ÖAW AUSTRIAN ACADEMY OF SCIENCES

Johann Radon Institute for
Computational and Applied Mathematics
Austrian Academy of Sciences (ÖAW)

RICAM JOHANN-RADON-INSTITUTE FOR COMPUTATIONAL AND APPLIED MATHEMATICS

State estimation approach to nonstationary inverse problems: discretization error and filtering problem

H.K. Pikkarainen

RICAM-Report 2005-20

State estimation approach to nonstationary inverse problems: discretization error and filtering problem

Hanna Katriina Pikkarainen[1]

October 12, 2005

Abstract. We examine a certain class of nonstationary inverse problems. We view them as a state estimation problem. The time evolution of the state of the system is modeled by a stochastic differential equation. The observation equation is linear with additive measurement noise. We introduce the time discrete infinite dimensional state estimation system concerning the problem. For computational reasons the space discretization is performed. The discretization errors in the discretized state evolution and observation equations are analysed stochastically. The solution to the corresponding finite dimensional filtering problem is presented.

Keywords: nonstationary inverse problem, state estimation, discretization error, filtering problem

AMS subject classifications: 62M20, 60G35, 60H15

1 Introduction

In practical measurements of physical quantities we have directly observable quantities and others that cannot be observed. If some of the unobservable quantities are of our primary interest, we are dealing with an inverse problem. In that case, we need to discover how to compute estimates for the quantities of primary interest from the observed values of the observable quantities, the measured data. In several applications one encounters a situation in which measurements are done in a nonstationary environment. More precisely, it may happen that the physical quantities that are in the focus of our primary interest are time dependent and the measured data depends on these quantities at different time instants. Inverse problems of this type are called nonstationary inverse problems. They are often viewed as a state estimation problem. Then the quantities in the measurement setting are treated as stochastic processes. The randomness describes our degree of knowledge concerning their realizations. Our information about their values is coded into their distributions. Therefore the randomness is due to the lack of information, not to the intrinsic randomness of the quantities in the measurement setting. The time evolution of the quantities of primary interest, the state of the system, is described by a stochastic differential equation referred to as the state evolution equation. The measurements are modeled by an observation equation containing the measurement noise. The solution to the state estimation problem is the conditional expectation of the quantities of primary interest with respect to the measured data. If our motive is, for instance, to monitor the quantities of primary interest in real time, we are dealing with a filtering problem in which the estimator is based on the current history of the measurement process. An example of nonstationary inverse problems is the process tomography problem where the consistence

[1]Johann Radon Institute for Computational and Applied Mathematics, Austrian Academy of Sciences, Altenbergerstrasse 69, A-4040 Linz, Austria. Email: hanna.pikkarainen@oeaw.ac.at

of a moving compound in a pipeline is estimated by doing measurements on the boundary of the pipe.

Often in state estimation approach the time variable is assumed to be discrete and the space variable to be finite dimensional. This is convenient from the practical point of view. Observations are usually done at discrete time instants and the computation requires space discretization. In many applications, it is assumed that the discretized version of the infinite dimensional state estimation system represents the reality. Nevertheless, discretization causes always an error, which should be included into the state estimation system. If we analyse the continuous infinite dimensional state evolution and observation equations, we may be able to present the distribution of the discretization error.

The state estimation approach to nonstationary inverse problems have been studied in the literature. In some of the publications the time evolution of the state of the system has not been in the main interest. In those the random walk model has been used as a state evolution equation and the observation equation is introduced more precisely. The use of the random walk model is justified by assuming that changes in the target during the measurements can be explained by an additional random variable. However, there are publications in which the time evolution of the target is discussed and modeled by a stochastic differential equation. At least papers involving dynamic electric wire tomography [1], electrical impedance process tomography [18, 14, 13, 16, 17, 15, 12, 3, 19] and optimal control in process tomography using EIT measurements [8, 9, 11, 10] has been published. The state evolution equation in those articles is a stochastic partial differential equation with a Brownian motion. The publications stated above are more related to the application and the mathematical accuracy in the presentation is slightly inadequate. In the article [2] some theoretical aspects concerning dynamical electric wire tomography and EIT with nonstationary object are presented but the error caused by space discretization is not taken into account. Nevertheless, discretization errors have been considered in the finite dimensional context in the book [3]. The discratization error is approximated by the error between the solution in a coarse and dense mesh. As far as we know, the doctoral dissertation of the author [6] is the only publication containing infinite dimensional state estimation systems and discretization errors. In that thesis the electrical impedance process tomography problem has been studied with mathematical accuracy. The time evolution of the concentration distribution is modeled by the stochastic convection–diffusion equation. The discretized state estimation system containing the discretization error is presented. In this article, we extend the results concerning electrical impedance process tomography to a certain class of nonstationary inverse problems. We introduce the time discrete infinite dimensional state estimation system concerning the problem. For computational reasons the space discretization is performed. The discretization errors in the discretized state evolution and observation equations are analysed stochastically. The solution to the corresponding finite dimensional filtering problem is presented.

2 State estimation

Let $D \subset \mathbb{R}^d$ be a domain that corresponds to the object of interest. We denote by $X = X(t,x)$, $x \in D$, a distributed parameter describing the state of the object – the unknown distribution of a physical target – at time $t \geq 0$. The parameter X is supposed to be unobservable. We assume that $X(t,\cdot)$ belong to a real separable Hilbert space H for every $t \geq 0$, for example $H = L^2(D)$. We suppose that instead of being a deterministic function X is a stochastic process $\{X(t)\}_{t\geq 0}$ in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with values in H. We assume that we have a model for the time evolution of the parameter X. The stochastic nature of the parameter X allows us to incorporate phenomena such as modelling uncertainties into the time evolution model. Let $T > 0$ and $\{\mathcal{F}_t\}_{t\in[0,T]}$ be a normal filtration in $(\Omega, \mathcal{F}, \mathbb{P})$ [7, Section 3.3]. Let Q be a positive self-adjoint trace class operator from H to itself with trivial null space, i.e., $\operatorname{Ker} Q = \{0\}$ and $W(t)$, $t \in [0,T]$, a Q-Wiener process in $(\Omega, \mathcal{F}, \mathbb{P})$ with values in H with respect to the filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$ [7, Section 4.1]. We assume that the time evolution of the parameter X can be modeled by a stochastic differential equation

$$dX(t) = AX(t)dt + dW(t)$$

for every $t > 0$ with the initial value $X(0) = X_0$ where $A : \mathcal{D}(A) \subset H \to H$ is a densely defined sectorial operator [4, Chapter 2]. Under some spesific assumptions the operator A can be, for example, a second order elliptic operator [4, Chapter 3]. Possible boundary conditions on the boundary of D are then included in the domain $\mathcal{D}(A)$ of the operator A. The term $dW(t)$ is a source term representing possible modelling errors in the time evolution model.

Let $Y = Y(t)$ denote a quantity that is directly observable at time $t \geq 0$. The observable quantity Y is described by the stochastic process $\{Y(t)\}_{t\in[0,T]}$ with values in $\mathbb{R}^L$. We assume that the dependence of Y upon the state X is known up to an observation noise. The measurement process is modeled by the equation

$$Y(t) = BX(t) + S(t)$$

where $B : H \to \mathbb{R}^L$ is a bounded linear operator and $S(t)$, $t \in [0,T]$, is an $\mathbb{R}^L$-valued stochastic process. For example, in X-ray tomography the measurement setting can be described by a linear mathematical model. The stochastic process S represents possible measurement errors.

Then the state estimation system we are interested in consists of the equations

$$dX(t) = AX(t)dt + dW(t), \quad t > 0, \tag{1}$$
$$X(0) = X_0, \tag{2}$$
$$Y(t) = BX(t) + S(t), \quad t > 0. \tag{3}$$

Equation (1) is called the state evolution equation and Equation (3) the observation equation. The state estimation problem can be formulated as follows: *Estimate the state X satisfying the state evolution equation (1) based on the observed values of Y.* Estimators of the state X are calculated by taking the conditional expectation of X with respect to the measurements.

We assume that the measurements are done in time instants $0 < t_1 < \ldots < t_n \leq T$. Then the state estimation system is

$$\begin{aligned} dX(t) &= AX(t)dt + dW(t), \quad t > 0, \\ X(0) &= X_0, \\ Y(t_k) &= BX(t_k) + S(t_k), \qquad k = 1, \ldots, n. \end{aligned}$$

We are interested in a real-time monitoring for the parameter X. Therefore we should be able to solve the filtering problem $\mathbb{E}(X(t_k) \mid Y(t_l),\ l \leq k)$ for all $k = 1, \ldots, n$ [7, Section 1.3]. For that reason we need to solve the stochastic differential equation (1) with the initial value (2) and to present the time discrete state evolution equation for the process X.

3 Time discrete state evolution equation

By the assumptions the operator A generates a strongly continuous analytic semigroup $\{U(t)\}_{t\geq 0}$ [4, Propositions 2.1.1 and 2.1.4]. Under some assumptions of the initial value X_0 the stochastic differential equation (1) has the weak solution [6, Definition 4.48].

Theorem 1. [7, Theorem 5.4] *If X_0 is $\mathcal{F}_0$-measurable, the stochastic differential equation (1) has the weak solution $X(t)$, $t \in [0,T]$, which is the predictable process given by the formula*

$$X(t) = U(t)X_0 + \int_0^t U(t-s)\, dW(s)$$

for all $t \in [0,T]$ almost surely.

We assume that the initial value X_0 is $\mathcal{F}_0$-measurable. Then the weak solution of the state evolution equation (1) is the predictable process given by the formula

$$X(t) = U(t)X_0 + \int_0^t U(t-s)\, dW(s) \tag{4}$$

for all $t \in [0,T]$ almost surely. If the initial value X_0 is a Gaussian random variable with mean x_0 and covariance Γ_0, the process X has a Gaussian modification by [6, Lemma 4.42] and [7, Theorem 5.2]. Furthermore, the mean of the Gaussian modification is $\mathbb{E}(X(t)) = U(t)x_0$ and the covariance operator is

$$\mathrm{Cov}(X(t)) = U(t)\Gamma_0 U^*(t) + \int_0^t U(t-s)QU^*(t-s)\, ds \tag{5}$$

for all $t \in [0,T]$ where $U^*(t)$ is the Hilbert adjoint of the operator $U(t)$.

The measurements are done in time instants $0 < t_1 < \ldots < t_n \leq T$. We use the notation $t_0 := 0$ and $X_k := X(t_k)$ and $\Delta_{k+1} := t_{k+1} - t_k$ for all $k = 0, \ldots, n-1$. Then the time discrete state evolution equation for the parameter X is

$$X_{k+1} = U(\Delta_{k+1})X_k + W_{k+1} \tag{6}$$

for all $k = 0, \ldots, n-1$ almost surely where

$$W_{k+1} := \int_{t_k}^{t_{k+1}} U(t_{k+1}-s)\, dW(s)$$

by [6, Theorem 4.39]. The term W_{k+1} can be seen as a state noise for all $k = 0, \ldots, n-1$. The state noise W_{k+1} is a Gaussian random variable with mean 0 and covariance operator

$$\mathrm{Cov}(W_{k+1}) = \int_{t_k}^{t_{k+1}} U(t_{k+1}-s)QU^*(t_{k+1}-s)\, ds \tag{7}$$

and it is independent of $\mathcal{F}_{t_k}$ for all $k = 0, \ldots, n-1$ by [6, Lemma 4.42] and [7, Theorem 5.2]. Thus X_k and W_{k+1} are independent for all $k = 0, \ldots, n-1$. Furthermore, the state noises at different time instants are uncorrelated since

$$\begin{aligned}&\mathrm{Cor}(W_{k+1}, W_{l+1})\\ &= \int_0^{t_{k+1}\wedge t_{l+1}} \chi_{[t_k,t_{k+1}]}(s)\chi_{[t_l,t_{l+1}]}(s)U(t_{k+1}-s)QU^*(t_{l+1}-s)\, ds = 0\end{aligned}$$

for all $k \neq l$ by [7, Proposition 4.13]. By the linearity of the state evolution equation (1) the time discrete state evolution equation is exact. Hence in time discretization we have not made an error.

There are some parameters in our time evolution model which we can choose rather freely and still have the time discrete state evolution equation (6). The operator A should be a densely defined sectorial operator. The covariance operator Q of the Wiener process can be an arbitrary positive self-adjoint trace class operator from H to itself with $\operatorname{Ker} Q = \{0\}$ [6, Proposition 4.17]. The natural choice of the filtration is the filtration defined by the Wiener process, i.e., $\mathcal{F}_t^W = \sigma(W(s),\ s \leq t)$ for all $t \in [0, T]$. Since the filtration should be normal, the augmented filtration $\{\mathcal{F}_t^{\overline{W,\mathbb{P}}}\}_{t\in[0,T]}$ is an appropriate choice assuming that the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is complete [6, Proposition 4.35]. We want that the initial value X_0 is a Gaussian H-valued $\mathcal{F}_0$-measurable function with mean x_0 and covariance Γ_0. Then the mean x_0 can be an arbitrary element in H and the covariance operator Γ_0 has same requirements as Q [6, Proposition 4.17].

4 Time discrete state estimation system

We use the notation $S^k := S(t_k)$ and $Y^k := Y(t_k)$ for all $k = 1, \ldots, n$. Then the time discrete state estimation system we are interested in is

$$X_{k+1} = U(\Delta_{k+1})X_k + W_{k+1}, \quad k = 0, \ldots, n-1, \tag{8}$$

$$Y^k = BX_k + S^k, \quad k = 1, \ldots, n. \tag{9}$$

Both the time discrete state evolution and observation equation (8) and (9), respectively, are linear. To be able to solve the filtering problem $\mathbb{E}(X_k \mid Y^l,\ l \leq k)$ computationally we need to discretize in space the time discrete state estimation system (8)–(9). The realizations of the parameter X are in the Hilbert space H. We need to choose a finite dimensional subspace of H and assume that the realizations of the parameter X are in that subspace. This causes a discretization error. Usually the discretization error is ignored in numerical implementations. The discretized state estimation system is assumed to represent the reality. We want to analyse the stochastic nature of the discretization error.

5 Space discretization

Let $\{V_m\}_{m=1}^{\infty}$ be a sequence of finite dimensional subspaces of H such that $V_m \subset V_{m+1}$ for all $m \in \mathbb{N}$ and $\overline{\cup V_m} = H$. Since H is a separable Hilbert space, there exists such a sequence, for example, V_m may be the subspace spanned by the m first functions in an orthonormal basis of H. Let $\{\varphi_l^m\}_{l=1}^{N_m}$ be an orthonormal basis of V_m for all $m \in \mathbb{N}$. We denote by $(\cdot, \cdot)$ the inner product in H. We define the orthogonal projection $P_m : H \to V_m$ for $m \in \mathbb{N}$ by

$$P_m f = \sum_{l=1}^{N_m} (f, \varphi_l^m)\varphi_l^m$$

for all $f \in H$. The subspaces V_m are appropriate discretization spaces if $P_m f \to f$ in H as $m \to \infty$ for all $f \in H$, i.e., the orthogonal projections P_m converge strongly to the identity operator.

Let $X : (\Omega, \mathcal{F}, \mathbb{P}) \to H$ be a random variable. Then for all $\omega \in \Omega$

$$P_m X(\omega) = \sum_{l=1}^{N_m} (X(\omega), \varphi_l^m)\varphi_l^m =: \sum_{l=1}^{N_m} (X^m(\omega))_l \varphi_l^m$$

where $X^m(\omega) := ((X(\omega), \varphi_1^m), \ldots, (X(\omega), \varphi_{N_m}^m))^T$ is an $\mathbb{R}^{N_m}$-valued random variable. We view X^m as a discretized version of the random variable X at the discretization level m. If X is a Gaussian random variable with mean $\bar{x}$ and covariance Γ, then X^m is also Gaussian [5, Theorem A.5]. Furthermore, the mean of X^m is $\mathbb{E}X^m = ((\bar{x}, \varphi_1^m), \ldots, (\bar{x}, \varphi_{N_m}^m))^T$ and the covariance matrix is defined by $(\operatorname{Cov} X^m)_{ij} := (\Gamma\varphi_i^m, \varphi_j^m)$ since

$$\mathbb{E}(X^m)_i (X^m)_j = \mathbb{E}(X, \varphi_i^m)(X, \varphi_j^m) = (\Gamma\varphi_i^m, \varphi_j^m) + (\bar{x}, \varphi_i^m)(\bar{x}, \varphi_j^m)$$

for all $i, j = 1, \ldots, N_m$ [7, Section 1.2].

5.1 State evolution equation

We want to discretize the time discrete state evolution equation (8). We use the discretization level m. We form an evolution equation for the discrete $\mathbb{R}^{N_m}$-valued stochastic process $\{X_k^m\}_{k=0}^n$ where $X_k^m := ((X_k, \varphi_1^m), \ldots, (X_k, \varphi_{N_m}^m))^T$ for all $k = 0, \ldots, n$. By using the time discrete state evolution equation (8),

$$\begin{aligned}
(X_{k+1}^m)_i &= (U(\Delta_{k+1})X_k + W_{k+1}, \varphi_i^m) \\
&= (U(\Delta_{k+1})P_m X_k, \varphi_i^m) + (U(\Delta_{k+1})(I - P_m)X_k, \varphi_i^m) + (W_{k+1}, \varphi_i^m) \\
&= \sum_{l=1}^{N_m} (U(\Delta_{k+1})\varphi_l^m, \varphi_i^m)(X_k^m)_l + (E_{k+1}^m)_i + (W_{k+1}^m)_i
\end{aligned}$$

for all $i = 1, \ldots, N_m$ and $k = 0, \ldots, n-1$ almost surely where the discrete stochastic process

$$E_{k+1}^m = ((X_k, (I - P_m)U^*(\Delta_{k+1})\varphi_1^m), \ldots, (X_k, (I - P_m)U^*(\Delta_{k+1})\varphi_{N_m}^m))^T$$

represent the discretization error and W_{k+1}^m is the state noise vector. Thus the discretized state evolution equation is

$$X_{k+1}^m = A_{k+1}^m X_k^m + E_{k+1}^m + W_{k+1}^m \tag{10}$$

for all $k = 0, \ldots, n-1$ almost surely where the matrix A_{k+1}^m is defined by

$$(A_{k+1}^m)_{ij} := (U(\Delta_{k+1})\varphi_j^m, \varphi_i^m) \tag{11}$$

for all $i, j = 1, \ldots, N_m$. We want to define the statistical quantities of the discrete stochastic processes $\{E_{k+1}^m\}_{k=0}^{n-1}$ and $\{W_{k+1}^m\}_{k=0}^{n-1}$.

The state noise W_{k+1} is a Gaussian random variable with mean 0 and covariance given by Formula (7) for all $k = 0, \ldots, n-1$. Hence the state noise vector W_{k+1}^m is Gaussian with mean 0 and covariance matrix

$$(\operatorname{Cov}(W_{k+1}^m))_{ij} = \int_{t_k}^{t_{k+1}} \left(U(t_{k+1} - s)QU^*(t_{k+1} - s)\varphi_i^m, \varphi_j^m\right) ds$$

for all $i, j = 1, \ldots, N_m$ and $k = 0, \ldots, n-1$. We define the matrix $Q_{k,l}^m(s)$ by

$$(Q_{k,l}^m(s))_{ij} := (U(t_k - s)QU^*(t_l - s)\varphi_i^m, \varphi_j^m)$$

for all $i, j = 1, \ldots, N_m$, $k, l = 0, \ldots, n-1$ and $s \in [0, t_k \wedge t_l]$ where $s \wedge t := \min(s,t)$. Then

$$\operatorname{Cov}(W_{k+1}^m) = \int_{t_k}^{t_{k+1}} Q_{k+1,k+1}^m(s)\, ds \tag{12}$$

for all $k = 0, \ldots, n-1$ where the integration is done componentwise. Since the state noises W_k and W_l are uncorrelated for all $k \neq l$, by the Gaussianity the state noise vectors W_k^m and W_l^m are independent if $k \neq l$.

We use our knowledge of the stochastic behaviour of the continuous state evolution equation (4) for the examination of the discretization error E_{k+1}^m for all $k = 0, \ldots, n-1$. The process X_k has a Gaussian modification with mean $U(t_k)x_0$ and covariance given by Formula (5) where $t = t_k$ for all $k = 0, \ldots, n$. Hence the discretization error E_{k+1}^m has a Gaussian version for all $k = 0, \ldots, n-1$. The mean of the Gaussian version is given by

$$(\mathbb{E}E_{k+1}^m)_i = (\mathbb{E}X_k, (I - P_m)U^*(\Delta_{k+1})\varphi_i^m) = (U(t_k)x_0, (I - P_m)U^*(\Delta_{k+1})\varphi_i^m)$$

for all $i = 1, \ldots, N_m$ and $k = 0, \ldots, n-1$. Since

$$(I - P_m)U^*(t)f = U^*(t)f - \sum_{l=1}^{N_m} (U^*(t)f, \varphi_l^m)\varphi_l^m \tag{13}$$

for all $f \in H$ and $t \in [0, T]$,

$$(\mathbb{E}E_{k+1}^m)_i = (U(t_{k+1})x_0, \varphi_i^m) - \sum_{l=1}^{N_m} (U(\Delta_{k+1})\varphi_l^m, \varphi_i^m)(U(t_k)x_0, \varphi_l^m)$$

for all $i = 1, \ldots, N_m$ and hence

$$\mathbb{E}E_{k+1}^m = \begin{bmatrix} (U(t_{k+1})x_0, \varphi_1^m) \\ \vdots \\ (U(t_{k+1})x_0, \varphi_{N_m}^m) \end{bmatrix} - A_{k+1}^m \begin{bmatrix} (U(t_k)x_0, \varphi_1^m) \\ \vdots \\ (U(t_k)x_0, \varphi_{N_m}^m) \end{bmatrix} \tag{14}$$

for all $k = 0, \ldots, n-1$. The covariance matrix of the Gaussian version is given by

$$\begin{aligned}
&(\operatorname{Cov} E_{k+1}^m)_{ij} \\
&= (\operatorname{Cov}(X_k)(I - P_m)U^*(\Delta_{k+1})\varphi_i^m, (I - P_m)U^*(\Delta_{k+1})\varphi_j^m) \\
&= (U(t_k)\Gamma_0 U^*(t_k)(I - P_m)U^*(\Delta_{k+1})\varphi_i^m, (I - P_m)U^*(\Delta_{k+1})\varphi_j^m) + \\
&\quad + \int_0^{t_k} \big(U(t_k - s)QU^*(t_k - s)(I - P_m)U^*(\Delta_{k+1})\varphi_i^m, (I - P_m)U^*(\Delta_{k+1})\varphi_j^m\big)\, ds
\end{aligned}$$

for all $i, j = 1, \ldots, N_m$ and $k = 0, \ldots, n-1$. By using Formula (13),

$$\begin{aligned}
&(U(t_k)\Gamma_0 U^*(t_k)(I - P_m)U^*(\Delta_{k+1})\varphi_i^m, (I - P_m)U^*(\Delta_{k+1})\varphi_j^m) \\
&= (U(t_{k+1})\Gamma_0 U^*(t_{k+1})\varphi_i^m, \varphi_j^m) + \\
&\quad - \sum_{l=1}^{N_m} (U(\Delta_{k+1})\varphi_l^m, \varphi_j^m)(U(t_k)\Gamma_0 U^*(t_{k+1})\varphi_i^m, \varphi_l^m) + \\
&\quad - \sum_{l=1}^{N_m} (U(\Delta_{k+1})\varphi_l^m, \varphi_i^m)(U(t_k)\Gamma_0 U^*(t_{k+1})\varphi_j^m, \varphi_l^m) + \\
&\quad + \sum_{l,p=1}^{N_m} (U(\Delta_{k+1})\varphi_l^m, \varphi_i^m)(U(\Delta_{k+1})\varphi_p^m, \varphi_j^m)(U(t_k)\Gamma_0 U^*(t_k)\varphi_l^m, \varphi_p^m)
\end{aligned}$$

and

$$\begin{aligned}
&\int_0^{t_k} \big(U(t_k-s)QU^*(t_k-s)(I-P_m)U^*(\Delta_{k+1})\varphi_i^m,(I-P_m)U^*(\Delta_{k+1})\varphi_j^m\big)\,ds\\
&=\int_0^{t_k} \big(U(t_{k+1}-s)QU^*(t_{k+1}-s)\varphi_i^m,\varphi_j^m\big)\,ds+\\
&-\sum_{l=1}^{N_m}(U(\Delta_{k+1})\varphi_l^m,\varphi_j^m)\int_0^{t_k}(U(t_k-s)QU^*(t_{k+1}-s)\varphi_i^m,\varphi_l^m)\,ds+\\
&-\sum_{l=1}^{N_m}(U(\Delta_{k+1})\varphi_l^m,\varphi_i^m)\int_0^{t_k}\big(U(t_k-s)QU^*(t_{k+1}-s)\varphi_j^m,\varphi_l^m\big)\,ds+\\
&+\sum_{l,p=1}^{N_m}(U(\Delta_{k+1})\varphi_l^m,\varphi_i^m)(U(\Delta_{k+1})\varphi_p^m,\varphi_j^m)\int_0^{t_k}\big(U(t_k-s)QU^*(t_k-s)\varphi_l^m,\varphi_p^m\big)\,ds
\end{aligned}$$

for all $i,j=1,\dots,N_m$ and $k=0,\dots,n-1$. We define the matrix $\Gamma_{0,l}^{m,k}$ by

$$(\Gamma_{0,l}^{m,k})_{ij}:=(U(t_k)\Gamma_0U^*(t_l)\varphi_i^m,\varphi_j^m)$$

for all $i,j=1,\dots,N_m$ and $k,l=0,\dots,n-1$. Then

$$\begin{aligned}
\operatorname{Cov}E_{k+1}^m=\Gamma_{0,k+1}^{m,k+1}-\Gamma_{0,k+1}^{m,k}(A_{k+1}^m)^T-A_{k+1}^m(\Gamma_{0,k+1}^{m,k})^T+A_{k+1}^m\Gamma_{0,k}^{m,k}(A_{k+1}^m)^T+\\
+\int_0^{t_k}Q_{k+1,k+1}^m(s)\,ds-\int_0^{t_k}Q_{k,k+1}^m(s)(A_{k+1}^m)^T\,ds+\\
-\int_0^{t_k}A_{k+1}^m(Q_{k,k+1}^m(s))^T\,ds+\int_0^{t_k}A_{k+1}^mQ_{k,k}^m(s)(A_{k+1}^m)^T\,ds
\end{aligned}\tag{15}$$

for all $k=0,\dots,n-1$ where the integration is done componentwise.

Since X_k and W_{k+1} are independent, also X_k^m and W_{k+1}^m as well as E_{k+1}^m and W_{k+1}^m are mutually independent for all $k=0,\dots,n-1$. On the other hand, X_k^m and E_{k+1}^m are correlated for all $k=0,\dots,n-1$. The correlation matrix of X_k^m and E_{k+1}^m can be calculated by using the continuous state evolution equation (4) for all $k=0,\dots,n-1$. Hence

$$\begin{aligned}
(\operatorname{Cor}(X_k^m,E_{k+1}^m))_{ij}&=(\operatorname{Cov}(X_k)\varphi_i^m,(I-P_m)U^*(\Delta_{k+1})\varphi_j^m)\\
&=(U(t_k)\Gamma_0U^*(t_k)\varphi_i^m,(I-P_m)U^*(\Delta_{k+1})\varphi_j^m)+\\
&\quad+\int_0^{t_k}\big(U(t_k-s)QU^*(t_k-s)\varphi_i^m,(I-P_m)U^*(\Delta_{k+1})\varphi_j^m\big)\,ds\\
&=(U(t_{k+1})\Gamma_0U^*(t_k)\varphi_i^m,\varphi_j^m)+\\
&\quad-\sum_{l=1}^{N_m}(U(\Delta_{k+1})\varphi_l^m,\varphi_j^m)(U(t_k)\Gamma_0U^*(t_k)\varphi_i^m,\varphi_l^m)+\\
&\quad+\int_0^{t_k}\big(U(t_{k+1}-s)QU^*(t_k-s)\varphi_i^m,\varphi_j^m\big)\,ds+\\
&\quad-\sum_{l=1}^{N_m}(U(\Delta_{k+1})\varphi_l^m,\varphi_j^m)\int_0^{t_k}(U(t_k-s)QU^*(t_k-s)\varphi_i^m,\varphi_l^m)\,ds
\end{aligned}$$

for all $i, j = 1, \ldots, N_m$ and therefore

$$\begin{aligned}
&\operatorname{Cor}(X_k^m, E_{k+1}^m) \\
&= \Gamma_{0,k}^{m,k+1} - \Gamma_{0,k}^{m,k}(A_{k+1}^m)^T + \int_0^{t_k} \left(Q_{k+1,k}^m(s) - Q_{k,k}^m(s)(A_{k+1}^m)^T\right)\, ds
\end{aligned} \tag{16}$$

for all $k = 0, \ldots, n-1$ where the integration is done componentwise.

5.2 Observation equation

The discretized observation equation is

$$Y^k = BP_m X_k + B(I - P_m)X_k + S^k = [B\varphi]X_k^m + \mathcal{E}_k^m + S^k \tag{17}$$

for all $k = 1, \ldots, n$ where $[B\varphi] := [B\varphi_1^m \ldots B\varphi_{N_m}^m]$ is the $L \times N_m$ matrix whose l^{th} column is $B\varphi_l^m$ for all $l = 1, \ldots, N_m$ and $\mathcal{E}_k^m := B(I - P_m)X_k$ represents the discretization error. We need to define the statistical quantities of the process $\{\mathcal{E}_k^m\}_{k=1}^n$.

We assume that the process $S(t)$, $t \geq 0$, is a Gaussian process independent of the process $X(t)$, $t \geq 0$. Then S^k is independent of X_k^m and $\mathcal{E}_k^m$ for all $k = 1, \ldots, n$. On the other hand, X_k^m and $\mathcal{E}_k^m$ are correlated for all $k = 1, \ldots, n$. We use our knowledge of the stochastic behaviour of the continuous state evolution equation (4). The process X_k has a Gaussian modification with mean $U(t_k)x_0$ and covariance given by Formula (5) where $t = t_k$ for all $k = 1, \ldots, n$. Hence the discretization error $\mathcal{E}_k^m$ has a Gaussian version for all $k = 1, \ldots, n$. The mean of the Gaussian version is

$$\mathbb{E}\mathcal{E}_k^m = B(I - P_m)U(t_k)x_0 = BU(t_k)x_0 + \sum_{i=1}^{N_m} (U(t_k)x_0, \varphi_i^m)B\varphi_i^m$$

and hence

$$\mathbb{E}\mathcal{E}_k^m = BU(t_k)x_0 + [B\varphi]\begin{bmatrix} (U(t_k)x_0, \varphi_1^m) \\ \vdots \\ (U(t_k)x_0, \varphi_{N_m}^m) \end{bmatrix} \tag{18}$$

for all $k = 1, \ldots, n$. The covariance matrix of the Gaussian version is

$$\begin{aligned}
\operatorname{Cov}\mathcal{E}_k^m &= B(I - P_m)\operatorname{Cov}(X_k)(I - P_m)B^* \\
&= B(I - P_m)U(t_k)\Gamma_0 U^*(t_k)(I - P_m)B^* + \\
&\quad + \int_0^{t_k} B(I - P_m)U(t_k - s)QU^*(t_k - s)(I - P_m)B^*\, ds \\
&= BU(t_k)\Gamma_0 U^*(t_k)B^* + \int_0^{t_k} BU(t_k - s)QU^*(t_k - s)B^*\, ds + \\
&\quad - BP_m U(t_k)\Gamma_0 U^*(t_k)B^* - \int_0^{t_k} BP_m U(t_k - s)QU^*(t_k - s)B^*\, ds + \\
&\quad - BU(t_k)\Gamma_0 U^*(t_k)P_m B^* - \int_0^{t_k} BU(t_k - s)QU^*(t_k - s)P_m B^*\, ds + \\
&\quad + BP_m U(t_k)\Gamma_0 U^*(t_k)P_m B^* + \int_0^{t_k} BP_m U(t_k - s)QU^*(t_k - s)P_m B^*\, ds
\end{aligned}$$

for all $k = 1, \ldots, n$. Let us denote for all $k, l = 1, \ldots, n$ by $[BU(t_k)\Gamma_0 U^*(t_l)\varphi]$ the $L \times N_m$ matrix whose i^{th} column is $BU(t_k)\Gamma_0 U(t_l)\varphi_i^m$ for all $i = 1, \ldots, N_m$ and by $[BU(t_k -$

$s)QU^*(t_l - s)\varphi]$ where $s \in [0, t_k \wedge t_l]$ the $L \times N_m$ matrix whose i^{th} column is $BU(t_k - s)QU(t_l - s)\varphi_i^m$ for all $i = 1, \ldots, N_m$. Let $f \in \mathbb{R}^L$. Then

$$\begin{aligned} BP_m U(t_k)\Gamma_0 U^*(t_k)B^* f &= \sum_{i=1}^{N_m} (B\varphi_i^m)(BU(t_k)\Gamma_0 U^*(t_k)\varphi_i^m)^T f \\ &= [B\varphi][BU(t_k)\Gamma_0 U^*(t_k)\varphi]^T f, \end{aligned}$$

$$\begin{aligned} BU(t_k)\Gamma_0 U^*(t_k)P_m B^* f &= \sum_{i=1}^{N_m} (BU(t_k)\Gamma_0 U^*(t_k)\varphi_i^m)(B\varphi_i^m)^T f \\ &= [BU(t_k)\Gamma_0 U^*(t_k)\varphi][B\varphi]^T f, \end{aligned}$$

and

$$\begin{aligned} BP_m U(t_k)\Gamma_0 U^*(t_k)P_m B^* f &= \sum_{i,j=1}^{N_m} (B\varphi_j^m)(U(t_k)\Gamma_0 U^*(t_k)\varphi_i^m, \varphi_j^m)(B\varphi_i^m)^T f \\ &= [B\varphi](\Gamma_{0,k}^{m,k})^T [B\varphi]^T f, \end{aligned}$$

for all $k = 1, \ldots, n$. Similarly,

$$\begin{aligned} \int_0^{t_k} BP_m U(t_k - s)QU^*(t_k - s)B^* \, ds &= \int_0^{t_k} [B\varphi][BU(t_k - s)QU^*(t_k - s)\varphi]^T \, ds, \\ \int_0^{t_k} BU(t_k - s)QU^*(t_k - s)P_m B^* \, ds &= \int_0^{t_k} [BU(t_k - s)QU^*(t_k - s)\varphi][B\varphi]^T \, ds, \\ \int_0^{t_k} BP_m U(t_k - s)QU^*(t_k - s)P_m B^* \, ds &= \int_0^{t_k} [B\varphi](Q_{k,k}^m(s))^T [B\varphi]^T \, ds \end{aligned}$$

for all $k = 1, \ldots, n$. Hence the covariance matrix is

$$\begin{aligned} \operatorname{Cov} \mathcal{E}_k^m = BU(t_k)\Gamma_0 U^*(t_k)B^* &+ \int_0^{t_k} BU(t_k - s)QU^*(t_k - s)B^* \, ds + \\ &- [B\varphi][BU(t_k)\Gamma_0 U^*(t_k)\varphi]^T - \int_0^{t_k} [B\varphi][BU(t_k - s)QU^*(t_k - s)\varphi]^T \, ds + \\ &- [BU(t_k)\Gamma_0 U^*(t_k)\varphi][B\varphi]^T - \int_0^{t_k} [BU(t_k - s)QU^*(t_k - s)\varphi][B\varphi]^T \, ds + \\ &+ [B\varphi](\Gamma_{0,k}^{m,k})^T [B\varphi]^T + \int_0^{t_k} [B\varphi](Q_{k,k}^m(s))^T [B\varphi]^T \, ds \end{aligned} \tag{19}$$

for all $k = 1, \ldots, n$.

The correlation matrix of X_k^m and $\mathcal{E}_k^m$ can be calculated by using the continuous state evolution equation (4) for all $k = 1, \ldots, n$. Then for all $i = 1, \ldots, N_m$, $j = 1, \ldots, L$ and $k = 1, \ldots, n$

$$\begin{aligned} (\operatorname{Cor}(X_k^m, \mathcal{E}_k^m))_{ij} &= \operatorname{Cor}((X_k, \varphi_i^m), (B(I - P_m)X_k, e_j)) \\ &= (B \operatorname{Cov}(X_k)\varphi_i^m, e_j) - (BP_m \operatorname{Cov}(X_k)\varphi_i^m, e_j) \\ &= (BU(t_k)\Gamma_0 U^*(t_k)\varphi_i^m, e_j) - (BP_m U(t_k)\Gamma_0 U^*(t_k)\varphi_i^m, e_j) + \\ &\quad + \int_0^{t_k} (BU(t_k - s)QU^*(t_k - s)\varphi_i^m, e_j) \, ds + \\ &\quad - \int_0^{t_k} (BP_m U(t_k - s)QU^*(t_k - s)\varphi_i^m, e_j) \, ds \end{aligned}$$

where e_j is the j^{th} coordinate vector in $\mathbb{R}^L$. Since for all $i = 1, \ldots, N_m$ and $j = 1, \ldots, L$

$$(BP_m U(t_k)\Gamma_0 U^*(t_k)\varphi_i^m, e_j) = \sum_{l=1}^{N_m} (U(t_k)\Gamma_0 U^*(t_k)\varphi_i^m, \varphi_l^m)(B\varphi_l^m, e_j) = (\Gamma_{0,k}^{m,k}[B\varphi]^T)_{ij}$$

and similarly,

$$\int_0^{t_k} (BP_m U(t_k - s)QU^*(t_k - s)\varphi_i^m, e_j)\, ds = \int_0^{t_k} (Q_{k,k}^m(s)[B\varphi]^T)_{ij}\, ds,$$

the correlation matrix is

$$\begin{aligned} \mathrm{Cor}(X_k^m, \mathcal{E}_k^m) = [BU(t_k)\Gamma_0 U^*(t_k)\varphi]^T + \int_0^{t_k} [BU(t_k - s)QU^*(t_k - s)\varphi]^T\, ds + \\ - \Gamma_{0,k}^{m,k}[B\varphi]^T - \int_0^{t_k} Q_{k,k}^m(s)[B\varphi]^T\, ds \end{aligned} \tag{20}$$

for all $k = 1, \ldots, n$.

6 Solution to the discretized filtering problem

The discretized state estimation system we are interested in is

$$X_{k+1}^m = A_{k+1}^m X_k^m + E_{k+1}^m + W_{k+1}^m, \quad k = 0, \ldots, n-1, \tag{21}$$

$$Y^k = [B\varphi]X_k^m + \mathcal{E}_k^m + S^k, \qquad k = 1, \ldots, n. \tag{22}$$

The state noise vectors W_k^m and W_l^m are mutually independent and also independent of X_0^m for all $k \neq l$. We assume that the observation noise vectors S^k are chosen such a way that S^k and S^l are mutually independent and also independent of X_0^m for all $k \neq l$ and S^k and W_l^m are mutually independent for all $k, l = 1, \ldots, n$. Usually, the filtering problem concerning a finite dimensional state estimation system where the discrete stochastic processes are Gaussian is solved by using the Kalman filtering method. In the proof of the formulas of the Kalman filter and predictor [3, Theorem 4,3] it is assumed that the noise terms in the state evolution and observation equations are mutually independent. In our case, the random variables $E_{k+1}^m + W_{k+1}^m$ and $\mathcal{E}_{k+1}^m + S^{k+1}$ are dependent for all $k = 0, \ldots, n-1$. Hence we cannot use the Kalman filtering method. However, by the Gaussianity we are able to solve the filtering problem.

Lemma 2. *The joint probability density of X_k^m, Y^k,..., and Y^1 is Gaussian for all $k = 1, \ldots, n$.*

Proof. If we are able to show that the sum

$$\sum_{i=1}^{N_m} \lambda_i (X_k^m)_i + \sum_{l=1}^{k} \sum_{i=1}^{L} \lambda_{N_m + (k-l)L + i} (Y^l)_i$$

where $k \in \{1, \ldots, n\}$ is fixed and $\lambda_i \in \mathbb{R}$ for all $i = 1, \ldots, N_m + kL$ is a Gaussian random variable, the joint probability distribution of X_k^m, Y^k,..., and Y^1 is Gaussian for all $k = 1, \ldots, n$ by [5, Theorem A.5]. Let $k \in \{1, \ldots, n\}$ be fixed and $\lambda_i \in \mathbb{R}$ for $i = 1, \ldots, N_m + kL$ be arbitrary. Since for all $l = 1, \ldots, k$

$$X_l = U(t_l - t_1)X_1 + \sum_{p=1}^{l} U(t_l - t_p)W_p$$

by [6, Theorem 4.39],

$$\sum_{i=1}^{N_m} \lambda_i (X_k, \varphi_i^m) + \sum_{l=1}^{k}\sum_{i=1}^{L} \lambda_{N_m+(k-l)L+i}(Y^l, e_i)$$
$$= \left(X_k, \sum_{i=1}^{N_m} \lambda_i \varphi_i^m \right) + \sum_{l=1}^{k} \left(X^l, \sum_{i=1}^{L} \lambda_{N_m+(k-l)L+i} B^* e_i \right) +$$
$$+ \sum_{l=1}^{k} \left(S^l, \sum_{i=1}^{L} \lambda_{N_m+(k-l)L+i} e_i \right)$$
$$= \left(X_1, \sum_{i=1}^{N_m} \lambda_i U^*(t_k - t_1)\varphi_i^m + \sum_{l=1}^{k}\sum_{i=1}^{L} \lambda_{N_m+(k-l)L+i} U^*(t_l - t_1) B^* e_i \right) +$$
$$+ \sum_{l=2}^{k} \left(W_l, \sum_{i=1}^{N_m} \lambda_i U^*(t_k - t_l)\varphi_i^m + \sum_{p=l}^{k}\sum_{i=1}^{L} \lambda_{N_m+(k-p)L+i} U^*(t_p - t_l) B^* e_i \right) +$$
$$+ \sum_{l=1}^{k} \left(S^l, \sum_{i=1}^{L} \lambda_{N_m+(k-l)L+i} e_i \right).$$

As a sum of independent real valued Gaussian random variables the sum

$$\sum_{i=1}^{N_m} \lambda_i (X_k^m)_i + \sum_{l=1}^{k}\sum_{i=1}^{L} \lambda_{N_m+(k-l)L+i}(Y^l)_i$$

is a Gaussian random variable. □

We denote the measured values of Y^1,..., and Y^k by y_1,..., and y_k for all $k = 1, \ldots, n$, respectively. We want to calculate the probability distribution of X_k^m conditioned on the measurement $\{y_1, \ldots, y_k\}$, i.e., the conditional probability distribution $\pi(x_k|y_k, \ldots, y_1)$ for all $k = 1, \ldots, n$. By the Gaussianity of the joint probability distribution of X_k^m, Y^k,..., and Y^1 for all $k = 1, \ldots, n$ it can be done. We need to know the expectation and covariance matrix of the joint probability distribution. They can be calculated by using our knowledge of the stochastic behaviour of the continuous state evolution equation (4). The process X_k has a Gaussian modification with mean $U(t_k)x_0$ and covariance given by Formula (5) where $t = t_k$ for all $k = 0, \ldots, n$. Thus

$$\mathbb{E}X_k^m = ((U(t_k)x_0, \varphi_1^m), \ldots, (U(t_k)x_0, \varphi_{N_m}^m))^T, \tag{23}$$
$$\mathbb{E}Y^k = \mathbb{E}(BX_k + S^k) = BU(t_k)x_0 + \mathbb{E}S^k \tag{24}$$

for all $k = 1, \ldots, n$. On the other hand, for all $i, j = 1, \ldots, N_m$ and $k = 1, \ldots, n$

$$(\operatorname{Cov} X_k^m)_{ij} = (U(t_k)\Gamma_0 U^*(t_k)\varphi_i^m, \varphi_j^m) + \int_0^{t_k} \left(U(t_k - s)QU^*(t_k - s)\varphi_i^m, \varphi_j^m\right) ds.$$

By the notation introduced earlier,

$$\operatorname{Cov} X_k^m = \Gamma_{0,k}^{m,k} + \int_0^{t_k} Q_{k,k}^m(s)\, ds \tag{25}$$

for all $k = 1, \ldots, n$. Since the process $S(t)$, $t \geq 0$, is independent of the process $X(t)$, $t \geq 0$,

$$\begin{aligned}\operatorname{Cov} Y^k &= \operatorname{Cov}(BX_k + S^k) = B \operatorname{Cov}(X_k) B^* + \operatorname{Cov} S^k \\ &= BU(t_k)\Gamma_0 U^*(t_k)B^* + \int_0^{t_k} BU(t_k - s)QU^*(t_k - s)B^*\, ds + \operatorname{Cov} S^k\end{aligned} \tag{26}$$

for all $k = 1, \ldots, n$.

To be able to calculate the cross correlation matrices we need the correlation operator of X_k and X_l for all $k = 1, \ldots, n$ and $l < k$. According to the continuous state evolution equation (4),

$$X_k = U(t_k - t_l)X_l + \int_{t_l}^{t_k} U(t_k - s)\, dW(s)$$

for all $k = 1, \ldots, n$ and $l < k$. Since $\int_{t_l}^{t_k} U(t_k - s)\, dW(s)$ is independent of $\mathcal{F}_{t_l}$ and X_l is $\mathcal{F}_{t_l}$-measurable,

$$\operatorname{Cor}(X_k, X_l) = U(t_k - t_l) \operatorname{Cov} X_l = U(t_k)\Gamma_0 U^*(t_l) + \int_0^{t_l} U(t_k - s)QU^*(t_l - s)\, ds$$

for all $k = 1, \ldots, n$ and $l < k$. Thus for all $i, j = 1, \ldots, N_m$

$$(\operatorname{Cor}(X_k^m, X_l^m))_{ij} = (U(t_k)\Gamma_0 U^*(t_l)\varphi_i^m, \varphi_j^m) + \int_0^{t_l} (U(t_k - s)QU^*(t_l - s)\varphi_i^m, \varphi_j^m)\, ds$$

and hence

$$\operatorname{Cor}(X_k^m, X_l^m) = \Gamma_{0,l}^{m,k} + \int_0^{t_l} Q_{k,l}^m(s)\, ds$$

for all $k = 1, \ldots, n$ and $l < k$. Similarly as the correlation matrix (20) of X_k^m and $\mathcal{E}_k^m$,

$$\begin{aligned}\operatorname{Cor}(X_k^m, \mathcal{E}_l^m) = [BU(t_k)\Gamma_0 U^*(t_l)\varphi]^T + \int_0^{t_l} [BU(t_k - s)QU^*(t_l - s)\varphi]^T\, ds + \\ - \Gamma_{0,l}^{m,k}[B\varphi]^T - \int_0^{t_l} Q_{k,l}^m(s)[B\varphi]^T\, ds\end{aligned}$$

for all $k = 1, \ldots, n$ and $l < k$. Therefore the correlation matrix of X_k^m and Y^l is

$$\begin{aligned}\operatorname{Cor}(X_k^m, Y^l) &= \operatorname{Cor}(X_k^m, [B\varphi]X_l^m + \mathcal{E}_l^m + S^l) \\ &= \operatorname{Cor}(X_k^m, X_l^m)[B\varphi]^T + \operatorname{Cor}(X_k^m, \mathcal{E}_l^m) \\ &= [BU(t_k)\Gamma_0 U^*(t_l)\varphi]^T + \int_0^{t_l} [BU(t_k - s)QU^*(t_l - s)\varphi]^T\, ds\end{aligned} \tag{27}$$

for all $k = 1, \ldots, n$ and $l < k$. The correlation matrix of Y^l and Y^k is

$$\begin{aligned}\operatorname{Cor}(Y^k, Y^l) &= \operatorname{Cor}(BX_k + S^k, BX_l + S^l) = B \operatorname{Cor}(X_k, X_l)B^* \\ &= BU(t_k)\Gamma_0 U^*(t_l)B^* + \int_0^{t_l} BU(t_k - s)QU^*(t_l - s)B^*\, ds\end{aligned} \tag{28}$$

for all $k = 1, \ldots, n$ and $l < k$.

The solution to the filtering problem we are interested in is given by the following theorem.

Theorem 3. [3, Theorem 3.5] *Let $X : \Omega \to \mathbb{R}^n$ and $Y : \Omega \to \mathbb{R}^m$ be two Gaussian random variables whose joint probability density $\pi : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}_+$ is of the form*

$$\pi(x, y) \propto \exp\left(-\frac{1}{2}\begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix}^T \begin{bmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{bmatrix}^{-1} \begin{bmatrix} x - x_0 \\ y - y_0 \end{bmatrix}\right)$$

where $x_0 \in \mathbb{R}^n$, $y_0 \in \mathbb{R}^m$, $\Gamma_{11} \in \mathbb{R}^{n\times n}$, $\Gamma_{22} \in \mathbb{R}^{m\times m}$ *and*

$$\Gamma = \begin{bmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{bmatrix}$$

is a positive definite symmetric $(n+m)\times(n+m)$ *matrix. Then the probability distribution of* X *conditioned on* $Y = y$ *is*

$$\pi(x \mid y) \propto \exp\left(-\frac{1}{2}(x-\bar{x})^T \widetilde{\Gamma}_{22}^{-1}(x-\bar{x})\right)$$

where $\bar{x} = x_0 + \Gamma_{12}\Gamma_{22}^{-1}(y - y_0)$ *and* $\widetilde{\Gamma}_{22} = \Gamma_{11} - \Gamma_{12}\Gamma_{22}^{-1}\Gamma_{21}$.

The joint probability distribution of X_k^m, Y^k,..., and Y^1 is Gaussian and we are able to calculate the expectation and correlation matrix. Hence for all $k = 1, \ldots, n$

$$\pi(x_k \mid y_k, \ldots, y_1) \propto \exp\left(-\frac{1}{2}(x_k - \bar{x}_k)^T \Gamma_k^{-1}(x_k - \bar{x}_k)\right) \tag{29}$$

where

$$\bar{x}_k = \mathbb{E}X_k^m + \mathrm{Cor}\left(X_k^m, \begin{bmatrix} Y^k \\ \vdots \\ Y^1 \end{bmatrix}\right) \mathrm{Cov}\left(\begin{bmatrix} Y^k \\ \vdots \\ Y^1 \end{bmatrix}\right)^{-1} \left(\begin{bmatrix} y_k \\ \vdots \\ y_1 \end{bmatrix} - \begin{bmatrix} \mathbb{E}Y^k \\ \vdots \\ \mathbb{E}Y^1 \end{bmatrix}\right) \tag{30}$$

and

$$\Gamma_k = \mathrm{Cov}\, X_k^m - \mathrm{Cor}\left(X_k^m, \begin{bmatrix} Y^k \\ \vdots \\ Y^1 \end{bmatrix}\right) \mathrm{Cov}\left(\begin{bmatrix} Y^k \\ \vdots \\ Y^1 \end{bmatrix}\right)^{-1} \mathrm{Cor}\left(\begin{bmatrix} Y^k \\ \vdots \\ Y^1 \end{bmatrix}, X_k^m\right). \tag{31}$$

The evaluation of the matrices needed in the filtering depends on the discretization space V_m, analytic semigroup $\{U(t)\}_{t\geq 0}$, element x_0, operators Γ_0, Q and B and statistics of the observation noise S. Usually the discretization space V_m is chosen such a way that the projection P_m is fairly easy to calculate, i.e., the inner products (f, φ_i^m) for $i = 1, \ldots, N_m$ can be easily calculated for $f \in H$. The element x_0 and operator Γ_0 represent our prior knowledge of the parameter X. The mean x_0 should illustrate the expected values and hence it depends heavily on the application. Our confidence of the expected value x_0 is indicated by the covariance operator Γ_0. Our certainty of the time evolution model is coded into the Wiener process and hence into the operator Q. The choice of Q depends on the application. The crucial factor in the evaluation of the matrices is the analytic semigroup $\{U(t)\}_{t\geq 0}$. Since it is defined as a Bochner integral of the resolvent operator of the operator A along a curve in the resolvent set [4, (2.0.2) p. 33], only in some special cases we can present the analytic semigroup in a closed form. The operator B is related to the measurement situation. The observation noise S represents the accuracy of the measurement equipment.

7 Conclusions

The examination of the continuous infinite dimensional state evolution equation is beneficial for solving nonstationary inverse problems by the state estimation method. The numerical computation requires space discretization. The discretization error can be considered only by knowing the stochastic nature of the time evolution of the object of interest. The knowledge of the continuous infinite dimensional state evolution equation of the quantity of primary interest allows us to calculate the probability distribution of the discretization errors both in the state evolution and observation equation.

The solution (29)–(31) to the discretized filtering problem is valid only in the Gaussian case. The assumption of Gaussianity seemed to be natural since the solution to the infinite dimensional state evolution equation is a Gaussian process if the initial value is assumed be a Gaussian random variable. Despite of the initial value the state noise is always a Gaussian process. Nonetheless, in some application non-Gaussian initial values may be reasonable.

Acknowledgements. This work has been supported by the Vilho, Yrjö and Kalle Väisälä Foundation of the Finnish Academy of Science and Letters, the Finnish Graduate School of Inverse Problems and the Austrian National Science Foundation FWF through the project SFB F1308. The author would like to thank Prof. Erkki Somersalo, Helsinki University of Technology, for his advice.

References

[1] Djebar Baroudi, Jari Kaipio, and Erkki Somersalo. Dynamical electric wire tomography: a time series approach. *Inverse Problems*, 14:799–813, 1998.

[2] J. Kaipio and E. Somersalo. Nonstationary inverse problems and state estimation. *Journal of Inverse and Ill-Posed Problems*, 7(3):273–282, 1999.

[3] Jari P. Kaipio and Erkki Somersalo. *Statistical and Computational Inverse Problems*, volume 160 of *Applied Mathematical Sciences*. Springer-Verlag, Berlin, Germany, 2004. ISBN 0-387-22073-9.

[4] Alessandra Lunardi. *Analytic Semigroups and Optimal Regularity in Parabolic Problems*, volume 16 of *Progress in Nonlinear Differential Equations and Their Applications*. Birkhäuser, Basel, Switzerland, 1995. ISBN 3-7643-5172-1.

[5] Bernt Øksendal. *Stochastic Differential Equations — An Introduction with Applications*. Springer-Verlag, fifth edition, 1998. ISBN 3-540-63720-6.

[6] Hanna Katriina Pikkarainen. *A Mathematical Model for Electrical Impedance Process Tomography*. Doctoral dissertation, Helsinki University of Technology, Espoo, Finland, 2005. ISBN 951-22-7651-8. http://lib.tkk.fi/Diss/2005/isbn9512276720/.

[7] Giuseppe Da Prato and Jerzy Zabczyk. *Stochastic Equations in Infinite Dimensions*, volume 44 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, United Kingdom, 1992. ISBN 0-521-38529-6.

[8] A.R. Ruuskanen, A. Seppänen, S. Duncan, E. Somersalo, and J.P. Kaipio. Optimal control in process tomography. In *Proceedings of 3rd World Congress on Industrial Process Tomography*, pages 245–251. The Virtual Centre for Industrial Process Tomography, 2003. ISBN 0-853-16240-9.

[9] A.R. Ruuskanen, A. Seppänen, S. Duncan, E. Somersalo, and J.P. Kaipio. Using tomographic measurements in process control. In *43rd IEEE Conference on Decision and Control*, 2004.

[10] A.R. Ruuskanen, A. Seppänen, S. Duncan, E. Somersalo, and J.P. Kaipio. Using Process Tomography as a Sensor for Optimal Control. *Applied Numerical Mathematics*, (in press).

[11] A.R. Ruuskanen, A. Seppänen, E. Somersalo, and J.P. Kaipio. The effect of linearizing the EIT observation model in concentration control problems. In *Proceedings of 4th World Congress on Industrial Process Tomography*, pages 873–878. The Virtual Centre for Industrial Process Tomography, 2005.

[12] A. Seppänen, L. Heikkinen, T. Savolainen, E. Somersalo, and J.P. Kaipio. An experimental evaluation of state estimation with fluid dynamical models in process tomography. In *Proceedings of 3rd World Congress on Industrial Process Tomography*, pages 541–546. The Virtual Centre for Industrial Process Tomography, 2003. ISBN 0-853-16240-9.

[13] A. Seppänen, M. Vauhkonen, E. Somersalo, and J.P. Kaipio. Effects of inaccuracies in fluid dynamical models in state estimation of process tomography. In *Process Imaging for Automatic Control*, volume 4188 of *Proceedings of SPIE*, pages 69–80, 2001.

[14] A. Seppänen, M. Vauhkonen, E. Somersalo, and J.P. Kaipio. State space models in process tomography – approximation of state noise covariance. *Inverse Problems in Engineering*, 9:561–585, 2001.

[15] A. Seppänen, M. Vauhkonen, E. Somersalo, and J.P. Kaipio. Inference of velocity fields based on tomographic measurements in process industry. In *4th International Conference on Inverse Problems in Engineering*, pages 453–460, 2002.

[16] A. Seppänen, M. Vauhkonen, P.J. Vauhkonen, E. Somersalo, and J.P. Kaipio. Fluid dynamical models and state estimation in process tomography: Effect due to inaccuracies in flow fields. *Journal of Electronic Imaging*, 10(3):630–640, 2001.

[17] A. Seppänen, M. Vauhkonen, P.J. Vauhkonen, E. Somersalo, and J.P. Kaipio. State estimation in three dimensional impedance imaging – Use of fluid dynamical evolution models. In *Proceedings of 2nd World Congress on Industrial Process Tomography*, pages 198–207. The Virtual Centre for Industrial Process Tomography, 2001.

[18] A. Seppänen, M. Vauhkonen, P.J. Vauhkonen, E. Somersalo, and J.P. Kaipio. State estimation with fluid dynamical evolution models in process tomography – an application with impedance tomography. *Inverse Problems*, 17:467–484, 2001.

[19] Aku Seppänen. *State Estimation in Process Tomography.* PhD thesis, University of Kuopio, Kuopio, Finland, 2005. ISBN 951-27-0343-2.