As shown in Figure 4.1, the FEM-mesh is 16×16 so the grid size is 5mm×5mm. The dimension of our state-space model is 16×16=256. Therefore the vector $\delta\mu_a$ or d we want to identify has 256 elements.

The detectors are assumed to locate all along the boundary of the medium, so for the 16×16 case there are 60 detectors totally. The source set-up is either single source or four sources located also on the boundary of the medium. We discretize the system with 10 pico second. The measurement data are read out from the detectors every 100 pico second. The simulation results of the extended Kalman filter (EKF) and the second order filter (SOF) are shown in Figure 4.2 and Figure 4.3.

As shown in the simulation results, the extended Kalman filter with one source performs poorly whereas the second order filter with four sources performs extremely well. Both extemded Kalman filter with four sources and second order filter with single source can improve the performance of the extended Kalman filter with single source while the improvement of using higher order filter is more significant than using multiple sources.

5 Conclusion

In this paper, we use the nonlinear filtering method to solve the inverse problem in the diffusion equation. This image reconstruction problem is presented using a state-space model and the associated signal processing method. Our method is tested against a two-dimensional real case and gives good results. One of the future research problems under investigation is the inverse problem of three-dimensional object.

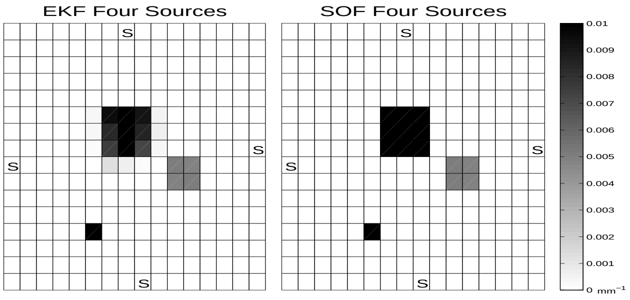

Figure 4.3: Comparison of Estimated $\delta\mu_a$ between EKF and SOF with Four Sources "S"

References

[1] B. D. O. Anderson and J. B. Moore, *Optimal Filtering*, Prentice Hall, Englewood Cliffs, New Jersey, 1979.

[2] K. J. Aström and P. Eykhoff, "System Identification-A Survey", *Automatica*, Vol. 7, pp.123-162, 1971.

[3] C. Johnson, *Numerical Solution of Partial Differential Equations by the Finite Element Method*, Cambridge University Press, Lund, Sweden, 1987.

[4] M. V. Klibanov, T. R. Lucas and R. M. Frank, "New Imaging Algorithm In Diffusion Tomography", *preprint.*

[5] F. L. Lewis, *Optimal Estimation with an Introduction to Stochastic Control Theory*, Wiley, New York, 1986.

[6] L. Ljung and T. Söderström, *Theory and Practice of Recursive Identification*, MIT Press, Cambridge, MA, 1982.

[7] L. Ljung, "Asymptotic Behavior of the Extended Kalman Filter as a Parameter Estimator for Linear Systems", *IEEE Trans. Automat. Control*, vol. AC-24, pp. 36-50, Feb. 1979.

[8] P. S. Maybeck, *Stochastic Models, Estimation, and Control*, Vol. 2, Mathematics in Science and Engineering, Series Vol. 141-2, Academic Press, New York, 1979.

[9] R. Model, R. Hunlich, D. Richter, H. Rinneberg, H. Wabnitz and M.Walzel, "Imaging in Random Media: Simulating Light Transport By Numerical Integration of the Diffusion Equation", *SPIE*, Vol. 2336, pp.11-22, Feb. 1995.

[10] A. P. Sage and J. L. Melsa, *Estimation Theory with Applications to Communications and Control*, McGraw-Hill Book Company, New York, 1971

[11] Q. Su, V. L. Syrmos and D. Y. Y. Yun "A Numerical Algorithm for the Diffusion Equation Using 3D FEM and the Arnoldi Method", *Proc. American Control Conference*, San Diego, CA, June 1999.